IMMIGRANT AND EMIGRANT EARNINGS: A LONGITUDINAL STUDY

GEORGE J. BORJAS*

This paper analyzes the relationship between earnings and the extent of assimilation, cohort quality change, and return migration experienced by the foreign-born population. The study uses the longitudinal data available in the Survey of Natural and Social Scientists and Engineers. The analysis reveals that there was a sizable decline in the skills of this population over the last two decades. In addition, the study shows that return migration is more likely among immigrants who did not perform well in the U.S. labor market.

I. INTRODUCTION

In the past decade, empirical research has found that immigrant earnings growth is remarkably rapid. For the most part, these studies used cross-section data sets to show that the number of years elapsed since immigration has an independent, positive, and numerically important impact on the earnings of immigrants.[1] Surprisingly, these cross-section studies find that the relative earnings of immigrants grow so rapidly that after ten to fifteen years immigrant earnings overtake the earnings of native workers. The steepness of the immigrant age/earnings profile was explained in terms of the human capital framework: immigrants have stronger incentives than native workers to invest in human capital investments. The result that, in the long run, immigrants earn more than natives was explained by assuming that immigrants are a relatively select group of individuals whose average "quality" exceeds that of the typical native worker.

Recent work [Borjas, 1985] raises serious doubts about the validity of the inferences drawn from the cross-section empirical results. Cross-section estimates of immigrant assimilation are biased if emigration (i.e., return migration) is not randomly distributed across the immigrant population or if the quality of successive immigrant cohorts changed over the sample period. Tracking specific immigrant cohorts between the 1970 and 1980 Public Use Samples of the U.S. Census reveals that most immigrant groups do *not* have substantially higher rates of earnings growth than their native counterparts.

* Professor of Economics, University of California, Santa Barbara, and Research Associate, National Bureau of Economic Research. This research was supported by the National Science Foundation (Grant SES-8603973).

1. See, for example, Borjas [1982], Carliner [1980], Chiswick [1978], DeFreitas [1974], and Long [1980]. Although most of the literature focuses on cross-section data sets, a few longitudinal analyses have been conducted. These include Chiswick [1980, Chapter 10], DeFreitas [1981], and Snipp and Tienda [1984]. The results of the longitudinal studies, however, are often inconclusive and contradictory. The entire literature has been recently surveyed by Greenwood and McDowell [1986].

Economic Inquiry
Vol. XXVII, January 1989, 21–37

Thus, cross-section results may have captured a secular decline in the quality of the immigrant cohorts admitted to the U.S. in the postwar period.

Unfortunately, Census data do not allow an exact matching of immigrant cohorts over time. Thus the study of immigrant earnings using the two most recent Censuses is not itself free of problems. Mismatches may arise because of sizable emigration rates among the foreign-born or because of changes in Census enumeration procedures which may undercount specific immigrant or ethnic groups in particular years. For example, if the 1970 Census missed large numbers of low-skilled workers (relative to the 1980 Census), then tracking cohorts over time may reveal little assimilation. This would lead to the statistical illusion that earnings capacities are declining across successive immigrant cohorts. The study of "synthetic cohorts," therefore, may build its own statistical problems into the analysis.[2]

One can avoid these problems by tracking specific individuals over time. This paper provides such a longitudinal study using the 1972–1978 Survey of Natural and Social Scientists and Engineers. Longitudinal data sets also allow an analysis of the emigration (i.e., return migration) of immigrants. It is well known that attrition rates in longitudinal data are relatively large. To the extent that the attrition of immigrants from the data is due to their return to the "home" country, the study of the Scientists and Engineers data set can be used to understand the selection mechanism that generates the return migrant flow.

Since the study is confined to a select subsample of the immigrant population, the results may not easily generalize to the less-skilled groups which form the bulk of immigration to the United States. However, despite the narrow focus of the samples, the analysis leads to results that are quite similar to those obtained in studies of synthetic Census cohorts. First, the longitudinal data reveals relatively small rates of assimilation or adaptation in the immigrant population, and second, the more recent waves of immigrants are less skilled than the earlier waves of immigrants.

Section II of the paper presents a conceptual framework for the analysis. It illustrates how longitudinal analyses of immigrant earnings are affected by assimilation and changes in cohort quality. Section III presents the basic empirical results, while section IV addresses the potential biases introduced into the analysis by nonrandom emigration of the immigrant population. Finally, section V summarizes the main findings of the study.

II. METHODOLOGY

In analyzing immigrant earnings, one must differentiate between cohort effects and aging or assimilation effects. Suppose that a longitudinal data set is available and that two earnings points are observed for each individual.

2. See Deaton [1985] for a theoretical treatment of the econometric problems associated with jointly estimating a series of cross-sections.

Let $w_{ij}(t)$ be the earnings of immigrant j at time t (t=0,1); $X_{ij}(t)$ a vector of variables describing his socioeconomic characteristics (including the individual's age); $y_j(t)$ the number of years the immigrant has resided in the United States; and C_j the calendar year in which the individual immigrated.[3] Finally, let $w_{nl}(t)$ and $X_{nl}(t)$ be the respective variables for native l. Consider the following two income generating equations:[4]

$$\ln w_{ij}(t) = X_{ij}(t)\theta_i + \alpha_1 y_j(t) + \alpha_2 y_j^2(t) + \beta_1 C_j + \beta_2 C_j^2 + \gamma_i t + \varepsilon_{ij}(t) \quad (1)$$

$$\ln w_{nl}(t) = X_{nl}(t)\theta_n + \gamma_n t + \varepsilon_{nl}(t). \quad (2)$$

The structural parameters in (1) and (2) identify three different sets of factors that can affect immigrant earnings over time: aging, cohort, and period effects. Earnings change as a result of the aging process and this growth is captured by the coefficients of the age variable (in the vector X) and of the years-since-migration variable y. From these coefficients one can trace the age/earnings profiles of immigrants and that of natives. An estimate of immigrant assimilation is obtained by calculating the rate at which the two age/earnings profiles converge.

Equation (1) also includes variables indicating the date in which immigration occurred. The coefficient vector (β_1,β_2), measuring cohort effects, gives the rate of change in earnings capacity across successive immigrant waves.[5] Finally, the coefficients γ_i and γ_n give the period effects, the impact of changes in aggregate economic conditions on immigrant and native earnings. In principle, the period effects may differ between immigrants and natives. These differences could arise, for instance, if immigrant earnings are more sensitive to changes in economic conditions than native earnings.

3. The variable C_j is scaled such that $C_j = 0$ if immigrant j migrated in 1970, -1 if he arrived in 1969, etc. In the data set analyzed below, the last possible date of entry is 1970.

4. The quadratic functional form used in (1) and (2) is somewhat restrictive but simplifies the estimation and presentation of the results. A more general functional form is used in the earlier Borjas [1985] analysis of Census data.

5. The native earnings function in (2) can be generalized to allow for cohort effects in the native population. The hypothesis that cohort effects may play an important role in determining native earnings has received careful study by Welch [1979] in the context of the earnings experience of the baby boom generation.

In general, aging, period, and cohort effects cannot be separately identified.[6] Consider the estimation of the various structural parameters using a single cross-section of data. Period effects are subsumed into the intercept since the variable t is constant in any given cross-section. In addition, the calendar year in which the cross-section is observed, say T, is defined by the identity $T \equiv C_j + y_j$, for every immigrant j. Substituting this identity into (1) yields

$$\ln w_{ij} = (\beta_1 T + \beta_2 T^2 + \gamma_i) + X_{ij}\theta_j + (\alpha_1 - \beta_1 - 2\beta_2 T)y_j + (\alpha_2 + \beta_2)y_j^2. \qquad (1')$$

Equation (1′) illustrates that a cross-section regression of immigrant earnings on years-since-migration does *not* identify any structural parameter. Moreover, (1′) shows that the intercept and the coefficient of the (linear) years-since-migration variable are functions of T, the year in which the cross-section is observed. Cross-section estimates of (1′), therefore, are inherently unstable and confound the assimilation effect with both cohort and period effects.

The availability of longitudinal data—so that an additional observation on earnings and the regressors is available k years after the first observation—breaks the identity that defines the year of the cross-section as the sum of C and Y. For any given calendar year of arrival, the variable y takes on two separate values. However, the right-hand-side variables in (1) *still* exhibit perfect collinearity. It is easy to show that

$$y_j = T - C_j + kt$$

The dummy variable measuring the period effect (t) is, therefore, a linear combination of the cohort and years-since-migration variables.

The identification of the structural parameters requires that a normalization be imposed on one of the three effects. The normalization used in this paper is that the period effect in the immigrant earnings function (γ_i) is identical to the period effect in the native earnings function (γ_n). This assumption does *not* state that labor markets are unchanged over the sample period. Instead, it assumes that the impact of any changes that did occur affect immigrant and native earnings by the same relative amounts. In a sense, this normalization uses information from outside the immigrant labor market (γ_n) to net out the period effect in the immigrant earnings function. One can then identify the two remaining sets of parameters from longitudinal data.

6. The question of whether aging, cohort, and period effects can be identified has a long history in demography (see Fienberg and Mason [1978]). The problem has also played a role in studies that measure quality change in capital stocks (for example, Hall [1968]). A recent exposition of the econometric issues is given by Heckman and Robb [1983].

The normalization imposes the substantive restriction that changes in aggregate economic conditions do not have a differential impact on immigrant and native earnings. Since little is known about the empirical validity of this hypothesis, it is useful to study periods in which aggregate labor market conditions changed only slightly. Most of the empirical analysis reported below analyzes earnings over the period 1974–78. In 1974, the adult unemployment rate was 4.8 percent, and had increased to 5.2 percent by 1978. In addition, the normalization is more likely to be valid if the native base resembles the immigrant population in significant ways. The greater the resemblance, the more likely that both groups will have experienced similar period effects. The analysis below is confined to persons in the same set of occupations (scientists and engineers), and makes it more likely that immigrants and natives faced the same changing labor market conditions.

The errors in equations (1) and (2) may be correlated over time for a given individual. The stochastic structure of these errors is given by

$$E[\varepsilon_{ij}(0) \cdot \varepsilon_{ij'}(1)] = E[\varepsilon_{nl}(0) \cdot \varepsilon_{nl'}(1)] = \sigma_{01}, \text{ for } j = j', l = l', \qquad (3)$$

$$= 0, \text{ otherwise,}$$

so that a constant covariance exists for immigrants and natives in the disturbances across the two time periods. The system in (1) and (2) is estimated by "stacking" the data in order to impose the cross-equation restriction that $\gamma_i = \gamma_n$, and using generalized least squares to allow for the covariance structure in (3).

III. BASIC EMPIRICAL RESULTS

The study uses the 1972–78 Survey of Natural and Social Scientists and Engineers. These data report the labor market and educational activities of about 50,000 persons engaged in natural or social sciences and engineering. The initial survey was conducted in 1972 and the respondents were then surveyed biennially until 1978. In addition, selected data from the 1970 Census (including 1969 earnings) was included for each individual in the survey.

The analysis is restricted to the random subsample of the respondents (about 52 percent) whose 1972 questionnaire included questions on the year of immigration to the United States. The study is further is restricted to men aged eighteen to fifty-six (as of 1972) who reported an unambiguous nativity status and a valid year of immigration (if foreign-born).[7] The earnings

7. Unfortunately, the survey only reports whether the individual was born in the U.S. or in a foreign country, without identifying the foreign country.

analysis reported in this section uses the 1974 and 1978 cross-sections of the survey.[8]

A key characteristic of the data is worth noting. The sample of scientists and engineers first contacted in 1972 was drawn, at random, by the Census Bureau from the population of persons who reported science/engineering jobs in the 1970 Census. Over time, however, persons may have gone on to different kinds of jobs or obtained additional degrees in the humanities. These persons were *not* deleted from the sample by the Census Bureau. Hence, some of the "scientists and engineers" in the data hold jobs or degrees which have little in common with the occupations usually associated with this population.

Table I summarizes the characteristics of the immigrant and native samples in both 1974 and 1978.[9] The statistics indicate that immigrants have slightly higher earnings than natives, and that the wage differential between the two groups did not change much over the period. Immigrants are somewhat more educated, older, more likely to be academics, and more likely to reside in the Northeastern region. Immigrants are also less likely to have left the scientific labor market.

Table II presents the generalized least squares estimates of the system in equations (1) and (2). As expected, higher education levels are associated with higher earnings, although the impact of education on earnings is smaller for immigrants. For instance, a Ph.D. degree increases the earnings of natives by 41.6 percent (relative to the omitted group of persons who lack a bachelor's degree), but increases the earnings of immigrants by only 31.4 percent. The regressions also show that persons employed in the academic sector earn about 13 to 17 percent less than persons in the nonacademic sector. Finally, the regressions reveal that wage differentials are generated by the type of highest degree obtained and by whether or not the current job is "science-related." Among natives, an engineering degree (the omitted category) is associated with higher earnings than the other types of degrees, and a science-related job pays about 5 percent higher earnings than a job that is not science-related.

Table II shows that the coefficients of years-since-migration are insignificantly different from zero, and, in fact, have perverse signs. This does *not*, however, imply that there is no assimilation or "catching-up" effect. The rate of earnings convergence between the two groups also depends on the

8. These two cross-sections are used initially because the earnings measure in the 1972 survey is not comparable. In particular, the 1972 survey gives annual earnings in 1971 (analogous to the 1969 earnings reported in the 1970 Census). Beginning with the 1974 survey, the respondents were asked their basic salary on the job which was converted by the Census Bureau into the annual salary rate used in this study.

9. The salary levels are not deflated. The semilogarithmic functional form of the earnings functions implies that the variables capturing the period effects automatically deflate the earnings data.

TABLE I
Means of Variables

Variables	Natives		Immigrants	
	1974	1978	1974	1978
LANNUAL	9.841	10.213	9.879	10.261
BA	.35	.34	.18	.17
MA	.41	.40	.40	.38
PH.D.	.21	.23	.40	.42
AGE	40.7	44.7	41.8	45.8
ACADEMIC	.17	.16	.26	.25
EAST	.25	.24	.31	.30
SOUTH	.29	.29	.19	.20
WEST	.22	.23	.26	.27
NONSCIENCE	.06	.08	.03	.04
PHYSICAL	.34	.32	.38	.37
SOCIAL	.08	.08	.08	.08
HUMANITIES	.10	.13	.06	.09
YSM	–	–	19.2	23.2
Sample Size	14196		1166	

Key to Variables:
LANNUAL = ln (annual salary rate)
BA = 1 if highest degree is a Bachelor's degree
MA = 1 if highest degree is a Master's degree
PH.D. = 1 if highest degree is a Doctoral degree (omitted group has less than a bachelor's degree)
ACADEMIC = 1 if individuals' current job is academic
EAST = 1 if individual resides in the Northeast
SOUTH = 1 if individual resides in the South
WEST = 1 if individual resides in the West (omitted region is the North-Central)
NONSCIENCE = 1 if current job is not science-related
PHYSICAL = 1 if field of highest degree is a physical science
SOCIAL = 1 if field of highest degree is a social science
HUMANITIES = 1 if field of highest degrees is in the humanities (omitted field is engineering)
YSM = years since migration.

TABLE II
Estimated Earnings Functions in 1974–78 Period*

Variables	Natives		Immigrants	
	Coefficient	t	Coefficient	t
CONSTANT	7.7371	(155.65)	7.1592	(36.44)
TIME78	.3143	(110.03)	.3143	(110.03)
BA	.1678	(11.60)	.0693	(1.08)
MA	.2275	(15.70)	.1593	(2.51)
PH.D.	.4158	(26.60)	.3452	(5.28)
AGE	.0799	(35.84)	.1021	(10.51)
AGE^2	−.0008	(−30.66)	−.0010	(−9.14)
ACADEMIC	−.1701	(−23.31)	−.1333	(−6.28)
EAST	.0367	(5.43)	.0433	(1.91)
SOUTH	.0332	(5.17)	.0508	(2.05)
WEST	.0070	(1.01)	−.0310	(−1.32)
NONSCIENCE	−.0524	(−6.51)	.0478	(1.39)
PHYSICAL	−.0575	(−10.13)	.0057	(.27)
SOCIAL	−.0178	(−1.83)	−.0454	(−1.33)
HUMANITIES	−.0378	(−5.63)	.0440	(1.62)
YSM	–	–	−.0028	(−.56)
YSM^2	–	–	.00005	(.55)
COHORT	–	–	−.0085	(−1.93)
$COHORT^2$	–	–	−.0002	(−1.47)

R^2 for system = .1334

Correlation between 1974 and 1978 equations = .586.

**TIME78* is a dummy variable set to unity if the observation is drawn from the 1978 survey. *COHORT* is the calendar year in which the immigrant arrived, scaled so that *COHORT*=0 if the immigrant arrived in 1970, −1 if he arrived in 1969, etc.

age coefficient. In particular, suppose the coefficients of age (A) and age squared are given by δ_{1i} and δ_{2i} in the immigrant earnings function, and by δ_{1n} and δ_{2n} in the native earnings function. The rate of assimilation experienced by immigrants is defined by

$$(\partial \ln w_i/\partial A) - (\partial \ln w_n/\partial A) = (\delta_{1i} + \alpha_1 - \delta_{1n}) + 2(\delta_{1i}A + \alpha_2 y - \delta_{2n}A). \quad (4)$$

TABLE III

Summary of Structural Parameters from 1974–78 Data*

Parameter	1974–78 system	1974 cross-section	1978 cross-section
$\ln(w_i/w_n)\mid_{y=0,\ C=0}$	−.2168	−.1178	−.1655
	(−4.43)	(−3.64)	(−3.90)
$\ln(w_i/w_n)\mid_{y=20,\ C=0}$	−.0844	−.0184	−.0041
	(−1.84)	(−1.37)	(.36)
$\partial \ln w/\partial A\mid_{y=0}$	.0097	.0064	.0114
	(1.74)	(2.26)	(3.64)
$\partial \ln w/\partial A\mid_{y=20}$	.0035	.0035	.0055
	(1.47)	(3.21)	(4.10)
$\partial \ln w/\partial C\mid_{C=0}$	−.0085	–	–
	(−1.93)		
Wage differential between 1970 and 1960 cohorts	−.0694	–	–
	(−1.91)		

* t-ratios are given in parentheses. The native/immigrant differentials are evaluated using the 1974 means of the socioeconomic variables in the immigrant sample.

Equation (4) accounts for the fact that as the immigrant ages two variables (A and y) are changing. Table II indicates that the age coefficient for the immigrant sample is numerically larger than the age coefficient for the native sample. The significance of this result will be discussed below.

The coefficient of *COHORT* is negative and statistically significant. The coefficient indicates that there has been a decline in the productivity of immigrants over the period. Since the variable *COHORT* is scaled such that it equals zero for persons who arrived in 1970, −1 for 1969 arrivals, etc., its coefficient measures the rate of change in cohort quality as of 1970. Table II shows that the wage of successive cohorts was declining at the annual rate of −.9 percent as of 1970.

The regressions in Table II (along with the means reported in Table I) are used to calculate the wage differential between immigrants and natives, the rate of convergence of the two age/earnings profiles, and the magnitude of the change in the earnings of immigrant cohorts. These statistics are reported in the first column of Table III. The first row of the table presents the wage differential between immigrants and natives calculated for the typical 1970 arrival *as of the date of arrival*, $\ln(w_i/w_n)\mid_{y=0,\ C=0}$. This calculation shows that the most recent wave of immigrants in the data starts out at a 21.7 percent wage disadvantage. The second row of the table predicts the wage differential twenty years after arrival, and it reveals that immigrants are still

earning 8.4 percent less than comparable natives. Recent immigrants in the science/engineering fields, therefore, will have substantially lower earnings than natives for a significant portion of their working lives.

The next two rows in the table use equation (4) to calculate the rate of convergence between immigrants and natives. This rate is positive and marginally significant upon arrival, but by twenty years after arrival the rate of convergence becomes insignificant. It is possible that the weak assimilation rates documented in Table III are related to the fact that the earnings functions are estimated within narrow skill groupings. Part of the assimilation of immigrants occurs as individuals search across jobs and occupations for better opportunities. The data, however, are restricted to men who were scientists and engineers *only* at the time of the initial survey. These men had the option of changing both jobs and occupations over time. The relatively low assimilation rates, therefore, cannot be explained away as an artifact of the sampling scheme.

The next two rows in Table III detail the magnitude of cohort quality changes. The rate of change in cohort productivity (as of 1970) is almost one percentage point per year. Equally striking, the wage disadvantage of the 1970 cohort relative to the 1960 cohort is 6.9 percent. Over a single decade, therefore, the productivity of immigrant cohorts declined by a substantial amount.

These findings are summarized in Figure 1, which illustrates the extent of wage differentials (relative to the comparable native population) for three immigrant cohorts (arrivals in 1950, 1960, and 1970) over the entire working life. Figure 1 shows the extent to which the earnings of immigrant waves have declined over the period. In addition, Figure 1 shows that the relative earnings of an immigrant cohort increase over the working life. For the earliest cohort (1950 arrivals), the aging effect is sufficiently strong so that within ten to fifteen years after arrival immigrant earnings overtake the earnings of native workers.[10] The 1970 cohort of immigrants, however, starts out at a much lower earnings level and the aging effect is insufficient for earnings parity, let alone overtaking, to occur over the working life.

It is instructive to compare these results with those obtained from the standard cross-section regressions that dominate the literature. The last two columns of Table III present estimates of the parameters estimated from the 1974 and 1978 cross-section, respectively. These results differ from the system estimates of column 1. The results also differ between the 1974 and 1978

10. The existence of overtaking for these early cohorts suggest that these cohorts were positively selected from the population of the countries of origin. Borjas [1987] presents a theoretical discussion of how economic and political factors determine the self-selection of the migrant pool. It should be noted that the overtaking points exhibited in Figure 1 are not a statistical artifact of the quadratic functional form used in the earnings functions. The overtaking result remains even when the years-since-migration variable is introduced as a vector of dummy variables.

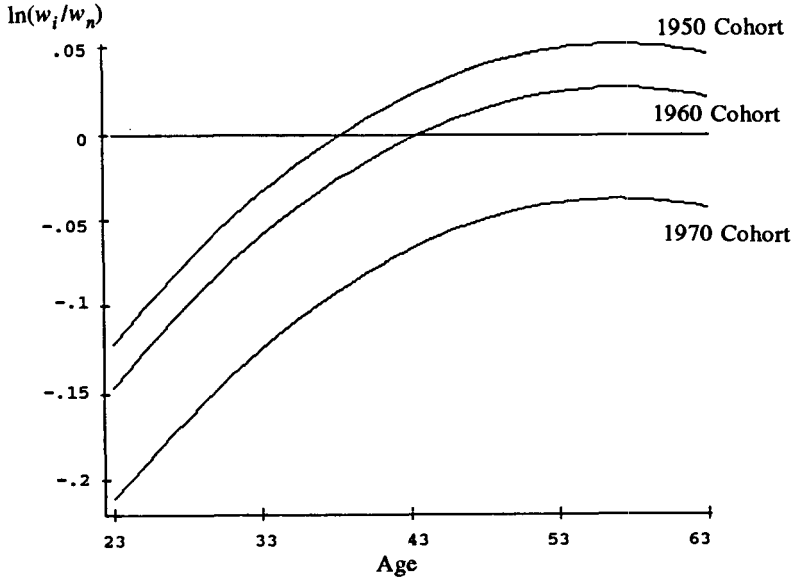

FIGURE 1

cross-sections. Newly arrived immigrants earn about 12 percent less than natives according to the 1974 cross-section and about 17 percent less according to the 1978 cross-section. Both of these estimates are lower than the 21.7 percent wage gap obtained from the longitudinal data. The cross-section regressions also suggest that the earnings gap between immigrants and natives disappears within twenty years. Figure 1, however, clearly shows that the earnings gap will remain throughout the working life of the most recent cohort of immigrants. The analysis of single cross-sections of data, therefore, provides unstable, unreliable, and misleading estimates of the structural parameters describing the immigrant experience.

The study presented in this paper focuses on a very small (and nonrandom) segment of the U.S. foreign-born population. This fact makes it all the more remarkable that the findings summarized in Table III so closely resemble the results obtained in the earlier Borjas [1985] study that uses the entire immigrant population. This similarity in results suggests that the trends in immigrant productivity in the last two decades characterize both the typical migrant as well as the segment of the foreign-born population that is highly educated and skilled.

IV. EMIGRATION SELECTIVITY

The results summarized in Table III are based on the sample of immigrants (and native workers) that survived up to the 1978 survey. Two kinds of selection biases affect these results. First, the construction of the sample ensures

TABLE IV
Logit Regression on Probability of Leaving the Sample by 1978*

Variable	Coefficient	Asymptotic t-ratio	$\partial P/\partial X$
CONSTANT	.6310	(1.69)	–
BA	−.0063	(.00)	−.0009
MA	−.1365	(−1.38)	−.0198
PH.D.	−.3944	(−3.59)	−.0571
AGE	−.0875	(−4.51)	−.0127
AGE^2	.0007	(3.00)	.0001
EAST	−.0144	(−.26)	−.0021
SOUTH	.1805	(3.60)	.0261
WEST	.2720	(5.19)	.0394
PHYSICAL	−.1039	(−2.49)	−.0150
SOCIAL	.2120	(3.23)	.0307
IMMIGRANT	1.4907	(10.74)	.2159
YSM	−.0985	(−6.47)	−.0143
YSM^2	.0013	(4.25)	.0002
$-2\ln L$	20409.1		

* The derivative $\partial P/\partial X$ is defined by $\beta P(1-P)$, where β is the logit coefficient and P is the attrition rate in the sample. The independent variables are defined as of 1972.

that only persons who are living in the United States as of the initial survey date are sampled. Since no information is available on which immigrants emigrated prior to the survey date (not even how many emigrants there are), little can be done about this type of selection bias.[11] A second kind of selection bias is generated by the high attrition rates of both native and foreign-born respondents *in* the survey. The attrition rate (i.e., not present in the survey by 1978) among natives was 15 percent, while among immigrants it was 22.6 percent. The large differential remains even after standardization for initial conditions in socioeconomic characteristics. Table IV presents a logit regression on the probability that the individual will drop out of the survey by 1978. Table IV shows that immigrants are more likely to leave the survey, and that this probability is strongly affected by the number of years that have elapsed since immigration. The most recent immigrants have an attrition rate that is twenty-two points higher than that of observationally equivalent natives. This gap drops to 9.3 percentage points for immigrants who have been in the U.S. for ten years, and vanishes for immigrants who have resided in the U.S. for twenty years.

11. To correct for this type of selection bias an entire *cohort* of immigrants would have to be tracked over time. This issue is further discussed in Jasso and Rosenzweig [1985].

TABLE V
Impact of Attrition Variable on Earnings by Immigration Status*

Dependent Variable	Native-born	Foreign-born
$\ln w_{69}$	−.0141	−.0550
	(−1.56)	(−1.91)
$\ln w_{71}$	−.0195	−.0579
	(−2.38)	(−2.23)
$\ln w_{71} - \ln w_{69}$	−.0054	−.0030
	(−.59)	(−.11)

* t-ratios are presented in parentheses. The regressions hold constant the individual's educational attainment, age, and years-since-migration (if immigrant). In the wage level equations, quadratics in age and years-since-migration are also included.

The difference in attrition rates between immigrants and natives may arise because immigrants are harder to track across surveys or because immigrants are likely to have higher emigration rates. Although the exact whereabouts of persons who drop from the survey are unknown, the summary statistics in Table I do not suggest that immigrants have higher rates of internal migration (across Census regions) than natives. It is reasonable to suppose, therefore, that part of the difference in attrition rates may be due to higher emigration rates in the foreign-born population.

There also exist sizable differences in weekly earnings between the samples of "stayers" and "leavers."[12] Table V presents the coefficient of the attrition variable (i.e., not present in the survey by 1978) on earnings level and growth regressions estimated by nativity status. The key result in this descriptive table is that the least skilled individuals disappear from the sample. This result is particularly strong among the foreign-born: future emigrants have significantly lower earnings than immigrants who stayed in the sample.

There are well-known (if not very robust) statistical techniques designed to handle the selection biases induced by nonrandom sample attrition.[13] Rather than rely on arbitrary identification restrictions, however, the

12. The use of weekly earnings (as opposed to annual earnings) makes the analysis below more comparable to the 1974–1978 regressions which used the annual salary rate.

13. These statistical techniques require the estimation of a first-stage attrition equations, and the use of predicted attrition probabilities in the earnings functions [see Heckman, 1979]. The methodology is most successful when the instruments in the first stage do not entirely overlap with the variables in the earnings function. Unfortunately, very few of these instruments exist in the Survey of Natural and Social Scientists and Engineers, and the selection parameter must be identified through non-linearities in the model. Preliminary estimates of the model (using the Heckman-two-stage estimator) were very unreliable.

construction of the data allows a straightforward analysis of the selection biases introduced by sample attrition. As noted earlier, the initial [1972] cross-section of the survey was based on a sample of scientists that had been enumerated in the 1970 Census. The Census record was then merged with the information collected in 1972. Thus for the sample of immigrants present in the U.S. as of 1972 (but who had immigrated in 1970 or earlier), key information on their economic status *prior* to any sample attrition is available. The relationship between attrition behavior in the 1972–78 period and the 1969 and 1971 earnings of individuals can be evaluated.

The system of equations in (1) and (2) was reestimated using the 1969–71 data in three alternative samples: (1) the sample of all immigrants present in the 1972 cross-section; (2) the sample of immigrants present in the 1972 cross-section who remained in the survey until the 1978 cross-section; and (3) the sample of immigrants present in the 1972 cross-section who had left the survey by 1978. The first of these samples contains the complete "population." The second sample (of "stayers") is a nonrandom subsample of the first, and the sample selection rule should induce selection biases due to nonrandom attrition. Finally, the third sample contains the group of individuals who, for various reasons (including emigration) left the survey at some point between 1972 and 1978. The dependent variables used are the 1969 and 1971 levels of (ln) weekly earnings.

Table VI summarizes the results for each of the three samples. The estimated "population" structural coefficients are presented in the first column of Table VI. These estimates reveal patterns that resemble those obtained earlier in the analysis of the 1974–78 period. In particular, immigrants start out their U.S. career at a significant wage disadvantage. The most recent cohort of immigrants (i.e., 1970 arrivals) enter the labor market earning approximately 21 percent less than comparable natives. Unlike the results presented in the previous section, the 1969–71 analysis reveals no assimilation or convergence effect, so that this wage differential is not narrowing over the working life cycle. The study of the 1969–71 data also reveals the existence of sizable cohort effects: As of the 1970 cross-section, immigrant earnings are declining at a rate of nearly 2 percent across successive annual waves (about twice the size of the estimate revealed by the 1974–78 data).[14]

The second column of Table VI presents the structural coefficients for the subsample of stayers—i.e., persons who were present in the survey as of 1978. This sample is identical in composition to the sample used in the pre-

14. A comparison of the results in Tables III and VI shows that the initial wage differential is roughly the same in the two runs of the model, but that the estimated assimilation effect differs between the two periods. In the 1974–78 model, the assimilation effect is positive but only marginally significant, while in the earlier period the assimilation effect is negative and insignificant. The differences in the estimates of the wage differential after twenty years in the U.S. are due to the fact that different point estimates of the (insignificant) assimilation effect are used in the calculations.

TABLE VI
Summary of Structural Parameters from 1969–71 Data*

Parameter	Population	Sample of "stayers"	Sample of "leavers"
$\ln(w_i/w_n)\mid_{y=0,\ C=0}$	−.2088 (−6.87)	−.1673 (−4.55)	−.2790 (−4.66)
$\ln(w_i/w_n)\mid_{y=20,\ C=0}$	−.2995 (−2.54)	−.2711 (−2.03)	−.3134 (−1.15)
$\partial \ln w/\partial A\mid_{y=0}$	−.0024 (−.27)	−.0045 (−.45)	.0068 (.35)
$\partial \ln w/\partial A\mid_{y=20}$	−.0067 (−1.05)	−.0059 (−.87)	−.0102 (−.55)
$\partial \ln w/\partial C\mid_{C=0}$	−.0199 (−2.28)	−.0168 (−1.68)	−.0254 (−1.30)
Wage differential between 1970 and 1960 cohorts	−.1850 (−2.59)	−.1610 (−1.97)	−.2233 (−1.42)

* t-ratios are given in parentheses.

vious section, and hence regressions estimated in this subsample will be biased in the same way as the regressions estimated in the 1974–78 earnings data. All of the substantive results of the study are evident in this subsample of the data. Immigrants start out at a significant wage disadvantage, there is no strong evidence of any assimilation or convergence effect, and there is a sizable cohort effect. Although these findings are qualitatively the same as those implied by the coefficients estimated in the "population," the numerical magnitude of the structural parameters differs somewhat across the two samples. Immigrants who survive throughout the period start out in the labor market with a smaller wage disadvantage than the average immigrant in the cohort. This, of course, reflects the fact that sample dropouts are chosen from the pool of "failures."

In addition, the extent of cohort quality change calculated from the sample of survivors is smaller (in absolute value) than the "true" extent of cohort quality change. At the margin, cohort quality is declining by about 2 percent in the entire population of immigrants, but by 1.7 percent in the sample of survivors. This result is predicted by standard specification error analysis if two assumptions are satisfied: (1) the probability of emigration is greater for more recent immigrants, and (2) the unsuccessful immigrants leave.

Finally, the third column in Table VI presents the estimated structural parameters for the sample of persons who dropped out from the survey. These coefficients reveal that future emigrants have substantially lower relative

wages than immigrants who presumably chose the U.S. as a permanent place of residence. The typical survivor, for example, starts out in the labor market with a 17 percent wave disadvantage, while the future dropout has an initial 28 percent wage disadvantage.[15]

V. SUMMARY

Do immigrants get a return for U.S. labor market experience above and beyond that received by native workers? This question, dealing with the concept of the adaptation or assimilation of immigrants in the labor market, is at the core of the research analyzing the earnings of immigrants. The literature answers the question in the affirmative, and shows that the payoff of U.S. labor market experience to immigrants is so large that many immigrant groups soon outperform the native population. Most of the literature, however, bases its findings on analyses of cross-section data sets. The results in these studies are suspect because nonrandom emigration propensities or systematic changes in the productivity or quality of immigrant cohorts can generate the empirical patterns observed in the cross-section studies.

This paper analyzes the extent of assimilation, cohort quality change, and emigration using the 1972–1978 Survey of Natural and Social Scientists and Engineers. The longitudinal study led to the following results.

1. The rate of convergence between the age/earnings profiles of immigrants and natives is relatively small. Recent immigrant scientists and engineers will find that earnings parity with natives, let alone overtaking, is not achieved within their working lives.

2. There has been a sizable drop in the skills of immigrant scientists and engineering cohorts in the last two decades. The average earnings capacity of the 1970 cohort is at least 7 percent lower than the earnings capacity of the wave that arrived in 1960.

3. To the extent that differences in the rates of sample attrition between immigrants and natives capture the emigration of foreign-born persons, the data indicate that emigration rates are sizable. The analysis also shows that the sample of emigrants is characterized by poor labor market outcomes. The earnings of future emigrants are about 11 percent lower than the earnings of foreign-born persons who choose the U.S. as a permanent place of residence.

These results are based on estimates of a structural model that uses longitudinal data to identify separately aging and cohort effects. It is well known that identification of these parameters hinges crucially on a normalization for period effects: how do secular changes in aggregate economic conditions affect the labor market for immigrants? In this paper, the impact of changes in aggregate economic conditions on earnings is assumed to be the same in

15. The difference in the entry wage differential between stayers and dropouts is only marginally significant ($t = 1.6$).

both the immigrant and the native labor market. This normalization leads to a framework where both aging and cohort effects are exactly identified. The importance of this assumption suggests that much can be learned from further research that analyzes different time periods, uses different normalization restrictions, and tries to determine how wage levels in the immigrant and native labor markets respond to changes in aggregate economic conditions.

REFERENCES

Borjas, George J. "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants." *Journal of Labor Economics*, October 1985, 463–89.

_____. "Self-Selection and the Earnings of Immigrants." *American Economic Review*, September 1987, 531–53.

Carliner, Geoffrey. "Wages, Earnings, and Hours of First, Second, and Third Generation American Males." *Economic Inquiry*, January 1980, 87–102.

Chiswick, Barry R. "The Effect of Americanization on the Earnings of Foreign-Born Men." *Journal of Political Economy*, October 1978, 897–921.

_____. "An Analysis of the Economic Progress and Impact of Immigrants." U.S. Department of Labor. Washington, D.C., 1980.

Deaton, Angus. "Panel Data from Time-Series of Cross-Sections." *Journal of Econometrics*, October/November 1985, 109–26.

DeFreitas, Gregory. "The Earnings of Immigrants in the American Labor Market." Ph.D. dissertation, Columbia University, 1980.

_____. "Occupational Mobility Among Recent Black Immigrants." *Proceedings of the 33rd Annual Meetings, Industrial Relations Research Association*, 1981, 41–47.

Fienberg, Stephen E. and William M. Mason. "Identification and Estimation of Age-Period-Cohort Models in the Analysis of Discrete Archival Data," in *Sociological Methodology*, edited by D. Schuessler. San Francisco: Josseu-Bass, 1978.

Greenwood, Michael J. and John M. McDowell. "The Potential Economic Consequences of U.S. Immigration: A Survey." *Journal of Economic Literature*, December 1986, 1738–72.

Hall, Robert E. "Technical Change and Capital from the Point of View of the Dual." *Review of Economic Studies*, January 1968, 35–46.

Heckman, James J. "Sample Selection Bias as a Specification Error." *Econometrica*, January 1979, 153–61.

Heckman, James J. and Richard Robb. "Using Longitudinal Data to Estimate Age, Period, and Cohort Effects in Earnings Equation," in *Analyzing Longitudinal Data for Age, Period, and Cohort Effects*, edited by H. Winsborough and O. Duncan. New York: Academic Press, 1983.

Jasso, Guillermina and Mark Rosenzweig. "How Well Do U.S. Immigrants Do? Vintage Effects, Emigration Selectivity, and the Occupational Mobility of Immigrants." Photocopy, University of Minnesota, June 1985.

Long, James E. "The Effect of Americanization on Earnings: Some Evidence For Women." *Journal of Political Economy*, June 1980, 620–29.

Snipp, C. Matthew and Marta Tienda. "Chicano Occupational Mobility." *Social Science Quarterly*, June 1984, 364–80.

U.S. National Science Foundation, Survey of Natural and Social Scientists and Engineers, 1972–1978. Washington, D.C., 1982.

Welch, Finis. "Effects of Cohort Size on Earnings: The Baby Boom Babies' Financial Bust." *Journal of Political Economy*, October 1979 Supplement, 565–98.