

# Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings

---

Darren Lubotsky

*University of Illinois at Urbana-Champaign*

I use longitudinal earnings data from Social Security records to study the effect of selective emigration on the measured progress of immigrants to the United States. The immigrant-native earnings gap closes by 10–15 percent during immigrants' first 20 years in the United States, or about half as fast as typical estimates from repeated cross sections of the decennial census. The divergent results indicate that emigration by low-wage immigrants has systematically led past researchers to overestimate the wage progress of immigrants who remain in the United States. Selective back-and-forth migration also leads typical estimates to overstate the measured decline in earnings among successive immigrant arrival cohorts between 1960 and 1980.

## I. Introduction

Immigration is one of the most important and most contentious policy debates in the United States. For many, support for high levels of im-

I especially wish to thank David Card for his advice and encouragement. I also thank John Abowd, Alan Auerbach, Richard Blundell, Kenneth Chay, Todd Elder, Kevin Hallock, Hilary Hoynes, Roger Koenker, Jonathan Leonard, Steven Levitt, Enrico Moretti, James Powell, two anonymous referees, and many seminar participants for their advice. I owe a special thanks to Bhashkar Mazumder, Lynn Riggs, and the staff members at the Center for Economic Studies at the U.S. Census Bureau and at 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, the Institute for Industrial Relations at the University of California, 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. Census Bureau. Research results and conclusions are my own and do not necessarily indicate concurrence by the U.S. Census Bureau, the Center for Economic Studies, or the CCRDC.

[*Journal of Political Economy*, 2007, vol. 115, no. 5]  
© 2007 by The University of Chicago. All rights reserved. 0022-3808/2007/11505-0004\$10.00

TABLE 1  
IMMIGRANT-NATIVE DIFFERENCES IN LOG EARNINGS IN THE  
1970–2000 DECENNIAL CENSUSES

IMMIGRANTS' PERIOD OF ENTRY	YEARS IN THE UNITED STATES	
	1st–5th Years	11th–15th Years
1965–69	–.377 (.013)	–.164 (.006)
1975–79	–.508 (.005)	–.249 (.007)
1985–89	–.545 (.006)	–.358 (.004)
1995–99	–.442 (.004)	. . .

NOTE.—Data are taken from the 1970–2000 decennial Censuses. The earnings gap during the first five years in the United States are computed from separate 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, using natives and recent immigrants, all of whom are aged 25–54. Earnings gaps during the eleventh through fifteenth years are computed from cross-sectional regressions using natives and immigrants who are aged 35–64. Immigrants who arrived before age 18 are excluded from the samples. Standard errors are in parentheses.

migration by lower-wage immigrants depends crucially on the ability of immigrants to assimilate into the U.S. workforce. Despite nearly a century of research, however, many of the basic facts about the immigrant experience in the American labor market remain in dispute, largely because of problems with the U.S. decennial census and other existing data sources. Table 1 shows tabulations from the 1970–2000 decennial Censuses of the average relative earnings of recent immigrant arrival cohorts during their first five years and during their eleventh through fifteenth years in the United States.<sup>1</sup> In the 1970 Census, immigrants who had been in the country for less than five years earned 38 percent less than native-born workers with similar labor market experience; in the 1990 Census, recent immigrants earned 55 percent less than natives. One area of agreement is that shifts in the national origin mix of immigrants away from developed, high-skill countries to industrializing, lower-skill countries have contributed to the decline in the average labor market earnings of immigrants over the past 30 years. 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, immi-

<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. Earnings gaps among new immigrants are computed from samples of men aged 25–54, whereas the earnings gaps among immigrants who have been in the country for 11–15 years are computed from samples of men aged 35–64. Immigrants who entered before the age of 18 are excluded.

grants 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> These disagreements about the degree to which immigrants assimilate into the U.S. labor market contribute to disagreements about whether or not the United States should be as welcoming to lower-skilled individuals as it traditionally has been.

Current policy initiatives and recent policy changes have largely been based on the view that immigrants increasingly enter the United States on the lower rungs of the economic ladder, feeding fears that inflows of unskilled immigrants lower the earnings and employment rate of native-born workers and impose large burdens on public transfer programs and services. For example, 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.<sup>3</sup> The 1996 welfare reform legislation 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 United States. A major policy shift currently being debated is whether the United States should adopt a point system, similar to that in Canada, that would make it easier for higher-skilled foreigners to obtain permanent residency status, perhaps at the expense of less skilled foreigners who qualify for permanent residency because they already have family in the United States.

I present evidence from new longitudinal earnings histories from Social Security records that selective out-migration by low-earning immigrants has led the evidence marshaled in favor of both the optimistic and the pessimistic views of past immigration to be overstated. As many as a third of immigrants to the United States eventually return to their home country. Since direct data on emigration are generally not available, little is known about whether it is the most successful or the least successful immigrants who return and how the selectivity of this flow

<sup>2</sup> For a statement of the former view, see, e.g., LaLonde and Topel (1992) and Duleep and Regets (1996, 1997). The latter view is expressed by Borjas (1985, 1995). Surveys of the economic literature on immigration are given in Borjas (1994, 1999), LaLonde and Topel (1997), and Smith and Edmonston (1997).

<sup>3</sup> By way of comparison, the Immigration Act of 1990 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), Bean and Fix (1992), and Smith and Edmonston (1997) provide additional summaries of U.S. immigration policies.

complicates inferences drawn from data on those who remain in the country.<sup>4</sup>

The earnings gap between native-born workers and a particular immigrant arrival cohort narrows sharply from one decennial census to the next. The earnings gap between natives and immigrants who arrived between 1965 and 1969 fell from 38 percent in the 1970 Census to 16 percent in the 1980 Census; the gap was eliminated in the 1990 Census, when this cohort had been in the United States for 21–25 years. The earnings gap between natives and immigrants who arrived in the late 1980s fell from 55 percent in the 1990 Census to 36 percent in the 2000 Census. Past researchers have interpreted the closing of the earnings gap from one census cross section to the next as evidence of rapid assimilation among first-generation immigrants, bolstering supporters who claim that assimilation lessens the negative fiscal and economic impacts from low-wage immigration. But the earnings gap from one census to the next could also close because of selective emigration from the United States by low-earning immigrants: If low-wage immigrants are more likely to leave the United States, then the earnings gaps measured in the censuses reflect assimilation among immigrants who stay, as well as the presence of fewer low-wage immigrants in the data over time.

Just as there may be selective permanent out-migration, there may also be selective back-and-forth migration between the United States and immigrants' home countries. The census and most other common household surveys, however, ask respondents when they arrived in the United States "to stay," and immigrants who made multiple trips are led to answer with the date of their most recent arrival rather than their initial arrival. Immigrants who report that they recently arrived are therefore disproportionately transient migrants, whereas immigrants who report that they arrived earlier are mainly permanent. If transient immigrants tend to have low earnings or slow earnings growth (which seems likely among migrant laborers or agricultural workers from Mexico), it will appear as though recent cohorts have lower earnings or slower earnings growth than earlier cohorts. The appearance of a downward trend in immigrant skills could also reflect, in part, recent U.S.

<sup>4</sup> Direct evidence is not available on the number of foreign- or American-born people who emigrate from the United States, and thus estimates of the rate of 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 United States 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 1970s had left the United States by 1980. Both studies find that Asian immigrants are less likely to emigrate from the United States than European and Latin American immigrants.

censuses doing a better job of finding and enumerating illegal or lower-skilled immigrants than earlier censuses had. Separating real changes in the skill or earnings level of immigrants over time from changes in census coverage and from churning among low-wage, transient immigrants is vitally important for evaluating claims that U.S. immigration policies should be adjusted to return the skill mix of immigrants back to what supposedly prevailed in earlier decades.

This paper uses a new sample of longitudinal earnings histories that helps overcome some of the limitations of the previous literature on immigrant earnings assimilation. Through a joint project of the Social Security Administration, the Internal Revenue Service, and the Census Bureau, the 1990 and 1991 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 United States.<sup>5</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 out-migration and its effect on measured immigrant earnings profiles can be examined.

Several important new results are found. Most important, the actual earnings growth of immigrants who remain in the United States is considerably slower than that implied by comparisons across decennial censuses. Estimates from the longitudinal earnings records indicate that immigrant earnings grow by about 10–15 percentage points more over their first 20 years in the United States than the earnings growth experienced by native-born workers. This is substantially slower than the 26-percentage-point growth implied by a comparison of immigrants and natives across decennial censuses. Selective emigration by immigrants with below-average earnings is qualitatively important and has systematically led analysts working with census data to overestimate assimilation into the U.S. labor market.

Temporary out-migration by low-wage immigrants also has 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 arrivals and 1985 arrivals is approximately one-third smaller when immigrants are classified by their initial date of arrival rather than their reported date. Thus analysis of the Social Security files

<sup>5</sup> Several previous studies have used longitudinal data to examine immigrant earnings. See Chiswick (1980), Borjas (1989), Hu (1999), and Duleep and Dowhan (2002). These and other related work are discussed below.

indicates that while immigrants do not assimilate nearly as rapidly as census-based estimates suggest, some of the concern that the United States is increasingly attracting lower-wage immigrants is perhaps misplaced.

Section II of the paper describes the matched Social Security earnings data and highlights some of their strengths and weaknesses. Section III describes the relationship between immigrant earnings measured in repeated cross sections from the decennial census and those measured in longitudinal data, in the presence of permanent and temporary selective out-migration. The main estimates are presented in Section IV. Section V presents conclusions.

## **II. How Comparable Are Matched Administrative Earnings Data and Household Surveys?**

The centerpiece of this study is new longitudinal earnings data formed by matching individual survey records from three large household surveys to respondents' annual Social Security earnings records. Specifically, respondents in the 1990 and 1991 SIPPs are linked by Social Security number to annual Social Security earnings records from 1951 to 1993; respondents in the March Supplement to the 1994 CPS are linked to earnings records from 1951 to 1997. The annual Social Security earnings records are employer reports to the Social Security Administration for the purpose of assessing Social Security and Medicare taxes and for determining future Social Security benefits. Though approximately 96 percent of workers are covered by the system today, some groups were not covered throughout the entire sample period of this study. Most self-employed professionals and members of the uniformed services entered the Social Security system between 1954 and 1956, whereas 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 unclear how much of their income goes unreported to the Social Security Administration.<sup>6</sup> In addition to annual earnings in each year, the Social Security records also contain information on the number of quarters of covered employment in each year,

<sup>6</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 primary groups not covered by the Social Security system today are self-employed, domestic, and farm workers who have particularly low earnings; railroad workers; and some employees of state and local governments. Statistics and short histories of key provisions of the Social Security program are given in Social Security Administration (1997, 2006). Studies by Card and Krueger (1993) and Chay (1995) used a similar match of the 1973 and 1978 March CPS to Social Security earnings records to examine black-white earnings differences. Bound and Krueger (1991) use those data to investigate the extent of measurement error in reported earnings in the CPS.

individuals' date of birth, race, and gender. Longitudinal information is not available on whether the individual was self-employed, weeks or hours worked, place of residence, schooling, and job training. Thus the focus of this study is annual earnings among workers, which capture both wage differences and labor supply choices at the intensive margin.<sup>7</sup>

Demographic information from the SIPP or CPS is matched to the longitudinal earnings records. The three household surveys include information on educational attainment, citizenship, and year and place of birth, among other things. Immigrants report the year they arrived in the United States "to stay," and answers are given in intervals: 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 SIPPs are pre-1960, five-year intervals from 1960–64 to 1975–79, 1980–81, 1982–84, and 1985–90 or 1985–91.<sup>8</sup> In what follows, I classify as an immigrant anyone born outside of the United States. People who were born abroad to American parents, were born in Puerto Rico or other outlying areas of the United States, or arrived in the United States prior to age 18 are dropped from the analysis. (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. Their labor market experience may be more similar to that of native-born workers than to that of immigrants who arrive in the United States as adults.)

I further restrict the sample to include only men born between 1930 and 1969. Men born before 1930 would be over 60 years old at the time of the 1990, 1991, or 1994 surveys; excluding these individuals reduces the risk that nonrandom mortality or nonrandom labor force participation decisions 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. I exclude women from the sample to avoid complications from selective labor force participation.

Several features of the matched data pose additional issues. The first problem is that the Social Security earnings are censored at the taxable

<sup>7</sup> Hu's (1999) analysis of longitudinal data from the Health and Retirement Study (HRS) indicates that repeated cross-sectional data may also lead to an overstatement of assimilation in employment rates among non-Hispanic white immigrants.

<sup>8</sup> Much of the analysis that follows consolidates some of the immigrant arrival cohorts. Except in table 2 below, where the CPS data are analyzed separately, the 1980–84 cohort used in this study includes observations from the SIPPs 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, discussed below, 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.

earnings ceiling in each year.<sup>9</sup> Figure 1 plots the real value of the top code (in 1997 dollars) and the fraction of the sample top-coded from 1951 to 1997. Average earnings increased between 1951 and 1965 and the tax ceiling remained relatively stable, which led to a steady rise in the proportion of the sample that is censored from 11 percent to 53 percent. Between 1966 and 1977 the tax ceiling increased several times and inflation eroded its real value, and the fraction of the sample top-coded increased and then decreased. Finally, between 1978 and 1997 the real value of the tax ceiling steadily increased, and thus between 1983 and 1997, about 11 percent of the sample is top-coded. The econometric procedures used below take into account the censored nature of the data.

A second problem is that the longitudinal earnings data record only an individual's earnings that were covered by the Social Security system. An earnings record of zero dollars in a given year may reflect someone who was not employed that year, who was outside of the United States for the year, or whose only earnings came from informal or other uncovered 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 household surveys. It is not possible to distinguish between immigrants who are legal residents of the United States and work in jobs not covered by the Social Security system, and immigrants in the United States legally or illegally who work "under the table" and do not pay taxes on their earnings. Other workers may have earnings in both the covered and uncovered sectors, and thus their earnings in the longitudinal data set are only a portion of their total earnings.

The final issue is that not all respondents in the three household surveys are matched to earnings records. The primary reason for a nonmatch is that an individual (or the proxy respondent) refused to give his Social Security number to the survey interviewer. Social Security numbers that are provided are verified by the Social Security Administration's Enumeration Verification System, a process that ensures survey respondents are matched only to their own Social Security earnings record. The Social Security Administration is also able to look up the correct Social Security number for survey respondents who did not know

<sup>9</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 overwithholding 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.

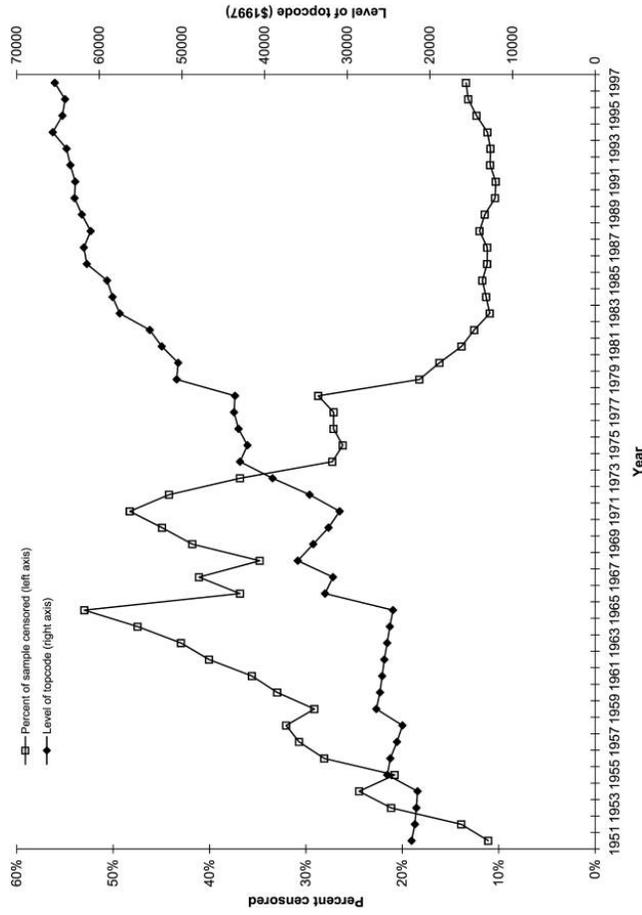


FIG. 1.—Level of top code and percentage of sample censored (Social Security earnings data, 1951–97). Data are taken from Social Security earnings records between 1951 and 1997 matched to individuals in the 1994 CPS and the 1990 and 1991 SIPP. Only observations with positive earnings are included in the calculation of the percentage top-coded in each year. The sample includes all men aged 17–64 who had positive potential experience and were born between 1930 and 1969. The annual top code is the Social Security annual maximum taxable earnings, given in Social Security Administration (2006).

their Social Security number or who provided an incorrect number.<sup>10</sup> Thus it is highly unlikely that an illegal immigrant who bought a Social Security card on the black market would be in the matched sample, much less matched to someone else's earnings history. Eighty-four percent of respondents in the 1994 March CPS are matched to Social Security earnings records. The match rates are 91 percent and 87 percent in the 1990 and 1991 SIPPs. The proportion of immigrants who are matched is lower, particularly among recent arrivals. The match rates in the CPS sample are 76 percent among immigrants who report arriving in the United States between 1950 and 1964, 72 percent among those who report arriving between 1965 and 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 nonrandom matches to Social Security earnings data, population weights are computed for the matched subsample to reflect the observable characteristics of the full sample in the household surveys. Specifically, let  $\omega_i$  denote the population weight provided in the household survey for individual  $i$  and  $p(x_i)$  denote the probability that a person with characteristics  $x_i$  is matched to an earnings record. If that observation is matched and used in the analysis, his final weight is given by  $\omega_i/\hat{p}(x_i)$ , the product of  $\omega_i$  and the inverse of the estimated match probability. The probability of a match is estimated with a logit model in which the match probability is a function of observables recorded in the household survey, including educational attainment; a square in potential labor market experience; weeks and hours worked; a square in reported earnings; and indicators for Hispanics, nonwhites, Hispanic nonwhites, Asians, and those who did not work in the survey year or were self-employed, worked in agriculture, worked for the government, or worked in the private sector.<sup>11</sup> The logit model is estimated separately for each of the three household surveys, for natives, and for five-year immigrant arrival cohorts. Note that since earnings reported in the cross-sectional survey are used to construct the match probabilities, unobservable factors, such as motivation or ability, that are correlated with earnings and

<sup>10</sup> No attempt is made by the Census Bureau or the Social Security Administration to match earnings data for individuals who refused to provide a Social Security number. I exclude from the sample a small number of additional respondents 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 data sources.

<sup>11</sup> 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); the SIPP variables refer to the month prior to the interview. Nevo (2003) analyzes a more general case of using weights to adjust for selection bias.

may be more prevalent among the matched subsample are incorporated to an extent into the final weights.

To gauge how well covered Social Security earnings reflect the more familiar earnings reported in the CPS and to explore how well the final sample weights adjust for nonmatches between the household surveys and the longitudinal earnings data, table 2 compares the level of earnings and the immigrant-native earnings gaps measured in the March Supplement to the 1994 CPS (which records annual earnings from 1993) with the corresponding measures from the 1993 Social Security earnings records. Columns 1 and 2 display unconditional average reported (log) earnings among the full sample of men in the CPS and the subsample of men who are matched to Social Security earnings records, separately for the native-born and immigrant arrival cohorts (as reported in the CPS). Below the level of log earnings is the raw earnings gap between each immigrant cohort and natives, the standard error of the earnings gap, and the (unweighted) sample size.<sup>12</sup>

From CPS earnings data, differences in reported earnings between natives and immigrants who arrived after 1980 are eight to 13 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 (col. 3), 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 disadvantage among the full sample is 19.7 percent and among the reweighted matched subsample is only 6.6 percent.

Column 4 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 (cols. 3 and 5). 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 11 percentage points. On the basis of these comparisons, the

<sup>12</sup> Because reported earnings in the CPS and census surveys are heavily clustered at round numbers, the comparison of medians is problematic. Table 2, therefore, reports means, after CPS earnings are censored at \$57,600, the Social Security taxable maximum in 1993. Top-coded observations in both the CPS and longitudinal earnings records are multiplied by 1.38. This factor approximates the uncensored mean among the artificially top-coded observations in the 1994 CPS. Observations with annual earnings below \$1,000 are dropped from this table.

TABLE 2  
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 (1)	Matched Subsample (2)	Reweightd Subsample (3)	Matched Subsample (4)	Reweightd Subsample (5)
A. 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
B. Immigrant Cohorts					
1950–69:					
Log earnings	10.18	10.18	10.20	10.09	10.13
Earnings gap	2.4%	1.2%	4.0%	3.2%	7.7%
Standard error	(5.2)	(5.5)	(5.2)	(6.5)	(6.6)
Sample size	340	254	254	234	234
1970–74:					
Log earnings	9.96	10.13	10.09	10.06	10.02
Earnings gap	–19.7%	–3.6%	–6.6%	–.2%	–2.7%
Standard error	(6.1)	(6.6)	(6.6)	(7.1)	(7.2)
Sample size	308	204	204	193	193
1975–79:					
Log earnings	9.98	10.01	9.98	9.91	9.89
Earnings gap	–17.8%	–16.1%	–17.6%	–14.9%	–16.2%
Standard error	(5.4)	(5.8)	(5.8)	(6.3)	(6.2)
Sample size	400	284	284	265	265
1980–85:					
Log earnings	9.79	9.87	9.81	9.79	9.73
Earnings gap	–36.6%	–29.5%	–34.8%	–27.1%	–32.1%
Standard error	(3.6)	(4.3)	(4.4)	(5.0)	(5.0)
Sample size	750	517	517	462	462
1986–89:					
Log earnings	9.66	9.74	9.65	9.71	9.65
Earnings gap	–49.6%	–43.0%	–50.6%	–35.1%	–39.7%
Standard error	(4.1)	(5.8)	(5.4)	(6.3)	(5.8)
Sample size	560	320	320	286	286
1990–94:					
Log earnings	9.58	9.71	9.60	9.57	9.49
Earnings gap	–58.2%	–45.3%	–55.5%	–49.1%	–55.5%
Standard error	(5.0)	(6.5)	(6.2)	(7.5)	(7.1)
Sample size	456	254	254	203	203

NOTE.—Data in cols. 1–3 are taken from the 1994 March CPS and reflect individuals' self-reported wage and salary, self-employment, and farm income from 1993. These earnings are top-coded at the 1993 Social Security maximum of \$57,600. Longitudinal data are taken from the 1993 Social Security earnings records among men matched in the 1994 March CPS. Top-coded observations in both data sets 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 natives and each immigrant cohort. Standard errors of the earnings gaps are given in parentheses. The means given in cols. 1, 2, and 4 are weighted by the March Supplement weights. The means in cols. 3 and 5 are weighted by the probability of each individual in the public use file being matched to earnings records, as described in Sec. II. Sample sizes are unweighted. Observations with annual earnings less than \$1,000 (1993 dollars) are dropped.

use of Social Security earnings data, rather than the familiar self-reported earnings in the CPS, does not systematically affect immigrant-native earnings comparisons.

### III. How Out-Migration Affects Immigrant Earnings Measured in Repeated Cross-Sectional and Longitudinal Data

This section begins with a discussion of how permanent out-migration from the United States affects measurement of within- and across-cohort wage changes in repeated cross-sectional data and in longitudinal data, and then summarizes existing empirical evidence on selective out-migration. Next, I describe how temporary out-migration, or migration back and forth between the United States and a migrant's home country, complicates measurement of immigrant wage changes and how data on immigrants' first year of covered earnings in the United States can help paint a more nuanced picture of immigrant wage growth. Finally, the section concludes by comparing sample statistics from the Social Security earnings records and from repeated cross-sectional data that empirically demonstrate the potential importance of selective permanent and temporary out-migration.

#### A. *Permanent Out-Migration*

Before I turn to direct estimates of immigrant earnings growth in longitudinal data, it is important to clarify the potential sources of any differences between earnings growth rates measured from longitudinal data and those from the more traditional repeated cross sections of the decennial census. The first issue is the *permanent* out-migration of immigrants. Permanent out-migration leads to several problems in measuring earnings assimilation. Most important, 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, by contrast, earnings data are available only 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 data sets, let  $E(w_{c,t}|x)$  denote the average earnings among immigrants who arrived in the United States at time  $c$ , measured at time  $t$ , conditional on the sample criterion  $x$ . Panel A of table 3 shows the average earnings of immigrants who arrived in 1967 and 1977, measured

TABLE 3  
MEASURES OF AVERAGE IMMIGRANT EARNINGS BY DATA SOURCE AND YEAR

	YEAR EARNINGS ARE MEASURED		
	1970	1980	1990
A. Earnings Measured in the Repeated Cross Section of the Censuses			
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		$E(w_{77,80}   \text{stay 3 years})$	$E(w_{77,90}   \text{stay 13 years})$
1987 arrivals			$E(w_{87,90}   \text{stay 3 years})$
B. Earnings Measured from Social Security Earnings Records Matched to the 1994 CPS			
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		$E(w_{77,80}   \text{stay 17 years})$	$E(w_{77,90}   \text{stay 17 years})$
1987 arrivals			$E(w_{87,90}   \text{stay 7 years})$

in the 1970, 1980, and 1990 decennial Censuses.<sup>13</sup> The longitudinal file contains data only for immigrants who remained in the United States until 1990, 1991, or 1994, depending on which of the three samples the individual appears in. The corresponding earnings measures from the longitudinal Social Security earnings records matched to the 1994 CPS are given in panel B of table 3. Thus the 1970 earnings of immigrants who arrived in 1967 and remained in the United States 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 United States until 1994 are available in the longitudinal data:  $E(w_{67,70} | \text{stay 27 years})$ .

The U.S. earnings at time  $t$  of an immigrant who arrived at time  $c$  can be decomposed as  $w_{ct} = \mu_{ct} + \varepsilon_{ct}$ , where  $\mu_{ct}$  is the average earnings that the *entire initial* arrival cohort would earn if they remained in the United States until time  $t$ , and  $\varepsilon_{ct}$  is the deviation of an immigrant's earnings from the group average. The average value of  $\varepsilon_{ct}$  is zero when the cohort arrives in the United States. Over time, however, if disproportionately less or more skilled workers leave the United States, then the average value becomes positive or negative.

<sup>13</sup> The 1970–90 Censuses, the CPS, and the SIPP record immigrants' arrival date by an interval (e.g., immigrants who arrived in 1967 would be recorded as arriving in the interval between 1965 and 1969). For simplicity, this discussion assumes that the exact year of arrival is known. Also for simplicity, the discussion that follows compares repeated cross sections of the census to the longitudinal data matched to the 1994 CPS. The actual sample used in the empirical work pools observations from the 1990 and 1991 SIPPs and the 1994 CPS that are matched to longitudinal data.

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

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

That is, their measured earnings growth is equal to the growth among the 1967 arrivals who remained in the United States at least until 1980, plus a bias term equal to the difference in earnings in 1970 between the immigrants who remained in the United States at least until 1970 and the subset who stayed until 1980. If emigration is concentrated among lower-earning immigrants, 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 United States until 1980. In addition, of course, the expected earnings growth of the original cohort if they all remained in the United States 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 provides an unbiased estimate of earnings growth *among the immigrants who are in the United States in 1994*. The growth in earnings of the 1967 arrival cohort between 1970 and 1980 in the longitudinal sample is given by

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

If out-migration is based on permanent earnings characteristics that are not related to immigrants' earnings growth over time (i.e., 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 out-migration. 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 out-migration 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

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

The difference in the last line of this expression will be zero if the 1967 and 1977 cohorts have similar selective out-migration patterns during their first three years in the United States. 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 out-migration 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 out-migration from differences due to the labor market skills of immigrants.

Out-migration also presents problems in making comparisons across cohorts using the longitudinal data. The difference in average earnings three years after entry between the 1967 and 1977 arrivals, measured in the longitudinal sample, is

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

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 out-migration 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 United States, then the additional 10 years of out-migration among the 1967 cohort mean that it will have more “successes” than the 1977 cohort will. This will give the appearance of a downward trend in earnings across successive arrival cohorts in the longitudinal data. This source of bias may be limited, however, for earlier arrival cohorts if most out-migration occurs within the first 10 years after entry.

Whether analyses based on repeated cross-sectional data over- or un-

derestimate actual immigrant earnings growth depends crucially on whether return migrants are systematically less or more skilled than migrants who remain in the United States. A few examples indicate that, in theory, return migrants could be positively or negatively selected: Some migrants plan to remain in the United States permanently but return to their home country because their earnings in the United States were lower than what they had expected. Perhaps counterintuitively, these “unlucky” migrants could come from the top or the bottom of the immigrant earnings distribution. A standard view of Mexican migration to the United States is that lower-skilled Mexicans earn a higher wage in the United States than what they would earn in Mexico, high-skilled workers earn a higher wage in Mexico, and those on the margin of moving tend to be Mexicans from the middle of the skill distribution. Migrants who were close to the margin tend to experience the smallest gains from migrating and are therefore the most likely to return to Mexico if their experiences in the U.S. labor market turn out to be slightly worse than expected. In this scenario return migration to Mexico is positively selected. By contrast, poor luck in the U.S. labor market will tend to generate negative selection among return migrants who come from countries that tend to send their most skilled individuals to the United States.<sup>14</sup>

Many migrants come to the United States with a plan to return to their home country after a limited period of time. This return flow may also be positively or negatively selected: If migrants plan to return after accumulating a target level of savings, success in the U.S. labor market implies that migrants hit their target and return home sooner. On the other hand, success in the United States may increase the return to staying longer and accumulating additional assets. Emigration may also be driven by the expiration of a visa that allows an immigrant to legally work in the United States, and many of these temporary visas are allocated to higher-skilled workers. For example, the H-1B visa program grants a temporary work permit to foreigners who have the equivalent of an American bachelor's degree and work in particular “specialty occupations.” Finally, migrants who face lower costs of migration may be more likely to return home; thus it should not be surprising that emigration rates tend to be quite high for migrants from Mexico, a country that supplies a large fraction of the low-earning migrants in the United States.

Although systematic data are not collected on the characteristics of emigrants from the United States, past research and tabulations in Sec-

<sup>14</sup> Self-selection and the decision to migrate from Mexico to the United States are explored in Chiquiar and Hanson (2005) and Ibarra and Lubotsky (2007). Self-selection among return migrants is modeled more formally in Borjas and Bratsberg (1996).

tion III.C below suggest that emigration is disproportionately more common among low-earning immigrants, and hence traditional measures of immigrant assimilation estimated from repeated cross sections of the census are likely to overestimate earnings growth among migrants who remain in the United States. For example, Chiswick (1980, 1986) compares the distribution of migrants' education, by arrival cohort, in the 1960 and 1970 Censuses and finds evidence that immigrants' educational attainment increased between censuses (especially among migrants from Mexico), which is consistent with selective emigration among less educated migrants or with immigrants acquiring additional schooling after arrival. My tabulations in Section III.C show a similar pattern in more recent census data. Lindstrom and Massey (1994) compare Mexicans in the 1990 U.S. Census with Mexicans surveyed between 1989 and 1991 who live in any of 11 Mexican communities with high rates of migration to the United States. Although there are some difficulties in comparing educational attainment across data sources, Lindstrom and Massey's evidence indicates that Mexicans in the United States are more educated, have higher earnings, and have greater English language ability than the Mexicans surveyed in Mexico who reported that they had moved to the United States and then returned to Mexico.<sup>15</sup>

Despite this limited evidence of selective migration from the United States, previous studies of immigrant earnings growth have not clearly shown whether earnings growth measured in repeated cross sections of the census is accurate. Hu (1999) (described in Hu [2000]) compares estimates of immigrant earnings assimilation from repeated cross sections from the 1970, 1980, and 1990 decennial Censuses and from longitudinal Social Security earnings records between 1951 and 1991 matched to respondents in the HRS. The HRS is a sample of individuals born between 1931 and 1941, and Hu selects a similar birth cohort from the censuses.<sup>16</sup> Hu presents evidence that repeated cross sections of the census overstate assimilation among non-Hispanic white migrants, but not among Hispanic migrants, conclusions similar to those below.

Two other studies are similar in spirit to this one: Chiswick (1980) (described in Chiswick [1986]) compares the relative wage growth of immigrants in the 1970 Census with longitudinal measures of wage growth between 1965 and 1973 in the original cohort of the National Longitudinal Survey of Older Men, respondents to which were aged

<sup>15</sup> Ibarra and Lubotsky (2007) provide evidence that Mexicans in the 2000 U.S. Census appear to overstate their educational attainment.

<sup>16</sup> Since the HRS sample is drawn from a narrowly defined birth cohort, it is difficult to separate the effects of immigrants' age or labor market experience at the time of migration from the effect of their period of arrival. More generally, most of the immigrants in the HRS sample arrived in the United States prior to 1970, and their experiences may be quite different from those of the more contemporary group of immigrants studied here.

45–59 in 1966. The National Longitudinal Survey contains only 98 migrants and does not record their year of arrival to the United States, making a direct comparison with census data difficult. Nevertheless, Chiswick's estimates indicate that immigrant earnings rose by about 1.1 percentage points per year between 1965 and 1973, which is similar to the rate of earnings growth he estimated in the 1970 Census. This similarity is surprising because one would think that falling cohort quality and selective migration would lead to a substantially larger estimate of immigrant wage growth in the census. Borjas (1989) studies earnings growth and emigration using a unique data set that tracked about 50,000 engineers and natural and social scientists between 1972 and 1978. Borjas's estimates indicate that sample attrition—which is most likely caused by emigration from the United States—is more common among lower-earning individuals in this interesting subpopulation of immigrants.

Finally, Duleep and Dowhan (2002) use longitudinal Social Security earnings records matched to the 1994 CPS (one of the three longitudinal files used in this study) and compare the change in the immigrant-native earnings gap at the thirtieth, fortieth, and fiftieth percentiles during the first 10 years after arrival for immigrants who arrived between 1960 and 1983. Although they find that the earnings gap decreases during the first 10 years in the United States, it is not possible to ascertain from their study which types of migrants leave the United States or how selective out-migration influences measures of immigrants' progress since the authors do not compute similar outcomes in repeated cross-sectional data.

In sum, if there is selective emigration, repeated cross sections of the decennial census do not identify wage growth among immigrants who stay in the United States. Past work indicates that low-earning immigrants are probably more likely to emigrate than higher-earning immigrants, and hence past measures of immigrant wage growth from the decennial census overstate assimilation among migrants who actually remain in the country. Previous attempts to accurately measure wage growth have generally relied on quite small samples that are not particularly representative of current immigration, an issue that is remedied in the work below.

#### *B. Temporary Out-Migration and Alternative Arrival Cohort Definitions*

While previous researchers have studied the potential effects of permanent out-migration, a less well-understood but equally important phenomenon is *temporary* out-migration. Many immigrants arrive in the United States, work for several years, return to their home country, and then reimmigrate to the United States at a later time; some migrants make this trip multiple times. Common household surveys are not de-

signed to ascertain when a migrant first arrived in the United States: The 1970–90 decennial Censuses, the CPS, and the SIPP ask the household respondent “When did this person come to the United States to stay?” The 1990 Census respondent instructions read “If the person has entered the United States (that is, the 50 states and the District of Columbia) more than once, fill the circle for the latest year he/she came to stay.” The 1980 Census instructions are somewhat different and read “If the person has entered the U.S. more than once, fill the circle for the year he or she came to stay permanently.” The 1970 Census does not appear to give instructions regarding how to respond for a migrant who has made multiple trips to the United States.<sup>17</sup> It is clear that the year of arrival data in the census, CPS, and SIPP do not necessarily correspond to migrants’ first year of arrival in the United States and potentially correspond to different concepts in different census years.

This misclassification of initial entry cohorts complicates estimation of both the earnings level and growth trajectories of immigrant cohorts derived from repeated cross-sectional data, particularly if the workers who migrate back and forth tend to be lower-skilled. For example, immigrants observed in the 1990 Census who report having arrived in the United States between 1985 and 1989 are composed of those who arrived in the United States for the first time during that period plus some of those who initially arrived earlier, left the United States, and then reentered between 1985 and 1989. The average earnings of those who report that they arrived between 1985 and 1989 may over- or underestimate the earnings of genuine new arrivals, depending on whether the return migrants tend to have higher or lower earnings than the new 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. If, as seems likely, reentrants have slower earnings growth than those who arrive for the first time, estimates of earnings growth will tend to understate the earnings growth of new arrivals, especially among recent cohorts that contain a particularly large number of reentrants. Finally, repeat migration by low-earning immigrants may also influence the trend in relative earnings across successive immigrant arrival cohorts. If the incidence of back-and-forth migration increases over time, or if it becomes more common among the least skilled immigrants, the relative earnings of successive cohorts measured in repeated cross-sectional data will tend to fall over time, all else equal.

An important benefit of using the matched longitudinal earnings data is that information is available on immigrants’ first year of covered

<sup>17</sup> Decennial census enumeration forms and instructions are available for download at <http://www.ipums.org/>. The 2000 Census asked “When did this person come to live in the United States?”

earnings and on their reported date of arrival in the household surveys. Discrepancies between immigrants' reported arrival date in the CPS or SIPP household surveys and their first year of earnings in the longitudinal Social Security data are consistent with a substantial degree of temporary, or back-and-forth, migration. Fourteen percent of immigrants in the longitudinal data have earnings prior to the year of their reported arrival year in the CPS or SIPP household survey. Two pieces of evidence suggest that many of these immigrants had temporarily left the United States 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 in the longitudinal data prior to their 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 United States for an entire calendar year before reimmigrating.<sup>18</sup>

These discrepancies are consistent with evidence from other studies that the census's year of arrival question does not accurately reflect immigrants' initial year of arrival. Ellis and Wright (1998) uncover additional evidence of back-and-forth migration by comparing immigrants' responses in the 1980 and 1990 Censuses to the questions "When did this person arrive in the U.S. to stay?" and "Where did this person live five years ago?" Over 16 percent of male immigrants in the 1980 Census who reported that they arrived between 1975 and 1980 also reported that they lived in the United States on April 1, 1975. This rate increased in the 1990 Census, where over 27 percent of male immigrants who reported that they arrived between 1985 and 1990 also reported that they lived in the United States on April 1, 1985. These discrepancies are most common among low-wage, less educated, and Mexican migrants. Although a small fraction of these migrants arrived between January and March of 1975 or 1985, Ellis and Wright argue that these discrepancies are largely caused by migrants who moved back and forth between their home country and the United States or by migrants who view the date they decided "to stay" in the United States as a distinct concept from when they physically arrived in the United States for the first time.

<sup>18</sup> Disruptions in earnings histories may not be definitive evidence of out-migration 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 United States.

Additional evidence of problems with the census arrival question comes from Redstone and Massey's (2004) analysis of the New Immigrant Survey Pilot. This data set includes 1,134 individuals who were admitted legally to the United States between July and August of 1996. The baseline survey asked respondents to report the date of their first entry, the date of their last entry, and the number of separate trips to the United States, among other things. The one-year follow-up survey asked immigrants to report the year they arrived in the United States to stay, a replica of the decennial census question. Redstone and Massey find that only 55 percent of immigrants respond to the census question with their initial date of arrival.

These discrepancies between immigrants' reported date of arrival in the household surveys and their first year of covered earnings in the longitudinal data raise the question of how to best classify immigrants into arrival cohorts in the empirical analyses. I explore three alternative measures of immigrants' arrival cohort, each of which has different strengths and weaknesses. The first cohort classification groups immigrants by their reported date of arrival, with any earnings prior to that date dropped from the analysis. This classification has the virtue that it is similar to the date of arrival measure available in the repeated cross-sectional data. One obvious drawback of this approach, however, is that for some immigrants the longitudinal data clearly establish an earlier date of initial arrival. This method of constructing arrival cohorts will tend to understate exposure to the U.S. labor market of immigrants who make multiple trips to the United States and report their most recent date of arrival in the survey.

A second drawback of classifying immigrants by their reported date of arrival is that many immigrants do not have earnings in the longitudinal data for some years after their reported arrival date. This likely reflects immigrants who initially worked in uncovered employment, attended school, or simply did not participate in the labor market when they entered the United States. A prominent example is immigrants who were legalized following the Immigration Reform and Control Act (IRCA) of 1986, which granted legal residency to about 3 million people who had been living in the United States illegally (see Bean, Edmonston, and Passel 1990; Smith and Edmonston 1997). Fourteen percent of immigrants who reported in the household surveys that they arrived as adults between 1975 and 1979 did not have covered earnings in the longitudinal data until after 1986; 27 percent of those who reported that they arrived between 1980 and 1985 did not have earnings until after 1986. A lag between the reported date of entry and the first year of covered earnings creates compositional changes in immigrant arrival cohorts over time as new immigrants enter the labor force or move from uncovered to covered work. This compositional change is likely to also

be present among arrival cohorts in census data if immigrants are more likely to participate in the census when they are in the United States legally. These compositional changes in cohorts over time can lead to an underestimate or an overestimate of individual earnings trajectories, depending on whether these delayed earners have relatively high or low earnings.

A second method to form arrival cohorts is to group immigrants by the earlier of their reported date of arrival and their first year of covered earnings. In the empirical work below I refer to this classification as immigrants' "adjusted date of arrival." Although this method also suffers from the problem of compositional changes in immigrant cohorts over time, I use the comparison of average earnings across cohorts when classified by "reported date of arrival" with that estimated when classified by "adjusted date of arrival" to shed light on which immigrants report an arrival date later than their first year of covered earnings. In particular, one of the striking comparisons below is that the average earnings of the 1960s and 1970s arrival cohorts are considerably higher with the reported date of arrival definition than with the adjusted date of arrival definition. This indicates that migrants who report that they arrived in these early cohorts tend to have earnings considerably higher than migrants who had earnings back then, but for whatever reason reported themselves to be more recent arrivals, and suggests that our inferences about the magnitude of changes in cohort quality are quite sensitive to the flow and earnings of these transient, or back-and-forth, migrants.

A final method of grouping immigrants minimizes the problem of compositional changes in cohorts over time by grouping immigrants solely by their first year of covered earnings in the Social Security data. Although this method understates the length of time spent in the United States by immigrants who initially worked in uncovered employment or who invested in U.S.-specific human capital in ways other than through working, it provides perhaps the most easily interpretable picture of immigrant earnings growth because it measures wage growth from the year of entry into the formal, or covered, U.S. labor market.

### *C. A Descriptive Picture of Immigrant Earnings*

Table 4 compares earnings and other characteristics of working immigrants and natives in the longitudinal and cross-sectional data sets in 1979, 1989, and 1995.<sup>19</sup> Two conclusions emerge from this table: first, the earnings and characteristics of immigrant cohorts are sensitive to

<sup>19</sup> I pool observations from a three-year window around 1979, 1989, and 1995 in the longitudinal sample to increase the precision of the estimates. Top-coded observations are multiplied by 1.38 to approximate the mean of the censored observations (see n. 12). All observations with annual earnings less than \$1,000 (1997 dollars) are dropped.

the choice of arrival cohort definition. The evidence is consistent with there being a relatively large number of low-earning immigrants who made multiple trips to the United States and report their most recent date of arrival in the household surveys. The second conclusion is that it is evident in “unconditional” sample means that immigrants’ measured earnings growth is faster in the repeated cross-sectional data than in the longitudinal data, consistent with selective out-migration among lower-earning immigrants.

The first three sets of columns display information from the Social Security records matched to individuals in the 1990 and 1991 SIPP and the 1994 CPS. Columns 1–3 show characteristics of workers measured in 1979, 1989, and 1995, where immigrants are classified by their reported date of arrival in the household surveys. In columns 4–6 immigrants are grouped by their adjusted date of arrival; that is, migrants with earnings prior to their reported period of arrival are moved to the earlier cohort. This simple change in how arrival cohorts are defined increases the number of migrants classified as 1960–74 arrivals and reduces the number classified as arriving between 1975 and 1991. For example, in 1979 there are 912 working migrants who reported that they arrived in the 1960s. Of these, 62 had earnings in the longitudinal data prior to 1960, whereas 148 migrants first appear in the longitudinal data during the 1960s but reported an arrival date of 1970 or later. Once these 210 migrants—or 23 percent of the original cohort—are reclassified, the adjusted 1960–69 arrival cohort has 998 migrants. Similarly, adjusting the 1970–74 cohort increases its size from 869 to 973 migrants, as 186 new migrants are moved into the cohort and 82 migrants are removed from it. These 268 migrants represent over 31 percent of the original cohort. Immigrants who are moved into earlier arrival cohorts tend to be less educated, so changing how arrival cohorts are defined tends to increase the fraction of 1960–74 arrivals who do not have a high school diploma and decreases their measured average earnings, while improving these measures among post-1975 arrivals.

The effects of the 1986 amnesty program are also clearly visible if one compares the sample sizes and characteristics when migrants are grouped by their adjusted date of arrival, in columns 4–6, and by their first year of covered earnings, in columns 7–9. Classifying immigrants solely by their first year of covered earnings increases the sample sizes of the 1985–89 and 1990–91 cohorts, at the expense of earlier cohorts, as migrants who reported that they arrived before the amnesty, but do not have Social Security earnings until after the amnesty, are moved to post-1985 cohorts. For example, there are 1,856 working immigrants in 1989 with an adjusted date of arrival of 1980–84 and another 662 with an adjusted date of arrival of 1985–89. When immigrants are classified solely by their first year of covered earnings, the 1980–84 cohort de-

TABLE 4  
DESCRIPTIVE STATISTICS FOR NATIVES AND IMMIGRANTS BY DATA SOURCE AND COHORT DEFINITION

	LONGITUDINAL EARNINGS DATA						REPEATED CROSS SECTIONS					
	1979 (1)	1989 (2)	1995 (3)	1979 (4)	1989 (5)	1995 (6)	1979 (7)	1989 (8)	1995 (9)	1979 (10)	1989 (11)	1995 (12)
A. Natives												
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%
Observations	61,955	97,469	61,961	61,955	97,469	61,961	61,955	97,469	61,961	160,051	235,312	36,174
B. Immigrant Cohorts by Period of Arrival												
	Reported Arrival Cohort			Adjusted Arrival Cohort			First Year of Covered Earnings			Reported Arrival Cohort		
1960-69:	10.34	10.31	10.20	10.32	10.29	10.18	10.34	10.29	10.16	10.39	10.46	10.23
Log wage	3.6%	7.8%	.6%	2.3%	5.4%	-1.4%	23.8%	6.1%	-2.8%	4.2%	13.7%	-.8%
Gap	39.9	49.9	53.9	40.6	50.4	54.8	45.5	51.0	55.5	40.5	50.3	54.9
Age	33.1%	32.9%	31.4%	35.2%	34.5%	31.4%	26.6%	36.7%	33.6%	37.1%	36.4%	27.2%
No HS	912	859	422	998	958	484	915	854	429	28,348	24,459	318
Observations	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	-19.4%	4.4%	-8.5%	-28.4%	-.9%	-10.9%	1.7%	-.5%	-6.1%	-23.7%	-2.0%	.7%
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	32.4%	33.7%	45.8%	35.2%	35.6%	47.4%	36.7%	34.6%	45.0%	40.5%	42.0%	30.8%
Observations	869	930	477	973	1,067	548	957	972	495	25,972	23,444	334

1975–79:												
Log wage	9.99	10.12	10.03	10.10	10.04	10.19	10.13	9.81	10.17	10.03		
Gap	-30.8%	-10.8%	-16.5%	-12.9%	-15.2%	-4.2%	-6.3%	-53.7%	-15.3%	-21.2%		
Age	32.9	38.7	33.3	38.1	33.3	39.0	44.9	33.4	40.1	45.6		
No HS	25.8%	34.0%	36.6%	34.1%	32.7%	30.2%	25.8%	34.2%	38.5%	35.5%		
Observations	551	1,266	700	1,261	708	972	504	28,533	32,729	536		
1980–84:												
Log wage	9.88	9.88	9.89	9.88	9.90	9.98	10.02	9.93	9.93	9.93		
Gap	-35.4%	-29.8%	-29.8%	-35.6%	-29.7%	-25.1%	-16.8%	-39.1%	-31.0%	-31.0%		
Age	35.5	40.3	40.3	34.8	39.6	36.4	41.7	36.0	42.1	42.1		
No HS	28.57%	34.74%	34.74%	27.94%	35.79%	25.74%	28.40%	38.60%	38.60%	33.00%		
Observations	1,937	1,299	1,299	1,856	1,261	1,431	733	48,795	48,795	929		
1985–89:												
Log wage	9.77	9.87	9.87	9.70	9.82	9.73	9.84	9.70	9.70	9.78		
Gap	-46.6%	-32.6%	-32.6%	-53.7%	-37.1%	-50.6%	-34.9%	-62.1%	-62.1%	-45.4%		
Age	35.5	35.3	35.3	34.4	34.1	33.9	36.0	34.7	34.7	36.9		
No HS	32.1%	33.4%	33.4%	29.9%	35.9%	34.9%	41.0%	34.7%	34.7%	33.3%		
Observations	809	847	847	662	897	1,599	1,344	43,661	43,661	755		
1990–91:												
Log wage	9.75	9.75	9.75	9.77	9.77	9.79	9.79	9.75	9.75	9.75		
Gap	-44.1%	-44.1%	-44.1%	-42.5%	-42.5%	-40.5%	-40.5%	-48.8%	-48.8%	-48.8%		
Age	36.5	36.5	36.5	36.5	36.5	36.2	36.2	37.3	37.3	37.3		
No HS	32.2%	32.2%	32.2%	28.3%	28.3%	37.8%	37.8%	38.0%	38.0%	38.0%		
Observations	355	355	355	282	282	420	420	304	304	304		

NOTE.—Data in cols. 1–9 are taken from Social Security earnings records matched to men in the 1990 and 1991 SIPP and the 1994 March CPS. Cohort definitions are described in Secs. III B and C. Data in cols. 10–12 are taken from the 1980 and 1990 decennial Censuses and the 1995–97 March CPS. Observations from the longitudinal data are weighted by the inverse of the probability of being matched to earnings records, as described in Sec. II. Observations from the 1990 Census and the 1995–97 March CPS are weighted by the sample weights provided in the data. Observations with annual earnings less than \$1,000 (1997 dollars) are dropped. The five statistics shown are the average of the log of annual earnings of group, the difference in log earnings between each immigrant cohort and natives (expressed as a percent), the average age, the percentage of each group without a high school diploma, and the unweighted sample count. The average is computed by first multiplying all top-coded observations by 1.38, which approximates the mean of the uncensored data.

creases to 1,431 and the 1985–89 cohort increases to 1,599. Since the migrants affected by the amnesty tend to be less educated, the fraction of the 1980–84 cohort that has less than a high school diploma falls from 28 percent to 26 percent, whereas the fraction of the 1985–89 cohort that has less than a high school education increases from 30 percent to 35 percent and the fraction of the 1990–91 cohort without a high school education increases from 28 percent to 38 percent.

Finally, columns 10–12 of table 4 show similar measures of earnings, age, and educational attainment measured in the 1980 and 1990 decennial Censuses and the pooled 1995–97 March CPS (which records annual earnings in 1994–96). These samples are described in more detail in the following section. The immigrant-native earnings gaps experienced by each cohort tend to fall over time in both the longitudinal data and the repeated cross-sectional data, but the rate of convergence is generally faster in the repeated cross-sectional 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. The comparable change for this cohort in the longitudinal data was only three to four percentage points, depending on the cohort definition. Similarly, the measured earnings gap among the 1975–79 arrivals narrowed by 38 percentage points between the 1980 and 1990 Censuses, whereas the corresponding change measured in the longitudinal data was between 14 and 22 percentage points, depending on the arrival cohort definition. This faster rate of earnings growth in repeated cross sections of the census is consistent with that data source conflating selective out-migration among low-earning immigrants and earnings growth among immigrants who remain in the United States. The following section more formally compares measures of earnings growth in repeated cross-sectional and longitudinal data.

#### **IV. Direct Estimates of Immigrant-Native Earnings Differences**

This section uses a median regression framework to do three things: First, I compare the relative earnings growth of immigrants in the longitudinal data with growth measured in repeated cross-sectional data drawn from the 1970–90 decennial Censuses and the 1995–97 CPS. Subsection *A* establishes the two main conclusions of this study: Selective out-migration by low-earning immigrants leads repeated cross-sectional data to overstate the relative earnings growth among immigrants who remain in the United States. The presence of low-earning immigrants who move back and forth between their home country and the United States, but report their more recent dates of arrival, leads to an overestimate of the secular deterioration in the level of earnings of newly arrived immigrants. Subsection *B* examines earnings progress among

Hispanic immigrants to the United States. Finally, in subsection *C*, I separately study the earnings progress of immigrants who arrived in the 1960s, the 1970s, and between 1980 and 1994.

*A. Immigrant Earnings in Repeated Cross Sections and Longitudinal Data*

In this subsection I contrast estimates of immigrant-native earnings gaps using two samples. The first sample is taken from the longitudinal Social Security records matched to respondents in the 1990 and 1991 SIPPs and the 1994 CPS (described in detail in Sec. II). The second sample is taken from repeated cross-sectional data and includes men drawn from the 1970 decennial Census 5 percent 1-in-100 state sample, the 1980 decennial Census 5 percent A sample, the 1990 decennial Census 5 percent sample, and the 1995–97 March CPS.<sup>20</sup> To match the end-points of the repeated cross-sectional sample, the longitudinal sample includes earnings between 1969 and 1997. Both samples include only men born between 1930 and 1969 and person-year observations with positive earnings and in which the individual was aged 25 or older. Earnings from the repeated cross-sectional files are measured as the sum of wage and salary, self-employment, and farm income and are artificially censored at the Social Security taxable maximum. Finally, immigrants who arrived after 1991 are excluded from both samples.

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's (1985, 1995) cohort-based analyses of the 1970–90 Censuses. Natives' log earnings are specified as a function of potential labor market experience and calendar time effects. Immigrants' earnings are further decomposed into the effects of arrival cohort ( $k_c$ ), potential experience at arrival ( $m_{ic}$ ), and the length of time since arrival in the United States ( $y_{ct}$ ).<sup>21</sup> Formally, 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

<sup>20</sup> Respondents in the 1995–97 March CPS files who were in the fifth through eighth rotation groups are dropped. This simplifies inferences because in principle households are surveyed in two consecutive March supplements. I use a 25 percent random sample of the natives in the longitudinal data and a 10 percent random sample of the natives in the 1980 and 1990 Census files to reduce the computational burden of the semiparametric models in this section.

<sup>21</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. Other examples of similar models of immigrant earnings are given by Funkhouser and Trejo (1995), Carliner (1996), Schoeni, McCarthy, and Vernez (1996), Schoeni (1997), and Hu (1999).

immigrant arrival cohorts, and  $t$  indexes time. The terms  $f(\text{Experience}_{ict})$  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>22</sup>

The immigrant arrival cohort effects ( $k_t$ ) are given by indicators for immigrants who arrived in 1950–59, 1960–64, 1965–69, 1970–74, 1975–79, 1980–84, 1985–89, and 1990–91 (the native-born are the excluded group). The effect of experience at arrival is captured by indicators for immigrants who arrived with 6–10, 11–15, 16–20, or more than 21 years of potential experience (immigrants who arrived with five or fewer years of experience are the excluded group). Finally, the effect of time in the United States is captured by indicators that an immigrant has been in the United States for 6–10, 11–15, 16–20, 21–25, 26–30, and more than 30 years (immigrants who have been in the United States fewer than five years are the excluded group).<sup>23</sup>

The standard interpretation of equation (5) is that the estimates capture differences in the labor market skills between immigrants and natives. This interpretation may be problematic, however, because in order to separately identify the effects on earnings of immigrants' period of arrival, experience at arrival, and time spent in the United States, the effects of experience and calendar time must be constrained to be equal among immigrants and natives. While these restrictions are somewhat standard in the literature, there are reasons to believe that they may be problematic. Most important, the assumption of common effects of calendar time is at odds with work by LaLonde and Topel (1992), Borjas (1995), and Lubotsky (2007), who demonstrate that widening wage inequality during the 1970s through 1990s reduced the relative earnings of migrants, as it did for lower-wage workers more generally. Those three studies propose alternative methods to separate immigrant-native differences in skills from differences in returns to skills. I do not make a similar adjustment in this paper since I focus on understanding differences in measured wage growth between longitudinal and repeated cross-sectional data during a common time period. An identical adjustment for changes in the structure of wages in both the longitudinal

<sup>22</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 postcollege education are assumed to have entered the labor market at ages 19, 20, 22, and 24. In addition, the calendar time effects for the repeated cross-sectional sample are simply a set of indicators for which cross section the observation is drawn from.

<sup>23</sup> Since period of arrival is reported with an interval in the household surveys, time in the United States and experience at arrival are measured from the midpoint of the arrival interval, where necessary.

and repeated cross-sectional data will not have any effect on the difference in estimates between the two samples.

The assumption of a common experience-earnings profile may also be problematic. Baker and Benjamin (1997) show that family composition affects immigrant husbands and wives' labor supply decisions. This suggests that differences in life cycle labor supply behavior may differ between immigrants and natives and thereby generate differences in their experience-earnings profiles.

In contrast to most previous studies, I do not control for individuals' educational attainment. This study focuses on accurately measuring unconditional differences in earnings between immigrants and natives and on better understanding how selective out-migration influences conclusions about immigrant assimilation. The immigrant arrival cohort fixed effects will therefore reflect differences between the native-born and successive arrival cohorts in educational attainment, as well as unobserved or unmeasured abilities (such as language skills).<sup>24</sup>

Because earnings are top-coded, ordinary least squares estimation of equation (5) is inappropriate. Instead, the model is estimated with Powell's (1984) semiparametric censored least absolute deviation (CLAD) 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>25</sup>

The standard errors are estimated from 50 bootstrap replications of a least absolute deviation, or median, regression using only observations not predicted to be censored using the final parameter estimates. In

<sup>24</sup> Excluding education from the model also has the virtue of avoiding several other potential difficulties. First, schooling is reported in different years in the two samples: schooling is reported in the longitudinal data at the time of the associated household survey (i.e., in 1990, 1991, or 1994), whereas it is reported contemporaneously in the repeated cross-sectional data (i.e., in 1970, 1980, 1990, or in 1995–97). Second, surveys conducted prior to 1990 asked respondents how many years of schooling they have acquired, whereas later surveys ask a degree-based question.

<sup>25</sup> Suppose that 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 top code in year  $t$ . The CLAD estimator is the value of  $\hat{\beta}$  that minimizes

$$\frac{1}{N} \sum_{i=1}^N |w_{it}^* - \min(x'_{it}\hat{\beta}, \bar{w}_t)|.$$

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, lognormal, or Weibull distribution). Chay and Honoré (1998) investigate the relative performance of 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. Deaton (1997) also discusses the CLAD estimator, and Biliias, Chen, and Ying (2000) discuss the bootstrap standard errors.

the longitudinal data, each replication contains a 50 percent random sample of individuals drawn (with replacement) from all persons who appear in the data at any time. All longitudinal earnings records associated with these individuals are included in the bootstrapped data set. In the repeated cross-sectional data, each replication contains a simple 50 percent random sample drawn (with replacement) from all observations. Equation (5) is estimated for each of the 50 data sets, and the standard deviation of the 50 parameter estimates (divided by the square root of two) is an unbiased estimate of the standard error of the parameter values.

One of the central results in the paper is that measured immigrant earnings growth is considerably faster in the repeated cross-sectional data than in the longitudinal data, regardless of how immigrant arrival cohorts are defined in the longitudinal data. This is apparent in table 5, which reports coefficient estimates of equation (5) based first on the repeated cross-sectional data and then on the longitudinal data with the three definitions of immigrants' arrival cohort. For convenience, the time in the United States and cohort effects ( $\hat{\gamma}_{ct}$  and  $\hat{k}_c$ ) are graphed in figures 2 and 3. The cross-sectional data suggest that immigrants' relative earnings grow by 20 percent in their first 10 years in the United States and by an additional 10–20 percent in each successive decade. By contrast, in the longitudinal data, immigrants' earnings grow by 12–15 percent during their first 15 years in the United States, but relatively little after that. The difference in earnings trajectories is consistent with low-earning immigrants being disproportionately more likely to emigrate from the United States over time. Since the repeated cross-sectional model is based on 280,411 immigrants and 598,515 natives, the coefficients in this model are quite precisely estimated, with standard errors on the order of two percentage points. Although the estimates from the longitudinal data are a bit less precise, the differences in estimates between the repeated cross-sectional data and the longitudinal data are generally statistically significant at conventional levels. For example, according to the cross-sectional data, the earnings gap between natives and immigrants closes by 33.5 percent during the first 21–25 years immigrants have been in the United States. This overestimates the gap found in longitudinal data by 15.4–20.5 percentage points, depending on the definition of the immigrant arrival cohort. The standard error of the difference is 3.9–5.1 percentage points, depending on cohort definition, and so the difference is clearly statistically significant.<sup>26</sup>

<sup>26</sup> The standard error of the difference in point estimates is based on the assumption that the two estimates are independent of one another. Another way to see that the estimates in the longitudinal data are statistically different from those in the repeated cross-sectional data is to note that the 95 percent confidence intervals generally do not overlap.

TABLE 5  
CLAD ESTIMATES OF IMMIGRANT EARNINGS GROWTH IN REPEATED CROSS-SECTIONAL  
AND LONGITUDINAL DATA

	REPEATED CROSS- SECTIONAL DATA (1)	LONGITUDINAL EARNINGS DATA, BY AR- RIVAL COHORT DEFINITION		
		Reported Period of Arrival (2)	Adjusted Period of Arrival (3)	First Year of Covered Earnings (4)
Years in the United States:				
6–10	.132 (.018)	.080 (.026)	.086 (.026)	.060 (.025)
11–15	.215 (.008)	.119 (.031)	.148 (.037)	.145 (.032)
16–20	.259 (.020)	.104 (.035)	.139 (.043)	.158 (.044)
21–25	.335 (.013)	.131 (.039)	.176 (.051)	.181 (.046)
26–30	.387 (.023)	.181 (.052)	.195 (.054)	.205 (.059)
31+	.534 (.027)	.308 (.095)	.284 (.131)	.128 (.119)
Experience at entry:				
6–10	-.111 (.005)	-.084 (.044)	-.076 (.044)	-.092 (.058)
11–15	-.164 (.005)	-.213 (.062)	-.213 (.071)	-.204 (.064)
16–20	-.215 (.007)	-.196 (.083)	-.160 (.090)	-.180 (.088)
21+	-.244 (.008)	-.303 (.088)	-.251 (.127)	-.314 (.106)
Period of arrival:				
1950–59	-.356 (.025)	-.140 (.076)	-.224 (.108)	-.194 (.073)
1960–64	-.257 (.020)	-.071 (.066)	-.202 (.059)	-.172 (.088)
1965–69	-.295 (.009)	-.176 (.060)	-.273 (.072)	-.265 (.060)
1970–74	-.355 (.020)	-.198 (.061)	-.282 (.056)	-.249 (.070)
1975–79	-.401 (.007)	-.276 (.057)	-.316 (.066)	-.216 (.075)
1980–84	-.492 (.019)	-.425 (.058)	-.460 (.049)	-.292 (.062)
1985–89	-.524 (.008)	-.389 (.065)	-.436 (.054)	-.407 (.046)
1990–91	-.543 (.067)	-.335 (.100)	-.378 (.176)	-.394 (.067)

NOTE.—The dependent variable is the log of annual earnings. Parameters are estimated by the method of censored least absolute deviation. In addition to the variables listed above, all models also include quartics in potential experience and calendar time. Repeated cross-sectional data are taken from the pooled 1970–90 decennial Censuses and the 1995–97 March CPS. Longitudinal data include earnings between 1969 and 1997 among men matched in the 1990 and 1991 SIPP and the 1994 March CPS. Additional information about sample construction is given in the text. Standard errors are based on 50 bootstrap replications, as described in the text. The repeated cross-sectional model uses 280,411 immigrants and 598,515 natives; the longitudinal models use about 3,069 immigrants and 10,772 natives, depending on the specification. Observations from the longitudinal data are weighted by the inverse of the probability of being matched to earnings records, as described in Sec. II. Observations from the repeated cross-sectional data are weighted by the sample weights provided in the data.

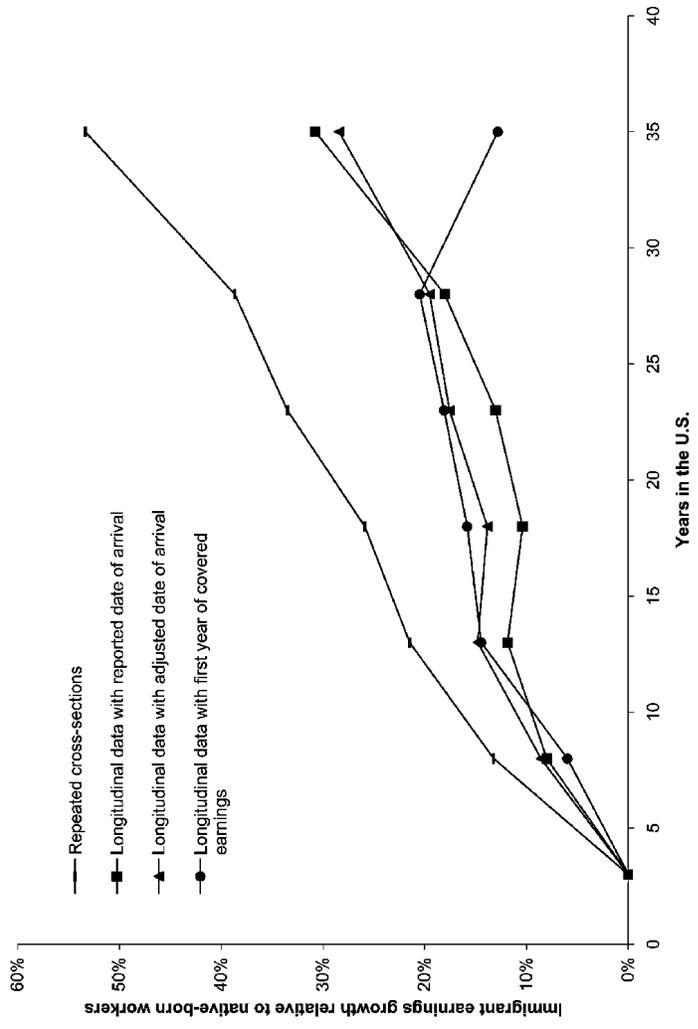


FIG. 2.—Immigrant earnings growth in repeated cross-sectional and longitudinal data. The figure plots the effect of immigrants' time in the United States on the immigrant-native earnings gap. Data are taken from estimates in table 5.

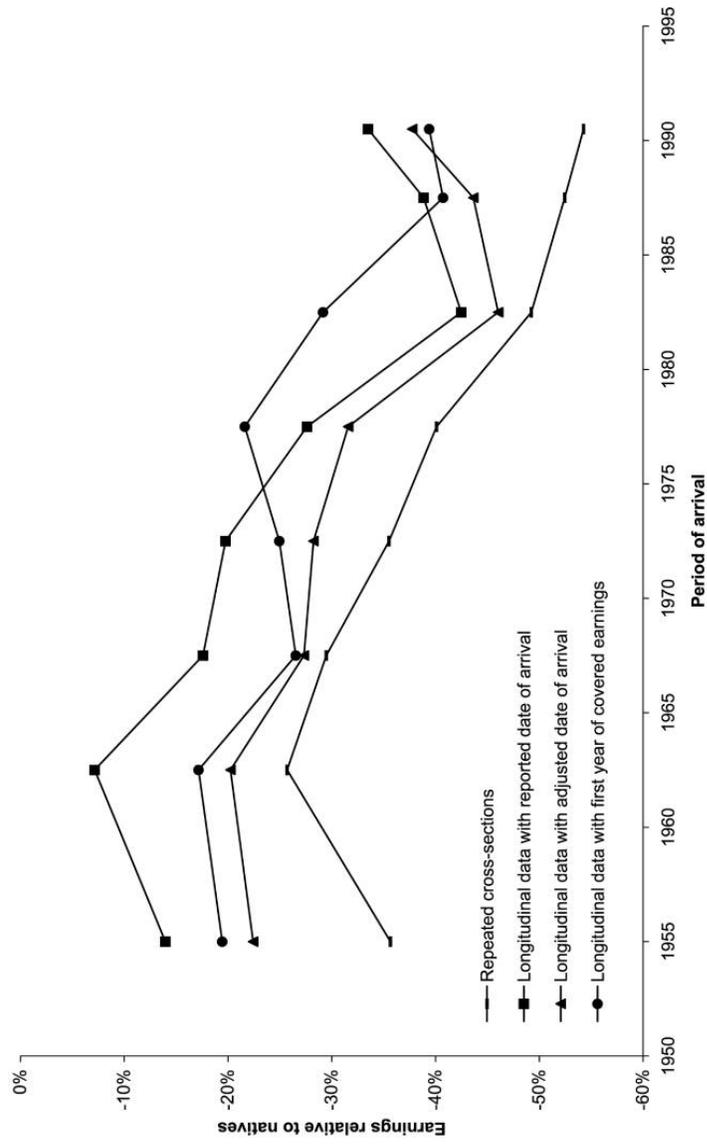


FIG. 3.—Immigrant arrival cohort fixed effects. The figure shows immigrant arrival cohort fixed effects; estimates are taken from table 5

Consistent with past research, both the repeated cross-sectional and the longitudinal earnings data indicate a marked decrease in immigrant arrival cohort fixed effects between 1960–64 and 1980–84 (see, e.g., Borjas 1985, 1995; LaLonde and Topel 1992; Funkhouser and Trejo 1995; Carliner 1996). In the repeated cross-sectional data, the level of earnings among immigrants who arrived between 1960 and 1964 was 26 percent below that of native-born workers. This gap increased to 49 percent among immigrants who arrived between 1980 and 1984. The arrival cohort fixed effects are smaller (in absolute value) in the longitudinal sample than in the repeated cross-sectional data, regardless of which cohort definition is used in the longitudinal data. This pattern is consistent with out-migration by the least successful immigrants over time. Indeed, if year of arrival were accurately measured in the repeated cross-sectional data, the divergence between the fixed-effect estimates from the longitudinal and repeated cross-sectional data would be a useful measure of the extent of nonrandom out-migration across cohorts.

Adjusting immigrants' arrival cohort for those who had earnings prior to their reported date of arrival lessens the secular decline in arrival cohort fixed effects. For example, the earnings of immigrants in the longitudinal data who report they arrived between 1960 and 1964 were 7.1 percentage points less than those of natives. If migrants who had earnings in the Social Security file prior to their reported date of arrival are reclassified into the arrival cohort of their first year of earnings (i.e., their adjusted date of arrival), the 1960–64 cohort fixed effect falls from  $-7.1$  to  $-20.2$ . Indeed, making this adjustment reduces the secular decline in the estimated cohort fixed effects between 1960–64 and 1980–84 by about one-third. This is consistent with there being a large number of low-wage immigrants who entered in the 1960s and 1970s, left for some time, and then reentered and are recorded in the data as having arrived in the 1970s and 1980s. Once these migrants are assigned back to their first arrival cohort, the apparent labor market "quality" of earlier immigrant cohorts is lowered.

The relative earnings of immigrants who arrived in the late 1970s and early 1980s are quite sensitive to how arrival cohorts are defined. Recall that many migrants do not have earnings in the longitudinal data until substantially later than their reported date of arrival, perhaps appearing only after they begin working in the formal sector. This is especially true for migrants who reported that they arrived between 1975 and 1984, many of whom do not have earnings until the 1986 IRCA granted citizenship to many illegal migrants. Since these migrants tend to be low-skilled and have low earnings, reclassifying them as post-1986 migrants raises the measured quality of the 1975–79 and 1980–85 cohorts, as seen by a comparison of the arrival cohort fixed effects when immigrants are

classified by their adjusted date of arrival and their first year of covered earnings.

The longitudinal earnings data (using the reported or adjusted arrival cohort definitions) in figure 3 also indicate a significant upsurge in the entry earnings among immigrants who arrived between 1985 and 1991. 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 1980s, 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. Finally, Borjas and Friedberg (2006) analyze data from the 1960–2000 decennial Censuses and find an upturn among migrants who arrived in the 1990s. Borjas and Friedberg attribute the relative increase in earnings among migrants who arrived in the 1990s to the increased availability of H1-B visas, which led to an increase in the number of higher-skilled migrants, and to the relative decline in earnings among low-skilled natives.

#### *B. Assimilation among Hispanic Migrants*

There are several important reasons for measuring earnings growth and changes in cohort quality specifically for Hispanic migrants to the United States. First, Hispanic immigration to the United States has grown from 30 percent of total immigration in 1970 to 45 percent in 2000, with the large bulk coming from Mexico. Along with this rising share has come an increase in public debate over illegal immigration and calls for changes to immigration policy. Some authors argue that Hispanic immigrants, and in particular those from Mexico, are increasingly less skilled than other immigrants and are also less likely to assimilate into the U.S. labor force (see, e.g., Borjas and Katz 2007). Accurately measuring the level and growth rate of earnings for this important subset of current migration is important for setting policy, especially policy toward Mexican immigration. Second, Mexican and other Hispanic immigrants have a particularly high rate of emigration, which suggests the possibility for a particularly large bias in earnings models that use repeated cross-sectional data. It bears noting, however, that the magnitude of bias depends on both the level of emigration and the degree of selectivity of the emigrants. There would be no bias from using repeated cross-sectional data to estimate earnings growth among a group with high, but random, emigration.

Table 6 shows estimates of equation (5) that include only immigrants who identify themselves as being Hispanic or of Spanish ethnicity plus

TABLE 6  
CLAD ESTIMATES OF HISPANIC IMMIGRANT EARNINGS GROWTH IN  
REPEATED CROSS-SECTIONAL AND LONGITUDINAL DATA

	Repeated Cross- Sectional Data (1)	Longitudinal Data (2)
Years in the United States:		
6–10	.104 (.028)	.069 (.036)
11–15	.095 (.011)	.082 (.047)
16–20	.101 (.029)	.051 (.058)
21–25	.144 (.019)	.077 (.058)
26–30	.209 (.035)	.136 (.073)
31+	.220 (.044)	.251 (.201)
Experience at entry:		
6–10	–.113 (.007)	–.161 (.062)
11–15	–.186 (.007)	–.182 (.081)
16–20	–.235 (.010)	–.377 (.135)
21+	–.251 (.010)	–.403 (.157)
Period of arrival:		
1950–59	–.467 (.037)	–.420 (.155)
1960–64	–.364 (.031)	–.357 (.076)
1965–69	–.436 (.014)	–.447 (.093)
1970–74	–.525 (.028)	–.492 (.076)
1975–79	–.634 (.010)	–.529 (.110)
1980–84	–.712 (.029)	–.624 (.067)
1985–89	–.780 (.008)	–.496 (.067)
1990–91	–.752 (.079)	–.439 (.193)

NOTE.—The dependent variable is the log of annual earnings. Parameters are estimated by the method of censored least absolute deviation. In addition to the estimates reported above, all models also include quartics in potential experience and calendar time. Repeated cross-sectional data are taken from the pooled 1970–90 decennial Censuses and the 1995–97 March CPS. Longitudinal data include earnings between 1969 and 1997 among men matched in the 1990 and 1991 SIPP and the 1994 March CPS. The sample of immigrants includes only men who report themselves to be of Hispanic or Spanish ethnicity. Immigrants are classified by their adjusted period of arrival. Additional information about sample construction is given in Sec. IV.A. Standard errors in the repeated cross-sectional data are based on 50 bootstrap replications; standard errors in the longitudinal data are approximated as described in this subsection. Observations from the longitudinal data are weighted by the inverse of the probability of being matched to earnings records, as described in Sec. II.

all native-born workers.<sup>27</sup> Migrants in the longitudinal data are grouped by their adjusted period of arrival. The longitudinal data reveal that the wage gap between Hispanic immigrants and U.S.-born workers closes by 6.9 percentage points during the first decade migrants are in the United States but does not close appreciably more during the following 15 years. The point estimates suggest some additional assimilation beyond 25 years in the United States, though these estimates are estimated quite imprecisely and are not statistically different from zero (or from the earnings level that prevailed during the first 25 years in the United

<sup>27</sup> Hispanic/Spanish ethnicity comes from respondent reports in the 1994 CPS or the 1990 or 1991 SIPP for those in the longitudinal sample and from respondent reports in the 1970, 1980, or 1990 Census or the 1995–97 CPS in the repeated cross-sectional sample. Those with imputed Hispanic/Spanish ethnicity are included in the sample.

States).<sup>28</sup> Earnings assimilation among Hispanics is therefore considerably slower than that of non-Hispanic migrants.

The rate of assimilation among Hispanic migrants in the longitudinal data is slightly slower than that measured in the repeated cross-sectional data. For example, Hispanic migrants in the repeated cross-sectional data who have been in the United States for six to 10 years have closed the earnings gap between themselves and natives by 10.4 percentage points, with a standard error of 2.8 percentage points. The corresponding estimate of earnings growth in the longitudinal data is 6.9 percentage points, 3.4 percentage points less than in the repeated cross-sectional data, with a standard error of 3.6 percentage points. Although the standard errors on these point estimates cannot rule out that the assimilation rates are equal in the two data sets and the difference in estimates between the two data sources is small in magnitude compared to the differences observed among all migrants, the pattern of point estimates is consistent with out-migration being more common among below-average-earning Hispanic migrants.<sup>29</sup>

The cohort effects in the longitudinal data indicate that the level of relative earnings of Hispanic migrants declined between 1960 and 1985, but then began to rise in the late 1980s. The cohort effect among the 1960–64 arrivals is  $-0.357$ . The cohort effect falls to a low of  $-0.624$  for the 1980–84 arrivals, before rebounding to  $-0.496$  among the 1985–89 arrivals. Migrants who arrived between 1990 and 1991 have relative earnings nearly equal to those of their counterparts who arrived in 1965–69. This dramatic increase in relative earnings among Hispanic migrants mirrors that found by Borjas and Friedberg (2006), though the increase appears to begin a few years earlier in the longitudinal data than in their census data.

<sup>28</sup> Restrictions on accessing the confidential data prevent me from computing bootstrap standard errors for estimates based on longitudinal data in table 6. Instead, the standard errors for Hispanics in the longitudinal data are approximated by

$$\sigma_{\text{Long,Hisp}} = \hat{\sigma}_{\text{Long,All}} \left( \frac{\hat{\sigma}_{\text{Cross,Hisp}}}{\hat{\sigma}_{\text{Cross,All}}} \right),$$

where  $\hat{\sigma}_{\text{Long,All}}$  is the standard error of the parameter estimates using all immigrants in the longitudinal data (col. 3 of table 5),  $\hat{\sigma}_{\text{Cross,Hisp}}$  is the standard error of the parameter estimates using Hispanic immigrants in the repeated cross-sectional data (col. 1 of table 6), and  $\hat{\sigma}_{\text{Cross,All}}$  is the standard error of the parameter estimates using all immigrants in the repeated cross-sectional data (col. 1 of table 5).

Since Hispanic migrants tend to earn less than non-Hispanic migrants, further analysis of the data is required to determine whether there is selective emigration within the non-Hispanic population or whether selective emigration overall is primarily driven by greater emigration among Hispanics. Unfortunately, data access restrictions also prevent me from addressing this further.

<sup>29</sup> Hu (1999, 2000) also finds similar assimilation profiles among Hispanics in his longitudinal and repeated cross-sectional samples.

*C. Longitudinal Estimates of Immigrant Earnings across Cohorts*

This subsection examines immigrant-native earnings gaps separately for three immigrant arrival cohorts, those who arrived in the 1960s, the 1970s, and between 1980 and 1994. There are a number of reasons why it is important to examine the cohorts separately: First, previous sections have shown that immigrant-native earnings gaps are sensitive to how immigrant arrival cohorts are defined. Since the incidence of discrepancies between immigrants' reported arrival date and their first year of covered earnings differs across cohorts, adjusting for these discrepancies could have bigger effects on some cohorts than on others. Second, the model in equation (5) estimates an average rate of earnings growth across all immigrant cohorts. If earnings growth rates differ across cohorts, the fixed-effects estimates will reflect differences in both the average level of earnings and earnings growth across cohorts.<sup>30</sup> Thus if successive immigrant cohorts have slower rates of earnings growth, a model that imposes a constant effect of time spent in the United States will tend to yield declining estimates of cohort fixed effects. Finally, and perhaps most important, if there are differences in selective out-migration across cohorts, the accuracy of estimating immigrant assimilation in repeated cross-sectional data may be better for some cohorts than for others.

The estimates underlying figures 4–6 are derived from CLAD models that control for a quartic in the number of years in the United States, a square in the potential experience at entry, an indicator for the foreign-born, and quartics in total potential experience and calendar time. Each immigrant arrival cohort is compared to natives who had entered 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

<sup>30</sup> To see this, consider a simple example in which there are two immigrant cohorts ( $c = 1, 2$ ) and earnings are given by  $w_{ict} = \alpha_c + \beta_c Y_{ict} + \varepsilon_{ict}$ , where  $Y_{ict}$  is the number of years the immigrant has been in the United States. This can be written as

$$w_{ict} = \alpha_1 + \alpha_2 + \bar{\beta} Y_{ict} + (\beta_1 - \bar{\beta}) \alpha_1 Y_{1it} + (\beta_2 - \bar{\beta}) \alpha_2 Y_{2it} + \varepsilon_{ict}$$

Suppose that the model given by  $w_{ict} = \alpha_1 + \alpha_2 + \bar{\beta} Y_{ict} + \varepsilon_{ict}$  is estimated, and assume that the same number of years of data is available for each cohort (so  $E(\alpha_c Y_{ict}) = 0$ ). Then the ordinary least squares estimate of  $\bar{\beta}$  is  $E(\hat{\beta}) = \beta_1 E(\alpha_1) + \beta_2 E(\alpha_2)$ , which is the average effect among both cohorts of time spent in the United States on earnings. The expected value of the indicator for the first cohort is  $E(\hat{\alpha}_1) = \alpha_1 + E(Y_{1it})(\beta_1 - \beta_2)$ . 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.

to the full sample of natives.<sup>31</sup> The figures show the level of the immigrant-native gap for an immigrant who arrives with five years of potential experience.

Figure 4 compares the relative earnings of immigrants in the longitudinal data based on immigrants' reported arrival date and based on their adjusted arrival date.<sup>32</sup> The pattern of results indicates that immigrants with earnings prior to their reported date of arrival to stay tend to have earnings above the average among their reported cohort but below the average among their original cohort. For example, immigrants who reported to have arrived between 1960 and 1969 entered the labor market with earnings 28 percent below those of natives. After 10 years the gap closed to 15 percent and was essentially stable after that point. When immigrants who reported arriving after 1969, but whose first year of earnings was during the 1960s, are included in this cohort, the earnings gap upon entry declines by five percentage points to 32 percent, and the earnings gap after 10 years is 23 percent. The level of earnings among the 1970–79 arrivals and the 1980–94 arrivals also declines when migrants are shifted to the earlier of their reported cohort or first year of earnings.

A problem documented in Section III is that many immigrants do not have earnings until some years after their reported date of arrival. If these immigrants tend to have below-average earnings, their lagged entry will tend to depress the earnings trajectory of the arrival cohort as a whole. To explore this effect, figure 5 compares the immigrant-native earnings gaps in the longitudinal data among all immigrants and then only among immigrants whose first year of earnings is within three years of their reported date of arrival. For all three arrival cohorts, the estimated rate of immigrant earnings growth is faster when immigrants whose first year of earnings is more than three years after their reported date of arrival are excluded. The immigrant-native earnings gap decreases among the 1960–69 arrivals who have been in the United States for 15 years from 24 to 20 percent. The gap after 15 years declines among the 1970–79 arrivals from 26 to 19 percent. The sample adjustment has perhaps the most noticeable effect on the 1980–94 cohort,

<sup>31</sup> The earnings history from the beginning of each immigrant cohort until 1997 is used. For example, the estimates for the 1960–69 cohort are based on earnings data for immigrants and natives from 1960 to 1997; the estimates for the 1980–94 cohort are based on data from 1980 to 1997.

<sup>32</sup> Immigrants' time in the United States is measured from the midpoint of the arrival interval. Earnings prior to and during the arrival interval are dropped; thus the first observed year corresponds to the third year in the United States. For example, suppose that an immigrant reports that he arrived between 1980 and 1984, but his first year of covered earnings is 1973. The adjusted cohort would then be the 1970–74 group, his time in the United States would be measured from 1972, and only his earnings beginning in 1975 would be used.

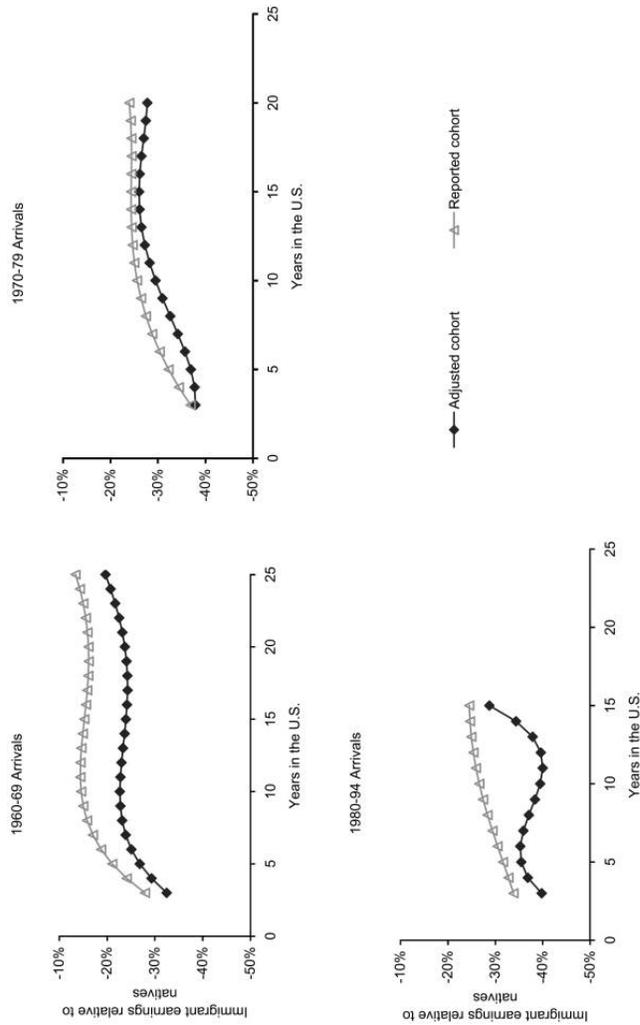


FIG. 4.—The effect of adjusting arrival cohorts by period of entry and cohort definition. Data are taken from Social Security earnings records matched to individuals in the 1990 and 1991 SIPP and in the 1994 CPS. Each line represents the predicted median earnings of immigrants relative to native-born workers based on different cohort definitions. The models also include controls for total potential experience and calendar time. The levels of the curves pertain to an immigrant who arrives in the United States with five years of potential experience. See the text for additional details.

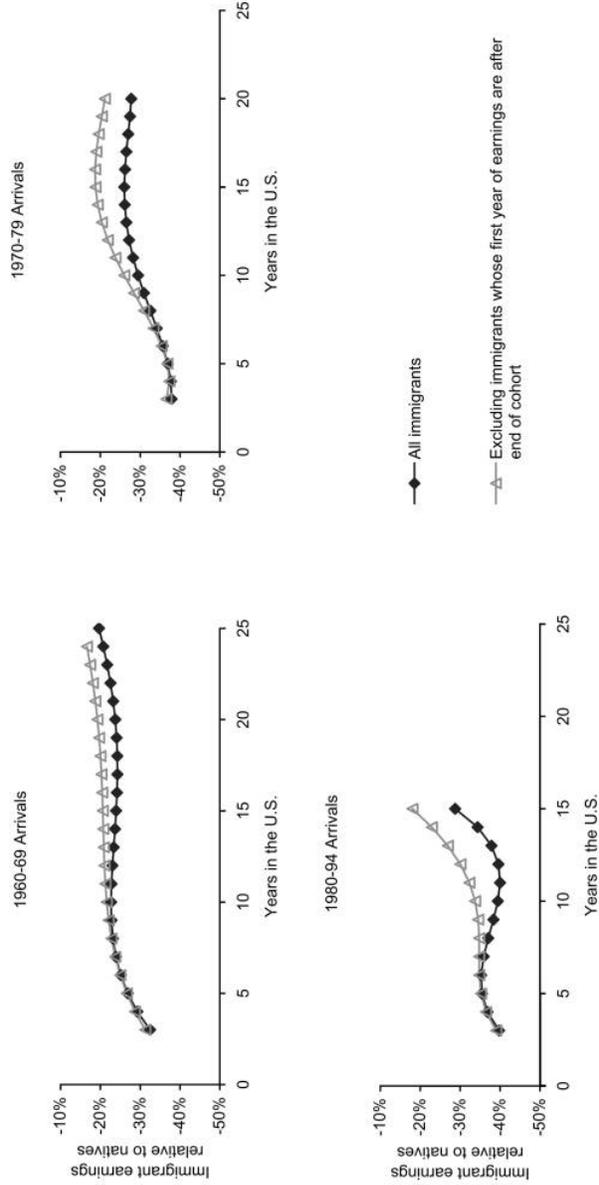


FIG. 5.—The effect of changes in cohort composition on estimated immigrant earnings. Data are taken from longitudinal earnings records matched to individuals in the 1990 and 1991 SIPP and in the 1994 CPS. Each line represents the predicted median earnings of immigrants relative to native-born workers. The cohorts are defined by the earlier of an immigrant's reported date of entry and his first year of earnings. The models also include controls for total potential experience and calendar time. The levels of the curves pertain to an immigrant who arrives in the United States with five years of potential experience. See the text for additional details.

where the gap after 10 years decreases from 39 to 34 percent, and the downward trend in immigrants' relative earnings is eliminated. It is important to note that a similar composition bias may exist in repeated cross-sectional data if immigrants are more likely to be included in the data when they transition from an illegal to a legal immigrant, or if successive censuses are more likely to pick up low-wage or illegal immigrants over time.

Finally, figure 6 compares immigrant-native earnings gaps, by cohort, in the repeated cross-sectional data and in the longitudinal data. To avoid the compositional changes in immigrant cohorts documented in figure 5, I grouped immigrants by their first year of covered earnings (and their time in the United States is measured from their first year of covered earnings). These pictures indicate that the discrepancies in immigrant assimilation between the longitudinal and repeated cross-sectional data appear to be concentrated in the immigrant arrival cohorts that arrived in the 1970s and, to a lesser extent, between 1980 and 1994. For example, the immigrant-native earnings gap among the 1960–69 arrivals closes by 15 percentage points during the first 10 years in the United States according to the longitudinal data and by 12 percentage points according to the repeated cross-sectional data. Yet, the earnings gap among the 1970–79 arrivals closes by 33 percentage points according to the repeated cross-sectional data, but by only 20 percentage points according to the longitudinal data. The earnings gap among the 1980–94 arrivals closes by 29 percentage points according to the repeated cross-sectional data, but by only 22 percentage points according to the longitudinal data. The longitudinal data also point to small differences in the rate of earnings growth across cohorts: during their first 10 years in the United States, the immigrant-native earnings gap among the 1960–69 arrivals closed by 15 percentage points, which is slightly slower than the 20- and 22-percentage-point change among the 1970–79 and 1980–84 arrivals.

## V. 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, and others reenter at a later date. Both types of migration decisions may bias the measured earnings progress of immigrants in repeated cross sections of the decennial census, by far the most common source of data used to examine immigrants' performance in the American labor market. Whereas permanent out-migration by low-earning immigrants will lead to an overstatement of earnings growth across census cross

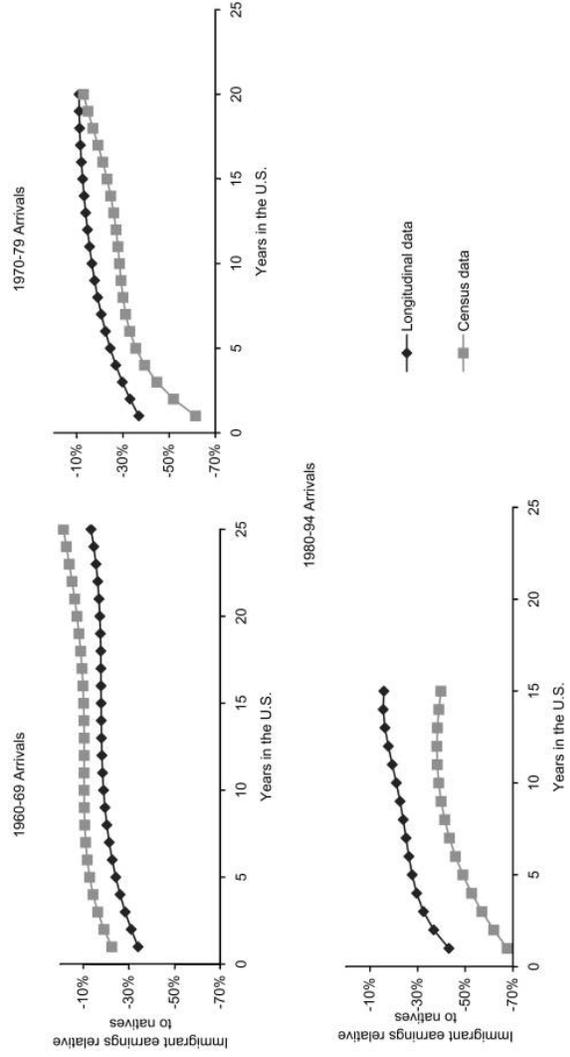


FIG. 6.—Immigrant-native earnings gaps in longitudinal and repeated cross-sectional data. Each line represents the predicted median earnings of immigrants relative to native-born workers. Samples are the same as those used in table 5. The cohorts in the longitudinal data are defined by the earlier of an immigrant's reported date of entry and his first year of earnings. The level of earnings pertains to immigrants who arrived in the United States with five years of potential experience.

sections, temporary out-migration 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 set is not ideal, it can correct some of the shortcomings of 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 United States until the 1990s was considerably slower than that implied by estimates from repeated cross sections of the decennial census. Measured in longitudinal data, over their first 20 years in the United States, immigrant earnings grow by 10–15 percent relative to the earnings of native-born workers. Repeated cross sections of the census suggest that immigrant earnings grow twice as fast, by about 26 percent over the same time period. Selective out-migration by low-earning immigrants, therefore, gives the misleading impression that the economic status of immigrants to the United States improves substantially faster as they assimilate into the labor market.

This result has two important implications for interpreting assimilation across generations. Selective emigration by low-earning immigrants means that the gains in educational attainment between the first and second generations will also be overestimated in repeated cross-sectional data, since the second generation is born largely to the more successful subset of immigrants who remain in the United States. Like computing earnings assimilation, correctly computing changes in educational attainment from immigrant parents to their native-born children requires knowing which immigrants stayed in the United States and gave birth, and which left. Second, previous research indicates that the earnings of children of immigrants tend to be equal to, or perhaps above, those of third-generation and higher Americans (see, e.g., Borjas 1993; Card 2005). To the extent that wage growth among first-generation immigrants is slower than previously reported in other research, more of the assimilation, therefore, occurs between the first and second generations rather than during first-generation immigrants' own working life.

The second key result of this paper is that there is a great deal of disagreement between immigrants' self-reported date of arrival "to stay" in the United States and their first year of covered Social Security earnings, a discrepancy caused in part by out-migration and the subsequent reentry into the United States. Nearly one-third of the decline in the level of earnings of immigrants between the 1960s and 1980s can be accounted for by the misclassification of many low-wage immigrants as more recent arrivals, when in fact they had entered the United States considerably earlier. A corollary of this finding is that the standard model in the economics literature in which immigrants remain in the United

States throughout their working life may not be appropriate for analyzing contemporary migration. Better understanding of the immigrant experience will require better data on migration patterns.

### References

- Baker, Michael, and Dwayne Benjamin. 1997. "The Role of the Family in Immigrants' Labor-Market Activity: An Evaluation of Alternative Explanations." *A.E.R.* 87 (September): 705–27.
- Bean, Frank D., Barry Edmonston, and Jeffrey S. Passel, eds. 1990. *Undocumented Migration to the United States: IRCA and the Experience of the 1980's*. Washington, DC: Urban Inst. Press.
- Bean, Frank D., and Michael Fix. 1992. "The Significance of Recent Immigration Policy Reforms in the United States." In *Nations of Immigrants: Australia, the United States, and International Migration*, edited by Gary P. Freeman and James Jupp. Melbourne: Oxford Univ. Press.
- Biliias, Yannis, Songnian Chen, and Zhiliang Ying. 2000. "Simple Resampling Methods for Censored Regression Quantiles." *J. Econometrics* 99 (December): 373–86.
- Borjas, George J. 1985. "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants." *J. Labor Econ.* 3 (October): 463–89.
- . 1989. "Immigrant and Emigrant Earnings: A Longitudinal Study." *Econ. Inquiry* 27 (January): 21–37.
- . 1993. "The Intergenerational Mobility of Immigrants." *J. Labor Econ.* 11, no. 1, pt. 1 (January): 113–35.
- . 1994. "The Economics of Immigration." *J. Econ. Literature* 32 (December): 1667–1717.
- . 1995. "Assimilation and Changes in Cohort Quality Revisited: What Happened to Immigrant Earnings in the 1980s?" *J. Labor Econ.* 13 (April): 201–45.
- . 1999. "The Economic Analysis of Immigration." In *Handbook of Labor Economics*, vol. 3A, edited by Orley Ashenfelter and David Card. Amsterdam: Elsevier Sci.
- Borjas, George J., and Bernt Bratsberg. 1996. "Who Leaves? The Outmigration of the Foreign-Born." *Rev. Econ. Statis.* 78 (February): 165–76.
- Borjas, George J., and Rachel Friedberg. 2006. "The Immigrant Earnings Turnaround of the 1990s." Manuscript (July), Harvard Univ.
- Borjas, George J., and Lawrence Katz. 2007. "The Evolution of the Mexican-Born Workforce in the United States." In *Mexican Immigration to the United States*, edited by George J. Borjas. Chicago: Univ. Chicago Press (for NBER).
- Bound, John, and Alan B. Krueger. 1991. "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *J. Labor Econ.* 9 (January): 1–24.
- Buchinsky, Moshe. 1994. "Changes in the U.S. Wage Structure 1963–1987: Application of Quantile Regression." *Econometrica* 62 (March): 405–58.
- Card, David. 2005. "Is the New Immigration Really So Bad?" *Econ. J.* 115 (November): F300–F323.
- Card, David, and Alan B. Krueger. 1993. "Trends in Relative Black-White Earnings Revisited." *A.E.R. Papers and Proc.* 83 (May): 85–91.
- Carliner, Geoffrey. 1996. "The Wages and Language Skills of U.S. Immigrants." Working Paper no. 5763 (September), NBER, Cambridge, MA.

- Chay, Kenneth Y. 1995. "Evaluating the Impact of the 1964 Civil Rights Act on the Economic Status of Black Men Using Censored Longitudinal Earnings Data." Manuscript (October), Princeton Univ.
- Chay, Kenneth Y., and Bo E. Honoré. 1998. "Estimation of Semiparametric Censored Regression Models: An Application to Changes in Black-White Earnings Inequality during the 1960s." *J. Human Resources* 33 (Winter): 4–38.
- Chiquiar, Daniel, and Gordon H. Hanson. 2005. "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States." *J.P.E.* 113 (April): 239–81.
- Chiswick, Barry R. 1978. "The Effect of Americanization on the Earnings of Foreign-Born Men." *J.P.E.* 86 (October): 897–921.
- . 1980. *An Analysis of the Economic Progress and Impact of Immigrants*. National Technical Information Service, no. PB80-200454. Washington, DC: U.S. Dept. Labor.
- . 1986. "Human Capital and the Labor Market Adjustment of Immigrants: Testing Alternative Hypotheses." In *Research in Human Capital and Development*, vol. 4, *Migration, Human Capital and Development*, edited by Oded Stark. Greenwich, CT: JAI.
- Deaton, Angus. 1997. *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. Baltimore: Johns Hopkins Univ. Press (for World Bank).
- Duleep, Harriet Orcutt, and Daniel J. Dowhan. 2002. "Insights from Longitudinal Data on the Earnings Growth of U.S. Foreign-Born Men." *Demography* 39 (August): 485–506.
- Duleep, Harriet Orcutt, and Mark C. Regets. 1996. "The Elusive Concept of Immigrant Quality: Evidence from 1970–1990." Program for Research on Immigration Policy PRIP-UI-41 (April), Urban Inst., Washington, DC
- . 1997. "Immigrant Entry Earnings and Human Capital Growth: Evidence from the 1960–1980 Censuses." In *Research in Labor Economics*, vol. 16, edited by Solomon W. Polachek. Greenwich, CT: JAI.
- Ellis, Mark, and Richard Wright. 1998. "When Immigrants Are Not Migrants: Counting Arrivals of the Foreign Born Using the U.S. Census." *Internat. Migration Rev.* 32 (Spring): 127–44.
- Friedberg, Rachel M. 1993. "The Labor Market Assimilation of Immigrants in the United States: The Role of Age at Arrival." Manuscript (March), Brown Univ.
- Funkhouser, Edward, and Stephen J. Trejo. 1995. "The Labor Market Skills of Recent Male Immigrants: Evidence from the Current Population Survey." *Indus. and Labor Relations Rev.* 48 (July): 792–811.
- Hu, Wei-Yin. 1999. "Assimilation and the Earnings of Immigrants: New Evidence from Longitudinal Data." Manuscript (August), Univ. California, Los Angeles.
- . 2000. "Immigrant Earnings Assimilation: Estimates from Longitudinal Data." *A.E.R. Papers and Proc.* 90 (May): 368–72.
- Ibarraran, Pablo, and Darren Lubotsky. 2007. "Mexican Immigration and Self-Selection: New Evidence from the 2000 Mexican Census." In *Mexican Immigration to the United States*, edited by George J. Borjas. Chicago: Univ. Chicago Press (for NBER).
- Jasso, Guillermina, and Mark R. Rosenzweig. 1990. *The New Chosen People: Immigrants in the United States*. New York: Sage Found.
- Jasso, Guillermina, Mark R. Rosenzweig, and James P. Smith. 1998. "The Changing Skills of New Immigrants to the United States: Recent Trends and Their Determinants." Working Paper no. 6764 (October), NBER, Cambridge, MA.
- LaLonde, Robert J., and Robert H. Topel. 1992. "The Assimilation of Immigrants

- in the U.S. Labor Market." In *Immigration and the Work Force*, edited by George J. Borjas and Richard B. Freeman. Chicago: Univ. Chicago Press (for NBER).
- . 1997. "Economic Impact of International Migration and the Economic Performance of Migrants." In *Handbook of Population and Family Economics*, vol. 1B, edited by Mark R. Rosenzweig and Oded Stark. Amsterdam: Elsevier Sci.
- Lindstrom, David P., and Douglas S. Massey. 1994. "Selective Emigration, Cohort Quality, and Models of Immigrant Assimilation." *Soc. Sci. Res.* 23 (December): 315–49.
- Lubotsky, Darren. 2007. "The Effect of Changes in the U.S. Wage Structure on Recent Immigrants' Earnings." Working paper, Univ. Illinois, Urbana-Champaign.
- Nevo, Aviv. 2003. "Using Weights to Adjust for Sample Selection When Auxiliary Information Is Available." *J. Bus. Econ. Statist.* 21 (January): 43–52.
- Powell, James L. 1984. "Least Absolute Deviations Estimation for the Censored Regression Model." *J. Econometrics* 25 (July): 303–25.
- Redstone, Ilana, and Douglas S. Massey. 2004. "Coming to Stay: An Analysis of the U.S. Census Question on Immigrants' Year of Arrival." *Demography* 41 (November): 721–38.
- Schoeni, Robert F. 1997. "New Evidence on the Economic Progress of Foreign-Born Men in the 1970s and 1980s." *J. Human Resources* 32 (Fall): 683–740.
- Schoeni, Robert F., Kevin F. McCarthy, and Georges Vernez. 1996. *The Mixed Economic Progress of Immigrants*. Santa Monica, CA: Center Res. Immigration Policy, Rand Corp.
- Smith, James P., and Barry Edmonston, eds. 1997. *The New Americans: Economic, Demographic, and Fiscal Effects of Immigration*. Washington, DC: Nat. Acad. Press.
- Social Security Administration. Office of the Chief Actuary. 1997. *History of the Provisions of Old-Age, Survivors, Disability, and Health Insurance, 1935–1996*. SSA Publication no. 11-11515 (October). Washington, DC: Soc. Security Admin.
- . Office Policy and Office of Research, Evaluation, and Statistics. 2006. *Annual Statistical Supplement to the Social Security Bulletin, 2005*. SSA Publication no. 13-11700 (February). Washington, DC: Soc. Security Admin.