"Longitudinal Data for the Analysis of Mobility in the U.S., 1850-1910"

Joseph P. Ferrie
DEPARTMENT OF ECONOMICS AND
INSTITUTE FOR POLICY RESEARCH
NORTHWESTERN UNIVERSITY
AND
NATIONAL BUREAU OF ECONOMIC RESEARCH

April 21, 2004

Abstract

The recent completion of nationally-representative samples from the late 19th and early 20th century U.S. decennial population censuses and a complete count file for 1880, together with the availability of nominal indexes and manuscripts from these censuses, has created a unique opportunity to construct longitudinal data by following individuals from one enumeration to another. This essay describes a series of new, linked datasets recently created and in progress that will provide longitudinal information on the geographic, occupational, and financial mobility of more than 41,000 Americans from the 1850s to the 1900s. Together with longitudinal data for more recent cohorts–such as the National Longitudinal Surveys (NLS) and the Panel Study of Income Dynamics (PSID)–these data make possible for the first time systematic comparisons of mobility over the last 150 years of American economic development.

Introduction

The history of the U.S. is, in many ways, a history of mobility. From the arrival of European settlers in the Western Hemisphere at the start of the seventeenth century and their progressive movement from a narrow band of colonies perched on the Atlantic seaboard to the conquest of a continent that spans two oceans; from a clusters of farmers working the land in

The National Science Foundation provided financial support for the collection of the 1850-60 and 1860-70 samples (Grant No. 9309689 and Grant No. 9730243).

isolated homesteads to a nation of factory and service workers in a thousand cities and suburbs; from a land rich in promise and poor in material wealth but with its hardships widely shared to a place of unequaled wealth and unprecedented concentration of wealth–in all these ways Americans have moved through geographic space, the occupational hierarchy, and the wealth distribution. But we have precious little data with which to assess the magnitude or the correlates of that movement for anything but the most recent decades of U.S. history.

Sources like tax lists, marriage resisters, and census enumerations provide snapshots of the population at discrete points in time, but it remains difficult to splice those snapshots together to form a coherent description of American mobility. For example, the 1860 federal population census reveals that among adult (age 20+), white native-born males, the top 5% owned 57% of all the wealth held by this group. By 1870, the top 5% owned only 54%.[1] This was a tumultuous decade, in which Southern slave owners saw a substantial fraction of their wealth erased by emancipation, in which much of a generation of prime-age males lost their lives on the battlefields of the Civil War, and in which rampant price inflation and nascent industrialization created new fortunes as it leveled others. Was the declining share of wealth held at the top of the distribution the result of an absolute decline in wealth at the top, or was it primarily caused by the acquisition of small amounts of wealth by those who had previously held none?

With a single set of cross sections, it is impossible to say. Nor can we say with any certainty whether the calculated decline captures all of the interesting movement within the

[1] Calculated using the 1860 and 1870 Integrated Public Use Microdata Series samples for 1860 and 1870 available at http://www.ipums.umn.edu/usa/index.html.

distribution. Two wealth distributions may be identical even if the individuals at each point in the distribution have seen enormous changes in their wealth; the loss of an entire fortune by the richest man in America and the gain of an equal fortune by someone previously without any wealth will result in the same change in the wealth distribution measured in two cross-sections as will absolute rigidity in individuals' relative wealth positions. The same problem applies in assessing the extent of geographic and occupational mobility.

This is not a trivial problem. Fundamental to understanding many attitudes toward the American political and economic systems is how Americans have viewed their own prospects for advancement and improvement. The perception that America's economic system is fluid and that its upper reaches are easily permeable has shaped a willingness to tolerate enormous disparities in economic outcomes. (Alesina and La Ferrara 2001) But we have little basis on which to assess the validity of this perception over much of the nation's history. We have even less capacity to assess the forces that have made easier such mobility as has occurred. For example, how important were geographic mobility and education in the process of movement within the occupational and wealth distributions? In the absence of data that follow for several years or decades the careers of thousands of specific individuals, we cannot say.

The data described here will address many of these shortcomings. With the availability of large, nationally-representative samples from the Integrated Public Use Microdata Series (IPUMS) as well as a complete-count file from the 1880 census, and the recent creation of nominal indexes for all of the extant population censuses 1850-1930, it is now possible at low cost to create samples that follow many individuals for up to thirty years. These data will make

it possible to compare the mobility experienced in a variety of dimensions–geographic, occupational, and financial–by Americans in the eight decades after 1850 to mobility in more recent years, and to assess how the sources of mobility have changed over the last 150 years. This essay provides a brief overview of the linked samples that now exist, describes those in progress, and offers some preliminary findings on long-run trends in mobility. Four types of comparison are pursued: (1) differences in mobility over a particular period for different demographic groups (defined by race); (2) differences in mobility rates over time (comparing mobility in different decades); (3) differences between the U.S. and Britain in mobility; and (4) differences between mobility in the 1850-1910 period and mobility in the 1969-99 period.

Previous Linked Samples

The earliest efforts to follow specific individuals across censuses were designed to measure the turnover in rural communities in the second half of the nineteenth century. James Malin's work in the 1930s on Kansas (1935), Merle Curti's in the 1950s on Trempealeau County, Wisconsin (1959), and Allan Bogue's in the 1960s on the Illinois and Iowa corn belt (1963) developed a methodology that was pursued by other scholars in one way or another for most of the period down through the middle of the 1980s. A sample was drawn from a county's or township's population in some base year, and those same individuals were sought in a subsequent enumeration for the same county. This yielded one piece of information immediately: the fraction of individuals who were successfully located within the second enumeration: the "persisters." For those individuals who survived to the second census and who were captured in the enumeration, the process also yielded information on how their fortunes

(the make-up of their families, where in the county or township they lived, their occupation and wealth) had changed over the intervening years.

By the 1960s, the work of Stephan Thernstrom (1964 and 1973) and others applied the same methodology to urban places as well. By the 1980s, studies had been completed for dozens of communities. One persistent shortcoming of this entire literature, though, stretching all the way back to Malin's work and applying with equal force to each of the studies conducted through the early 1980s, was the inability to say anything about the fate of the "non-persisters"–the sixty percent or so of the base year population who could not be found in a subsequent enumeration.[2] Surely some of them had died, and some had remained in the community and escaped enumeration. But a substantial fraction were no doubt out-migrants. We could have greater confidence in drawing conclusions solely on the basis of persisters if the non-persisters were a simple random draw from the base year population. But they were seldom so cooperative. Non-persisters compared to persisters were invariably younger, less skilled, and less likely to have property or familial connections to the communities in which they failed to re-appear. Given that migration had at least some economic and psychic costs, we would also be surprised to find that persisters and non-persisters behaved similarly over a decade or more: the non-persisters might have expected at least some benefit from their movement to compensate them for the cost they have borne in undertaking it.

This difficulty with community-based persistence studies was recognized by Thernstrom and others. But in the absence of a means of sifting through the enumerations for the rest of the

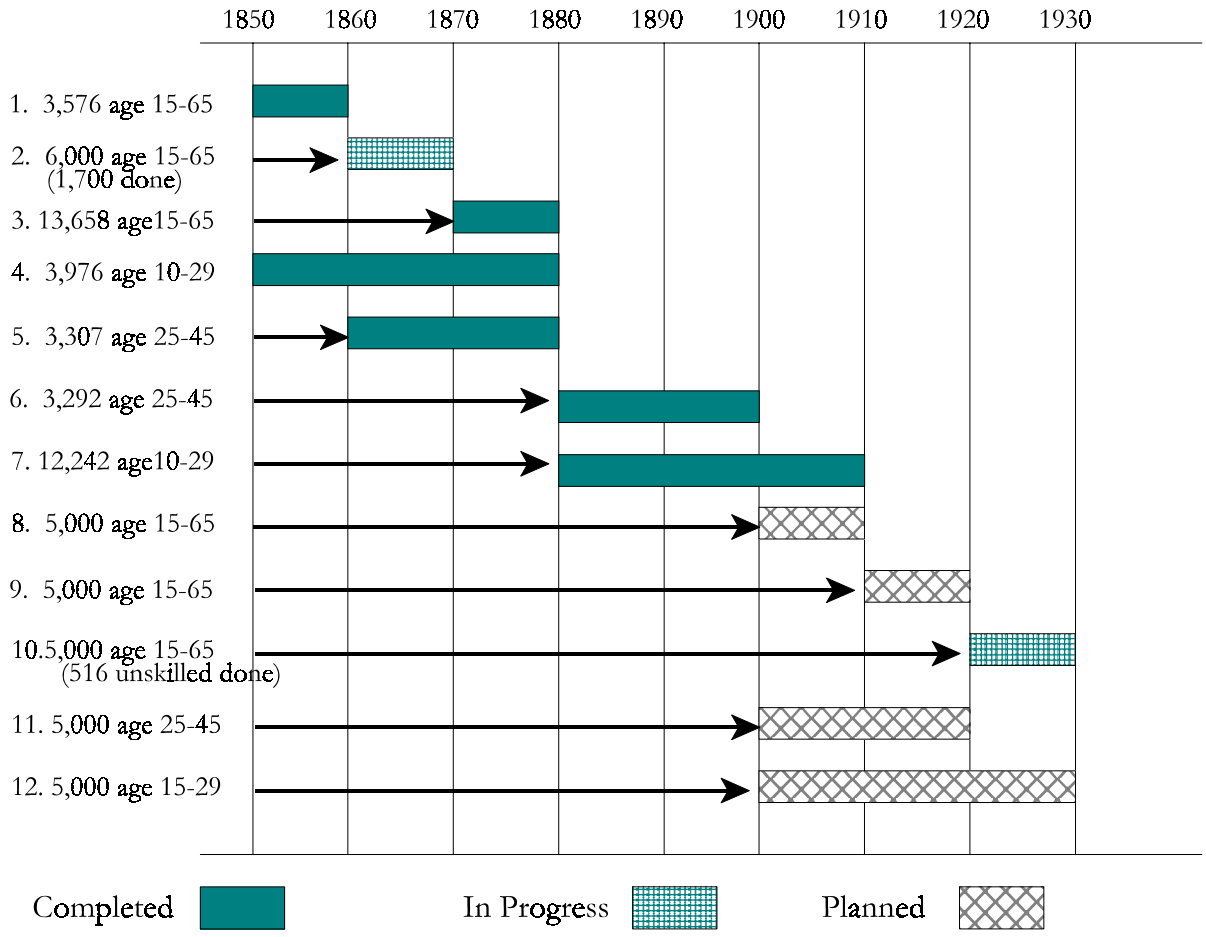
[2] Parkerson reports the average persistence rate in 63 such community-based studies was 38.3% (1982, 102).

country to locate the non-persisters, there was little to be done. By the late 1980s, however, the ability to follow non-persisters regardless of how far they moved was improved by the appearance of nominal indexes to the 1850 population censuses. These indexes made possible studies by Steckel (1988) and Schaefer (1985) in which samples from the 1860 population census were located in the 1850 census a decade earlier. Because the individual's reported state of birth (which was not included in the census until 1850) was essential in the linkage process, while the 1850 census was the only enumeration completely indexed at the time, backward linkage was the only feasible strategy. Further, because a separate index existed for each state, it was necessary to limit the 1860 sample to males with a child ten or more years in age, so the child's birthplace could indicate the state index most likely to yield a positive match. As modern studies find economic mobility concentrated among those in their twenties and early thirties, this strategy unfortunately eliminated what was likely the most mobile segment of the population. Subsequent work by Knights on Boston (1991), Herscovici on Newburyport (1998), Davenport on Schoharie County, New York (1984), and Guest *et al.* on the National Panel Study (1987) followed populations forward into later censuses, but each study was limited either by just one point of origin or low rates of linkage. For example, the National Panel Study that traced young men forward from the 1880 census to the 1900 census located only 160 of the 4,041 (4.0 percent) individuals sought outside their 1880 state of residence; the linked 1880-1900 sample described below shows that 32.6 percent of males age 25 to 45 in 1880 had moved to another state by 1900.

In 1996, Ferrie took advantage of an early version of the 1850 IPUMS and a new index to the 1860 population census to create a new sample that overcame many of these difficulties. It used a nationally-representative base-year population drawn from thousands of locations, and included nearly 5,000 men of all ages linked forward from 1850 to 1860. The release of subsequent IPUMS units, together with the release of the complete-count file from the 1880 census and indexes for all of the extant censuses from 1850 to 1930 now makes possible the creation of similar linked data, and the comparison of mobility trends over time.

New Linked Samples Completed and In Progress

Figure 1 shows the progress to date in creating new linked samples. The samples that use the 1880 sample for either the initial or terminal year were constructed electronically, by taking advantage of the complete-count file for 1880 and the IPUMS files for the other year. The other samples employ some hand linkage using the existing census indexes and the IPUMS files for the other year. For example, the 1860-70 files uses the 1860 IPUMS and the 1870 index. The intervals were chosen to yield the maximum number of ten, twenty, and thirty year linkages as possible at the minimum cost. In each linkage, individuals were identified as possible links on the basis of Soundex-coded surname and a consistently-coded given name (so "Jas." and "James" were both assigned the same code, for example), age (with an allowance for up to three years between reported age in the base year and predicted age in the target year plus the length of the interval, so a 30 year old in 1850 could be matched to someone 57 to 63 in 1880, for example), state of birth, race, and parents' birthplaces (where it was available as a census question itself or where it could be constructed for some individuals based on the birthplaces



of parents listed prior to the individual, as in the 1850 census).[3] For the linkages done by hand, additional information on the names, ages, and birthplaces of spouses and children was used to reject unlikely matches (though this eliminates from the sample those individuals who entered a completely different family between enumerations).

[3] Though it is possible to link immigrants and they have already been linked in the 1850-60 and 1880-1900 samples using these techniques, for the sake of consistency the results reported here will pertain only to the native-born. Subsequent versions of all the linked data will contain immigrants. Beginning in 1870 when free blacks first appear in the population schedules in large number, this group is included in the results that follow.

A detailed examination of the sources of linkage failure for the 1850-60 linkage is contained in Ferrie (1996). The greatest difficulty was common names, though the frequency with which a particular combination of surname and given name appeared was unrelated to any observable characteristics, so this did not impart a substantial bias to the linked sample. The newer samples constructed in this project were less susceptible to even this difficulty. The indexes now, in most cases, contain the individual's birthplace and age, making it far faster to search multiple matches. For the samples created using IPUMS files and the 1880 complete-count data, it was possible to partition individuals into very small groups based on full name, age, race, birthplace, and parents' birthplaces, eliminating most multiple matches in the process.

Other sources of linkage failure were mortality, omission from or incorrect enumeration in the target census, separation from the parental household, and omission from the index for the target census. For the linked samples created more recently using IPUMS files and census indexes (1860-70 and 1920-30), these are also the most likely sources of linkage failure, with mortality likely a particular problem for the 1860-70 linkage, as a result of the Civil War. For the sample linked backward to the 1880 complete count file (1880-1900), only correct inclusion in the target census remains an issue, as mortality and departure from the parental household were overcome by searching the entire file and searching only individuals who had survived to the terminal census; for the samples linked forward to 1880 (1850-80, 1860-80, and 1870-80), mortality remains a source of linkage failure.

The availability of Public Use Samples makes it possible to overcome some of these biases. By partitioning the linked data into cells and constructing weights for those cells based

	(1) Unweighted	(2) Weighted	(3) Unweighted	(4) Weighted
Age	0.001	0.001	0.001	0.001
	(2.95)**	(1.92)	(2.90)**	(1.88)
$Age^2 \times 10^{-2}$	-0.001	-0.001	-0.001	-0.001
	(2.20)*	(1.83)	(2.16)*	(1.80)
Midwest	-0.009	-0.003	-0.009	-0.003
	(5.13)**	(1.52)	(5.15)**	(1.54)
South	-0.015	-0.003	-0.015	-0.003
	(7.69)**	(1.40)	(7.62)**	(1.34)
Mountain and Pacific	-0.019	-0.016	-0.019	-0.016
	(4.47)**	(3.17)**	(4.48)**	(3.18)**
Farmer	0.008	-0.001	0.008	-0.001
	(4.70)**	(0.72)	(4.62)**	(0.79)
Laborer	-0.007	-0.006	-0.007	-0.006
	(2.85)**	(2.15)*	(2.81)**	(2.12)*
Real Estate $\times 10^{-6}$			0.012	0.015
			(1.71)	(2.46)*
Personal Estate $\times 10^{-6}$			-0.015	-0.019
			(0.94)	(1.67)
Observations	49,164	49,164	49,164	49,164
Pseudo-R^2	0.011	0.002	0.012	0.002
Predicted Probability	0.033	0.035	0.033	0.035

Absolute value of z-statistics in parentheses
* significant at 5%; ** significant at 1%

Table **1**. Probit Regression (Partial Effects) on Linkage 1860-70.

on nationally-representative cross-sectional samples, the linked samples can be made more representative of the population of males who survived the years that they span. For example, the 1860-70 linked data was partitioned by characteristics in 1870: age group (under 45, 45 and over), occupation (farmer, non-farmer), and birthplace (Northeast, Midwest, South & Far West). The representation of these cells in the 1870 population was determined from the Public Use Sample, and weights were constructed so these cells in the linked data represented the same proportions of the U.S. white, native-born male population age 25-75 in 1870.

	(1) Unweighted	(2) Weighted
Age.	0.011	0.011
	(1.10)	(1.07)
Age^2	-0.009	-0.010
	(1.03)	(1.07)
Midwest	-0.022	-0.002
	(3.10)**	(0.23)
Souh	-0.041	0.014
	(4.97)**	(1.60)
Mountain and Pacific	-0.037	-0.018
	(2.80)**	(1.28)
Farmer	0.032	0.005
	(4.82)**	(0.71)
Laborer	-0.013	-0.013
	(1.45)	(1.32)
Black	-0.089	-0.022
	(9.37)**	(1.92)
Observations	17,746	17,746
Pseudo-R^2	0.013	0.001
Predicted Probability	0.182	0.185

Absolute value of z-statistics in parentheses
* significant at 5%; ** significant at 1%

Table **2**. Probit Regression (Partial Effects) on Linkage 1880-1900.

Tables 1 and 2 show the results of probit regressions for the 1860-70 and 1880-1900 linkages, using the observed characteristics in the later census to explain linkage success. Results with and without weights are shown. The figures shown are marginal effects. For example, the marginal effect of -0.009 for "Midwest" in Column 1 of Table 1 means that a resident of the Midwest was 0.9 percentage points less likely to be successfully linked from 1860 to 1870 than a resident of the Northeast. The general pattern is one of decreased importance for the terminal year characteristics in explaining linkage as weights are introduced. For example, the effect of residence in the Midwest is reduced by a factor of three and the effect of residence in the South

by a factor of five. Fewer characteristics exhibit statistical significance when weights are employed. The only exceptions to this pattern are the wealth variables which, if anything, are more important at distinguishing linked from unlinked individuals when weights are used. This suggests the need to introduce an additional cell in the weighting process to take account of differential linkage by wealth.

Table 2 shows a similar pattern for 3,200 individuals followed for twenty years, from 1880 to 1900. Each of the characteristics that was substantively or statistically significant in explaining linkage with unweighted data was reduced to substantive and statistical insignificance by the introduction of weights. For example, blacks were half as likely as whites to be linked according to Column (1), but the weighted results in Column (2) reduce the gap to one ninth in size and statistical insignificance.

Geographic Mobility

The first use to which these new linked samples can be put is the calculation of rates of geographic mobility. Though synthetic cohorts created with successive census cross-sections can reveal patterns of net mobility across states, linked data can reveal both gross flows across state boundaries and flows across even smaller units, such as counties, townships, and wards. Table 3 shows rates of interstate and intercounty migration for native-born males over ten, twenty, and thirty year intervals. Though migration rates generally increase in going from ten to twenty to thirty year rates, this is no doubt the result of the increased time that individuals are at risk to make a change in location as the interval between the dates when their locations are observed

increases. More meaningful comparisons are those holding constant the length of th interval and varying the dates of the end-points.

Mobility rates fell over the Civil War decade, but increased dramatically in the 1870s in the war's wake. The 1860-80 period saw somewhat greater migration, particularly interstate migration, than the 1880-1900 period. In the three decades before 1880, the 1870s show the highest migration rate. The twenty-year and thirty-year migration rates all show some decrease over time, with the fraction remaining in the same county higher in 1880-1900 and 1880-1910 than in 1860-80 and 1850-1880. Both rates show a shift away from interstate migration and toward intercounty migration, probably as a result of the growth of cities and reduced westward

	N	Same County	Different County, Same State	Different Sate
Ten-Year Rate				
1850-60*	3,457	59.6%	19.3%	21.1%
1860-70*	1,711	66.5	18.1	15.5
1870-80				
White	11,681	49.9	22.5	27.6
Black	1,971	26.6	34.2	39.2
Twenty-Year Rate				
1860-80*	3,307	41.9	21.5	36.6
1880-1900				
White	2,839	47.9	23.4	28.7
Black	446	25.2	38.4	36.4
Thirty-Year Rate				
1850-80*	3,976	38.1	26.1	35.8
1880-1910				
White	10,836	39.6	29.8	30.6
Black	1,396	26.5	42.7	30.8

Table 3. Migration rates for native-born males, age 15-65 (1850-60, 1860-70, 1870-80), age 25-45 (1860-80, and 1880-1900), and age 10-29 (1850-80, 1880-1910).

migration. Most frontier-bound migrants needed to cross both state and county boundaries to reach their destination, unlike city-bound migrants who could often travel to a nearby city in a nearby county without crossing a state border. The migration rates for blacks are uniformly higher than for whites, even in the years immediately following the Civil War, well before the Great Migration of the twentieth century.

For purposes of comparison, young native-born males (age 20-29) in 1971 drawn from the National Longitudinal Survey (NLS) Young Men Cohort were substantially less likely to make interstate or intercounty moves than males in any of the linked samples described here. Between 1971 and 1981, 36.3% of the NLS males made an intercounty move, and 18.9% made an interstate move (N=2,430). When the ten-year intervals in Table 3 are limited to males age 20-29 at the start of the interval, intercounty migration rates were 50% (1850-60), 41% (1860-70), and 56% (1870-80) for whites, and 77% (1870-80) for blacks; interstate migration rates were 27% (1850-60), 19% (1860-70), and 30% (1870-80) for whites, and 39% (1870-80) for blacks. Compared to the NLS, males in the mid-nineteenth century was roughly a third more geographically mobile.

The PSID allows some additional comparisons. Though intercounty moves can be examined only with special permission, the publicly available PSID files can be used to examine interstate mobility over intervals as long as 30 years. Age ranges that correspond to those in Table 3 were used. Ten-year interstate migration rates ranged from 11 to 13 percent, twenty year rates ranged from 14 to 19 percent, and the thirty-year rate (1969-99) was 30 percent. Only for

the thirty-year rate was recent mobility as frequent as historical mobility: the rate for 1880-1910 was also 30% for both blacks and whites.

Net migration calculations can obscure interesting patterns of migration only apparent in gross flows. Using the 1850-80 linked sample, Figures 2 and 3 show gross out-migration and in-migration by states. Clearly, the states located on the 1850 frontier were places of significant

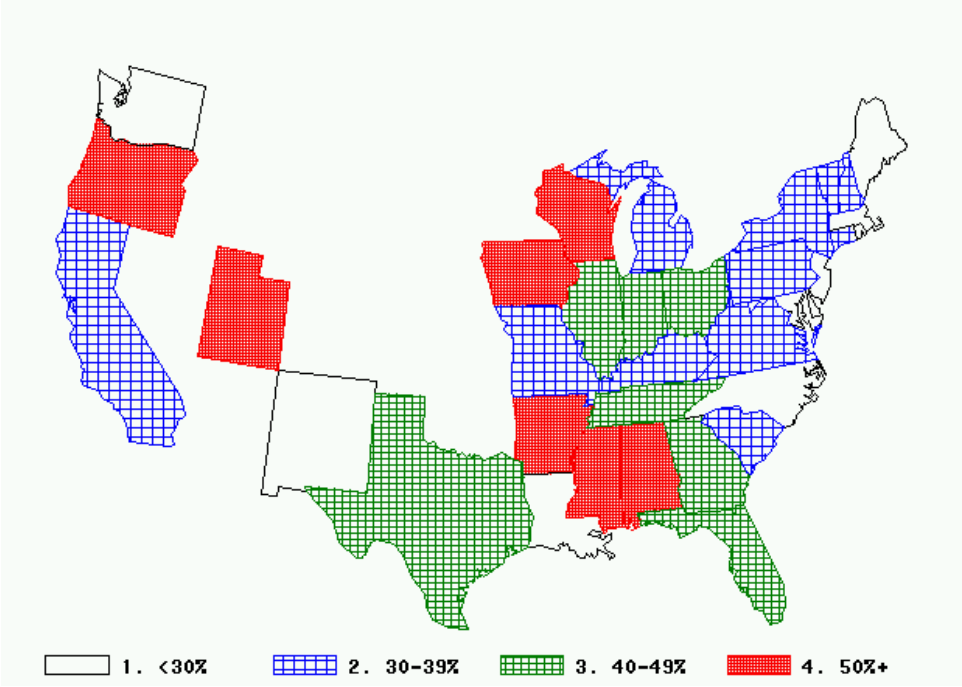


Figure 2: U.S. Rates of Out-Migration, 1850-80.

population turnover: they both received large numbers of in-migrants and supplied large numbers of out-migrants to states still farther west.

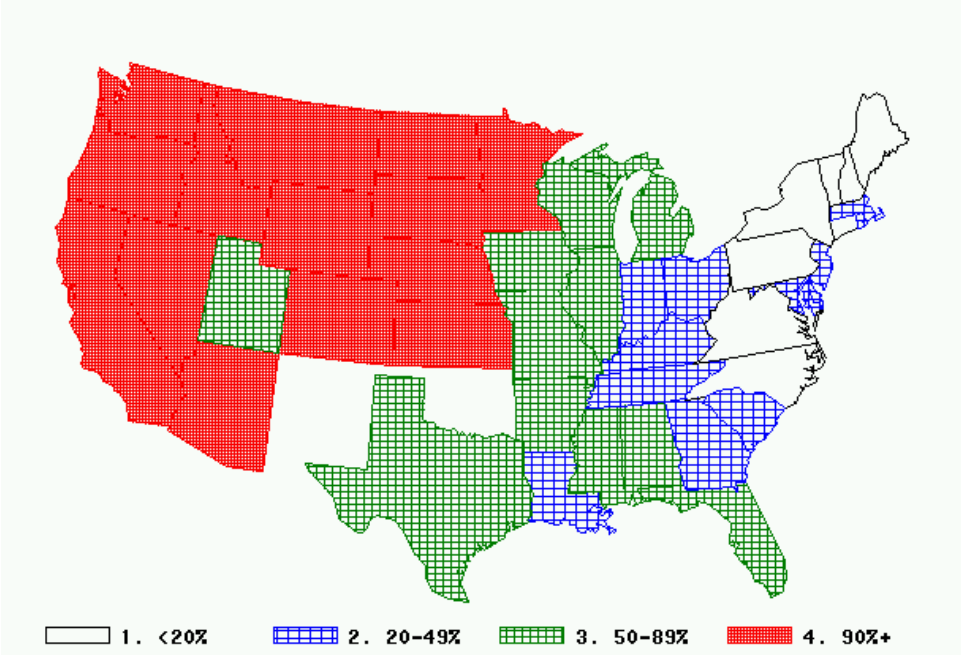


Figure 3: U.S. Rates of In-Migration, 1850-80.

Occupational Mobility

The linked samples can also be used to explore occupational mobility, as occupation was reported in each of the census years used. The 1850-80 sample has been used in Ferrie and Long (2003) to compare occupational mobility in the U.S. to mobility in Britain (for which a comparable 1851-81 linked sample of nearly 30,000 young males has now been created).[4] Figures 4 and 5 show rates of occupational mobility between fathers and sons, and between the sons' first and last jobs. The comparison reveals considerably more mobility in the U.S. than in Britain.

[4] Weights have not yet been created for the British linked sample, so the comparisons between the U.S. and Britain here employ the unweighted data. The weighted data is again used in Tables 4 and 5.

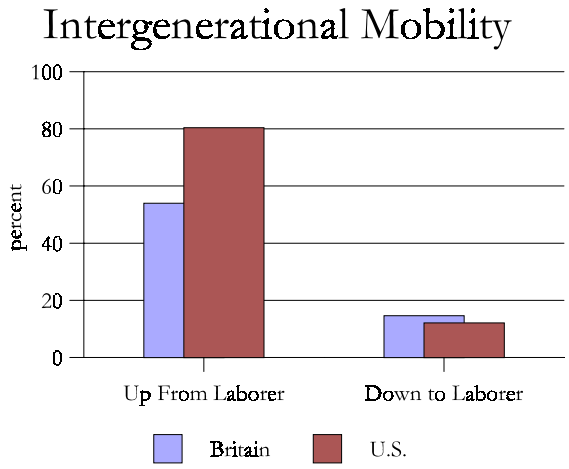


Figure 4. Occupational mobility, 1850-80.

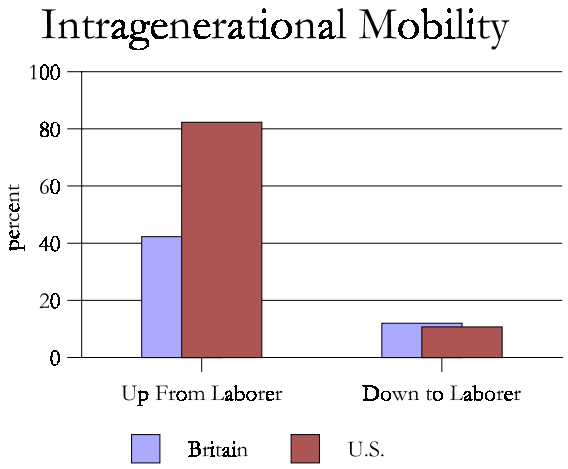


Figure 5. Occupational mobility, 1850-80.

Using the simplest concept of occupational mobility (up from unskilled laborer *to* any other job or down to unskilled laborer *from* any other job), the comparison reveals more mobility upward both across and within generations in the U.S., and roughly the same chances of downward mobility in both places. Though the economies of the U.S. and Britain were similar in many respects, their occupational structures differed markedly (with far fewer farm owners, and a great deal more skilled and semi-skilled workers in Britain). To reduce this source of incompatibility, two strategies were followed. The first examined occupational mobility while ignoring the farm sector completely, so upward moves from unskilled laborer could occur only through entry into white collar, skilled, or semi-skilled jobs. The second compared mobility in Britain in the period 1851-81 to mobility in the U.S. in the period 1880-1910 by which time more urban and industrial growth had occurred in the U.S. Neither strategy altered the strong finding of roughly twice as much upward occupational mobility in the U.S. as in Britain.

	Up from Laborer/Semi-Skilled		Down to Laborer/Semi-Skilled	
	N	%	N	%
1850-60*	651	57.3	2,314	10.9
1860-70*	325	53.6	1,128	14.6
1870-80				
White	3,456	63.5	6,486	14.4
Black	1,266	36.9	276	48.1
1860-80*	592	71.3	2,426	13.3
1880-1900				
White	692	62.5	2,128	16.3
Black	124	41.9	88	26.0
1850-80*	190	71.4	611	14.9
1880-1910				
White	1,602	68.7	1,988	17.7
Black	275	54.9	62	30.0

Note: "Laborer" includes 1950 occupation codes 700-970 and higher (service workers, farm laborers, and laborers). "Semi-Skilled includes occupations 600-690 (operatives and kindred)

Table 4. Occupational mobility rates for native-born males, age 15-65 (1850-60, 1860-70, 1870-80), age 25-45 (1860-80, and 1880-1900), and age 20-29 (1850-80, 1880-1910).

Similar calculations for the other linked samples are shown in Table 4. The mobility measured is between white collar (professionals, merchants, and clerks), highly skilled blue collar jobs (carpenters, blacksmiths, jewelers), and farming on the one hand and less skilled blue collar workers (factory operatives) and unskilled common laborers on the other. Fewer than a third of the least skilled workers in the labor force in 1860 failed to move into higher status occupations by 1880, with the majority of the upward moves into farming. Half of the unskilled workers in 1860 were farmers by 1880. Rates of downward mobility were quite modest. For whites, rates of upward mobility diminished between 1860-1880 and 1880-1900, and rates of downward mobility increased somewhat. Entry into farming became less frequent, though it still accounted for most of the upward mobility.

Comparison of mobility rates for blacks and whites in the 1870-80 and 1880-1900 samples show less upward mobility for blacks than for whites, though more than a third of unskilled blacks in 1880 had become farmers by 1900. The gap between black and white upward mobility rates narrowed over the late nineteenth century, from 27 percentage points in the 1870s to 20 percentage points in the 1860-80 period, to 14 percentage points in the 1880-1910 period. Downward mobility rates were nearly three times as great for blacks as for whites in this period, though, suggesting that such gains as blacks achieved were somewhat tenuous.

As was mentioned above, for the 1850-80 and 1880-1910 samples, it is possible to measure intergenerational mobility as well as the mobility an individual experienced over his own career. Of white males whose fathers were semi-skilled or unskilled in 1850, 62% (N=316) achieved white collar or skilled jobs or entered farming by 1880; over the following three decades, an identical fraction of sons of semi-skilled and unskilled fathers (N=825) moved up to white collar or skilled job or entered farming. For blacks, this measure of intergenerational mobility was 42% (N=136) 1880-1910. Downward intergenerational mobility for whites rose slightly from 18% (N=2,152) 1850-80 to 20% (N=3,978) 1880-1910. A third of blacks whose fathers were in white collar or skilled jobs or were farmers in 1880 were in semi-skilled or unskilled jobs themselves in 1910.

The NLS and PSID again provide modern mobility data to compare to the linked historical samples. When native-born males age 20-29 in 1971 are examined in 1981, 44% of those who were in semi-skilled or unskilled jobs in 1971 had attained white collar or skilled

jobs (50% of whites and 32% of blacks). When the ten-year historical samples were limited to males age 20-29 in the initial year, upward mobility rates for whites rose from 56% (N=271) in the 1850s to 57% (N=131) in the 1860 to 66% (N=1,380) in the 1870s. For blacks, 38% moved up over the 1870s. At younger ages, then, upward occupational mobility was more likely in the 1850-80 period than in the 1970s, though the importance of farming in the second half of the nineteenth century accounts for much of the difference.

The PSID allows comparisons to the historical linked data that span ten, twenty, and thirty years; it also allows intergenerational mobility calculations for the interval 1969-99. In the 1960s, 1970s, and 1980s, 20 to 24% of those who began a decade in a semi-skilled or unskilled job moved up to a white collar or skilled job by the end of the decade. When the historical samples are restricted to males age 30 at the start of a decade, the rates of upward mobility for the 1850s and 1860s were more than twice as great, while the rate in the 1870s was three times the modern rate over a decade. Over twenty years in the PSID (1969-89 and 1979-99), 28-34% moved up, compared to 63-71% in the historical samples. Over thirty years, modern upward intragenerational mobility was 44% 1969-99, more than 37% lower than the historical rates for 1850-80 and 1880-1910. Intergenerational upward mobility 1969-99 was 43% in the PSID, 30% lower than in the historical thirty-year samples.

Wealth Accumulation

Finally, the linked samples make possible investigation of the mobility within the wealth distribution experienced by American males.[5] This is particularly illuminating for the

[5] Steckel (1990) has examined movement within the real estate wealth distribution 1850-60, and Steckel and Krishnan (1992) have compared mobility in the real estate wealth distribution 1850-60 to

1860-70 linkage, as the census reported both real and personal property in both years, making this the only time in the 19th century for which wealth mobility can be assessed using a comprehensive definition of wealth, similar to that available in more recent surveys from the late 20th century. Though the size of the 1860-70 sample is still quite small, it is clear that mobility can be discerned: Table 5 shows probabilities for making the transition from quartiles of the 1860 distribution to quartiles of the 1870 distribution for males age 30-65 in 1860.

Total Wealth Quartile in 1860	Total Wealth Quartile in 1870 Top	2nd	3rd	Bottom
Top	64.5%	19.9%	10.5%	5.1%
2nd	21.8	43.1	27.2	7.9
3rd	8.6	22.3	44.0	25.0
Bottom	4.8	14.9	24.0	56.2

Table 5. Transition probabilities from 1860 total wealth quartiles to 1870 total wealth quartiles, native-born males age 30-65 in 1860. (N=877)

Males who began the 1860s below the median in total wealth ($1,500) had a better than one-in-four chance of moving into one of the top two quartiles by the end of the decade. Conversely, those in the top decile had a better than one-in-three change of descending to a lower quartile over the decade. It remains to be seen how these transitions differed across the

mobility in the National Longitudinal Survey. But because of the age of the linked sample (limited to men old enough to have a child ten or more years in age by 1860) and the absence of personal wealth in the 1850 census, these comparisons can be improved upon with the linked data offered here.

country (remember that substantial personal wealth in slaves was erased by emancipation) and over the life cycle.

Comparisons with wealth transitions in modern data like that contained in the NLS and PSID are also now possible. Table 6 presents similar wealth transitions for the 1984-94 PSID for males age 30-65 in 1984. In order to enhance comparability with the 1860-70 data, the PSID data has been weighted to reflect the 1860-70 age distribution (which was somewhat younger, with a median age ten years below that in the more recent data). The most striking differences between Table 5 and Table 6 is the apparent sharp drop in the prospects for upward wealth mobility for the very poor over the century and a quarter after 1860, with the lowest quartile's chances of moving above the median falling by a factor of two (from roughly 0.18 to 0.09). The biggest change was not the drop in the odds of moving to the top quartile, but the drop in the odds of moving up into the second quartile. Even as the prospects for those at the bottom have shrunk, though, the mobility of those closer to the median has increased from the 1860s to the 1980s: 43% of those in the third quartile rose above the median in recent years, compared to only 30% in the 1860s.

Total Wealth Quartile in 1984	Total Wealth Quartile in 1994			
	Top	2nd	3rd	Bottom
Top	61.9%	20.3%	14.4%	3.8%
2nd	25.5	39.4	28.6	6.4
3rd	10.5	32.6	37.8	19.2
Bottom	3.7	5.2	28.7	62.4

Table 6. Transition probabilities from 1984 total wealth quartiles to 1994 total wealth quartiles, males age 30-65 in 1984. (N=700) Source: PSID.

Conclusions

The samples described here will be of great use in assessing patterns of geographic, occupational, and financial mobility. They will make possible the analysis of these patterns over the long-run as well, in that they will permit comparisons to more recent surveys like the NLS and PSID. Future work will focus on: 1) completion of the samples listed as "In Progress" or "Planned" in Figure 1; 2) strengthening the matches in the 1850-80, 1870-80, 1860-80, and 1880-1900 samples by examining the names, ages, and birthplaces of the family members of the linked individuals to eliminate "false positives;" 3) adding more young males to the 1860-80 and 1880-1900 samples; 4) adding intergenerational information for samples that include young males still living at home; and 5) providing consistently coded versions of all samples and making them publicly available.

The analysis performed here has revealed a substantial decline in geographic and occupational mobility between the historical samples and modern samples (both the NLS and PSID). Wealth mobility has also changed over the last 150 years, though in a more subtle manner: movement from the bottom quarter of the wealth distribution to the upper half is now less common than it was in the past, but more modest movement from just below the median to just above it has improved. The sources of these changes, and of the changes in geographic and occupational mobility, wil be explored in subsequent work.

References

Alesina, Alberto, and Elina La Ferrara (2001). "Preferences for Redistribution in the Land of Opportunities." National Bureau of Economic Research Working Paper No. w8267 (May).

Bogue, Allan (1963). *From Prairie to Corn Belt; Farming on the Illinois and Iowa Prairies in the Nineteenth Century*. (Chicago: University of Chicago Press).

Curti, Merle (1959). *The Making of an American Community; a Case Study of Democracy in a Frontier County*. (Stanford: Stanford University Press).

Davenport, David (1984). "Tracing Rural New York's Out-Migrants, 1855-1860." *Historical Methods* (17: 59-67).

Ferrie, Joseph (1996). "A New Sample of Males Linked from the Public Use Micro Sample of the 1850 U.S. Federal Census of Population to the 1860 U.S. Federal Census Manuscript Schedules." *Historical Methods* (29: 141-156).

Ferrie, Joseph, and Jason Long (2003). "A Tale of Two Labor Markets: Career Mobility in the U.S. and Britain." Unpublished.

Guest, Avery, *et al.* (1987). "Notes from the National Panel Study." *Historical Methods* (20: 63-77).

Herscovici, Steven (1998). "Migration and Economic Mobility: Wealth Accumulation and Occupational Change Among Antebellum Migrants and Persisters." *Journal of Economic History* (58: 927-56).

Knights, Peter (1991). *Yankee Destinies : the Lives of Ordinary Nineteenth-century Bostonians.* (Chapel Hill: University of North Carolina Press).

Malin, James (1935). "The Turnover of Farm Population in Kansas." *Kansas Historical Quarterly* (20: 339-372).

Parkerson, Donald (1982). "How Mobile Were Nineteenth Century Americans? *Historical Methods* (15: 99-109).

Thernstrom, Stephan (1964). *Poverty and Progress; Social Mobility in a Nineteenth Century City*. (Cambridge: Harvard University Press).

Thernstrom, Stephan (1973). *The Other Bostonians; Poverty and Progress in the American Metropolis, 1880-1970*. (Cambridge: Harvard University Press).

Schaefer, Donald (1985). "A Statistical Profile of Frontier and New South Migration, 1850-1860." *Agricultural History* (59: 563-578).

Steckel, Richard (1988). "Census Matching and Migration: A Research Strategy." *Historical Methods* (21: 52-60).

Steckel, Richard (1990). "Poverty and Prosperity: A Longitudinal Study of Wealth Accumulation." *Review of Economics and Statistics* (72: 275-85).

Steckel, Richard, and Jayanthi Krishnan (1992). "Wealth Mobility in America: A View from the National Longitudinal Survey." National Bureau of Economic Research Working Paper No. w4137 (August).