Intergenerational Occupational Mobility in the U.S., 1850-2013

Joseph Ferrie, Northwestern University
Catherine Massey, Census Bureau
Jonathan Rothbaum, Census Bureau

Abstract

Intergenerational mobility in the U.S. is generally studied using only two generations. We construct and analyze new three-generation samples spanning 1850 to the present: one linking the 1850, 1880, and 1910 censuses; and another linking the 1940 Census, the 1973-1990 Current Population Surveys, and Census 2000 and the 2001-2013 American Community Surveys. We also break up the three-generation samples into multiple parent-child observations to explore changes in mobility across two-generations over time. We find a statistically significant correlation between outcomes of grandchildren and their grandparents. From 1850-1910, occupational wealth of adult grandchildren was 0.15 percent higher for every 1 percent increase in the occupational wealth of the grandparent. From 1940 to the present, the educational attainment of an adult grandchild is 0.29 years greater for every 1 year increase in their grandparent’s 1940 educational attainment. These relationships remain after accounting for the direct parent-child relationship.

Extended Abstract

INTRODUCTION/BACKGROUND

Intergenerational mobility and its determinants have been the focus of extensive academic research and public discussion. One fact clearly emerges from the literature (summarized in Black and Devereaux 2011): parental outcomes matter for child economic success. However, we still do not know how status persists across more than two generations or how mobility has changed over time.

An abundance of recent research examines intergenerational mobility from parent to child. There is strong evidence that mobility is the same for children entering the work force today as it was for those born in the 1970s (Chetty et al. 2014, Hertz 2007, Lee and Solon 2009). However, estimates over longer periods suggest that mobility has not been stable, including some evidence of decline (Aaronson and Mazumder 2008) or of large gains from 1940 to 1970 with potential declines afterwards (Hilger 2015). In addition, research suggests that the mobility of children observed from 1850 to 1880 was greater than the mobility of children observed in the 1970s (Long and Ferrie 2013). These analyses, however, cannot speak to changes in mobility over the course of the nineteenth and twentieth centuries or how accounting for more generations of a family may affect our understanding of intergenerational transmission of status.
Due to data limitations, analyses of mobility are often limited to two generations. Attempts to examine more than two generations have been limited to small samples, non-U.S. data, and comparisons of income averages by surname over time (Clark 2014). To overcome these issues we link individuals across U.S. census data to examine changes in intergenerational and multigenerational mobility in the U.S. from 1850 to the present. Our analysis examines both occupational and income mobility, as well as mobility in education attainment.

We create two separate three-generational links: 1) using the 1850, 1880, and 1910 decennial censuses and 2) using the 1940 Census, the 1973-1990 Current Population Survey (CPS) Annual Social and Economic Supplements (ASEC), and Census 2000 and 2003-2013 American Community Surveys (ACS) to compare multigenerational mobility over time. Lastly, we observe and compare mobility across multiple two-generational links from our two three-generation samples. Our analysis examines both occupational and income mobility, as well as mobility in education attainment.

DATA AND METHODS

We use several sources of census decennial and survey data to measure mobility from 1850 to 2013. To create three-generational linkages spanning 1850 to 1910, we use the 1850, 1880, and 1910 decennial censuses. We begin with a sample of male children co-residing with a male household head in the full-count 1850 U.S. Population Census. We observe the father’s occupation and construct a parent-child link for these children. We then link these children to the full-count 1880 U.S. Population Census to observe their adult occupation and to construct a parent-child link with their sons. We link these sons to their observations in the 1910 Census to observe the son’s adult occupation, which is the occupation of the grandchild of the household head originally observed in 1850. To link records across censuses, we use first and last name, age, birthplace, and both parents’ birthplaces. We require that each individual’s own birthplace and the birthplaces of both parents match exactly, but we allow a difference of up to three years in anticipated age across census observations. For names, we require that phonetic codes for first and last name match exactly, keeping the potential match with the closest name distance, as measured by the SAS SPEDIS function.

To create three-generational links across the twentieth century, we use the complete count 1940 Census, the 1973-1990 CPS ASEC, the 2000 Long Form Census, and the 2003-2013 ACS. We employ one of the Census Bureau’s unique linkage keys, called Person Identification Keys (PIK), to link individuals across data sources. PIKs are assigned by the Person Identification Validation System (PVS), which uses a probabilistic matching algorithm to compare characteristics of records in census and survey data to characteristics of records in a reference file constructed from the Social Security Administration (SSA) NUMIDENT file and other administrative data. These characteristics may include Social Security Number (SSNs), full name, date of birth, address, and parents’ names depending on the information available in the
census or survey. The PIK uniquely identifies a particular person and allows us to link individuals across PIKed data sources.

For the contemporary linkages, we begin with a sample of both male and female children assigned PIKs in the 1940 Census. We observe these children living with their parents in 1940, creating the first parent-child link. We then search for these PIKs in the 1973, 1978, and 1981-1990 CPS ASEC. From this linkage, we observe the outcomes of the second generation as adults. We create the parent-child link from the second to the third generation by observing the second generation residing with their own children. We then locate the PIKs of the third generation observation in the 2000 Long Form Census or the 2003-2013 ACS.

To measure mobility between two generations, we regress child outcome \( Y_{it} \) on parent outcome \( Y_{it-1} \)

\[
Y_{it} = \alpha + \beta_1 Y_{it-1} + \epsilon_{it}.
\]

We also include fixed effects for the year we observe the child’s adult outcome and fixed effects for the year we observe the parent’s outcome as well as age and age-squared terms. We consider several outcomes, including log income, log earnings, log occupational prestige scores, log income rank, and educational attainment. Education attainment and income are not available for censuses before 1940, so we calculate average wealth by occupation using the 1860-70 Censuses and assign the average occupational wealth to the 1850-1910 censuses. We also construct occupational prestige scores using the 1950 Census IPUMS to calculate average wage and salary income by occupation.

These regressions produce a simple summary statistic that measures mobility across two generations. To measure mobility across three generations, we regress the child’s outcome on the parent’s outcome while controlling for the grandparent’s outcome \( Y_{it-2} \):

\[
Y_{it} = \alpha + \beta_1 Y_{it-1} + \beta_2 Y_{it-2} + \epsilon_{it}.
\]

We control for age and age squared of the child, parent, and grandparent, and we include year fixed effects. We report the coefficients on both the parent and the grandparent outcomes.

**Preliminary Findings**

Tables 1 and 2 show some preliminary results on how status persists beyond two generations. In both cases, the status of grandparents is related to the status of children independent of the direct transmission from grandparent to parent and then from parent to child. In Table 1, from 1850-1910 the occupational wealth of grandparents is highly correlated with child outcomes even after controlling for parents. In Table 2, we show that an impact of grandparents on grandchildren holds today in education, even after controlling for parent education.
Table 1: Occupational Wealth Elasticities Across Three Generations
Grandfathers in 1850, Fathers in 1880, Sons in 1910

<table>
<thead>
<tr>
<th></th>
<th>Column (1)</th>
<th>Column (2)</th>
<th>Column (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(Father)</td>
<td>0.191***</td>
<td>0.162***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.027)</td>
<td></td>
</tr>
<tr>
<td>ln(Grandfather)</td>
<td>0.152***</td>
<td>0.101***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.028)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>6.023</td>
<td>6.315</td>
<td>5.464</td>
</tr>
<tr>
<td>Observations</td>
<td>2,141</td>
<td>2,141</td>
<td>2,141</td>
</tr>
</tbody>
</table>

Note: Standard errors in parentheses. p-values *** <0.01 ** <0.05 * <0.10

Table 2: Education Elasticities Across Three Generations

<table>
<thead>
<tr>
<th></th>
<th>Column (1)</th>
<th>Column (2)</th>
<th>Column (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(Parent)</td>
<td>0.402***</td>
<td>0.382***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.009)</td>
<td></td>
</tr>
<tr>
<td>ln(Grandparent)</td>
<td>0.291***</td>
<td>0.035***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.006)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>8,952</td>
<td>8,952</td>
<td>8,952</td>
</tr>
</tbody>
</table>

Note: Standard errors in parentheses. p-values *** <0.01 ** <0.05 * <0.10

ONGOING RESEARCH AND EXPECTED OUTCOMES

First, we are expanding the contemporary comparisons (1940-2013) to include more measures of status, include income, earnings, and occupational prestige. This will allow us to better understand how the multigenerational transmission of status differs along different dimensions of achievement, and also allow us to make direct comparisons across the full period spanned by our data, 1850-2013.

Second, we are linking our two three-generation panels together to create a set of four, five, and six generation panels. With these longer panels, we can explore how status is passed through families over even longer periods of time.
REFERENCES


