

# Chutes or Ladders?

## A Longitudinal Analysis of Immigrant Earnings

Darren Lubotsky\*  
339 Wallace Hall  
Research Program in Development Studies  
Woodrow Wilson School of Public and International Affairs  
Princeton University  
Princeton, NJ 08544  
*lubotsky@princeton.edu*

August, 2000

---

\*I especially wish to thank David Card for his advice and encouragement. I also thank Alan Auerbach, Richard Blundell, Kenneth Chay, Hilary Hoynes, Jonathan Leonard, Enrico Moretti, James Powell, and participants in several seminars for their advice; and the staff of the California Census Research Data Center (CCRDC) for their help in facilitating my use of the confidential data. Financial support from the Robert Burch Center for Tax Policy Analysis, the Center for Labor Economics, and the Institute for Industrial Relations at U.C. Berkeley; and the CCRDC is gratefully acknowledged. The research in this paper was conducted while I was a research associate at the CCRDC in association with the Center for Economic Studies of the U.S. Bureau of the Census. Research results and conclusions are my own and do not necessarily indicate concurrence by the Bureau of the Census, the Center for Economic Studies, or the CCRDC.

## Abstract

This study uses Social Security earnings records matched to recent cross-sections of the SIPP and CPS to study the earnings progress of U.S. immigrants. The data show that immigrants' earnings grow 10 to 13 percent during their first twenty years in the U.S. relative to the earnings of natives with similar labor market experience. By comparison, estimates of immigrants' relative wage growth from cross-sections of the decennial Census are substantially higher. The divergent results reflect the selective outmigration of low-earning immigrants. The longitudinal earnings histories also show that 14 percent of immigrants have earnings in the U.S. prior to their most recent date of arrival, which points to a significant amount of back-and-forth migration between the U.S. and immigrants' home countries. The misclassification in previous work of these largely low-wage immigrants as recent arrivals accounts for close to one-third of the measured decline in the level of earnings of immigrant arrival cohorts between 1960 and 1980. The new evidence presented here, therefore, suggests that previous analyses had overestimated both the rate of earnings growth among immigrants who remain in the U.S. and the secular decline in the level of earnings across arrival cohorts.

*JEL Classifications:* C24, J1, J31, J61.

# 1 Introduction

The relative earnings of newly arrived immigrant cohorts to the United States have steadily declined over the past thirty years. In the 1970 Census, for example, immigrants who had been in the country for less than five years earned 39 percent less than similarly aged native-born workers; in the 1990 Census, recent immigrants earned 55 percent less than natives.<sup>1</sup> Most researchers agree that shifts in the national origin mix of immigrants away from developed, high-skill countries to industrializing, lower-skill countries contributed to the decline in the average labor market earnings of immigrants over this period. Beyond this, however, there is sharp disagreement over the experience of immigrants in the U.S. labor market. In one view, immigrants quickly develop English language and other skills necessary to move up the American earnings distribution. Thus, despite their low initial earnings, immigrants quickly assimilate into the U.S. labor market. Other researchers take a more pessimistic view and argue that immigrants – particularly recent arrivals – tend to earn significantly less than natives throughout their working life, and thus do not assimilate in any meaningful way.<sup>2</sup>

Recent U.S. immigration policy has been influenced by the latter view. Based both on concerns over the fiscal impacts of lower-wage immigrants who pay little in taxes yet draw heavily on public transfers and services, and on fears that inflows of unskilled immigrants lower the earnings and employment rate of native-born workers, the Immigration Act of 1990 increased the number of visas allocated on the basis of occupational skills from 54,000 to 140,000 per year, and restricted the number of visas for unskilled workers to 10,000 per year.<sup>3</sup> Recent welfare reform legislation has similarly sought to discourage immigrants from migrating in order to receive benefits by severely restricting the ability of new immigrants to obtain cash transfers, food stamps, and Medicaid during their first five years in the U.S.

The evidence on the economic progress of immigrants is, however, decidedly mixed.<sup>4</sup>

---

<sup>1</sup>These figures are computed from cross-sectional regressions of the log of annual wage and salary, self-employment, and farm earnings on a quartic in potential experience and an indicator for immigrants. The sample includes men aged 25 to 64, and excludes immigrants who entered before the age of 18.

<sup>2</sup>For a statement of the former view see, for example, LaLonde and Topel (1992) and Duleep and Regets (1996, 1997). The latter view is expressed by Borjas (1985, 1995).

<sup>3</sup>The Act also provided for 480,000 visas allocated on the basis of family reunification and 55,000 “diversity visas,” which were given to immigrants from countries underrepresented in the 1965 visa allocation. Jasso and Rosenzweig (1990) and Smith and Edmonston, eds (1997) provide summaries of the history of U.S. immigration policy; for additional details on the 1986 and 1990 immigration laws see Bean and Fix (1992).

<sup>4</sup>Surveys of the economic literature on immigration are given in Borjas (1994, 1999), LaLonde and Topel (1997), and Smith and Edmonston, eds (1997).

The earnings gap between native-born workers and a particular arrival cohort of immigrants narrows sharply from one decennial Census to the next. For example, in the 1990 Census, immigrants who arrived between 1965 and 1969 had earnings equal to those of natives. However, there are problems with this type of repeated cross-sectional analysis that previous work has been unable to address: first, a sizable fraction of immigrants eventually leave the United States. Thus, the sample of immigrants from an arrival cohort changes across Census datasets.<sup>5</sup> To the extent that those who choose to leave the U.S. have systematically better or worse labor market outcomes than those who remain, immigrant earnings growth measured across Censuses reflects both increased earnings among those who stay, as well as compositional changes brought about by selective outmigration. In particular, if immigrants with below-average earnings are more likely to emigrate, then the average earnings of the group who remain in the U.S. will tend increase over time.

Related to this effect, some immigrants leave the U.S. and then reenter at a later date. The Census, however, asks respondents when they arrived in the U.S. “to stay,” and many immigrants appear to answer with the date of their most recent arrival. If the group of transient immigrants tends to have low earnings (for example, migrant laborers or agricultural workers from Mexico), then recently enumerated immigrants in the Census will appear to have lower earnings than earlier immigrants. In addition, because the reentrants may have already assimilated into the U.S. labor market when they arrived the first time, their measured earnings growth may appear to be slower than that of previous immigrant arrival cohorts. Therefore, because of both permanent and temporary outmigration, immigrant earnings growth measured in the Census may give a misleading impression of the actual earnings trajectories of immigrants who stay in the United States.

This paper uses a new sample of longitudinal earnings histories that help overcome the limitations of the previous literature on immigrant earnings assimilation. Through a joint project of the Social Security Administration and the Census Bureau, the 1990 and 1991

---

<sup>5</sup>Direct evidence is not available on the number of foreign- or American-born people who emigrate from the U.S., and thus estimates of the rate emigration have to be inferred from other data sources. Jasso and Rosenzweig (1990) use annual counts of resident aliens from the Immigration and Naturalization Service, and estimate that the number of immigrants who left the U.S. between 1960 and 1980 was 41 percent of the number of new immigrants during that period. Borjas and Bratsberg (1996) compare INS data on immigrant inflows with population estimates from the 1980 Census and conclude that about 20 percent of immigrants who arrived in the 1970’s had left the U.S. by 1980. Both studies find that Asian immigrants are less likely to emigrate from the U.S. than are European and Latin American immigrants. In addition to outmigration, changes in the coverage of illegal immigrants may lead to differences in the sample of immigrants across Censuses.

Survey of Income and Program Participation (SIPP) and the 1994 March Supplement to the Current Population Survey (CPS) have been matched to annual earnings records from 1951 to 1997. The first contribution of this work is to compare the implied earnings growth of immigrants found in repeated cross-sections of the decennial Census with estimates from this fixed panel of individuals that is free of any bias caused by nonrandom emigration out of the U.S.<sup>6</sup> In addition, since immigrants' reported date of arrival in the CPS and SIPP can be compared to their first year of covered earnings, the extent of temporary outmigration and its effect on measured immigrant earnings profiles can be examined.

Several important new results are found. Most importantly, the actual earnings growth of immigrants who remain in the U.S. is considerably slower than that implied by comparisons across decennial Censuses. For example, estimates from the longitudinal earnings records indicate that immigrant earnings grow by about 10 percentage points more over their first twenty years in the U.S. than the earnings growth experienced by native-born workers. This is substantially slower than the 25 percentage point growth implied by a comparison of immigrants and natives across decennial Censuses. The temporary outmigration of some immigrants also has a significant effects on measures of the trend in earnings levels across successive immigrant arrival cohorts. In particular, the decline in the level of earnings between 1960 and 1980 arrivals is approximately one-third smaller when immigrants are classified by their initial date of arrival, rather than their reported date.

The next section of the paper describes the matched Social Security earnings data and highlights some of their strengths and weaknesses. Section 3 describes the relationship between immigrant earnings measured in repeated cross-sections from the decennial Census and in longitudinal data, in the presence of permanent and temporary selective outmigration. Section 4 provides evidence on the extent to which estimates from the Census overstate the relative growth of immigrant earnings and compares the earnings of immigrants who arrived in the 1960's, 1970's, and from 1980 to 1994. Finally, Section 5 concludes.

---

<sup>6</sup>Several previous studies have used longitudinal data to examine immigrant earnings. See Chiswick (1980), Borjas (1989), Duleep (1999), and Hu (1999). With the exception of the work by Duleep, these studies focus primarily on earlier immigrant cohorts than those examined here. None of these studies consider the effect of temporary migration on estimates of immigrants' relative earnings.

## 2 How Comparable are Matched Administrative Earnings Data and Household Surveys?

The new data sources used in this analysis are the 1990 and 1991 Survey of Income and Program Participation linked by individuals' Social Security number to annual Social Security earnings records from 1951 to 1993, and the March Supplement to the 1994 Current Population Survey linked to earnings records from 1951 to 1997.<sup>7</sup> Only men born between 1930 and 1969 are included in the sample. Since most men work, the labor force participation decision does not pose as significant an issue as it would for an analysis of women. Men born before 1930 would be over 60 years old at the time of the 1990, 1991, or 1994 surveys; thus there is a risk that nonrandom mortality would bias the sample of older workers in favor of the more healthy. Those born after 1969 would be under 25 years old at the time of the 1990, 1991, and 1994 cross-sectional surveys, and may have not completed their schooling. Finally, a small number of people whose reported gender in the cross-sectional survey does not match that in the Social Security record or whose year of birth differed by more than two years in the two sources are not included in this analysis. Thus, an illegal immigrant who bought a Social Security card on the black market would not be in this sample unless the gender and birth year associated with the Social Security number matched their own.

The Social Security records contain longitudinal information on an individual's annual earnings and the number of quarters of covered employment in the year. Other information, such as educational attainment and the date and place of birth, are taken from the CPS or SIPP cross-sectional surveys.<sup>8</sup> Individuals born outside of the United States are classified as

---

<sup>7</sup>The earnings records are confidential and are used through an arrangement with the Center for Economic Studies of the U.S. Census Bureau. The earnings data are from reports made by employers to the Social Security Administration for the purpose of assessing Social Security taxes on employees. Though nearly all workers are covered by the system today, some groups were not covered throughout the entire sample period of this study. In particular, most self-employed professionals and members of the uniformed services entered the Social Security system between 1954 and 1956; and employees of the federal government hired before 1984 had the option of participating in the system. Although self-employed individuals are covered by the system, it is not known how much of their income goes unreported to the Social Security Administration. A short history of key provisions of the Social Security law is given in Social Security Administration (1997). Studies by Card and Krueger (1993) and Chay (1995) used a similar match of the 1973 and 1978 March CPS's to Social Security earnings records to examine black-white earnings differences. Bound and Krueger (1991) use that data to investigate the extent of measurement error in reported earnings in the CPS.

<sup>8</sup>Longitudinal information is not available on whether the individual was self-employed, how many weeks or hours they worked during the year, where they lived, and when and where they obtained schooling or job training. Thus, although most previous work analyses hourly or weekly earnings, this study can only examine annual earnings. Hu (1999) uses Social Security earnings records matched to the Health and Retirement Survey and notes that the use of annual earnings may be the more relevant measure for policy applications

immigrants. People born abroad to American parents, born in Puerto Rico or other outlying areas of the U.S., or who arrived in the U.S. prior to age 18 are not included in the sample. (Immigrants who arrived as children and attended school in the United States are likely to assimilate to a large extent prior to entering the labor market. Thus their labor market experience may be more similar to native-born workers than to immigrants who arrive in the U.S. as adults.) Immigrants' date of arrival reported in the CPS or SIPP is given by an interval. The 1994 CPS intervals are pre-1950, 1950-59, five-year intervals from 1960-64 to 1975-79, and then two-year intervals from 1980-81 to 1992-93. The intervals in the 1990 and 1991 SIPP are pre-1960, five-year intervals from 1960-64 to 1975-79, 1980-81, 1982-1984, and 1985-1990 or 1991.<sup>9</sup>

Several features of the matched data pose additional issues. The first problem is that the Social Security earnings are censored at the taxable earnings ceiling in each year.<sup>10</sup> Figure 1 plots the real value of the topcode (in 1997 dollars) and the fraction of the sample topcoded from 1951 to 1997. From 1951 to 1965, average earnings increased while the tax ceiling remained relatively stable, leading to a steady rise in the proportion of the sample that is censored from 11 percent to 53 percent. Between 1966 and 1977, as the tax ceiling was increased several times and inflation eroded its real value, the fraction of the sample topcoded was not stable. Finally, between 1978 to 1997 the real value of the tax ceiling steadily increased and, thus, from 1983 to 1997 only about 11 percent of the sample is topcoded. The econometric procedures used below take into account the censored nature of the data.

The second problem with the Social Security earnings data is that only an individual's annual earnings that are covered by the the Social Security system are recorded. An earnings record of zero dollars in a given year reflects someone who was not employed that year, was outside of the United States, or whose only earnings were from informal or other uncovered and analyses of overall well-being.

---

<sup>9</sup>Much of the analysis that follows consolidates some of the immigrant arrival cohorts. Except in Table 1 below, where the CPS data is analyzed separately, the 1980-84 cohort used in this study includes observations from the SIPP's in which the individual reported to have arrived between 1980 and 1984, and CPS observations in which the individual reported to have arrived between 1980 and 1985. The exclusion of immigrants who were under 18 years of age at the time of arrival is based on their age at the midpoint of the arrival interval.

<sup>10</sup>A small number of observations from the two SIPP sources are above the tax ceiling in a few years. This may have arisen from people working two jobs during the year and the second employer over-withholding income for Social Security taxes. Since the reported earnings may still be censored, though at a higher level, earnings for these observations are replaced with the taxable maximum in that year.

employment. Some immigrants (and natives) work entirely in the uncovered sector and thus will not have any Social Security earnings, though they may report their uncovered earnings in the Census, CPS, or SIPP surveys.<sup>11</sup> Other workers may have earnings in both the covered and uncovered sector, and thus their earnings in the longitudinal dataset is only a portion of their total earnings.

The final issue is that not all respondents in the three cross-sectional datasets are matched to earnings records. The primary reason for this is that an individual refused to give their Social Security number – or gave the wrong number – to the survey interviewer. The match rate for the native born population is 84 percent in the CPS, and 91 and 87 percent in the two SIPP’s. The proportion of immigrants who are matched is lower, particularly among recent arrivals. In the CPS sample, 76 percent of immigrants who reported arriving in the U.S. between 1950 and 1964 are matched to earnings records. The match rates are 72 percent among those report arriving between 1965 to 1979, and 62 percent among those who report arriving between 1980 and 1994. The match rate for immigrants in the SIPP samples is 87 percent among those who arrived before 1980, and 75 percent among those who arrived between 1980 and 1991.

To partially correct for potential selection bias induced by non-random matches to Social Security earnings data, population weights are computed for the matched subsample to reflect the observable characteristics of the full cross-sectional population. Specifically, let  $p(x)$  denote the probability that a person with characteristics  $x$  is matched to an earnings record. If that observation is matched and used in the analysis, their weight is given by the inverse of the estimated match probability,  $1/\hat{p}(x)$ . The probability of a match is estimated with a logit model that includes controls for educational attainment, a square in potential labor market experience, weeks and hours worked, a square in reported earnings (in the CPS or SIPP), and indicators indicators for Hispanics, nonwhites, Hispanic nonwhites, Asians, as well as for those who did not work in the survey year or were self-employed, worked in agriculture, for the government, or in the private sector.<sup>12</sup> Since earnings reported in the

---

<sup>11</sup>It is not possible to distinguish between immigrants who are legal residents of the U.S. and work in jobs not covered by the Social Security system, and immigrants in the U.S. legally or illegally who work “under the table” and do pay taxes on their earnings.

<sup>12</sup>The logit model is estimated separately for people in the three cross-sectional datasets, and for natives and five-year immigrant arrival cohorts. Some of the variable definitions differ in the CPS and SIPP samples. For example, the CPS model is based on earnings, weeks, and hours in the past year (i.e. 1993); while the SIPP variables refer to the month prior to the interview. Since the three cross-sectional surveys already have population weights, the combined weight to correct for nonmatches is given by the product of the population

cross-sectional survey are used to construct the match probabilities, unobservable factors (such as motivation or ability) that are correlated with earnings and may be more prevalent among the matched subsample are, to an extent, incorporated into the weights.

To gauge how well covered Social Security earnings reflect the more familiar earnings reported in the CPS, Table 1 compares the full sample of workers in the March Supplement to the 1994 CPS (which records total earnings from 1993) with the 1993 Social Security earnings record. The first two columns display the average reported (log) earnings among the full sample and the subsample that is matched to Social Security earnings records, as well as the earnings gap between each immigrant cohort and natives, the standard error of the earnings gap, and the (unweighted) sample size.<sup>13</sup>

Using CPS earnings data, differences in reported earnings between natives and immigrants who arrived after 1980 are eight to thirteen percentage points smaller among the matched subsample than among the full CPS sample. However, when the matched immigrants are reweighted to reflect the observable characteristics of the full CPS sample (column three), the earnings of the matched sample are fairly similar to those of the full sample. An exception is the group of immigrants who arrived between 1970 and 1974, for whom the earnings gap among the full sample is 19.6 percent and among the reweighted matched subsample is only 6.6 percent.

The fourth column reports the earnings and earnings gaps based on the 1993 Social Security earnings record for the matched subsample. About 6.5 percent fewer natives have Social Security earnings than report earnings in the CPS ( $= 1 - 21,296/22,781$ ) and average Social Security earnings among natives are about 11 percent lower than the reported earnings in the CPS. However, the earnings gaps between natives and immigrants based on Social Security earnings records are quite similar to those based on CPS-reported earnings (columns three and five). Indeed, for most immigrant cohorts the gaps are within two percentage points of each other, though for immigrants who arrived in 1986–89 the earnings gap differs by ten percentage points. Based on these comparisons, the use of Social Security earnings data,

---

weight and the inverse probability of being matched to Social Security earnings data. Nevo (1998) analyzes a more general case of using weights to adjust for selection bias.

<sup>13</sup>Because reported earnings in the CPS and Census surveys are heavily clustered at round numbers, the comparison of medians is problematic. Table 1, therefore, reports means, after censoring CPS earnings at \$57,600, the Social Security taxable maximum in 1993. Topcoded observations in both the CPS and longitudinal earnings records are multiplied by 1.38. This factor approximates the uncensored mean among the artificially topcoded observations in the 1994 CPS. In addition, observations with annual earnings below \$1000 are dropped from this table.

rather than the familiar self-reported earnings in the CPS, does not systematically affect immigrant-native earnings comparisons.

In the subsequent analysis, immigrant earnings in the longitudinal data are compared to earnings measured in repeated cross-sections of the 1970 through 1990 decennial Census and the 1995 through 1997 March Supplement to the Current Population Survey. (Though they contain a vastly smaller sample size than cross-sections from the Census, the three CPS cross-sections provide an additional time period of data approximately five years after the 1990 Census.) The Census sample includes men born between 1930 and 1969, drawn from the 1970 5% 1-in-100 state sample, the 1980 5% A sample, and 1990 5% sample. Due to the very large samples of natives available in 1980 and 1990, a 10% random sample of the native men (but all of the immigrants) from these files is used. Respondents in the March CPS files who were in the fifth through eighth rotation group are dropped.<sup>14</sup> Earnings from the Census and CPS cross-sectional files are measured as the sum of wage and salary, self-employment, and farm income.

### 3 How Outmigration Affects Immigrant Earnings Measured in Repeated Cross-sectional and Longitudinal Data

Before turning to estimates of immigrant earnings growth, it is important to clarify the potential sources of any differences between earnings growth rates measured from repeated cross-sections and from longitudinal earnings records. The first issue is the *permanent* outmigration of immigrants. Such behavior leads to several types of biases. Most importantly, the average earnings of an immigrant cohort will appear to increase from one Census cross-section to the next if those who leave the country tend to have below-average earnings. In the matched longitudinal sample, in contrast, earnings data is only available for workers who remain in the country until the 1990, 1991, or 1994 cross-sectional surveys are administered. To see the relationship between estimates of immigrant earnings growth from the two datasets, let  $E(w_{c,t}|x)$  denote the expected value of earnings among immigrants who arrived in the U.S. at time  $c$ , measured at time  $t$ , conditional on the sample criterion  $x$ . The following table shows the average earnings of immigrants who arrived in 1967 and 1977,

---

<sup>14</sup>This simplifies inferences because in principle households are surveyed in two consecutive March Supplements.

measured in the 1970, 1980, and 1990 decennial Censuses:<sup>15</sup>

	Year		
	1970	1980	1990
1967 Arrivals	$E(w_{67,70} \text{stay 3 years})$	$E(w_{67,80} \text{stay 13 years})$	$E(w_{67,90} \text{stay 23 years})$
1977 Arrivals	–	$E(w_{77,80} \text{stay 3 years})$	$E(w_{77,90} \text{stay 13 years})$

The corresponding measures from the longitudinal Social Security earnings records matched to the 1994 CPS are given by:

	Year		
	1970	1980	1990
1967 Arrivals	$E(w_{67,70} \text{stay 27 years})$	$E(w_{67,80} \text{stay 27 years})$	$E(w_{67,90} \text{stay 27 years})$
1977 Arrivals	–	$E(w_{77,80} \text{stay 17 years})$	$E(w_{77,90} \text{stay 17 years})$

Thus, the 1970 earnings of immigrants who arrived in 1967 and remained in the U.S. at least until 1970 can be estimated from the 1970 Census,  $E(w_{67,70}|\text{stay 3 years})$ ; however only the earnings in 1970 among members of this cohort who remained in the U.S. until 1994 are available in the longitudinal data,  $E(w_{67,70}|\text{stay 27 years})$ .

Suppose the earnings at time  $t$  of an immigrant who arrived at time  $c$  are given by  $w_{ct} = \mu_{ct} + \varepsilon_{ct}$ , where  $\mu_{ct}$  is the average earnings that the *entire initial* arrival cohort would earn if they remained in the U.S. until time  $t$  and  $\varepsilon_{ct}$  is the deviation of an immigrant's earnings from the group average. Upon the cohort's arrival, the expected value of  $\varepsilon_{ct}$  is zero. Over time, however, if disproportionately less- or more-skilled workers leave the U.S., then the expected value becomes positive or negative.

The earnings growth of the 1967 arrivals measured between the 1970 and 1980 Censuses is given by

$$\Delta_{70,80}^{\text{Census}} = E(w_{67,80}|\text{stay 13 years}) - E(w_{67,70}|\text{stay 3 years})$$

<sup>15</sup>The Census, CPS, and SIPP record immigrants' arrival date by an interval (for example, immigrants who arrived in 1967 would be recorded as arriving between 1965 and 1969). For simplicity, this discussion assumes the exact year of arrival is known.

$$\begin{aligned}
&= (\mu_{67,80} - \mu_{67,70}) + (E(\varepsilon_{67,80}|\text{stay 13 years}) - E(\varepsilon_{67,70}|\text{stay 3 years})) \\
&= (\mu_{67,80} - \mu_{67,70}) \\
&\quad + E(\varepsilon_{67,80}|\text{stay 13 years}) - E(\varepsilon_{67,70}|\text{stay 13 years}) \\
&\quad + E(\varepsilon_{67,70}|\text{stay 13 years}) - E(\varepsilon_{67,70}|\text{stay 3 years}) \\
&= E(w_{67,80}|\text{stay 13 years}) - E(w_{67,70}|\text{stay 13 years}) \\
&\quad + E(\varepsilon_{67,70}|\text{stay 13 years}) - E(\varepsilon_{67,70}|\text{stay 3 years}) \tag{1}
\end{aligned}$$

That is, their measured earnings growth is equal to the growth among the 1967 arrivals who remained in the U.S. at least until 1980, plus a bias term equal to the difference in earnings in 1970 between the immigrants who remained in the U.S. at least until 1970 and the subset who stayed until 1980. If lower-earning immigrants tend to leave the U.S., this last term is positive and the observed change in earnings between Censuses overstates the true increase in earnings experienced by immigrants who remained in the U.S. until 1980. In addition, of course, the expected earnings growth of the original cohort if they all remained in the U.S. until 1980,  $\mu_{67,80} - \mu_{67,70}$ , cannot be estimated from the 1970 and 1980 Censuses.

In contrast to estimates from the Census, immigrant earnings growth measured in longitudinal data provide an unbiased estimate of earnings growth *among the immigrants who are in the U.S. in 1994*. The growth in earnings of the 1967 arrival cohort between 1970 and 1980 in the longitudinal sample is given by

$$\begin{aligned}
\Delta_{70,80}^{\text{Longitudinal}} &= E(w_{67,80}|\text{stay 27 years}) - E(w_{67,70}|\text{stay 27 years}) \\
&= (\mu_{67,80} - \mu_{67,70}) \\
&\quad + E(\varepsilon_{67,80}|\text{stay 27 years}) - E(\varepsilon_{67,70}|\text{stay 27 years}) \tag{2}
\end{aligned}$$

If outmigration is based on permanent earnings characteristics that are not related to immigrants' earnings growth over time (that is, if  $E(\varepsilon_{c,t}|\text{stay } k \text{ years}) = E(\varepsilon_{c,t'}|\text{stay } k \text{ years})$ , for all  $t$  and  $t'$ ), then the estimate from the longitudinal data is also equal to the potential earnings growth the initial cohort would have experienced in the absence of outmigration. In this case, only the level of earnings of the original cohort cannot be identified from longitudinal data on those who remained in the country.

The Census provides meaningful comparisons of across-cohort earnings differences only if the outmigration process is constant across cohorts. In particular, the difference in average earnings three years after entry between the 1967 and 1977 arrivals, measured in the 1970

and 1980 Censuses, is

$$\begin{aligned}
\hat{Y}_{67,77}^{\text{Census}} &= E(w_{77,80}|\text{stay 3 years}) - E(w_{67,70}|\text{stay 3 years}) \\
&= (\mu_{77,80} - \mu_{67,70}) \\
&\quad + E(\varepsilon_{77,80}|\text{stay 3 years}) - E(\varepsilon_{67,70}|\text{stay 3 years})
\end{aligned} \tag{3}$$

The difference in the last line of this expression will be zero if the 1967 and 1977 cohorts have similar selective outmigration patterns during their first three years in the U.S. Of course, changes in immigrant characteristics, such as the fraction of immigrants who arrive from nearby countries or arrive as refugees, will likely change the outmigration process. In this case, it is not possible to separately identify across-cohort differences in earnings – in levels or growth rates – that are due to differential selective outmigration from differences due to the labor market skills of immigrants.

Outmigration leads to a difficulty with the longitudinal data when earnings comparisons are made across cohorts. The difference in average earnings three years after entry between the 1967 and 1977 arrivals, measured in the longitudinal sample, is

$$\begin{aligned}
\hat{Y}_{67,77}^{\text{Longitudinal}} &= E(w_{77,80}|\text{stay 17 years}) - E(w_{67,70}|\text{stay 27 years}) \\
&= (\mu_{77,80} - \mu_{67,70}) \\
&\quad + E(\varepsilon_{77,80}|\text{stay 17 years}) - E(\varepsilon_{67,70}|\text{stay 27 years})
\end{aligned} \tag{4}$$

The first term in the second line of this expression,  $\mu_{77,80} - \mu_{67,70}$ , is the difference in earnings three years after arrival among all immigrants in the two cohorts. The second term is the difference in earnings between the two cohorts caused by differential outmigration between the year they entered and 1994, when the CPS survey was administered. If, for example, the least successful immigrants in each year tend to leave the U.S., then the additional ten years of outmigration among the 1967 cohort means that it will have more “successes” than will the 1977 cohort. However, this source of biases may be limited for earlier arrival cohorts if most outmigration occurs within the first ten years after entry.

While previous researchers have remarked on the potential effects of permanent outmigration, a less well-understood but equally important phenomenon is *temporary* outmigration. Specifically, many immigrants arrive in the U.S., work for several years, return to their home country, and then re-immigrate to the U.S. Since the Census, CPS, and SIPP ask immigrants when they arrived in the U.S. “to stay,” a fraction of those who entered the country

multiple times will have reported their most recent date of arrival. This introduces additional bias into estimates of immigrant earnings derived from repeated cross-section of the Census, particularly if the group of workers who temporarily emigrate tend to be low-skilled immigrants.

For example, immigrants observed in the 1990 Census who report having arrived in the U.S. between 1985 and 1989 are composed of those who arrived in the U.S. for the first time, plus those who initially arrived earlier, left the U.S., and then reentered between 1985 and 1989. This affects the measured average earnings of the earlier and new arrival cohorts: if the reentrants have below-average earnings relative to their initial cohort, the earnings of the initial cohort will appear to rise over time (as discussed above). Having already acquired some U.S.-specific skills during their previous work in the U.S., the reentrants may have higher or lower earnings than the first-time arrivals in the new arrival cohort. Thus, the measured average earnings of those who report being recent arrivals in the 1990 Census may be higher or lower than the average earnings of first-time arrivals.

Similarly, the observed earnings *growth* of an arrival cohort is a mixture of the earnings growth among immigrants who arrived for the first time and the growth among reentrants. Since the group of reentrants will have assimilated when they arrived in the U.S. the first time, the growth among all immigrants who report themselves to be recent arrivals may be slower than the growth among those who arrived for the first time.<sup>16</sup> Thus, analyses of repeated cross-sectional data may give a misleading picture of both the earnings level and growth trajectories of immigrant cohorts.

The first year an immigrant worked in the Social Security covered sector in the United States can be easily found in the longitudinal earnings data. Fourteen percent of immigrants have such earnings prior to their date of entry reported in the CPS or SIPP. Two pieces of evidence suggest that many of these immigrants had temporarily left the U.S. and reported their most recent date of arrival in the CPS or SIPP survey: first, the country-of-origin pattern is similar to the pattern of emigration rates found by Jasso and Rosenzweig (1990) and Borjas and Bratsberg (1996). 10 percent of Asian immigrants, 13 percent of Europeans, 17 percent of Latin Americans, and 19 percent of Canadians had earnings prior to their

---

<sup>16</sup>The shape of the relative immigrant earnings growth path is typically found to be convex. Hence, the mixture of earnings growth between early and new arrivals will be slower than the growth of the new arrivals themselves. This may be complicated by cohort effects in the rate of earnings growth or by an earnings effect from the transient immigrants' time spent outside the U.S.

reported date of arrival. A second, more direct piece of evidence on temporary emigration is that 50 percent of the immigrants who had earnings prior to their reported date of entry had a year without any covered earnings prior to their reported date of entry, which is exactly what would occur if the immigrant left the U.S. for an entire calendar year before re-immigrating.<sup>17</sup>

To the extent that immigrants initially work in uncovered employment, attend school, or do not participate in the labor market when they enter the United States, their first year of covered earnings may be significantly *later* than their reported date of entry. For example, the Immigration Reform and Control Act (IRCA) of 1986 is estimated to have granted citizenship to about three million people who had been living in the U.S. illegally.<sup>18</sup> In the longitudinal data, 14 percent of the immigrants who reported in the CPS or SIPP surveys that they arrived as adults between 1975 and 1979 did not have covered earnings until after 1986; the proportion is 27 percent among those who reported to have arrived between 1980 and 1985.<sup>19</sup> The consequence of this type of discrepancy is that the composition of the set of workers in a given arrival cohort changes as new immigrants enter the labor force or move from uncovered to covered work. This compositional change is also present among similar arrival cohorts in Census data, as shown below, which likely reflects a higher propensity among immigrants to participate in the Census when they are in the U.S. legally.

Because of the possible discrepancies between immigrants' reported date of arrival and their first year of covered earnings, three definitions of immigrants' arrival cohort are used in the analysis below that take advantage of the availability of longitudinal earnings data: first, immigrants are grouped by their reported date of arrival, with any earnings prior to that date dropped from the analysis. Second, immigrants are grouped by the earlier of their reported date of arrival and their first year of covered earnings. Third, immigrants are grouped solely by their first year of covered earnings, ignoring their reported date of entry. None of these definitions is perfect. In particular, as immigrants enter the labor force or move from

---

<sup>17</sup>46 of the 50 percent had a year without covered earnings in the period from five years before until the end of their reported arrival interval. For example, the window for the 1975–79 cohort would be from 1970 to 1979. Disruptions in earnings histories may not be definitive evidence of outmigration since 27 percent of native-born workers had a year without covered earnings between the ages of 25 and 35. Nevertheless, the evidence is suggestive that there may be a significant number of immigrants who leave and then reenter the U.S.

<sup>18</sup>See Bean, Edmonston and Passel, eds (1990) and Smith and Edmonston, eds (1997).

<sup>19</sup>Most of the immigrants who report a date of arrival prior to their first year of covered earnings are from Latin America. Such discrepancies may also arise from reporting errors.

uncovered to covered employment, the composition of workers in a cohort based on either of the first two definitions may change over time, as described in the preceding paragraph. Grouping immigrants based solely on their first year of earnings has the advantage that it follows a consistent cohort of immigrants over time, though it understates the length of time spent in the U.S. by immigrants who initially worked in uncovered employment.

To provide an initial assessment of the effect of alternative definitions of immigrants arrival cohorts, Table 2 compares earnings in 1979, 1989, and 1995 in the longitudinal and repeated cross-sectional datasets based on the three definitions described above. In the first set of columns immigrants are grouped by their reported date of arrival. In the second set of columns, immigrants who had earnings prior to their reported date of arrival are moved into the cohort that reflects their first year of earnings. The table also presents the average age, percent without a high school degree, and the sample size in each cell.<sup>20</sup> In 1979, there are 869 observations of people who reported to have arrived between 1970 and 1974. Of these, 82 had earnings prior to 1970; and 186 people who reported to have arrived after 1974 had their first year of earnings between 1970 and 1974. In changing the arrival period for these 268 people (nearly 31 percent of the original 869 people in the cell), the adjusted 1970–74 arrival cohort has 973 people. This new cohort of immigrants is about one and a half years younger, slightly less educated, and earned about nine percent less than the original group who reported that they came in 1970–74.<sup>21</sup>

The composition of the other immigrant arrival cohorts also change considerably: the 1960–69 cohort increases from 912 to 998 immigrants in 1979, as 62 people from this group had earnings prior to 1960, and 148 people had their first year of covered earnings in that period, but reported a date of arrival later than 1969. The size of the 1975–79 cohort decreases from 551 to 436 people, as 153 are moved to earlier cohorts and 38 people are included added to it. Finally, an example of changes among a recent immigrant cohort is the group who arrived between 1980 and 1984. In 1989, the cohort adjustment decreases the

---

<sup>20</sup>Observations from a three-year window around 1979, 1989, and 1995 are used in the longitudinal sample. Topcoded observations are multiplied by 1.38 to approximate the mean of the censored observations (see footnote 13). All observations with annual earnings less than \$1000 (\$1997) are dropped.

<sup>21</sup>It is also interesting to note that about 30 percent of immigrants do not have a high school degree, and this proportion does not change much over time or across cohorts (compared to 15 percent among natives in 1979, and 9 percent in 1995). Tabulations not reported here indicate that the primary difference in the educational attainment of immigrants and natives is whether or not individuals have completed high school. Beyond high school, the educational distributions are quite similar, though immigrants are slightly more likely than natives to have post-graduate education.

sample size by 77 people, from 1937 to 1856. This is composed of 195 people being moved to earlier cohorts, and 114 people being moved from post-1985 cohorts into the 1980-84 group. The earnings disadvantage with natives in 1989 increases from 46.6 to 53.7 percent as a result of the change in the composition of the cohort. Clearly, important changes in the earnings and characteristics of arrival cohorts are brought about by reclassifying immigrants who had earnings prior to their reported date of arrival.

The effect of the 1986 amnesty program can be seen from the change in the composition of the 1985-89 arrival cohort between the second and third sets of columns. The latter groups immigrants by the period in which they first had covered Social Security earnings, ignoring their self-reported date of arrival. The second set of columns shows 662 observations of immigrants in which the earlier of their first year of earnings or their reported date of arrival falls between 1985 and 1989. However, in the third set of columns there are 1599 immigrants in that cohort. The difference, 937 observations, is from immigrants who reported to have arrived before 1985, but did not have any covered earnings until the 1985-89 period. Even though the redefined cohort is considerably less educated and slightly younger than the cohort defined in columns 4 through 6, many of the immigrants who were given amnesty had already been in the country a few years and had likely assimilated to an extent.<sup>22</sup> The result is that the average earnings of the 1985-89 cohort rises by three percent from the second to the third set of columns.

The effect of changing the arrival cohort for immigrants who were presumably affected by the amnesty can also be seen by examining the average earnings of the 1975-79 and 1980-84 cohorts. In columns 4 and 5 the earnings of the 1975-79 cohort increase by about seven percent from 1979 to 1989 (and their earnings gap with native-born works decreases by 14.4 percentage points). However, clearly the composition of the group changed over the decade with inclusion of the immigrants who entered the labor market after the 1986 amnesty program. In columns 7 through 9, the newly legalized immigrants are included in the 1985-89 cohort. Here the remaining 1975-79 group show a considerably larger increase in earnings between 1979 and 1989 of 15 percent, and the immigrant-native earnings gap decreases by 21.9 percentage points.

The fourth set of columns presents similar estimates from the 1980 and 1990 decennial

---

<sup>22</sup>About half of the three million immigrants who were given amnesty did so under a special program for agricultural workers that had less strict residency requirements than the general amnesty program. See Bean et al., eds (1990).

Census, and the 1995 through 1997 March Supplement to the CPS. (Immigrants in these datasets can only be grouped by their reported date of arrival, of course.) Here there is also evidence of a significant change in the composition of the 1975–79 cohort between the 1980 and 1990 Census, as the sample count increases by about 15 percent. In contrast, the sample count decreases among earlier cohorts, which is most likely the result of outmigration. These changes in the composition of arrival cohorts that likely result from the legalization of many immigrants after 1986 suggest there is value in maintaining consistent cohorts over time by grouping immigrants by their first year of covered earnings.

Finally, the decreases in immigrant–native earnings gaps between the 1980 and 1990 Census are generally larger than those in the longitudinal data. For example, in the 1980 Census, the 1960–69 arrivals had 4.2 percent higher earnings than natives; this gap increased by 9.5 percentage points to 13.7 percent in the 1990 Census. In the longitudinal data (columns 1–2 and 4–5), the comparable change in the earnings gap was only three to four percentage points; In columns 7 and 8, this cohort’s relative earnings actually declined between 1979 and 1989. Similarly, the measured earnings gap among the 1975–79 arrivals narrowed by 38.4 percentage points between the 1980 and 1990 Census. The change in the longitudinal data were smaller in magnitude. This is suggestive that the Census results capture both selective outmigration among low–earning immigrants and earnings growth among immigrants who remain in the U.S. The following section addresses this issue more formally.

Differences in earnings measured in repeated cross–sectional data and longitudinal data may be informative about the magnitude of permanent and temporary outmigration. In addition, given the significant number of discrepancies between immigrants’ reported date of arrival and their first year of covered earnings, it is important to examine the sensitivity of immigrant–native earnings gaps to how immigrants are classified into arrival cohorts.

## **4 Direct Estimates of Immigrant–Native Earnings Differences**

This section uses a median regression framework to first compare the relative earnings gaps of immigrants in the longitudinal data with the gaps in the 1970 through 1990 decennial

Census and the 1995 through 1997 Current Population Surveys.<sup>23</sup> The second part of this section examines in more detail the earnings progress of immigrants who arrived in the 1960's, the 1970's, and between 1980 and 1994, using only the longitudinal data. The goal is to document the effect of permanent and temporary outmigration on estimates of the level and growth rate of immigrant earnings, and to shed light on differences across immigrant arrival cohorts in the level and growth rate of earnings.

#### 4.1 Immigrant earnings in repeated cross-sections and longitudinal data

There is now a fairly standard human capital framework to compare immigrant and native earnings, principally associated with Chiswick's (1978) cross-sectional analysis of the 1970 Census and Borjas' (1985, 1995) cohort-based analyses of the 1970 through 1990 Censuses. For natives, log earnings are specified as a function of potential labor market experience and calendar time effects. Immigrants' earnings are further decomposed into permanent arrival cohort effects ( $k_c$ ), potential experience-at-arrival effects ( $m_{ic}$ ), and indicators for the length of time spent in the U.S. ( $y_{ct}$ ).<sup>24</sup> Specifically, the model of the log of annual earnings is given by

$$w_{ict} = \alpha + \beta f(\text{Experience}_{ict}) + \gamma f(\text{Year}_{ict}) + k_c + m_{ic} + y_{ct} + \varepsilon_{ict} \quad (5)$$

where  $i$  indexes the individual,  $c$  indexes native-born workers and immigrant arrival cohorts, and  $t$  indexes time.  $f(\text{Experience}_{ict})$  and  $f(\text{Year}_{ict})$  are quartics in potential experience and calendar time, and  $\varepsilon_{ict}$  represents unobserved influences on earnings and measurement error.<sup>25</sup>

---

<sup>23</sup>A 25 percent random sample of the natives in the longitudinal data is used to reduce the computational burden of the semi-parametric models used in this section. In addition, the analyses in subsection 4.1 excludes observations for people under 25 years old, since such individuals in the repeated cross-sectional samples may have not yet completed their schooling.

<sup>24</sup>Friedberg (1993) demonstrates the importance of controlling for the age at which an immigrant enters the U.S. labor market in a similar model of earnings. The sample requirement that a person be in the labor force generates a negative correlation between the age at which an immigrant arrived and the years since migration. In this study, the additional requirement that individuals be born between 1930 and 1969 exacerbates this correlation. For example, an immigrant who arrived in 1960 was at most 30 years old at the time of arrival, while someone who arrived in 1980 could have been up to 50 years old. Other examples of similar models of immigrant earnings are given by Carliner (1996), Funkhouser and Trejo (1995), Hu (1999), Schoeni, McCarthy and Vernez (1996), and Schoeni (1997).

<sup>25</sup>Potential experience is the number of years an individual has been out of school. People who did not finish high school are assumed to have entered the labor market at age 18. High school graduates, people with some college, those with a four-year college degree, and those with any post-college education are assumed to have entered the labor market at ages 19, 20, 22, and 24. In addition, the calendar time effects

The immigrant arrival cohort effects ( $y_{ct}$ ) are given by indicators for those who arrived in 1950–59, 1960–64, 1965–69, 1970–74, 1975–79, 1980–84, 1985–89, and 1990–91 (natives are the excluded group).<sup>26</sup> The experience at migration effect is captured by indicators for immigrants who arrived with 6 to 10, 11 to 15, 16 to 20, or more than 21 years of potential experience (immigrants who arrived with five or fewer years of experience are the excluded group). Finally, the years in the U.S. effect is captured by indicators that an immigrant has been in the U.S. for 6 to 10, 11 to 15, 16 to 20, 21 to 25, 26 to 30, and more than 30 years (immigrants who have been in the U.S. fewer than five years are the excluded group).<sup>27</sup>

Educational attainment is not explicitly included in the model. Thus, differences in educational attainment between natives and immigrant cohorts are subsumed in the permanent cohort fixed effects. The focus of this study is on correctly measuring differences in earnings between immigrants and natives, in the presence of selective emigration. Differences in educational attainment, language ability, unobserved skills, and changes in the return to skills, may explain changes in the immigrant–native earnings gap, but exploring this is beyond the scope of this study.

In order to separately identify the effect on earnings of immigrants’ period of arrival, experience at arrival, and time spent in the U.S, the experience and calendar time effects are constrained to be equal among immigrants and natives. While this identification assumption is standard in the literature, there are reasons to believe it may be problematic: Baker and Benjamin (1997) show that family composition affects immigrant husbands and wives’ labor supply decisions. This suggests that differences in lifecycle labor supply behavior may differ between immigrants and natives, which may generate differences in their experience–earnings profiles. Calendar–time effects may differ between immigrants and natives if, for example, changes in the wage structure during the 1980’s and 1990’s disproportionately decreased the average earnings of immigrants. In light of these issues, it may not be appropriate to interpret the coefficient estimates from equation 5 as a purely causal effect of increased human capital investment by immigrants on their relative earnings.<sup>28</sup> A more conservative interpretation is

---

for the repeated cross-sectional sample are simply a set of indicators for which cross-section the observation is drawn from.

<sup>26</sup>Because of differences between the date of arrival intervals in the 1994 through 1997 CPS, immigrants who arrived after 1991 are not included in this subsection.

<sup>27</sup>These effects are measured from the midpoint of the arrival interval for the for longitudinal sample. Such an imputation is not necessary for the Census sample.

<sup>28</sup>Chiswick (1978) proposed a model in which immigrants arrive in the U.S. with few U.S.–specific skills, and hence have lower earnings than observationally similar natives. Immigrants, though, have a greater

that the model is a descriptive device for comparing immigrant and native earnings, with the focus being on how the parameter estimates are different based on longitudinal and repeated cross-sectional data.

The reported earnings data in the Census and CPS cross-sectional samples are artificially censored at the Social Security taxable maximum in each year to be consistent with the longitudinal earnings data. Because earnings are topcoded, ordinary least squares estimation of equation 5 is inappropriate. Instead, the model is estimated with Powell’s (1984) semi-parametric censored least absolute deviation estimator, which is robust to heteroscedasticity and does not require knowledge of the underlying distribution of the unobservable error component. The identifying assumption is that the median of  $\varepsilon_{ict}$  is zero conditional on the regressors.<sup>29</sup>

To account for the longitudinal structure of the data, the standard errors of the coefficient estimates are estimated from fifty bootstrap replications. In each replication, a fifty-percent random sample is drawn (with replacement) from all persons who appear in the data at any time. All longitudinal earnings records associated with these observations are included in the bootstrapped dataset. Equation 5 is estimated for each of the 50 datasets, and the standard deviation of the fifty parameter estimates (divided by the square root of two) is an unbiased estimate of the standard error of the parameter values.<sup>30</sup>

Table 3 reports coefficient estimates of equation 5 using repeated cross-sections of the Census and CPS, and the longitudinal earnings data.<sup>31</sup> For convenience, the time-in-the-

---

incentive than natives to invest in human capital upon arrival, and thus experience faster earnings growth. The theoretical underpinnings of immigrants’ investment in human capital is further discussed in Borjas (1998) and Duleep and Regets (1999).

<sup>29</sup>Suppose the true data generating process is given by  $w_{it} = x'_{it}\beta + \varepsilon_{it}$ , and the observed value of earnings is  $w^*_{it} = \min(w_{it}, \bar{w}_t)$ , where  $\bar{w}_t$  is the topcode in year  $t$ . The censored least absolute deviation estimator (CLAD) is the value of  $\hat{\beta}$  that minimizes  $\frac{1}{N} \sum_{i=1}^N |w^*_{it} - \min(x'_{it}\beta, \bar{w}_t)|$ . This is implemented by the iterative method suggested by Buchinsky (1994). An alternative estimator is the tobit model, which is based on the assumption that the distribution of  $\varepsilon_{it}$  is known (i.e. homoscedastic with a normal, log normal or Weibull distribution). Chay and Honoré (1998) investigate the relative performance several estimators of the censored regression model and conclude that nonnormality in the distribution of log earnings may lead to significant biases in tobit estimates. The qualitative conclusions of this section are not sensitive to whether the CLAD or tobit model is used.

<sup>30</sup>The fifty sets of parameter estimates are obtained from least absolute deviation estimation of equation 5 using only those observations which are not predicted to be censored based on the CLAD coefficient estimates.

<sup>31</sup>There are relatively few immigrants in the longitudinal sample who arrived prior to 1960. The results in the this section are not substantively affected by the exclusion of immigrants in both the cross-sectional and longitudinal samples who arrived prior to 1960. The results are also not different if the time period from 1979 to 1997 is used, rather than 1969 to 1997.

U.S. effects ( $\hat{y}_{ct}$ ) are graphed in Figure 2. Given the significant discrepancies between immigrants' report of their date of arrival and their first year of covered earnings, immigrants are first grouped based on their reported date of arrival (with any earnings data prior to that date dropped from the analysis). Immigrants are then grouped by the earlier of their reported date and their first year of earnings (this is referred to as the "adjusted cohort").

Measured immigrant earnings growth is considerably faster in the repeated cross-sectional data than in the longitudinal data, suggesting greater outmigration among low-earning immigrants. The Census and 1995–97 CPS cross-sections suggest that immigrant relative earnings grow by 20 percent in their first ten years in the U.S., and by an additional 10 to 20 percent in each successive decade. By contrast, in the longitudinal data immigrant earnings grow by 12 to 15 percent during their first 15 years in the U.S., but relatively little after that.

The estimated arrival cohort effects ( $\hat{k}_c$ ) are plotted in Figure 3.<sup>32</sup> Consistent with most past research, both the repeated cross-sections and the longitudinal earnings data indicate a marked decrease in the level of earnings of successive immigrant cohorts between 1960–64 and 1980–84.<sup>33</sup> In the repeated cross-sectional data, the level of earnings among immigrants who arrived between 1960 and 1964 was 26 percent below native-born workers. This gap increased to 49 percent among immigrants who arrived between 1980 and 1984. The secular decline in earnings is even larger in the longitudinal data when immigrants are grouped by their reported date of arrival. The earnings gap among the 1960–64 and 1980–84 cohorts in that data are 7 and 42 percent below natives.

Adjusting immigrants' arrival cohort for those who had covered earnings prior to their reported date of entry reduces the secular decline in the estimated cohort fixed effects between 1960–64 and 1980–84 by about one-third. The earnings gap between natives and the 1960–64 cohort, for example, increases from 7 percent to 20 percent, while the gap among the 1980–84 cohort only increases by 4 percentage points. This is consistent with there being a large number of low-wage immigrants who entered the 1960's and 1970's, left for some time, and then reentered and are recorded in the data as having arrived in the 1970's and 1980's. Once these people are assigned back to their first arrival cohort, the apparent labor market

---

<sup>32</sup>Since the years-in-the-U.S. effects exclude immigrants' first five years, the cohort fixed effects can be interpreted as each cohorts' relative earnings upon entry. As discussed below, however, this interpretation is not valid if the rate of earnings growth differs across cohorts.

<sup>33</sup>See, for example, Borjas (1985, 1995), Carliner (1996), LaLonde and Topel (1992), and Funkhouser and Trejo (1995).

“quality” of earlier immigrant cohorts is lowered.

The level of earnings for each arrival cohort are higher in the longitudinal sample than in the repeated cross-sections. This is consistent with outmigration by the least successful immigrants over time, coupled with the data requirement that people who are in the longitudinal sample had to stay in the U.S. until the early 1990’s. A potential measure of the extent of nonrandom outmigration, therefore, is the divergence between the fixed effect estimates from the longitudinal and repeated cross-sectional data. However, this comparison is complicated by the fact that arrival dates in the Census may be substantially misreported. If the reported date of arrival in the Census could be adjusted, presumably the secular decline in entry earnings in that sample would be flatter than that estimated here.

The longitudinal earnings data in Figure 3 also indicate a significant upsurge in the entry earnings of immigrants who arrived between 1985 and 1991, relative to those who arrived in the first half of the 1980’s.<sup>34</sup> Whether grouped by reported or adjusted arrival cohorts, the earnings level of immigrants were between 2.4 and 3.6 percentage points higher among 1985–89 arrivals than among 1980–84 arrivals. This upward trend in the earnings of very recent immigrants is also found by Funkhouser and Trejo (1995), who examine hourly earnings in several supplements to the CPS in the 1980’s, and found in a slightly different form by Jasso, Rosenzweig and Smith (1998), who use INS data to examine the occupational structure of immigrants who arrived between 1972 and 1995. The improved labor market status of recent immigrants may reflect a decrease in illegal immigration following the Immigration Reform and Control Act in 1986, and the increase in occupational skills-based visas that were awarded after the Immigration Act of 1990.<sup>35</sup> The fact that this upturn in immigrant earnings is not found in the repeated cross-sectional estimates may be because there exists a large number of recently arrived illegal immigrants in the 1990 Census that are not present in longitudinal sample.

## 4.2 Longitudinal estimates of immigrant earnings across cohorts

This section addresses whether there are differences in the level and rate of growth of earnings among successive immigrant arrival cohorts. The model given by equation 5 esti-

---

<sup>34</sup>Note, however, that because of the smaller sample size, the standard errors for the 1990–91 cohort’s fixed effect are quite a bit larger than those for earlier cohorts.

<sup>35</sup>The Immigration Reform and Control Act sought to decrease the flow of illegal entrants through sanctions on employers who knowingly hire illegal immigrants and by increasing border enforcement.

mates an average rate of earnings growth across all immigrant cohorts. If, however, earnings growth rates differ by arrival cohort, the fixed effects estimates will reflect both differences in the average level of earnings and in earnings growth across cohorts.<sup>36</sup> Thus if successive immigrant cohorts have slower rates of earnings growth, a model that imposes a constant effect of time spent in the U.S. will tend to yield declining estimates of cohort fixed effects.

To address this issue, the earnings model given by equation 5 is estimated separately for immigrants in the longitudinal data who arrived in the 1960's, the 1970's, and between 1980 and 1994. The immigrant effects are parameterized as a quartic in the number of years in the U.S., and a square in the potential experience at entry (rather than with sets of dummy variables). Since the arrival cohorts are analyzed separately, the indicators for the period of arrival are replaced by an immigrant dummy variable. Each arrival cohort is compared to the population of natives who were already in the labor market at the time the immigrant group arrived. That is, the 1960–69 arrivals are compared to natives who entered the labor market by 1969; the 1970–79 arrivals are compared to natives who entered the labor market by 1979; and the 1980–94 arrivals are compared to the full sample of natives.<sup>37</sup>

Figure 4 compares the relative median earnings of immigrants using two arrival cohort definitions, one based on their reported arrival date and the other based on the earlier of their reported date and their first year of earnings.<sup>38</sup> The level of the earnings profiles correspond to immigrants who arrived in the U.S. with five years of potential experience. Immigrants who reported to have arrived between 1960 and 1969 entered the labor market with earnings 28 percent below natives. After ten years the gap closed to 15 percent and was essentially

---

<sup>36</sup>To see this, consider a simple example in which there are two immigrant cohorts ( $c = 1, 2$ ) and earnings are given by  $w_{ict} = \alpha_c + \beta_c Y_{ict} + \epsilon_{ict}$ , where  $Y_{ict}$  is the number of years the immigrant has been in the U.S. This can be written as  $w_{ict} = \alpha_1 + \alpha_2 + \bar{\beta} Y_{ict} + (\beta_1 - \bar{\beta}) \alpha_1 Y_{i1t} + (\beta_2 - \bar{\beta}) \alpha_2 Y_{i2t} + \epsilon_{ict}$ . Suppose the model given by  $w_{ict} = \alpha_1 + \alpha_2 + \bar{\beta} Y_{ict} + \epsilon_{ict}$  is estimated, and assume the same number of years of data are available for each cohort (so  $E(\alpha_c Y_{ict}) = 0$ ). Then the OLS estimate of  $\bar{\beta}$  is  $E(\hat{\bar{\beta}}) = \beta_1 E(\alpha_1) + \beta_2 E(\alpha_2)$ , which is the average effect among both cohorts of time spent in the U.S. on earnings. The expected value of the indicator for the first cohort is  $E(\hat{\alpha}_1) = \alpha_1 + E(Y_{i1t})(\beta_1 - \beta_2)$ . Thus, if the first cohort has faster earnings growth than the second cohort,  $(\beta_1 - \beta_2)$  is positive and  $E(\hat{\alpha}_1) > \alpha_1$ . In addition, the fixed effect estimates cannot be interpreted as the initial earnings level for each cohort. An additional bias is present if more years of data are available for cohorts that entered earlier.

<sup>37</sup>The earnings history from the beginning of each immigrant cohort until 1997 is used (for example, for the 1960–69 cohort, the estimates are based on earnings data for immigrants and natives from 1960 to 1997); and all men with positive potential experience are included in the sample.

<sup>38</sup>Immigrants' time in the U.S. is measured from the midpoint of the arrival cohort. Earnings prior to and during the arrival interval are dropped; thus the first observed year corresponds to the third year in the U.S. For example, suppose an immigrant reports that he arrived between 1980 and 1984, but his first year of covered earnings is in 1973. The adjusted cohort would then be the 1970–74 group, his time in the U.S. would be measured from 1972, and only his earnings beginning in 1975 would be used.

stable after that point. When immigrants who reported arriving after 1969, but whose first year of earnings was during the 1960's, are included in this cohort, the earnings gap upon entry declines by five percentage points to 32 percent, and the earnings gap after ten years is 23 percent. A similar decline occurs when the compositions of the 1970–79 and 1980–94 cohorts are adjusted, again suggesting the presence of low-earning immigrants who enter the U.S., leave, and then return. The misclassification of these transient immigrants as more recent arrivals than they actually were leads to a significant overstatement of the earnings of earlier immigrant groups, and consequently exaggerates the decline in immigrant earnings across arrival cohorts.

As documented in Section 3, many immigrants do not have earnings until some years after their reported date of arrival. This introduces changes in the composition of cohorts over time as new immigrants enter the covered sector. To explore this effect, Figure 5 examines the sensitivity of the immigrant earnings estimates to the exclusion of immigrants who did not have covered earnings until three or more years after their reported arrival period. Also shown are the earnings trajectories from Figure 4 in which immigrants are grouped by the earlier of their first year of covered earnings and their reported date of arrival. For all three arrival cohorts, the exclusion of immigrants who entered the covered sector after their reported date of arrival increases the estimated rate of immigrant earnings growth. The immigrant–native earnings gap decreases among the 1960–69 arrivals who have been in the U.S. for 15 years from 24 to 20 percent. Among the 1970–79 arrivals the decline is from 26 to 19 percent. This adjustment also eliminates the anomalous downward trend in the relative earnings of immigrants who arrived between 1980 and 1994. After the sample adjustment, the immigrant–native earnings gap among this most recent cohort declines from 39 percent upon arrival to 34 percent after 10 years. These results indicate that the labor–market entry of low-earning immigrants who did not enter the covered sector until years after their reported date of arrival may lead to significant changes in the composition of immigrant cohorts and, consequently, an understatement in their measured relative earnings growth. It is important to note that to the extent that the earnings of workers in uncovered employment – in particular, immigrants in the country illegally – are not recorded in the Census, a similar bias may exist in that data as well.

Finally, Figure 6 compares the earnings of immigrants grouped solely by their first year

of covered earnings.<sup>39</sup> The entry earnings of immigrants who arrived in the 1960's are 34 percent below those of native-born workers. The earnings gaps upon entry for the 1970–79 and 1980–94 cohorts are 37 and 43 percent. There are also small differences in the rate of earnings growth across cohorts: during their first ten years in the U.S., the immigrant–native earnings gap among the 1960–69 arrivals closed by 15 percentage points. Among the 1970–79 and 1980–94 cohorts, the gap closed by 20 and 22 percentage points. Thus, while there is evidence of a decline in the relative earnings among successive immigrant cohorts, the change is considerably smaller than that suggested by the trend in the cohort fixed effects in Section 4, and certainly does not support the idea of a dramatic decline in the earnings of recent immigrants.

## 5 Conclusion

Using longitudinal earnings records from 1951 to 1997, this work has addressed several important issues in the study of immigrant earnings. Many immigrants to the United States do not remain in the country throughout their working lives. Some permanently leave the country, while others reenter at a later date. Both types of migration decisions may bias the measured earnings of immigrants in repeated cross-sections of the decennial Census. Whereas permanent outmigration by low-earning immigrants will lead to an overstatement of earnings growth across Census cross-sections, temporary outmigration by low-earning immigrants may lead to the appearance of a decline in the level of earnings of successive arrival cohorts if reentrants tend to report the date of their most recent arrival in the Census questionnaire. Though the longitudinal data is not ideal, it provides a different perspective on these issues from repeated cross-sections of the Census.

This study provided evidence for several new empirical results: first, actual earnings growth among immigrants who remained in the U.S. until the 1990's was considerably slower than that implied by estimates from repeated cross-sections of the decennial Census. Measured in longitudinal data, over their first twenty years in the U.S., immigrant earnings grow by 10 to 13 percent relative to the earnings of native-born workers. Repeated cross-sections of the Census suggest immigrant earnings grow twice as fast, by about 25 percent over the same time period. Selective outmigration by low-earning immigrants may, therefore, give the misleading impression that the economic status of immigrants to the United States improves

---

<sup>39</sup>Immigrants' time in the U.S. is measured from their first year of earnings.

substantially as they assimilate into the labor market.

The second key result is that there is a substantial amount of disagreement between immigrants' self-reported date of arrival and their first year of covered Social Security earnings, caused in part by outmigration and the subsequent reentry into the U.S. Nearly one-third of the decline in the level of earnings of immigrants between the 1960's and 1980's can be accounted for by the misclassification of many low-wage immigrants as more recent arrivals, when in fact they had entered the U.S. considerably earlier. A corollary of this finding is that the standard model in the economics literature in which immigrants remain in the U.S. throughout their working life may not be appropriate for analyzing contemporary migration. The decision to emigrate from the U.S., and possibly return, has received little scholarly attention.

## References

- Baker, Michael and Dwayne Benjamin**, “The Role of the Family in Immigrants’ Labor–Market Activity: An Evaluation of Alternative Explanations,” *The American Economic Review*, September 1997, 87 (4), 705–727.
- Bean, Frank D. and Michael Fix**, “The Significance of Recent Immigration Policy Reforms in the United States,” in Gary P. Freeman and James Jupp, eds., *Nations of Immigrants: Australia, The United States, and International Migration*, Melbourne: Oxford University Press, 1992, pp. 41–55.
- , **Barry Edmonston, and Jeffrey S. Passel, eds**, *Undocumented Migration to the United States: IRCA and the Experience of the 1980’s*, Washington, D.C.: Urban Institute Press, 1990.
- Borjas, George J.**, “Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants,” *Journal of Labor Economics*, 1985, 3, 463–489.
- , “Immigrant and Emigrant Earnings: A Longitudinal Study,” *Economic Inquiry*, January 1989, 27, 21–37.
- , “The Economics of Immigration,” *Journal of Economic Literature*, December 1994, 32, 1667–1717.
- , “Assimilation and Cohort Quality Revisited: What Happened to Immigrant Earnings in the 1980s?,” *Journal of Labor Economics*, 1995, 13, 201–245.
- , “The Economic Progress of Immigrants,” Working Paper 6506, National Bureau of Economic Research April 1998.
- , “The Economic Analysis of Immigration,” in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 3, Amsterdam, The Netherlands: Elsevier Science B.V., 1999, chapter 28.
- **and Bernt Bratsberg**, “Who leaves? The Outmigration of the Foreign–Born,” *The Review of Economics and Statistics*, February 1996, 78 (1), 165–176.
- Bound, John and Alan B. Krueger**, “The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?,” *Journal of Labor Economics*, 1991, 9 (1), 1–24.
- Buchinsky, Moshe**, “Changes in the U.S. Wage Structure 1963–1987: Application of Quantile Regression,” *Econometrica*, March 1994, 62 (2), 405–458.
- Card, David and Alan B. Krueger**, “Trends in Relative Black–White Earnings Revisited,” *American Economic Review: Papers and Proceedings*, May 1993, 83 (2), 85–91.
- Carliner, Geoffrey**, “The Wages and Language Skills of U.S. Immigrants,” Working Paper 5763, National Bureau of Economic Research September 1996.

- Chay, Kenneth Y.**, “Evaluating the Impact of the 1964 Civil Rights Act on the Economic Status of Black Men Using Censored Longitudinal Earnings Data,” October 1995. Unpublished working paper.
- and **Bo E. Honoré**, “Estimation of Semiparametric Censored Regression Models: An Application to Changes in Black–White Earnings Inequality During the 1960s,” *The Journal of Human Resources*, 1998, 33 (1), 4–38.
- Chiswick, Barry R.**, “The Effect of Americanization on the Earnings of Foreign–born Men,” *Journal of Political Economy*, October 1978, 86 (5), 897–921.
- , *An Analysis of the Economic Progress and Impact of Immigrants*, Washington, D.C.: U.S. Department of Labor, 1980. National Technical Information Service, No. PB80–200454.
- Duleep, Harriet Orcutt**, “Immigrant Earnings Growth 1960–1990: Initial Insights from Longitudinal Data on Individuals,” 1999. Unpublished manuscript, Urban Institute.
- and **Mark C. Regets**, “The Elusive Concept of Immigrant Quality: Evidence from 1970–1990,” Program for Research on Immigration Policy PRIP–UI–41, The Urban Institute April 1996.
- and —, “Immigrant Entry Earnings and Human Capital Growth: Evidence from the 1960–1980 Censuses,” in Solomon W. Polachek, ed., *Research in Labor Economics*, Vol. 16, Greenwich, Conn.: JAI Press, 1997, pp. 297–317.
- and —, “Immigrants and Human–Capital Investment,” *The American Economic Review: Papers and Proceedings*, May 1999, 89 (2), 186–197.
- Friedberg, Rachel M.**, “The Labor Market Assimilation of Immigrants in the United States: The Role of Age at Arrival,” March 1993. Unpublished manuscript, Brown University.
- Funkhouser, Edward and Stephen J. Trejo**, “The Labor Market Skills of Recent Male Immigrants: Evidence from the Current Population Survey,” *Industrial and Labor Relations Review*, July 1995, 48 (4), 792–811.
- Hu, Wei-Yin**, “Assimilation and the Earnings of Immigrants: New Evidence from Longitudinal Data,” August 1999. Unpublished manuscript, University of California, Los Angeles.
- Jasso, Guillermina and Mark R. Rosenzweig**, *The New Chosen People: Immigrants in the United States*, New York: Russell Sage Foundation, 1990.
- , —, and **James P. Smith**, “The Changing Skills of New Immigrants to the United States: Recent Trends and Their Determinants,” Working Paper 6764, National Bureau of Economic Research October 1998.

- LaLonde, Robert J. and Robert H. Topel**, “The Assimilation of Immigrants in the U.S. Labor Market,” in George J. Borjas and Richard B. Freeman, eds., *Immigration and the Work Force*, Chicago: The University of Chicago Press, 1992, pp. 67–92.
- and —, “Economic Impact of International Migration and The Economic Performance of Migrants,” in Mark R. Rosenzweig and Oded Stark, eds., *Handbook of Population and Family Economics*, Vol. 1B, Amsterdam, The Netherlands: Elsevier Science B.V., 1997, chapter 14.
- Nevo, Aviv**, “Using Weights to Adjust for Sample Selection,” August 1998. Unpublished manuscript, University of California, Berkeley.
- Powell, James L.**, “Least Absolute Deviations Estimation for the Censored Regression Model,” *Journal of Econometrics*, 1984, 25, 303–325.
- Schoeni, Robert F.**, “New Evidence on the Economic Progress of Foreign-Born Men in the 1970’s and 1980’s,” *The Journal of Human Resources*, 1997, 32 (4), 683–740.
- , **Kevin F. McCarthy, and Georges Vernez**, *The Mixed Economic Progress of Immigrants*, Santa Monica, CA: Center for Research on Immigration Policy, RAND, 1996.
- Smith, James P. and Barry Edmonston, eds**, *The New Americans: Economic, Demographic, and Fiscal Effects of Immigration*, Washington, D.C.: National Academy Press, 1997.
- Social Security Administration, Office of the Chief Actuary**, “History of the Provisions of Old-Age, Survivors, Disability, and Health Insurance, 1935–1996,” October 1997. SSA Pub. No. 11–11515.

Table 1: Average Immigrant and Native Earnings in the March 1994 CPS and Social Security Earnings Data

		1993 CPS Reported Earnings			1993 Social Security Earnings	
		Full sample	Matched subsample	Reweighted subsample	Matched subsample	Reweighted subsample
Natives	Log earnings	10.16	10.17	10.16	10.06	10.05
	Sample size	27,052	22,781	22,781	21,296	21,296
Immigrant cohorts						
1950-69	Log earnings	10.18	10.18	10.20	10.09	10.13
	Earnings gap	2.45%	1.20%	4.01%	3.22%	7.73%
	Standard error	(5.18)	(5.45)	(5.25)	(6.48)	(6.56)
	Sample size	340	254	254	234	234
1970-74		9.96	10.13	10.09	10.06	10.02
		-19.75%	-3.61%	-6.60%	-0.20%	-2.73%
		(6.11)	(6.64)	(6.58)	(7.10)	(7.18)
		308	204	204	193	193
1975-79		9.98	10.01	9.98	9.91	9.89
		-17.82%	-16.11%	-17.60%	-14.94%	-16.20%
		(5.36)	(5.83)	(5.78)	(6.27)	(6.22)
		400	284	284	265	265
1980-85		9.79	9.87	9.81	9.79	9.73
		-36.60%	-29.46%	-34.81%	-27.08%	-32.05%
		(3.63)	(4.32)	(4.44)	(5.03)	(5.00)
		750	517	517	462	462
1986-89		9.66	9.74	9.65	9.71	9.65
		-49.64%	-42.96%	-50.59%	-35.14%	-39.74%
		(4.10)	(5.81)	(5.43)	(6.33)	(5.78)
		560	320	320	286	286
1990-94		9.58	9.71	9.60	9.57	9.49
		-58.17%	-45.33%	-55.46%	-49.12%	-55.48%
		(5.02)	(6.47)	(6.17)	(7.49)	(7.10)
		456	254	254	203	203

Note: 1993 CPS earnings are individuals' self-reported wage and salary, self-employment, and farm income from the 1994 CPS. These earnings are topcoded at the 1993 Social Security maximum of \$57,600. Topcoded observations in both datasets are multiplied by 1.38 to approximate the uncensored mean. Only observations with positive earnings are used. Earnings gaps are computed as the difference in log earnings between each immigrant cohort and all natives. The means given in the first, second, and fourth columns are weighted by the March Supplement weights. The means in the third and fifth columns are reweighted by the probability of each individual in the public use file being matched to earnings records. See text for details. Sample sizes are unweighted. Observations with annual earnings less than \$1000 (\$1993) are dropped. Standard error of the earnings gap given in parentheses.

Table 2: Native and Immigrant Earnings by Cohort Definition

	Longitudinal Earnings Data						Repeated Cross-Sections						
	1979	1989	1995	1979	1989	1995	1979	1989	1995	1979	1989	1995	
	1	2	3	4	5	6	7	8	9	10	11	12	
<i>Natives:</i>													
Log wage	10.30	10.23	10.19	10.30	10.23	10.19	10.30	10.23	10.19	10.35	10.32	10.24	
Age	35.0	38.6	41.1	35.0	38.6	41.1	35.0	38.6	41.1	35.7	39.4	42.1	
No HS	14.6%	11.8%	9.4%	14.6%	11.8%	9.4%	14.6%	11.8%	9.4%	17.6%	14.0%	9.4%	
N	61,955	97,469	61,961	61,955	97,469	61,961	61,955	97,469	61,961	160,051	235,312	36,174	
<i>Immigrant Cohorts by period of arrival:</i>													
	Reported Arrival Cohort			Adjusted Arrival Cohort			First Year of Covered Earnings			Reported Arrival Cohort			
1960-69	Log wage	10.34	10.31	10.20	10.32	10.29	10.18	10.54	10.29	10.16	10.39	10.46	10.23
	Gap	3.58%	7.77%	0.61%	2.28%	5.38%	-1.40%	23.84%	6.08%	-2.82%	4.24%	13.69%	-0.84%
	Age	39.9	49.9	53.9	40.6	50.4	54.8	45.5	51.0	55.5	40.5	50.3	54.9
	No HS	33.1%	32.9%	31.4%	35.2%	34.5%	31.4%	26.6%	36.7%	33.6%	37.1%	36.4%	27.2%
	N	912	859	422	998	958	484	915	854	429	28,348	24,459	318
1970-74	Log wage	10.11	10.28	10.11	10.02	10.22	10.08	10.32	10.23	10.13	10.12	10.30	10.25
	Gap	-19.43%	4.45%	-8.49%	-28.41%	-0.93%	-10.92%	1.68%	-0.49%	-6.15%	-23.71%	-2.00%	0.75%
	Age	34.2	43.6	48.8	32.7	42.1	47.7	41.1	42.9	48.4	35.0	44.6	49.7
	No HS	32.37%	33.75%	45.79%	35.16%	35.58%	47.38%	36.71%	34.64%	45.00%	40.49%	41.96%	30.81%
	N	869	930	477	973	1067	548	957	972	495	25,972	23,444	334
1975-79	Log wage	9.99	10.12	10.03	10.03	10.10	10.04	10.04	10.19	10.13	9.81	10.17	10.03
	Gap	-30.82%	-10.84%	-16.45%	-27.34%	-12.92%	-15.18%	-26.10%	-4.23%	-6.30%	-53.73%	-15.27%	-21.21%
	Age	32.9	38.7	44.4	33.3	38.1	43.6	33.3	39.0	44.9	33.4	40.1	45.6
	No HS	25.77%	33.97%	36.58%	25.06%	34.11%	32.71%	35.73%	30.16%	25.80%	34.23%	38.52%	35.48%
	N	551	1266	700	436	1261	708	561	972	504	28,533	32,729	536
1980-84	Log wage		9.88	9.89		9.88	9.90		9.98	10.02		9.93	9.93
	Gap		-35.35%	-29.75%		-35.60%	-29.68%		-25.15%	-16.81%		-39.07%	-31.01%
	Age		35.5	40.3		34.8	39.6		36.4	41.7		36.0	42.1
	No HS		28.57%	34.74%		27.94%	35.79%		25.74%	28.40%		38.60%	33.00%
	N		1937	1299		1856	1261		1431	733		48,795	929
1985-89	Log wage		9.77	9.87		9.70	9.82		9.73	9.84		9.70	9.78
	Gap		-46.62%	-32.56%		-53.67%	-37.11%		-50.64%	-34.86%		-62.09%	-45.42%
	Age		35.5	35.3		34.4	34.1		33.9	36.0		34.7	36.9
	No HS		32.07%	33.41%		29.85%	35.93%		34.85%	41.05%		34.65%	33.29%
	N		809	847		662	897		1599	1344		43,661	755
1990-91	Log wage			9.75			9.77			9.79			9.75
	Gap			-44.15%			-42.49%			-40.51%			-48.79%
	Age			36.5			36.5			36.2			37.3
	No HS			32.21%			28.27%			37.82%			38.00%
	N			355			282			420			304

Note: The first three panels are immigrants from the matched Social Security earnings data. The first panel groups immigrants by their reported date of arrival; the second by the earlier of their reported date of arrival and their first year of covered earnings; and the third groups immigrants solely by their first year of covered earnings. The fourth panel are estimates from the 1980 and 1990 Census, and the 1995-1997 March CPS. Observations from the longitudinal data are weighted by the inverse of the probability of being matched to earnings records, as described in the text. Observations from the 1990 Census and the 1995-1997 March CPS are weighted by the sample weights. See text for additional details. Observations with annual earnings less than \$1000 (\$1997) are dropped.

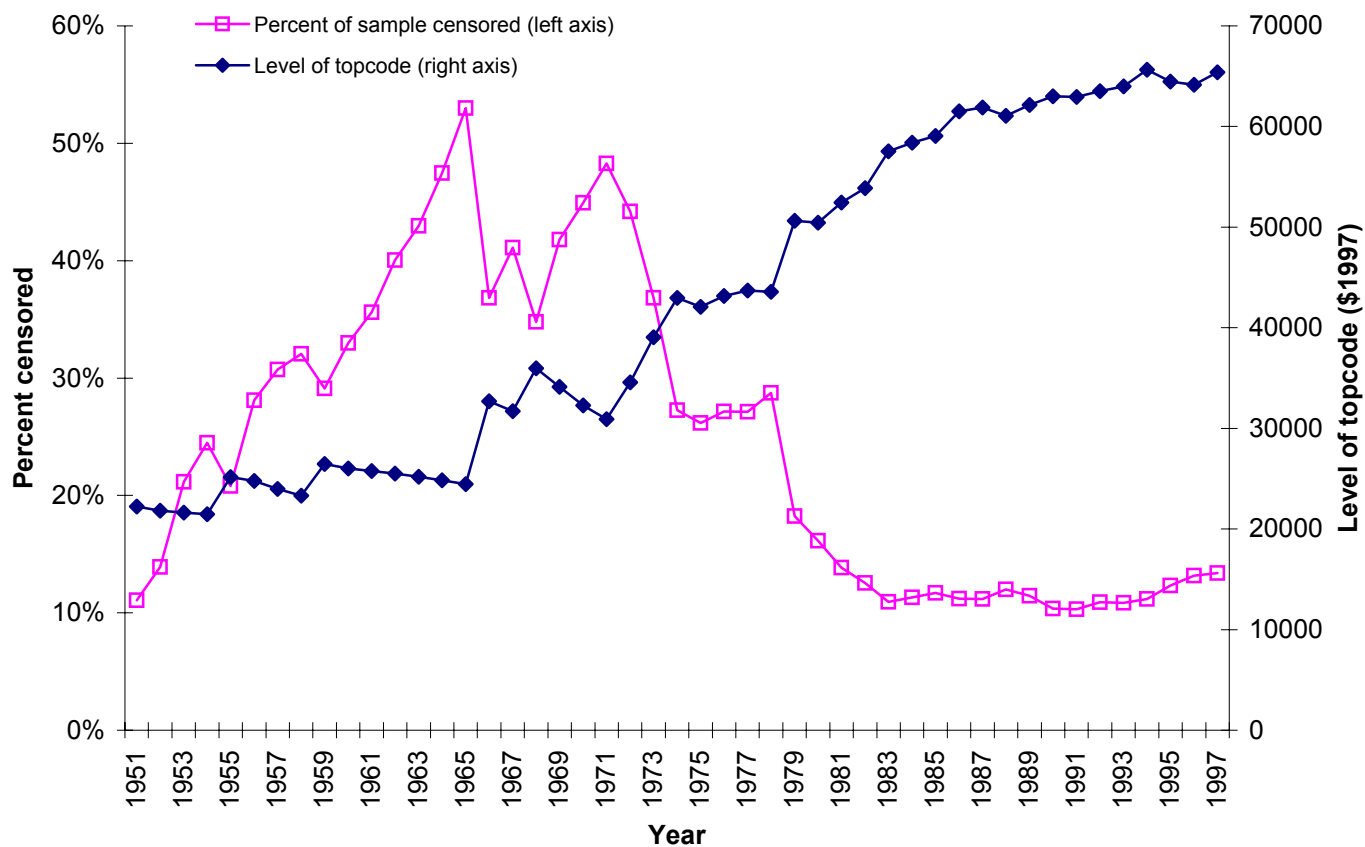
The statistics shown are the average of the log of annual earnings, the difference in log earnings between each immigrant cohort and natives (expressed as a percent), the average age, the percent of each cell without a high school degree, and the unweighted sample count. The average is computed by first multiplying all topcoded observations by 1.38, which approximates the mean of the uncensored data. See text for details.

Table 3: Censored Least Absolute Deviation Estimates of Immigrant Earnings Growth in the Census and Social Security Earnings Records

	Repeated Cross-sections	Longitudinal earnings data	
		Reported period of arrival	Adjusted period of arrival
Years in the U.S.			
6--10	0.1325 (0.0178)	0.0803 (0.0260)	0.0860 (0.0261)
11--15	0.2150 (0.0078)	0.1188 (0.0307)	0.1480 (0.0371)
16--20	0.2590 (0.0202)	0.1040 (0.0352)	0.1387 (0.0428)
21--25	0.3351 (0.0128)	0.1306 (0.0393)	0.1759 (0.0514)
26--30	0.3867 (0.0230)	0.1805 (0.0519)	0.1952 (0.0543)
31 +	0.5336 (0.0269)	0.3076 (0.0954)	0.2843 (0.1315)
Experience at entry			
6--10	-0.1106 (0.0049)	-0.0837 (0.0443)	-0.0765 (0.0444)
11--15	-0.1643 (0.0053)	-0.2127 (0.0618)	-0.2126 (0.0709)
16--20	-0.2148 (0.0067)	-0.1956 (0.0829)	-0.1601 (0.0896)
21+	-0.2438 (0.0081)	-0.3025 (0.0884)	-0.2514 (0.1272)
Period of arrival			
1950--59	-0.3565 (0.0248)	-0.1396 (0.0759)	-0.2242 (0.1078)
1960--64	-0.2570 (0.0204)	-0.0714 (0.0656)	-0.2024 (0.0587)
1965--69	-0.2947 (0.0089)	-0.1759 (0.0604)	-0.2732 (0.0716)
1970--74	-0.3552 (0.0203)	-0.1976 (0.0610)	-0.2824 (0.0564)
1975--79	-0.4010 (0.0068)	-0.2762 (0.0568)	-0.3158 (0.0656)
1980--84	-0.4924 (0.0187)	-0.4247 (0.0578)	-0.4604 (0.0488)
1985--89	-0.5245 (0.0076)	-0.3885 (0.0645)	-0.4364 (0.0543)
1990--91	-0.5427 (0.0665)	-0.3349 (0.0999)	-0.3776 (0.1758)
Number of immigrants	280,411	3,069	2,917
Number of natives	598,515	10,772	10,772
Total observations	878,926	202,561	202,172
Noncensored obs.	660,773	163,702	163,582

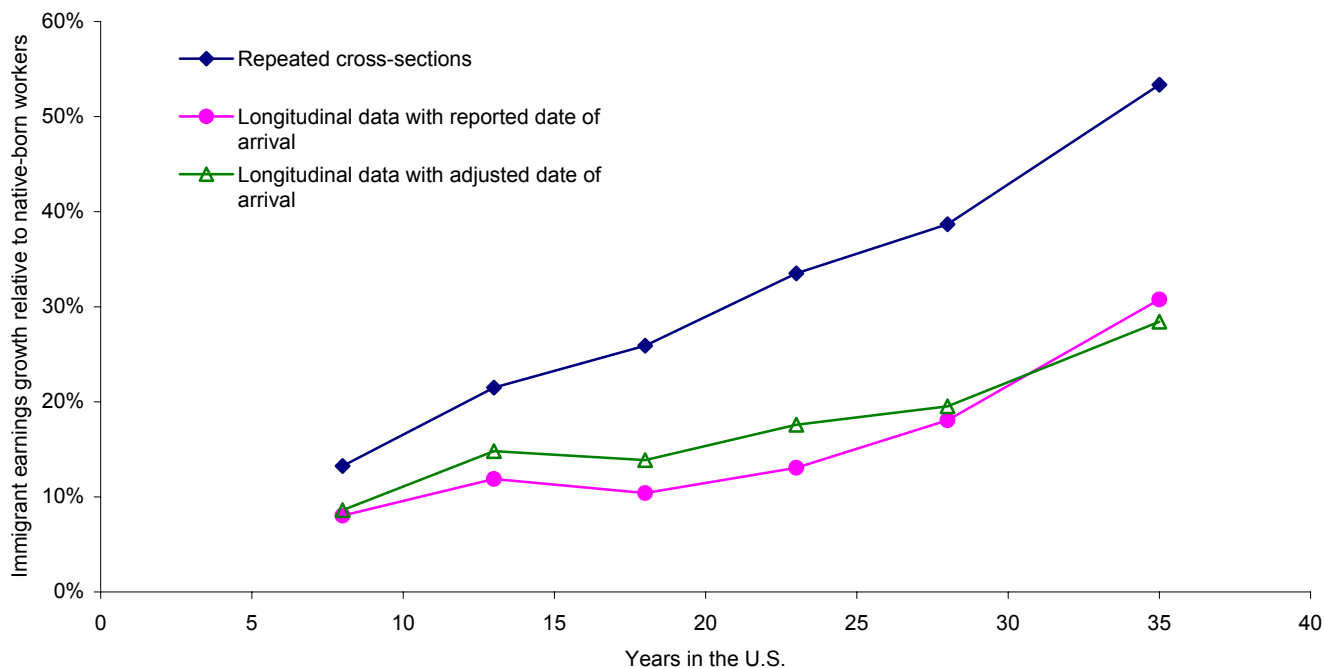
Note: Dependent variable is the log of annual earnings. Parameters are estimated by the method of censored least absolute deviation. Repeated cross-sectional data are from the 1970 through 1990 decennial Census, and the 1995-1997 March CPS. Standard errors are based on 50 bootstrap replications. All models include quartics in potential experience and calendar time. See text for additional details.

Figure 1: Level of Topcode and Percent of Sample Censored  
1951-1997



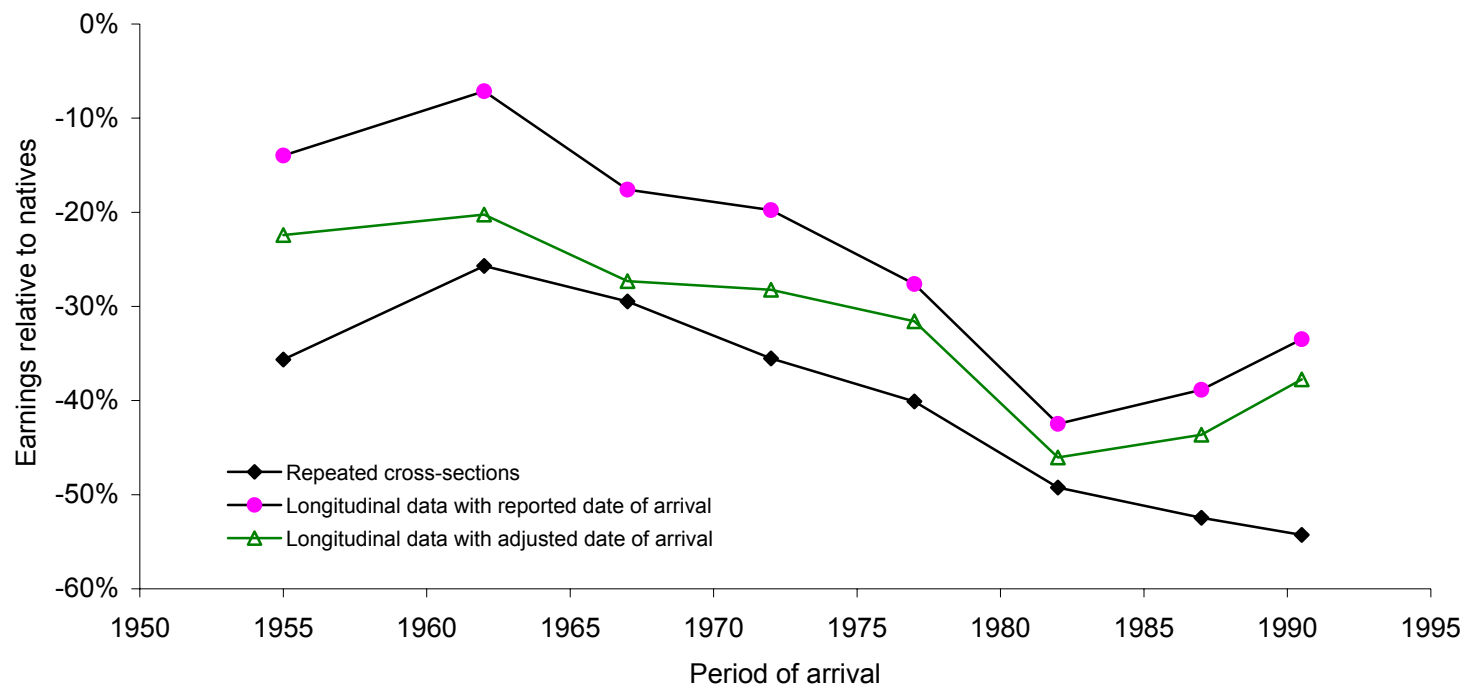
Note: Only observations with positive earnings in the year are included in the calculation of the percent topcoded. Sample includes all men aged 17 to 64 with positive potential experience, born between 1930 and 1969, in each year. See text for additional details.

Figure 2: Immigrant Earnings Growth in Repeated Cross-Sectional and Longitudinal Data



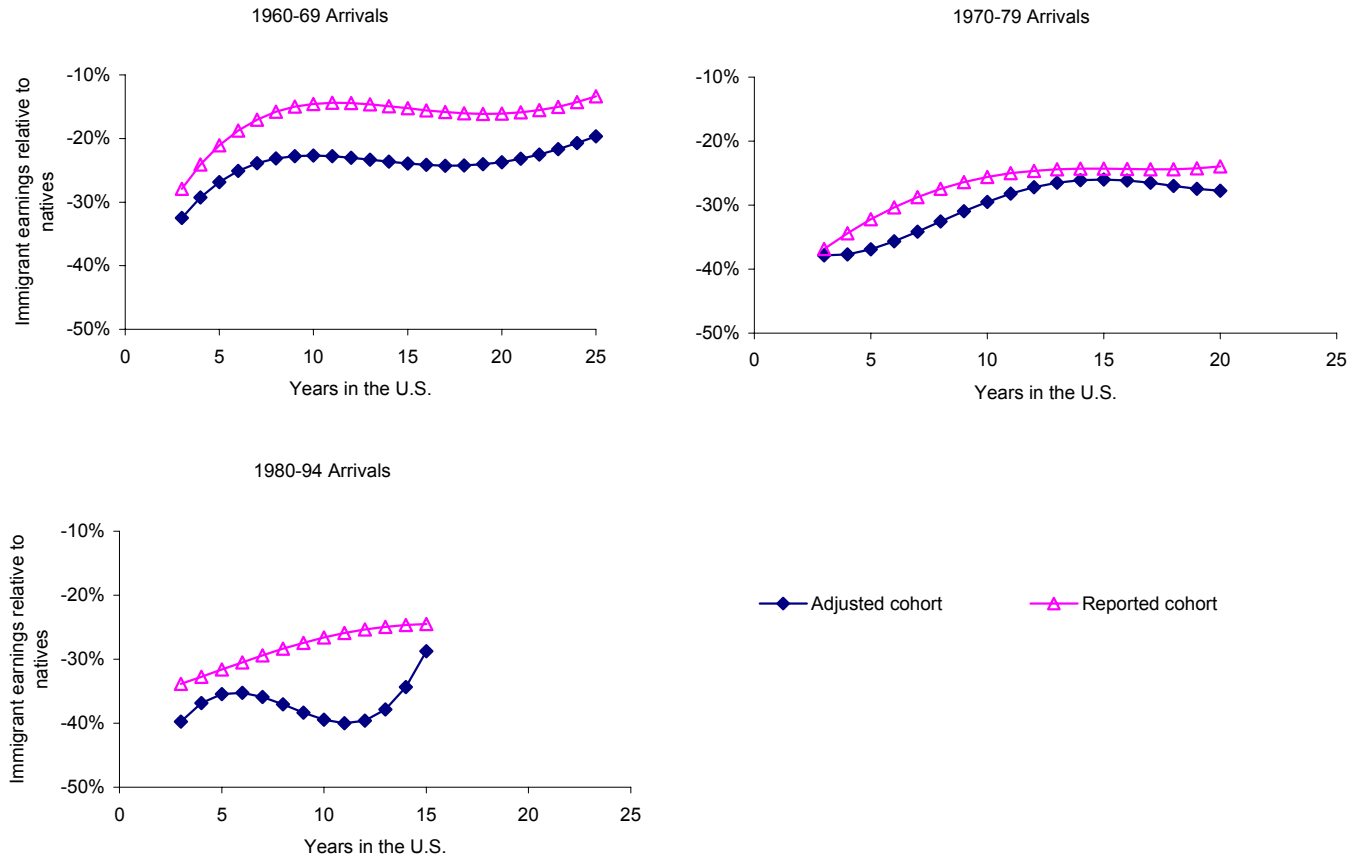
Note: Data points are parameter estimates for the increase in the earnings of immigrants, relative to natives, compared to their earnings during their first five years in the U.S. The model also controls for quartics in potential labor market experience and calendar time, immigrant arrival cohort fixed effects, and immigrants' labor market experience upon entry. The adjusted date of arrival classifies immigrants based on the earlier of their reported date of arrival and their first year of covered earnings. See text and Table 3 for details.

Figure 3: Immigrant Cohort Fixed Effects



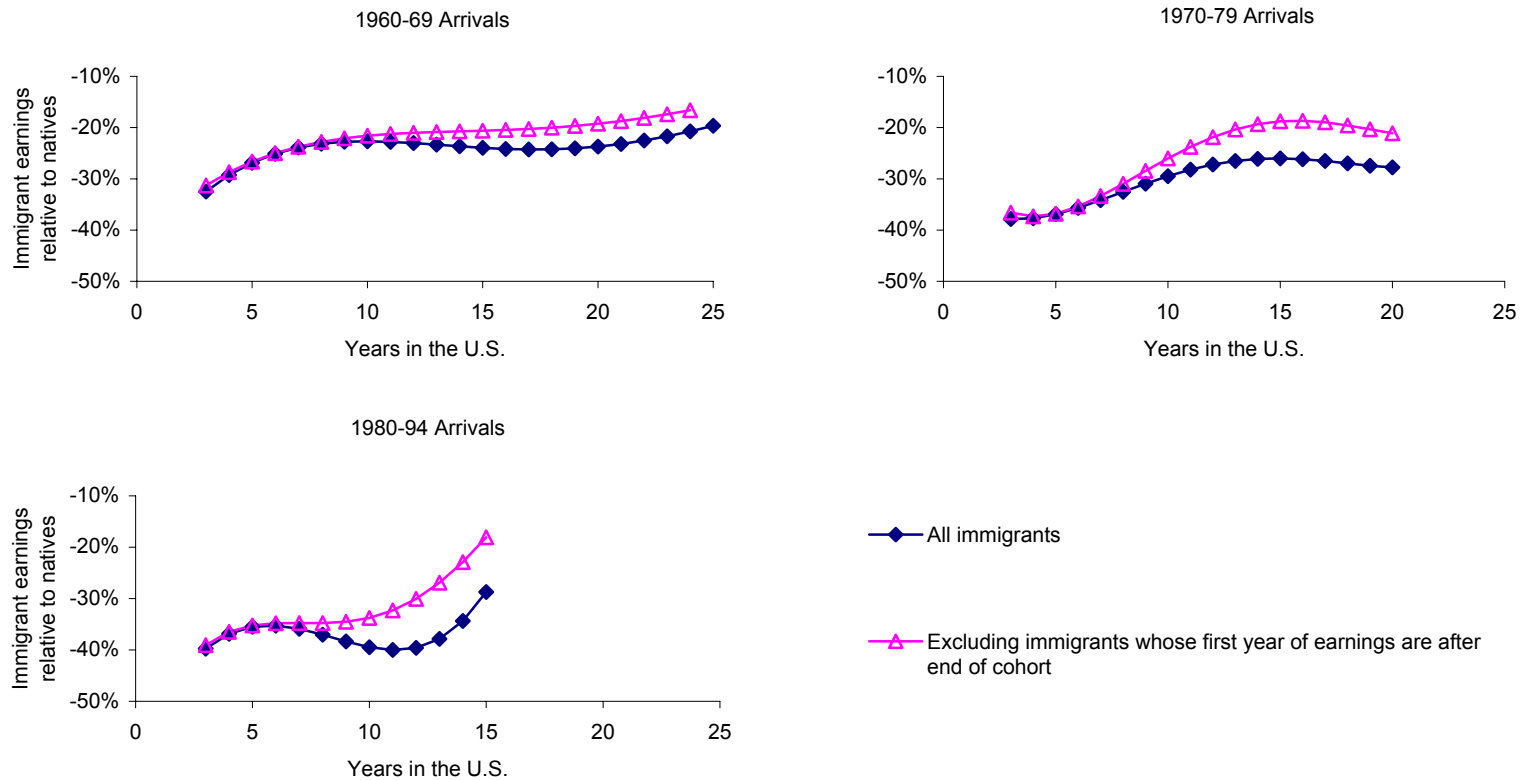
Note: Data points are parameter estimates of immigrant arrival cohort fixed effects. The model also controls for quartics in potential labor market experience and calendar time, immigrants' time in the U.S., and immigrants' labor market experience upon entry. The adjusted date of arrival classifies immigrants based on the earlier of their reported date of arrival and their first year of covered earnings. See text and Table 3 for details.

Figure 4: The Effect of Adjusting Arrival Cohorts  
By Period of entry and cohort definition



Note: Each line represents the predicted median earnings of immigrants relative to native-born workers. The line marked by the black diamonds is based on grouping immigrants by their reported date of arrival. The line marked by the open triangles is based on grouping immigrants by the earlier of their reported date of arrival and their first year of earnings. The models also include controls for total potential experience and calendar time. The levels of the curves are for an immigrant who arrives in the U.S. with five years of potential experience. The sample of natives for each immigrant cohort consists of those who entered the labor market prior to the end of the immigrant arrival cohort. Models are estimated by the method of censored least absolute deviation.

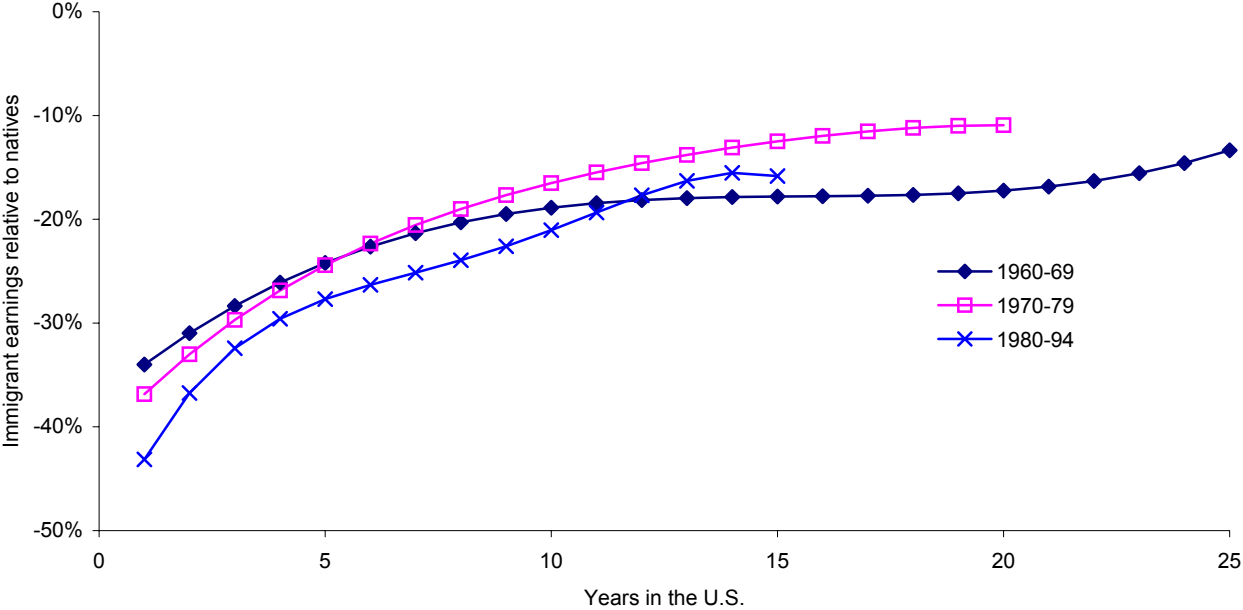
Figure 5: The Effect of Changes in Cohort Composition on Estimated Immigrant Earnings  
 Cohorts defined as the earlier of reported date of entry and first year of covered earnings



Note: Each line represents the predicted median earnings of immigrants relative to native-born workers. The cohorts are defined by the earlier of an immigrant's reported date of entry and their first year of earnings. The line marked by the open triangles is the predicted earnings when immigrants who reported to have arrived in the given time period, but whose first year of earnings was more than three years after the end of the cohort, are excluded from the sample. The models also include controls for total potential experience and calendar time. The levels of the curves are for an immigrant who arrives in the U.S. with five years of potential experience. The sample of natives for each immigrant cohort consists of those who entered the labor market prior to the end of the immigrant arrival cohort. Models are estimated by the method of censored least absolute deviation.

Figure 6: Immigrant Earnings by Period of Arrival

Cohorts defined by first year of covered earnings



Note: Each line represents the predicted median earnings of immigrants relative to native-born workers. The models also include controls for total potential experience and calendar time. The levels of the curves are for an immigrant who arrives in the U.S. with five years of potential experience. The sample of natives for each immigrant cohort consists of those who entered the labor market prior to the end of the immigrant arrival cohort. Models are estimated by the method of censored least absolute deviation.