Intergenerational Mobility of Immigrants in the US over Two Centuries

By RAN ABRAMITZKY, LEAH BOUSTAN, ELISA JÁCOME, AND SANTIAGO PÉREZ*

Using millions of father-son pairs spanning more than 100 years of US history, we find that children of immigrants from nearly every sending country have higher rates of upward mobility than children of the US-born. Immigrants' advantage is similar historically and today despite dramatic shifts in sending countries and US immigration policy. Immigrants achieve this advantage in part by choosing to settle in locations that offer better prospects for their children.

Immigrants who move to the United States aspire to offer a better future for their children. Both today and in the past, many immigrants earn less than US-born workers upon first arrival and do not completely catch up within a single generation.[1] However, a defining feature of the "American Dream" is the view that even immigrants who come to the United States with few resources and little skills have a real chance at improving their children's prospects.

This paper studies the intergenerational mobility of the children of immigrants (the "second-generation"). We ask whether the sons of immigrants achieve earnings parity with the sons of the US-born, and how their relative intergenerational mobility changed between the late 19th century and today.[2] On the one hand, children of immigrants might be in a particularly good position to

* Abramitzky: Department of Economics, Stanford University, 579 Serra Mall, Stanford, CA 94305, and NBER (e-mail: ranabr@stanford.edu); Boustan: Princeton University, Department of Economics, Princeton, NJ 08544 and NBER (email: lboustan@princeton.edu); Jácome: Princeton University, Department of Economics, Princeton, NJ 08544 (email: ejacome@princeton.edu); Pérez: Department of Economics, University of California, Davis, 1 Shields Ave, Davis, CA 95616, and NBER (e-mail: seperez@ucdavis.edu). We thank the editor, Esther Duflo, for her most valuable guidance, and three anonymous referees for their very useful feedback. We also thank the feedback of Chris Becker, Alvaro Calderón, Bruno Caprettini, Raj Chetty, Jennifer Hunt, Michele Rosenberg, Martin Saavedra, Kjell Salvanes, seminar participants at Copenhagen Business School, Copenhagen University, Munich, Princeton, Norwegian School of Economics in Bergen, Rotterdam, Stanford, Tilburg, and UC Davis, as well as conference participants at the 12th International conference on Migration and Development in Madrid, the "Intergenerational Mobility, Gender, and Family Formation in the Long Run" conference in Oslo, the Social Science History Association annual meeting, and SITE conference on immigration at Stanford. A longer version of the paper including additional results is available as NBER working paper #26408.

[1] See Abramitzky, Boustan, and Eriksson (2014) on earnings convergence for immigrants in the past and Lubotsky (2007) for more recent immigrants.

[2] In historical data, we are only able to link sons across Census waves, given that daughters often change their names at marriage. Thus, for consistency, we mostly focus on father-son pairs throughout our analysis.

move up the ladder, given that their parents may earn less than their true talent or ability would suggest (for example if they had little exposure to US education or faced discrimination in the labor market). Immigrant families who recently arrived in the US might also be more footloose, thus allowing them to settle in areas with better prospects for their children. On the other hand, children of immigrants might grow up in segregated neighborhoods, suffer from discrimination themselves, and otherwise enjoy more limited opportunities than the children of the US-born.

We find that, both historically and today, children of immigrants at the bottom of the income distribution have higher rates of upward mobility than children of similarly ranked US-born fathers, and to a strikingly similar degree in each time period. Second-generation immigrants growing up at the 25th percentile of the income distribution end up 5–6 percentile rank points higher than the children of the US-born who were also raised at the same income rank. Second-generation immigrants today exhibit a similar degree of upward mobility, despite notable shifts in countries of origin away from Europe toward Latin America and Asia, as well as major changes in US immigration policy from a regime of nearly open borders (to European immigrants) to one of substantial restrictions.

Our analysis spans 130 years of US history and focuses on three cohorts of immigrants who entered the US during the 19th or 20th centuries. The first two cohorts consist of four million first-generation immigrants observed with their children in the US in either the 1880 or 1910 Censuses.[3] The 1880 cohort contains mostly immigrants from Northern and Western Europe (e.g., Ireland, Germany, and the UK), whereas the 1910 cohort includes more immigrants from Southern and Eastern Europe thought to have faced greater initial disadvantages in the US labor market.[4] We follow the children of these immigrants to the 1910 and 1940 Censuses, respectively—using information on their name, year of birth, and birthplace—and compare their adult outcomes to those of the children of US-born whites.[5] Because the Censuses did not collect income data before 1940, our analysis of these cohorts relies on computing a number of alternative proxies for individual-level income ("income scores") based on a person's detailed occupation, age, and state of residence.

[3] Immigrants in these cohorts entered the US before the imposition of strict quotas based on country of origin imposed in 1921, which ended the era of nearly open borders for European migrants and dramatically reduced the inflow of immigrants from Eastern and Southern Europe.
[4] For more background on immigration during the Age of Mass Migration, see Abramitzky and Boustan (2017).
[5] More than 95 percent of immigrants in this period were white. Focusing on US-born whites as the comparison group ensures that the higher mobility of second-generation immigrants that we observe is not due to black-white differences in mobility. Collins and Wanamaker (2017) find that African Americans had much lower rates of upward mobility than whites historically. Thus, including African Americans in the sample would make immigrants' mobility advantage appear even larger.

The third cohort includes children of immigrants born around 1980. Unlike the historical cohorts, immigrants in this modern cohort entered the US during an era of substantial immigration policy restrictions and came mainly from poorer and more ethnically diverse countries in Latin America and Asia. To study this cohort, we use aggregate administrative data made public by the Opportunity Insights project, which is based on links of parents and their nearly six million children (Chetty et al. 2018a, 2018b). In these data, we observe children's outcomes in 2014–15 and the outcomes of their parents in 1994–2000. In addition, we use data from the General Social Surveys (GSS), in which we observe the outcomes of children of immigrants and the US-born in 2000–2018. The GSS have smaller sample sizes but include some children of undocumented immigrants as well as information on occupations to facilitate comparisons with the historical data.

We start by documenting income score gaps between immigrants and US-born workers. In both the past and present, there is a wide variation in labor market outcomes among first-generation immigrants from different countries. Immigrants from countries like Finland and Norway in the past and Vietnam and the Dominican Republic today earn below the US-born on average, whereas immigrants from countries like England (historical) and India (today) out-earn their US-born counterparts. With a few exceptions, earnings differences for immigrant groups range between +20 and -40 log points. However, for most sending countries in which first-generation immigrants earned less than the US-born, we find that second-generation immigrants catch up or even overtake the earnings of the children of the US-born.

Our main analysis then uses father-son links to compare the average rank that children can expect to reach in the national income distribution, conditional on their parent's income or income-score rank and immigration status (a similar approach to Chetty et al. 2014 and Chetty et al. 2018a). In both the past and present, we find that the children of immigrants are more upwardly mobile than the children of the US-born: conditional on the rank of their parents, children of immigrants have a higher expected rank in adulthood. The higher level of upward mobility among children of immigrants is especially meaningful in relatively poor families: The estimated gaps imply that children of immigrants with parents in the 25th percentile have a similar expected rank as children of US-born individuals whose parents were ranked about 15–20 percentiles higher, which is about one half of the contemporaneous gap between blacks and whites in the US (Chetty et al. 2018a). We find a smaller mobility gap between children of immigrants and children of the US-born at the top of the income distribution.

Our data enable us to estimate rates of intergenerational mobility separately by an individual's country of origin. When doing so, we find that, both in the past and today, immigrants at the bottom of the income distribution from nearly every sending country have higher rates of upward mobility than the children of the US-born. Moreover, despite all the changes that have occurred over the last century—more restricted borders, a shift to poorer sending countries relative to the US, the growing importance of services relative to manufacturing—children of immigrants today achieve strikingly similar rates of upward mobility to those of the past. For example, children of poor Mexican and Dominican immigrants today have a similar advantage relative to the children of the US-born as did the children of poor Danish or Swiss immigrants in 1910. This finding suggests that the descendants of immigrants of all backgrounds can eventually integrate into the US economy.

In the last part of the paper, we explore the question of why children of immigrants are more upwardly mobile, focusing primarily on the historical data but comparing with the modern data whenever possible. First, geography matters. Both today and in the past, we find that immigrant parents were more likely than US-born parents to move to areas offering better prospects for their children. In the historical cohorts, the intergenerational gap between immigrants and the US-born is reduced by about 50 percent when comparing children growing up in the same US region (Census division) and disappears entirely when comparing children growing up in the same county. In the modern cohort, the immigrant mobility advantage is 25 percent smaller in the average county than at the national level. Finally, we discuss the role of immigrants' self-selection, and provide suggestive evidence that immigrant fathers were "under-placed" in the income distribution (that is, their earnings did not fully reflect their abilities), thereby giving their children more room to improve.

I. Literature

Our paper contributes to the understanding of immigrant assimilation and intergenerational mobility in the US. Closest to our paper is Card, DiNardo, and Estes (2000), which uses cross-sectional data to compare average outcomes by country of origin for first- and second-generation immigrants in the labor market from 1940 until the mid-1990s. Like this paper, we also find that the children of immigrants fare better on average than the children of the US-born conditional on the average outcomes of the first generation. Our paper broadens the analysis to study immigrants

and their children over more than a century (1880 to 2015). More importantly, our paper adds actual parent-child linkages, enabling us to assess the view that even poor immigrants can substantially improve their children's prospects, as well as explore the channels through which immigrant families achieve their mobility advantage.

Most existing work on economic outcomes during the Age of Mass Migration focuses on assimilation *within* a generation (see for instance, Ferrie 1997; Hatton 1997; Minns 2000; and Abramitzky, Boustan, and Eriksson 2014). One exception is Ward (2019), who documents persistent outcomes across generations by ancestry (national origin of grandfather) in the early 20th century. We go beyond Ward by contrasting mobility rates for the children of immigrants with those of the US-born, and by comparing mobility rates in both the past and present.[6]

Finally, our paper adds to the broader literature on intergenerational mobility in the United States. A number of papers estimate contemporary levels of intergenerational mobility (e.g., Lee and Solon 2009; Chetty et al. 2014, 2017; Mazumder 2015; Hilger 2016; and Davis and Mazumder 2019).[7] A series of related studies document historical rates of intergenerational mobility, and compare these historical rates to present-day levels (e.g., Long and Ferrie 2013; Feigenbaum 2015, 2018; Olivetti and Paserman 2015; Ferrie et al. 2016; Card, Domnisoru, and Taylor 2018; Olivetti, Paserman, and Salisbury 2018; Tan 2018; Derenoncourt 2019; and Pérez 2019). Borjas (1992) develops a theory of ethnic capital, which suggests that the social mobility process may differ between immigrants and the US-born.[8]

II. Data

A. Historical Datasets for Studying Economic Mobility

We measure historical mobility rates using two new datasets of linked Census records. The first dataset links sons observed in the 1880 Census to the 1910 Census, and the second links sons observed in the 1910 Census to the 1940 Census.[9] These data allow us to observe an individual's

[6] Borjas (1993) uses cohort-level data from the 1940 and 1970 Censuses to study the intergenerational mobility of immigrants in the US. Perlmann and Waldinger (1997) and Perlmann (2005) compare the outcomes of second-generation immigrants in the past and the present but do not have parent-child linkages. Our work also relates to studies that focus on the intergenerational progress of specific immigrant groups, including the Irish famine immigrants (Collins and Zimran 2019), and Mexican Americans (Duncan et al. 2017; Kosack and Ward 2018).

[7] Bhattacharya and Mazumder (2011), Chetty et al. (2018a) and Collins and Wanamaker (2017) focus on black-white differences in mobility, and Hilger (2017) focuses on Asian Americans.

[8] Outside of the US, Aydemir, Chen, and Corak (2009) focus on the intergenerational mobility of immigrants to Canada, Bolotnyy and Bratu (2018) on immigrants to Sweden, and Gielen and Webbink (2019) on the Netherlands.

[9] For reasons of both data availability and historical interest, our analysis does not include immigrants in the 1940–1970 cohorts. First, the 1940 Census is the last Census for which information on names is publicly available, allowing for backward but not forward links from 1940. Second, neither of the sources that we use for the modern period include immigrants in these cohorts: the GSS started collecting data in 1972 and the

own labor market outcomes during adulthood and his father's labor market outcomes during his childhood. Women cannot be systematically matched across historical Censuses because they typically change their last names after marriage, so the past-present comparison focuses on men. Because the modern data define immigrant status based on father's birthplace, we use the same definition in the historical data for comparability.[10]

Linking Fathers and Sons. — To create each linked sample, we matched all males aged 0–16 in a childhood Census (that is, either in 1880 or in 1910) to a later Census (1910 or 1940, respectively) using information on first and last names, age, and state of birth. We use the linking algorithm developed in Abramitzky, Boustan, and Eriksson (2012, 2014), which is explained and evaluated in Abramitzky et al. (2019). We restrict the analysis to father-son pairs in which: (1) both the father and the son were white, (2) the son was living with his father at the time of the earliest Census (so that we can observe a father's labor market outcomes),[11] (3) the father was born in the US or in one of the 17 largest sending countries during this time period,[12] and (4) both the father and the son were aged 30–50 by the time we measure their labor market outcomes.[13] Finally, we exclude father-son pairs with missing information on occupation.[14] Appendix A provides further details on the matching procedure as well as sensitivity checks for possible "false positives" and lack of representativeness.[15]

Online Appendix Table A1 shows the sample size in each of the historical cohorts as we impose these restrictions. In our baseline samples, we match 23 percent of individuals in the 1880 Census to 1910 and 29 percent of individuals in the 1910 Census to 1940, standard match rates for historical samples given the presence of common names and various causes of non-matches

Opportunity Insights data covers children of immigrants born after 1978. In addition, these cohorts coincide with a period of limited migration into the US: in 1970, only 4.7 percent of the US population was born abroad (compared to nearly 15 percent in the past [1880–1920] and today).

[10] Appendix A and the Online Appendix in the NBER working version of this paper show that the results in both the past and today are similar when defining immigrant status using the mother's place of birth (or using both parents' birthplaces in the historical data when such information is available).

[11] Most children in this age group were living with their father regardless of immigration status: 88 and 90 percent among children with a US-born father in the 1880 and 1910 Census, respectively, and 89 and 90 percent of children with a foreign-born father.

[12] The countries we consider are Austria, Belgium, Canada, Denmark, England, Finland, France, Germany, Ireland, Italy, Norway, Portugal, Russia, Scotland, Sweden, Switzerland, and Wales. These countries are the largest source of immigrants in the 1850–1910 period, accounting for 93 and 88 percent of the foreign-born population in the 1880 and 1910 Censuses, respectively. We restrict the analysis to large sending countries so that we have enough observations to estimate our country-by-country results. Although we refer to the children of immigrants as "the second generation," we do not distinguish between sons born inside and outside of the US, because the Opportunity Insight Data does not include such information (that is, we include both the "1.5 generation" and the "second generation" in our samples of children). The vast majority of sons in our data are born in the US (98 percent in the historical cohorts and 95 percent in the GSS modern cohort).

[13] We restrict our sample to men ages 30–50 in order to measure fathers and sons at the same moment in the earnings lifecycle, and to avoid measuring income when children or parents are too young or old (Mazumder 2005).

[14] Our findings are similar when we assign an income of zero to those with missing occupation (Online Appendix Figure B2).

[15] In the Online Appendix, we also show that our results are similar when we instead implement the linking approach developed in Abramitzky, Mill and Pérez (2020). Finally, in the Online Appendix of the NBER version of the paper we show that the results are similar when we construct the historical samples to match the characteristics of the Opportunity Insights sample.

(mortality, return migration, under-enumeration, transcription error; see Abramitzky et al. 2019). After imposing the sample restrictions, our baseline samples include 1.3 million men in the 1880–1910 (30 percent of whom are second-generation immigrants) and 2.7 million men in 1910–1940 (24 percent of whom are second-generation immigrants). The largest origin countries are Germany, Ireland, and England in the first cohort, and Germany, Russia, and Canada in the second.

Assigning Income Measures in Historical Data. — Because the 1940 Census is the first US Census to include information on individual earnings, we need to construct proxies of individual income for fathers and sons in our pre-1940 historical samples (we refer to these proxies as "income scores").[16] Our preferred approach is to use a statistical model to predict income from a rich set of covariates for white men aged 30–50 in the 1940 Census, and then use this model to predict income for men in earlier years (we discuss farmers below, whom we handle separately because there is limited information on farm income in the 1940 Census). In particular, we regress log income in 1940 on a number of fixed effects and complete set of interaction terms using 3-digit occupation, age, and current state of residence as our explanatory variables.[17]

The 1940 income variable excludes income from self-employment. Because the vast majority of farmers are self-employed, we compute income for farmers using a method developed by Collins and Wanamaker (2017). Specifically, we make use of the facts that the 1940 Census records the incomes of farm laborers, and that later Census years record how much farmers earn relative to farm laborers. We thus compute farmer incomes by multiplying the income of farm laborers in 1940 with the ratio of earnings for farmers versus farm laborers in the 1960 Census, by region and immigration status.[18]

One limitation of using the 1940 Census to compute income scores for individuals in earlier Censuses is that we assume that the relative rank between income cells is stable from 1880 to 1940.

[16] In our baseline analysis, we also use predicted (rather than actual) income for the sons we observe in the 1940 Census (so as to use a consistent measure across all of our historical samples). However, the results are similar when we instead use individual income data from the 1940 Census (see Online Appendix Figure B2).

[17] In Appendix B, we show that results are similar when we use a person's country of birth as an additional variable in the income prediction. In all interaction terms, we interact covariates with 1-digit Census occupations and with Census region, instead of state. This method is similar to the machine-learning approach for computing income scores proposed by Saavedra and Twinam (2018). Indeed, the correlation between our predicted earnings and those estimated using their approach is 0.99 for fathers in the 1910–1940 cohort.

[18] When performing this calculation, we focus on US-born whites and immigrants from the sending countries we include in the paper. We note that Collins and Wanamaker use the 1960 Census (rather than 1950) because the 1950 Census only asked the income question to a small subset of the population and is therefore not useful for this exercise. In addition to adjusting for farmer income, we also follow Collins and Wanamaker (2017) and scale up farm managers' and farm laborers' incomes to account for in-kind transfers. Finally, when using actual rather than predicted income for the sons we observe in the 1940 Census (second row in panel (b) of Online Appendix Figure B2), we also adjust the income of self-employed non-farmer sons following Collins and Wanamaker (2017). To do so, we calculate the ratio of mean earnings for self-employed non-wage workers in an occupation to the mean earnings of wage workers in the same occupation from the 1960 Census and then use this ratio to scale wage earnings of the self-employed in the 1940 Census.

Appendix B shows that results are robust to using more contemporaneous sources to compute income scores, namely the 1901 Cost of Living Survey and the 1900 Census of Agriculture. The appendix also documents that immigrants have a similar mobility advantage when using a sample of father-son pairs for whom we can observe fathers in the 1940 Census and thus can measure *actual*, rather than only predicted, income for both father and son.[19] Overall, our conclusions are unaffected by the specific income proxies we use.[20] Appendix B provides further details on the construction of the income scores and sensitivity checks.

B. Modern Datasets for Studying Economic Mobility

Our linked Census data cannot be extended beyond 1940 because confidentiality restrictions dictate that individuals' names are only released to the public 72 years after a Census is taken. Instead, we rely on two other sources of data to follow father-son pairs in the modern labor market.

First, we use publicly available aggregate data from the Opportunity Insights project (Chetty et al. 2018a). These data are based on nearly six million children born between 1978 and 1983, out of whom 310,000 have an immigrant father born in one of the 21 largest sending countries to the US.[21] These data are built by the Opportunity Insights team in two steps: In the first step, the 2000 US Census and the 2005–2015 American Community Survey (ACS) are linked to federal income tax returns (using a person identifier assigned by the Census Bureau). This procedure enables the researchers to observe a person's income from the tax records as well as his or her race and place of birth from the Census or the ACS. In the second step, the Opportunity Insights team obtained information on the family income of an individual's parents during his/her childhood. A parent is defined as the first person who claims an individual as a dependent in the federal income tax records between the years 1994 and 2015. Note that this procedure for linking parents to children will exclude pairs in which either the child or the parents lack a Social Security Number. As a consequence, this sample only includes children who are either US citizens or authorized immigrants, and whose parents are also US citizens or authorized immigrants.

[19] We do so in Appendix B by considering the *fathers* who are young enough in 1910 that they can be linked to the 1940 Census.
[20] The one exception to this claim is the set of results based on the IPUMS occupational income score, which maps occupations to median income in the 1950 Census. Reliance on the 1950 income distribution is problematic because the incomes of farmers deteriorated between 1880 and 1950 due to secular changes in the agricultural sector, leading farm income derived from 1950 data to be too low for fathers in 1880 or 1910. Online Appendix Figure B2 shows that once farmers' incomes in the IPUMS occupational score are adjusted so that farmers have the same average rank as in our baseline score, we again see a large immigrant mobility advantage.
[21] We use the 21 countries for which the Opportunity Insights project reports children outcomes by ventile of the parental income distribution. The countries are Canada, China, Colombia, Cuba, the Dominican Republic, Ecuador, El Salvador, Germany, Greece, Guatemala, Haiti, India, Israel, Italy, Jamaica, Japan, Mexico, the Philippines, South Korea, the UK, and Vietnam.

Income for this cohort is measured as the average annual income from 2014 to 2015, when these children were 31 to 37 years old, whereas parental income is measured as the average household income from 1994 to 2000 (father's income is not reported separately in this dataset). We use the publicly available version of the data, which reports information collapsed by ventile bin in the national income distribution for immigrants from 21 countries of origin. Hence, for each parental ranking, we observe the average income rank of children whose parents fall into that ventile, by a father's country of origin.[22] When focusing on country-by-country outcomes at the 25th and 75th percentiles (rather than on outcomes throughout the full parental income distribution), we are able to expand our analysis to 51 countries of origin.

The Opportunity Insights dataset has several attractive features for the study of the intergenerational mobility of immigrants: It is very large relative to other panel datasets (no other longitudinal dataset would enable us to look at mobility differences by country of origin in the modern period), and includes administrative income data both for parents and their children.[23] However, the dataset has two limitations for our purposes. First, it does not include information on individual-level occupations—the main labor market outcome we observe in the historical data—thus complicating our past-present comparisons of intergenerational mobility. Second, because unauthorized immigrants do not have a Social Security Number (SSN), these data exclude the children of unauthorized immigrants. This limitation affects Hispanic immigrants more than other groups. Specifically, Chetty et al. (2018a) compares the Opportunity Insights data to the American Community Survey and reports that they include about 79 percent of Hispanics, whereas the coverage is close to 100 percent for other ethnic groups. We add data from the General Social Survey (described below) as well as cross-sectional data from the Current Population Survey, both of which include unauthorized immigrants, to partially address this concern. We also note that the Opportunity Insights data *do* include first-generation migrants (and their children) who were at some point undocumented, but who later on became authorized and hence received a Social Security Number. For example, the data include first-generation immigrants who were in the country by 1982 and benefited from the 1986 Immigration Reform and Control Act amnesty of

[22] These data correspond to the "Non-Parametric Estimates of Income Ranks for Second Generation Immigrant Children by Parent Income, Country of Origin, and Gender" and "Parametric Estimates of Income Ranks for Second Generation Immigrant Children by Parent Income, Country of Origin, and Gender" files available on the Opportunity Insights website at https://opportunityinsights.org/data/.

[23] For instance, the PSID (which includes information on parental income) started in 1968 and thus does not have much coverage of immigrant families who arrived in the US after the border reopened to immigration in 1965. Other panel datasets (for instance, the NLSY used in Borjas 1993) include retrospective questions (asked of the children) about the occupation and education in the parent's generation, but do not contain information on parental income.

2.7 million previously unauthorized immigrants, a number that represents nearly all of the undocumented immigrants living in the US at the time (Passel, 1986).[24]

We supplement our analysis of the Opportunity Insights data with data from the General Social Survey (GSS). Although the GSS has substantially smaller sample sizes, it includes questions on occupation and covers unauthorized immigrants, both of which facilitate comparison with the historical data. The survey includes information on a respondent's occupation, parents' places of birth (in the US or abroad), and retrospective questions on parents' occupations when the respondent was 16 years old. We use this information to assign income proxies for fathers and their sons using an analogous procedure to the one we implement with the historical data.[25] To make this sample as comparable as possible to the historical data, we focus on respondents who: (1) are males, (2) lived in the US by age 16, (3) were 30 to 50 years old at the time of the survey, and (4) have available information on their and their fathers' occupations. Finally, to maximize comparability with the Opportunity Insights sample while retaining enough observations, we restrict the analysis to survey years 2000 to 2018. After imposing these sample restrictions, our GSS sample includes 3,100 respondents, nine percent of whom are second-generation immigrants.

III. Preliminary Evidence: Earning Gaps between Immigrants and US-born since 1880

We start by estimating earning gaps between US-born men and first- and second-generation immigrants by country of origin. For this exercise, we use cross-sectional samples, rather than father-son pairs.[26] Our interest is in documenting the average earnings disadvantage faced by first-generation immigrants by country of origin, and then to ask whether the average child of these

[24] Conceptually, we might expect the children of undocumented immigrants to be even more upwardly mobile than the children of authorized migrants. As we document below, lack of US-specific human capital among first-generation immigrants partly explains the high mobility among their children (i.e., first-generation immigrants are "under-placed" in the US income distribution). Given the restrictions that unauthorized immigrants face in the labor market, we might expect them to be even more "under-placed" in the US labor market relative to their true abilities, making their children (who will often be born in the US and have citizenship) especially likely to be upwardly mobile.

[25] We use the 2006–2015 CPS to create income proxies for fathers in the GSS sample. To do so, we regress logged total earned income on a number of fixed effects and complete set of interaction terms using occupation, age, region, and immigrant status (and their interaction terms) as our explanatory variables. We then use this regression to predict income scores for GSS respondents and their fathers. We note that the GSS reports parental immigrant status but does not record country of origin and that the most precise measure of geography in the GSS is region. In addition, the GSS does not include information on parental age.

[26] For our historical cohorts, we use the 1880, 1910, and 1940 Censuses of Population. For the modern cohort, we use the 1980 Census for the fathers' generation and the CPS's Annual Social and Economic Supplement for 2006–2015 for the sons (Flood et al. 2018, Ruggles et al. 2018). We restrict the 1980 Census sample to men aged 30–50, who have a son present in the household, and who were born in the US or in one of the top sending countries identified in the Opportunity Insights sample. We restrict the CPS sample to US-born men aged 30–50 and whose fathers were either born in the US or in one of the top sending countries. For both of these datasets, we construct occupation-based income scores using an analogous procedure to the one that we use in both the GSS and the historical data. In the appendix to the NBER version of the paper, we show that the results are similar when using actual income whenever possible (Figure A1) and when excluding the US South from the historical samples (Figure A4).

immigrants was able to erase some or all of these gaps by the next generation. Specifically, we estimate the following equation, first for the father's generation and then for the son's:

$$(1)\ Y_{iac} = \alpha + \sum_{c=1}^{N} \beta_c \,\text{Country}_c + \ \delta_a + \epsilon_{iac}$$

where Y_{iac} is the logged income score of father or son i of age a from sending country c. The β_c coefficients reveal the average difference in income score between US-born individuals and immigrants from country c. All regressions include a quadratic in father's or son's age δ_a.

Figure 1 plots the β_c coefficients for first-generation (in black) and second-generation (in gray) immigrants in each cohort. For most countries of origin and time periods, the gaps in income scores between first-generation immigrants and US-born men range between -40 log points and +20 log points. In the modern cohorts, four countries faced larger gaps of about -60 log points (the Dominican Republic, Haiti, Mexico, and Vietnam). In the past, immigrants from Scandinavian countries faced the largest earnings penalties. In all cohorts, immigrants from a set of sending countries outperform the US-born in the labor market. Today, these high-earning countries include Canada, Germany, India, Japan, and the United Kingdom, whereas in the past the highest-earning immigrants were from countries such as England and Scotland.

In each of the three cohorts, second-generation immigrants from countries in which first-generation immigrants earned less than the US-born close or even reverse the corresponding earnings gap in the second generation, leading children of immigrants to out-earn children of US-born individuals. There are only a few cases in which the gap remained above 5 log points by the second generation (Norwegians in 1910 and 1940, Finns in 1940, and Jamaicans, Haitians, and Mexicans in the modern data), but for most sending countries, the earnings gap is substantially smaller by the second generation.

For countries whose first-generation immigrants already out-earned US-born individuals in the labor market, the pattern is more mixed, with the corresponding earnings advantage becoming smaller for some countries, remaining similar in size for others, or even growing by the second generation in some cases. Although earnings gaps for countries that start below the US-born diminish substantially across generations, it is still the case that the children of low-earning immigrants tend to earn less than the children of high-earning immigrants, consistent with the persistence results in Abramitzky, Boustan, and Eriksson (2014) and Ward (Forthcoming).

IV. Main Results: Intergenerational Mobility of Immigrants since 1880

We have so far considered average earnings gaps, without any regard for the parents' rank in the income distribution. The higher earnings growth for the children of immigrants relative to the children of US-born parents in Figure 1 could be driven by higher rates of upward mobility from the bottom of the income distribution, lower rates of downward mobility from the top of the income distribution, or both. The essence of the American Dream is that even immigrants at the bottom of the income distribution can have children that succeed. Thus, we turn to the linked data, both historical and modern, to study the intergenerational mobility of children of immigrants and US-born individuals whose parents had comparable labor market outcomes, as measured by their rank in the national income distribution (following Chetty et al. 2014 and Chetty et al. 2018a).

To do so, we first rank each son based on his income score (or actual income when using the Opportunity Insights data), relative to other sons born in the same year. Similarly, we rank fathers relative to all other fathers with sons born in the same year.[27] Both when ranking sons and when ranking fathers, the ranks are based on a person's position in the *national* income distribution.[28] We then estimate a model in which we regress a son's rank on his father's rank, allowing both the slope and intercept to differ for sons of immigrants and US-born individuals:

$$(2) \quad Rank\ Son = \alpha + \beta_0 2nd\ Gen. + \beta_1 Rank\ Father + \beta_2 2nd\ Gen. \times Rank\ Father\ + \epsilon.$$

The constant term α measures *absolute* rank mobility for the children of the US-born, that is, the expected rank of children of US-born individuals with fathers at the very bottom of the income distribution. The parameter β_0 captures the degree to which this expected rank is different for the children of immigrants. The parameter β_1 measures the rate of *relative* mobility of children of US-born individuals, or the association between the ranks of children and those of their fathers. Finally, β_2 measures the degree to which this association is different for immigrants and their children.

[27] In the Opportunity Insights data, we use family income rank because we do not have separate information on a father's income rank.
[28] If there are individuals in the sample who have the same father, we assign the father the average rank across his sons. Due to the GSS's much more limited sample size compared to the historical linked data and the Opportunity Insights data, we pool all survey years (2000 to 2018) and include survey-year fixed effects in the regressions.

We compare the economic mobility of the sons of immigrants to the sons of the US-born in Figure 2. To do so, we plot the regression lines corresponding to equation (2), separately for each of our cohorts, as well as a binned scatterplot showing the mean income rank of the children of immigrants and of the US-born by parental income ventile rank. Panels (a) and (b) suggest that in the past, the sons of immigrants on average out-earned comparable sons of US-born individuals throughout the income distribution, and particularly for sons of fathers with below-median rank. Second-generation immigrant sons with fathers at the 25th percentile in 1880 and 1910 reached an average income rank that was 6 percentiles above the average rank reached by sons of the US-born with similar fathers' incomes.[29] Panels (c) and (d) show the mobility patterns for the modern cohort when using the GSS and Opportunity Insights data, respectively. Today, the sons of immigrants remain more upwardly mobile than their counterparts with US-born parents, and this pattern also mainly holds in the bottom half of the income distribution. In particular, and similar to the historical results, sons of immigrants in the 25th percentile rank about 5 percentile points higher than sons of US-born individuals, regardless of whether we use the GSS or the Opportunity Insights data.

Table 1 reports the estimated values of the coefficients corresponding to equation (2) for each of the samples. For all three cohorts, we find a higher intercept (a positive β_0) for the sons of immigrants relative to the children of the US-born, indicating higher levels of absolute mobility. The difference in intercepts ranges from 7 to 8 percentiles in all cohorts. The slopes for immigrant families are in all cases smaller than the slopes for the US-born (a negative β_2), suggesting a weaker association between a father's and his son's rank among immigrant families.

In Figure 3, we allow the intercept and the slope of the rank-rank relationship to differ by country of origin, both in the historical and the modern data. We then use these estimates to compute the expected income rank for sons born at the 25th percentile for each country of origin.[30] The figure shows that sons from nearly all sending countries in the past and the present (except Norway and Belgium in 1880, Norway in 1910, and Haiti, Trinidad and Tobago, and Jamaica in the modern cohort) had a higher expected rank than the children of US-born individuals with

[29] We also find that the sons of unskilled immigrant fathers moved into white-collar and skilled blue-collar jobs at higher rates than the sons of unskilled US-born fathers, who were more likely to transition into farming (see the Online Appendix of the NBER version of the paper).
[30] In this exercise, we focus on the Opportunity Insights for the modern data because the GSS data does not record the father's country of origin (only whether the father is foreign-born).

parents at the 25th percentile.[31] Immigrants from Asian countries, notably China and India, appear especially upwardly mobile in the current period.[32]

Figure 4 shows that the sons of immigrants and the US-born appear more similar when looking at those who grew up with parents in the 75th percentile of the income distribution, particularly in the modern period. In the past, sons of immigrants had slightly higher expected ranks, and today the two groups achieve similar expected ranks. At the top of the income distribution, sons of immigrants from some sending countries like Portugal and Belgium in the past and Jamaica, El Salvador, and Mexico today did worse than the sons of the US-born, whereas those from other countries like Russia and Ireland in the past and India and China today did better.

The modern data also enable us to look at the intergenerational mobility of the *daughters* of immigrants. Panel (a) of Figure 5 suggests that daughters of immigrants on average out-earned daughters of US-born individuals throughout the income distribution. Panel (b) shows the expected income rank for daughters born at the 25th percentile for each country of origin. As for sons, we find that daughters of immigrants from nearly every sending country achieve higher levels of upward mobility than the daughters of the US-born.

Overall, we conclude that both in the past and today, second-generation immigrants are more upwardly mobile than the children of the US-born. This advantage is not driven by children of immigrants from certain ethnicities or country-of-origin groups: there is higher upward mobility among children of immigrants from nearly every sending country, and particularly at the bottom of the income distribution.

V. Why are children of immigrants more upwardly mobile?

A. Location Choices in the First Generation

[31] One noticeable pattern from Figure 3 is that the only three countries where sons of immigrants at the 25th percentile have a lower expected rank than the sons of the US-born are countries where a significant share of the population is black. This pattern is consistent with a broader pattern of black-white mobility gaps in the US (see for example Chetty et al. 2018a). In Figure 5 of the NBER version of this paper, we also show mobility patterns by immigrant race and ethnicity. In contrast, when focusing on daughters, we no longer observe that the children of immigrants from Caribbean countries with a significant black population have a lower expected rank than the children of the US-born (see Figure 5 in this paper).

[32] While this paper cannot explain why immigrants from some countries are more mobile than others, it is natural to ask if factors such as linguistic, physical, and religious distance from the US play a role in immigrants' mobility. For each of the countries in our data, we calculated: (1) linguistic distance to the US using the measure from Chiswick and Miller (2005), (2) physical distance to the US, and (3) an indicator that takes a value of one for predominantly Protestant countries. In the modern data, there is a clear positive correlation between all distance measures and mobility. This correlation is largely driven by Asian countries, which are both very distanced in all measures from the US and whose immigrants are very upwardly mobile. In contrast, in the past there is little correlation between the various distance measures and mobility. This is mainly because in the past all sending countries (besides Canada) were in Europe, so there is very little variation in their physical distance to the US. The main variation in linguistic distance in the past is driven by whether the country is English speaking, but immigrants from English-speaking countries did not seem to be more or less upwardly mobile than the average immigrant.

Figure 6 documents that both historically and today, immigrants tend to settle in areas that offer higher levels of upward mobility for their children. Specifically, the correlation between the share of migrants in an area and the share of upwardly mobile sons is 0.36 for the 1880 cohort, 0.51 for the 1910 cohort, and 0.23 for the most recent cohort (upward mobility is defined here as the share of sons of US-born fathers in the bottom fifth of the national income distribution who reach the top fifth of the distribution).[33] Immigrants in the modern cohort continue to settle in places that offer strong prospects for upward mobility, but there are also a number of high-mobility locations (e.g., Wyoming and Utah) that do not attract many immigrants today.

We next illustrate more explicitly that differences in the geographic distribution of immigrants and US-born individuals might explain the mobility gap between the two groups. To do so, we re-run our rank-based specification from equation (2), and sequentially add fixed effects for the US South, Census division, state, state-by-urban status, and county, all of which are based on an individual's childhood location (we only have data to run these specifications for the historical cohorts). Regressions without location fixed effects reveal the mobility gap between children of immigrants and children of the US-born at the national level. Regressions with location fixed effects measure the gap between children *who grew up in the same location*. If immigrant fathers chose to settle in areas that offered high prospects for mobility overall, we expect the gap to be smaller in the regressions with location fixed effects. Figure 7 shows that indeed including fixed effects for the US South or for Census division diminishes the intergenerational gaps at the 25th percentile by roughly 50 percent in both historical cohorts.[34] The intergenerational gaps are further reduced when we add childhood state fixed effects (fourth bar of each graph), and they fully close when we further include either childhood state-by-urban status (fifth bar) or childhood county fixed effects (sixth bar). We thus conclude that immigrant parents' location choice was an important driver of the immigrant mobility advantage we see in the past.

In the modern data, we are more limited in our ability to illustrate the role of geography in explaining the mobility advantage for the children of immigrants. Specifically, the GSS survey only contains information on Census division of residence at age 16 and the Opportunity Insights data allows for some aggregate comparisons by county of residence. Similar to the past, the

[33] Rates of intergenerational mobility both historically and today vary greatly by region (Chetty et al. 2017; Tan 2018; Connor 2019). Previous research has also found that first-generation immigrants in the modern period are more likely than the US-born to settle in regions that offer them higher wages (Borjas 2001; Cadena and Kovak 2016).

[34] The gaps also diminish substantially when we focus on children at the 50th and 75th percentile of the income distribution (see Appendix Figure A12 in the NBER WP version).

intergenerational gap is reduced when comparing children growing up in the same region in the GSS, although the magnitude of the reduction is smaller than in the past (Figure 7, panel (c)). The Opportunity Insights data report the average ranks of children who grew up at the various percentiles of the income distribution by parental nativity for 1,053 counties. For each of these counties, we calculated the gap in expected rank between children of immigrants and children of the US-born at the 25th percentile who grew up in the same county. We find that the average gap is 4.1 percentiles in these 1,053 counties, which is smaller than the national gap of 5.8 percentiles, suggesting that location choice still matters today but to a somewhat lesser degree than in the past.

Overall, our findings suggest an important role for migrants' location choices in explaining the higher mobility of their children. There are two main reasons why immigrants might have been in a better position than the US born to choose an area with more opportunities for their children. First, immigrants who have just arrived in the US likely have greater flexibility than the US born (who are already settled in an area) to locate in a specific region of the country. Second, by leaving their home country, immigrants have already revealed themselves to be willing to move, which might further help them avoid low-mobility regions once in the US.

B. Father's income not reflecting earnings potential

Another factor that could explain the higher rates of upward mobility for the children of immigrants is that immigrant fathers did not earn up to their potential in the US labor market, thereby giving their children more room to improve. In other words, if immigrant fathers are ranked lower in the distribution than their skills or ability would suggest (e.g., due to a lack of language ability, discrimination, or more limited labor market networks), then the children of these immigrants are expected to perform better than their fathers' outcomes would predict.

One way to test this mechanism is to compare the children of immigrant fathers who arrived in the US at different ages. Immigrant fathers who arrived as children likely assimilated into the labor market more easily (for instance, by being exposed to US education from an early age), so that their observed income is likely closer to their potential earnings.[35] Panel (a) of Figure 8 shows that the mobility gap at the 25th percentile is indeed four times larger for the children of immigrants who arrived as adults (17 years old or older) than for the children of immigrants who arrived before schooling age (7 years old or younger). Indeed, the latter group exhibits only a small mobility gap

[35] Ward and Alexander (2018) shows that age of arrival was a key driver of immigrant earnings during the Age of Mass Migration, with immigrants arriving at younger ages out earning those arriving later in life.

relative to the children of the US-born.[36] Panels (b) and (c) of Figure 8 show that the mobility gap grows with fathers' age of arrival, but especially for individuals with fathers from non-English-speaking countries. This finding is consistent with the idea that immigrant fathers from non-English speaking countries end up lower in the income distribution partly due to an imperfect command of English, thus leaving their children with more opportunity to move up the income distribution (see also Bleakley and Chin 2010).

The children of US-born fathers who migrated *internally* within the US are a useful comparison group that can shed light on the relative roles of location choices and father's under-placement in the income distribution. On the one hand, internal migrants were likely to sort into areas with better prospects for their children. On the other hand, these fathers were raised in the US and so would not likely suffer from under-placement in the income distribution in their new location. Indeed, Online Appendix Figure C2 shows that the upward mobility of children of US-born fathers who migrated internally falls *between* that of children of immigrants and children of US-born fathers who remained in their state of birth, suggesting an important role for both geographic choices and under-placement.

C. Immigrant self-selection

A study of the long-term implications of immigration for the destination country requires a focus on the individuals who chose to migrate and raise children in the host country. Individuals who choose to move to the US are self-selected, which could be one reason why the children of immigrants are more upwardly mobile (for example, if immigrants are more forward looking or care more about the upward mobility of their children, and are thus more likely to invest in them). Interestingly, we find that the immigrant mobility advantage is present even for children of immigrants from countries where most of the migration flow is comprised of refugees, for whom selection is arguably less important (for instance, the Vietnamese in the modern data).

Yet immigrant selection may have manifested itself in many ways. Our historical data allow us to study one aspect of self-selection as an explanation for immigrants' higher upward mobility: namely, that immigrant families invested more in their children's educational attainment, which

[36] We focus this analysis on the 1910–1940 cohort because the 1880 Census did not include information on year of arrival to the US. Parental age of arrival to the US is also not available in the GSS. We also find evidence that fathers are less under-placed as they spend more time in the US labor market. Specifically, Online Appendix Figure C1 shows that the immigrant-native gap at the 25th percentile is about 25% (40%) smaller for children whose immigrant fathers are aged 40–50 relative to children whose immigrant fathers are aged 30–40 in the 1880–1910 (1910–1940) cohort.

in turn served as a vehicle toward achieving better labor market outcomes. Figure 9 considers the 1910–1940 cohort—the only cohort for which we have data on completed years of schooling for the sons—and shows that, if anything, sons of US-born fathers were more likely to graduate from high school (panel (a)) and reach higher grades in school (panel (b)). Similarly, panel (c) shows that sons of US-born fathers who were aged 12–16 were more likely to be attending school in 1910 than the sons of immigrant fathers.[37] Hence, we conclude that historically, higher educational investments among immigrant families were unlikely to be the explanation for their higher upward mobility, although education may play a larger role today.[38]

VI. Conclusions

We use newly constructed linked datasets of fathers and sons to study whether the children of immigrants achieve earnings parity with the children of the US-born, and how the intergenerational mobility of immigrants has changed over the last two centuries. We find that, both in the past and today, children of immigrants, especially those growing up in poorer families, had greater chances of moving up in the income distribution relative to the children of US-born parents. This finding is not driven by immigrants from any particular country or ethnic origin. Rather, when estimating mobility rates by country of origin, we find that children of immigrants from nearly every sending country have higher rates of upward mobility than the children of the US-born.

That immigrants were similarly upwardly mobile both a hundred years ago and today is remarkable. For example, whereas immigrants in the past predominantly hailed from Europe, immigrants today have more diverse ethnic backgrounds, coming from various Latin American and Asian countries. To the extent that members of ethnic minorities might have different levels of intergenerational mobility (as shown in a number of studies including Chetty et al. 2018a), we might have expected this change in the ethnic mix to have an influence on the rates of intergenerational assimilation. The past-present similarity is also remarkable given that immigrants today come from countries with lower income levels relative to the US, and given the

[37] Similarly, using the 10 percent sample of the 1880 Census, we see that sons of immigrant fathers aged 12–16 are 8 percentage points less likely to be attending school than sons of US-born fathers (57 versus 65 percent).

[38] Despite these lower levels of educational attainment, panel (d) confirms that the children of immigrants enjoyed higher income mobility because they earned a higher income at any given level of education. The fact that educational differences do not explain the mobility gap is not entirely surprising because, as Goldin (1998) shows, the returns to schooling in 1940 were lower than in recent years. However, we also note that the children of immigrants enjoyed a faster rate of *educational mobility:* For a given level of father's education, the children of immigrants in the bottom half of the distribution tended to achieve more years of schooling than the children of the US-born, consistent with the finding by Card, Domnisoru, and Taylor 2018 (see panel (e)).

dramatic changes in US immigration policy over the last century (the ending of open borders for European immigrants and the imposition of additional regulations for immigrant entry).[39] Overall, our findings stand in contrast to the nostalgic view that it was easier for immigrants in the past to integrate into the US economy and society.

We find that an important explanation for why the children of immigrants are more upwardly mobile is that immigrant families are more likely than the US-born to move to areas that offer better prospects for their children. This finding is especially relevant in light of the decline in both geographical (Molloy, Smith, and Wozniak 2011) and economic (Chetty et al. 2017) mobility that has taken place in recent decades in the US. It suggests that to the extent that "pockets of opportunity" remain available in the US, immigrants might be able to enjoy high levels of mobility even if overall mobility declines.

The success of second-generation immigrants gives a more optimistic view of immigrant assimilation than previous studies that have focused only on the first generation. Indeed, we find that second-generation immigrants overtake, rather than just catch up with, the children of the US-born with comparable family incomes. Although some politicians have a short-term perspective on immigrant assimilation, our findings suggest that this view might underestimate the long-run success of immigrants. Our findings are more consistent with the idea of the "American Dream," by which even immigrants who come to the United States with few resources and little skills have a real chance at improving their children's prospects.

APPENDIX A: SENSITIVITY OF RESULTS TO LINKING PROCEDURE AND SAMPLE CONSTRUCTION

i. Linking Algorithm

A first concern with Census links based on the available information on names, age, and place of birth is the potential for linking children to the wrong adults ("false positives"). Online Appendix Figure A1 shows that the results are similar when using more conservative versions of the Abramitzky, Boustan, and Eriksson linking algorithm, which reduces sample size but also likely reduces matching errors. First, we restrict the sample to *exact* matches by full (non-standardized) name and age. Second, we restrict the sample to individuals whose first and last

[39] US GDP per capita was more than 7 times higher than in Mexico or China when migration flows from these countries took off in 1970, whereas it that was only 2–3 times higher than in European sending countries circa 1900. Lower initial income level of a country-of-origin group might reduce their mobility if a group's average "ethnic capital" affects their upward mobility (Borjas 1992).

names are unique within a five-year age band. Third, we require matches to be both exact *and* unique within a five-year age band.[40]

A second concern is the extent to which we are able to generate representative samples. It is well-known that linked samples are not fully representative of the population, as individuals with uncommon names and other attributes are more likely to match between Census waves (Abramitzky et al. 2019). Online Appendix Table A2 compares the sons in our linked samples to the relevant population of US-born men in the 1910 and 1940 Censuses.[41] Our 1880–1910 sample of linked sons has a similar share of second-generation immigrants as the full population (26 vs. 29 percent), and has an average income score that is roughly 3 percent higher than the population. Our 1910–1940 linked sample has fewer second-generation immigrants than the full population (22 vs. 24 percent), and has an average income score that is 5 percent higher than the population. We address the non-representativeness of our linked data by running a version of our main specification in which we reweight the matched data so as to mimic the population on characteristics such as age, state of residence, birthplace, and occupation. Online Appendix Figure A1 shows that the results when reweighting the data are similar to those in the main specification.[42]

ii. Sample Construction

Our baseline approach defines a son's immigrant status using his father's birthplace. In Online Appendix Figure A1 we show that our main results remain unchanged when defining immigrant status using mother's birthplace, or the immigrant status of both parents. Second, unlike with the historical data, our main specification for the modern period compares immigrants to the entire US-born population regardless of race.[43] Figure A29 in the NBER working version of this paper shows that results are similar when comparing white immigrants to the white population in the modern period or comparing all immigrants to the white population in the modern period.[44]

[40] Our main results are similar if we instead implement the linking algorithm developed in Abramitzky, Mill and Pérez (2020), although the exact ordering of countries is slightly different. We provide further details on the results using this approach in the Online Appendix.

[41] That is, we focus on white men aged 30 to 50 whose fathers were either born in the US or in one of the source countries of interest.

[42] These weights are computed using two alternative procedures. In the first procedure, observations are weighted using earlier Census (i.e., childhood) characteristics: To construct this weight, we use the cross-sectional version of the earlier Census and identify the individuals (i.e., the sons) we are able to link forward. We then regress the indicator for being a "matched" individual on fixed effects for the state of residence, the son's age, the birthplace, as well as the father's birthplace and broad occupational category. In the second procedure, observations are weighted using later Census (i.e., adulthood) characteristics: To construct this weight, we use the cross-sectional version of the later Census and identify the individuals (i.e., the sons) we are able to link backward. We then regress the indicator for being a "matched" individual on fixed effects for the son's state of residence, age, birthplace, and broad occupational category.
In both cases, the weights we use to re-weight observations are the inverse probability that an individual can be linked across Censuses.

[43] In the past, immigrants were overwhelmingly white, and so we compare immigrants to the white US-born population in the historical data. Today, in contrast, immigrants are more ethnically diverse.

[44] We note that in this case we define second-generation status in the historical cohorts using the mother's birthplace for comparability with the corresponding modern data.

APPENDIX B: SENSITIVITY OF RESULTS TO INCOME ASSIGNMENT

Online Appendix Figure B1 shows the main results for father-son pairs for whom we observe *actual* rather than predicted income in the 1940 Census. We do so by linking *fathers* who were relatively young in 1910 (20–30 years old) from 1910 to 1940, which enables us to observe these fathers' actual income in 1940 when they were 50–60 years old.[45] Combining this sample with our linked sample of sons, we are able to observe actual income of both fathers and sons for 62,000 pairs, of which 7,800 are immigrants (i.e., those pairs for which we are able to link both the son and his father from 1910 to 1940). For this set of father-son pairs, we show a similar immigrants' mobility advantage when using either actual income or income scores. We also continue to find that children of immigrants from most sending countries are more upwardly mobile than the children of the US-born (but naturally, the exact ranking of countries changes once we divide the 7,800 immigrant fathers into 17 sending countries).

Online Appendix Figure B2 summarizes the robustness of our estimates for the 1880-1910 and 1910-1940 cohorts to using various proxies of income. Each row in this figure shows the values of the intercept and the slope of the rank-rank relationship for both sons of immigrant and US-born fathers. First, we show that the results are similar when we use actual rather than predicted income for the sons we observe in the 1940 Census (second row of panel (b)). Second, we continue to find an immigrant mobility advantage when we use alternative measures of income for farmers (second and third row in panel (a) and the third and fourth row in panel (b)). Specifically, we use the 1900 Census of Agriculture to calculate average farmers' net earnings (farm income minus expenses) for each US county.[46] Alternatively, given the prominence of farming and the difficulty in assigning farmers an income measure, we show that the main results do not change when we drop

[45] We remind readers that the limitations of this exercise are: (1) that actual income in the 1940 Census is only available for wage earners and is notably missing for farmers, and (2) fathers are observed in a different age range than sons (50–60 years old, instead of 30–50 years old). When performing this exercise, we rank fathers and sons only relative to other fathers and sons who are also wage earners.

[46] Income is calculated as the sum of the value of farm output not fed to livestock and the value of house rent and food/fuel produced on farm and consumed by the family. Expenses are calculated as the sum of expenditures for farm labor; fertilizer; feed, seed, and threshing; taxes; depreciation of value of buildings; and depreciation of machinery. For more details on this calculation, we refer the reader to Goldenweiser (1916) and Abramitzky et al. (2012). We note that for farmers living in counties for which we could not calculate farmer net earnings, we assigned them the state estimate for farmer net earnings. For farmers living in counties for which this approach yielded negative net earnings, we assigned them the minimum positive value of county-level farmer earnings in their state. Finally, for farmers living in counties for which this approach yielded very large net earnings (i.e., obvious outliers above the 99th percentile of the farmer earnings distribution), we assigned them the median value of farmer earnings in their state.

fathers and sons who are farmers from the sample and compute ranks based on the non-farming national income distribution.[47]

Our baseline income score is based on the 1940 Census. We show that results are similar when we use more contemporaneous income sources for 1880 and 1910 fathers and sons (fourth row of panel (a) and fifth row of panel (b)). Specifically, for non-farmers, we use information on average earnings by occupation from the 1901 Cost of Living Survey (Preston and Haines 1991).[48] For farmers, we use the 1900 Census of Agriculture (as described above).

We next use a measure of income that also incorporates a respondent's country of birth in the income prediction (i.e., allows for an immigrant penalty in earnings), and find that sons of immigrant fathers continue to be more upwardly mobile than comparable children of US-born fathers.[49]

Unlike when we use other income scores, children of immigrants and children of the US-born look similarly mobile if we use the IPUMS occupation score, which is based on assigning each occupation its median earnings in 1950 (sixth and seventh rows in panels (a) and (b) of Online Appendix Figure B2, respectively). However, the IPUMS occupation score is not well suited to our setting mainly because the score for farmers is too low: Farming was a relatively high-paid occupation in the early 1900s, but it became a low-paid occupation by 1950.[50] That means that the IPUMS occupation score is too low for farmers in 1880 and 1910. This assignment is problematic in our context because US-born individuals were much more likely than immigrants to be farmers (and the children of both US-born and immigrants tended to shift away from farming). Therefore, assigning farmers an income that is too low results in an underestimation of incomes of US-born fathers, and to an overestimation of the rates at which children of the US-born moved up the income ladder relative to children of immigrants. Indeed, if we use the IPUMS occupation score and only adjust farmers' incomes in 1880 and 1910 to correspond to their appropriate rank in the 1900s (based on our baseline income score), we again see that immigrants have a mobility

[47] Of course, moving into or out of farming was by itself an avenue for mobility, which is why we prefer to include farmers in our baseline sample despite the challenges in measuring farmer income.
[48] The 1901 Cost of Living Survey is closer in time to the earlier Censuses but less detailed than our baseline measure (based on the 1940 Census): it includes roughly 150 non-agricultural occupations, whereas the 1950 occupational classification in our linked sample has around 220 occupations.
[49] The NBER working paper shows all results are similar using this measure. There, we also present a specification that incorporates the father's country of origin into the son's income prediction. In that specification, the average income rank of children of immigrants is higher than the average income rank of children of the US-born except at the very bottom of the income distribution (and as a result, the estimated intercept of the rank-rank relationship ends up being similar for both groups).
[50] For instance, farmers were at the median of the income distribution in the Iowa 1915 Census (which included information on farmer's income) but were at about the bottom 10th percentile nationally in 1950 (Feigenbaum 2018).

advantage in both historical cohorts (seventh and eighth rows in panels (a) and (b) of Online Appendix Figure B2, respectively) .

REFERENCES

Abramitzky, Ran, and Leah Boustan. 2017. “Immigration in American Economic History.” *Journal of Economic Literature* 55(4): 1311–1345.

Abramitzky, Ran, Leah Boustan, Katherine Eriksson, James Feigenbaum, and Santiago Pérez. 2019. “Automated Linking of Historical Data.” Working Paper.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2012. “Europe’s Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration.” *American Economic Review* 102(5): 1832–1856.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2014. “A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration.” *Journal of Political Economy* 122(3): 467–506.

Abramitzky, Ran, Roy Mill, and Santiago Pérez. "Linking individuals across historical sources: A fully automated approach." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 53.2 (2020): 94-111.

Aydemir, Abdurrahman, Wen-Hao Chen, and Miles Corak. 2009. “Intergenerational Earnings Mobility Among the Children of Canadian Immigrants.” *The Review of Economics and Statistics* 91(2): 377–397.

Bhattacharya, Debopam and Bhashkar Mazumder. 2011. “A Nonparametric Analysis of Black-White Differences in Intergenerational Income Mobility in the United States” *Quantitative Economics* 2: 335–379.

Bleakley, Hoyt and Aimee Chin. 2010. “Age at Arrival, English Proficiency, and Social Assimilation among US Immigrants.” *American Economic Journal: Applied Economics* 2(1): 165–92.

Bolotnyy, Valentin and Cristina Bratu. 2018. “The Intergenerational Mobility of Immigrants and the Native-Born: Evidence from Sweden.” Working Paper.

Borjas, George. 1992. “Ethnic Capital and Intergenerational Mobility.” *The Quarterly Journal of Economics* 107(1): 123–150.

Borjas, George. 1993. “The Intergenerational Mobility of Immigrants.” *Journal of Labor Economics* 11(1): 113–135.

Cadena, Brian C. and Brian K. Kovak. “Immigrants Equilibrate Local Labor Markets: Evidence from the Great Recession.” *American Economic Journal: Applied Economics* 8(1): 257–290.

Card, David, John DiNardo, and Eugenia Estes. 2000. “The More Things Change: Immigrants and the Children of Immigrants in the 1940s, the 1970s, and the 1990s.” in George J. Borjas (Ed.), *Issues in the Economics of Immigration*. Chicago: University of Chicago Press

Card, David, Ciprian Domnisoru, and Lowell Taylor. 2018. “The Intergenerational Transmission of Human Capital: Evidence from the Golden Age of Upward Mobility.” Working Paper 25000, National Bureau of Economic Research.

Chetty, Raj, David Grusky, Nathaniel Hendren, Maximilian Hell, Robert Manduca, and Jimmy Narang. 2017. “The Fading American Dream: Trends in Absolute Income Mobility Since 1940.” *Science*.

Chetty, Raj, John N. Friedman, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter. 2018a. “Race and Economic Opportunity in the United States: An Intergenerational Perspective.” Working Paper 24441, National Bureau of Economic Research.

Chetty, Raj, John N. Friedman, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter. 2018b. “The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility.” Working Paper 25147, National Bureau of Economic Research.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States". *Quarterly Journal of Economics*, 129(4): 1553–1623.

Collins, William J. and Marianne H. Wanamaker. 2017. "Up from Slavery? African American Intergenerational Economic Mobility Since 1880." Working Paper 23395, National Bureau of Economic Research.

Collins, William J. and Ariell Zimran. 2019. "The Economic Assimilation of Irish Famine Migrants to the United States." Working Paper 25287, National Bureau of Economic Research.

Connor, Dylan Shane. 2019. "Class Background, Reception Context, and Intergenerational Mobility: A Record Linkage and Surname Analysis of the Children of Irish Immigrants." *International Migration Review.*

Derenoncourt, Ellora. 2019. "Can you move to opportunity? Evidence from the Great Migration." Working Paper.

Davis, Jonathan and Bhash Mazumder. 2019. "The Decline in Intergenerational Mobility After 1980." Federal Reserve Bank of Chicago Working Paper No. 2017-05.

Duncan, Brian, Jeffrey Grogger, Ana Sofia Leon and Stephen J. Trejo. 2017. "New Evidence of Generational Progress for Mexican Americans." Working Paper 24067, National Bureau of Economic Research.

Feigenbaum, James J. 2018. "Multiple Measures of Historical Intergenerational Mobility: Iowa 1915 to 1940." *Economic Journal* 128(612): F446–F481.

Feigenbaum, James J. 2015. "Intergenerational Mobility during the Great Depression." Working Paper.

Ferrie, Joseph P. 1997. “The Entry Into The US Labor Market Of Antebellum European Immigrants, 1840–1860.” *Explorations in Economic History* 31(3): 295–330.

Ferrie, Joseph, Catherine Massey, and Jonathan Rothbaum. 2016. “Do Grandparents and Great-Grandparents Matter? Multigenerational Mobility in the US, 1910–2013.” Working Paper 22635, National Bureau of Economic Research.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, and J. Robert Warren. Integrated Public Use Microdata Series, Current Population Survey: Version 6.0 [dataset]. Minneapolis, MN: IPUMS, 2018. https://doi.org/10.18128/D030.V6.0.

Gielen, Anne and Dinand Webbink. 2019. “Migration Selection and Assimilation: Evidence from the Great Migration from Surinam to the Netherlands” Working Paper.

Goldin, Claudia. 1998. “America's Graduation from High School: The Evolution and Spread of Secondary Schooling in the Twentieth Century.” *Journal of Economic History* 58(2): 345–374.

Goldenweiser, E. A. 1916. “The Farmer's Income." *American Economic Review* 6(1): 42–48.

Hatton, Timothy. 1997. “The Immigrant Assimilation Puzzle in Late Nineteenth Century America.” *Journal of Economic History* 57(1): 34–62.

Hilger, Nathaniel. 2016. “The Great Escape: Intergenerational Mobility in the United States Since 1940.” Working Paper 21217, National Bureau of Economic Research.

Hilger, Nathaniel. 2017. “Upward Mobility and Discrimination: The Case of Asian Americans.” Working Paper 22748, National Bureau of Economic Research.

Kosack, Edward and Zachary Ward. 2018. “El sueño americano? The generational progress of Mexican Americans prior to World War II.” Working Paper.

Lee, Chul-In and Gary Solon. 2009. “Trends in Intergenerational Income Mobility.” *Review of Economics and Statistics* 91(4): 766–772.

Lubotsky, Darren. 2007. “Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings.” *Journal of Political Economy* 115(5): 820–867.

Manson, Steven, Jonathan Schroeder, David Van Riper, and Steven Ruggles. 2019. IPUMS National Historical Geographic Information System: Version 14.0 [Database]. Minneapolis, MN: IPUMS. http://doi.org/10.18128/D050.V14.0

Mazumder, Bhashkar. 2005. “Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data.” *The Review of Economics and Statistics* 87(2): 235–255.

Mazumder, Bhashkar. 2015. “Estimating the Intergenerational Elasticity and Rank Association in the US: Overcoming the Current Limitations of Tax Data.” Federal Reserve Bank of Chicago Working Paper No. 2015-04.

Minns, Chris. 2000. “Income, Cohort Effects and Occupational Mobility: A New Look at Immigration to the United States at the Turn of the 20th Century.” *Explorations Economic History* 37(4): 326–50.

Molloy, Raven, Christopher L. Smith, and Abigail Wozniak. "Internal migration in the United States." *Journal of Economic perspectives* 25, no. 3 (2011): 173-96.

Olivetti, Claudia and M. Daniele Paserman. 2015. “In the Name of the Son (and the Daughter): Intergenerational Mobility in the United States, 1850-1940.” *American Economic Review* 105(8): 2695–2724.

Olivetti, Claudia, M. Daniele Paserman, and Laura Salisbury. 2018. "Three-generation Mobility in the United States, 1850–1940: The Role of Maternal and Paternal Grandparents." *Explorations in Economic History* Vol. 70: 73–90.

Passel, Jeffrey. 1986. "Estimating the Number of Undocumented Aliens." *Monthly Labor Review*, September. Washington DC: Bureau of Labor Statistics.

Perlmann, Joel. 2005. *Italians Then, Mexicans Now: Immigrant Origins and Second-Generation Progress, 1890–2000.* Russell Sage Foundation.

Perlmann, Joel and Roger Waldinger. 1997. "Second generation decline? Children of immigrants, past and present--a reconsideration." *The International Migration Review* 31(4): 893–922.

Pérez, Santiago. 2019. "Intergenerational Occupational Mobility Across Three Continents." *The Journal of Economic History* 79(2) 383–416.

Preston, Samuel H., and Michael R. Haines. 1991. *Fatal Years: Child Mortality in Late Nineteenth Century America.* Princeton, NJ: Princeton University Press

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. IPUMS USA: Version 8.0 [dataset]. Minneapolis, MN: IPUMS, 2018. https://doi.org/10.18128/D010.V8.0

Saavedra, Martin and Tate Twinam. 2018. "A Machine Learning Approach to Improving Occupational Income Scores." Working Paper.

Tan, Hui Ren. 2018. "A Different Land of Opportunity: The Geography of Intergenerational Mobility in the Early 20th-Century US." Working Paper.

Ward, Zachary. Forthcoming. "The Not-So-Hot Melting Pot: The Persistence of Outcomes for Descendants of the Age of Mass Migration." *American Economic Journal: Applied Economics*.

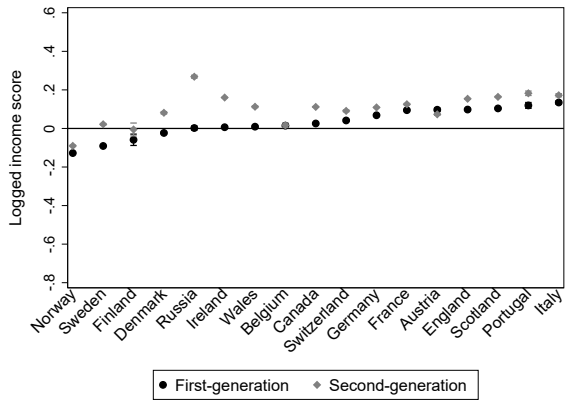
Ward, Zachary and Rohan Alexander. 2018. “Age at Arrival and Assimilation during the Age of Mass Migration.” *The Journal of Economic History* 78(3) 904–937.

Ward, Zachary. 2019. “The Low Return to English Fluency during the Age of Mass Migration.” Working Paper.

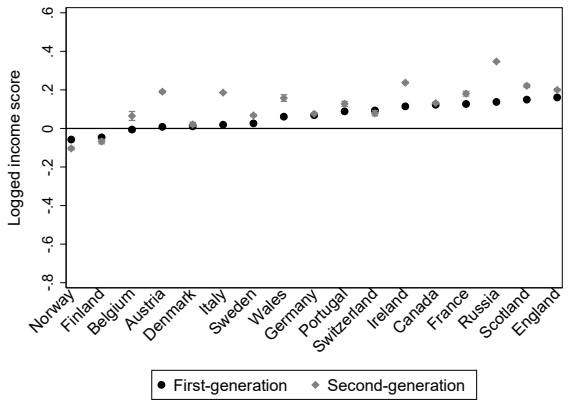
Figures and Tables

Figure 1: Cross-Sectional Earning Gaps for First- and Second-Generation Immigrants relative to the US-Born, by Country of Origin

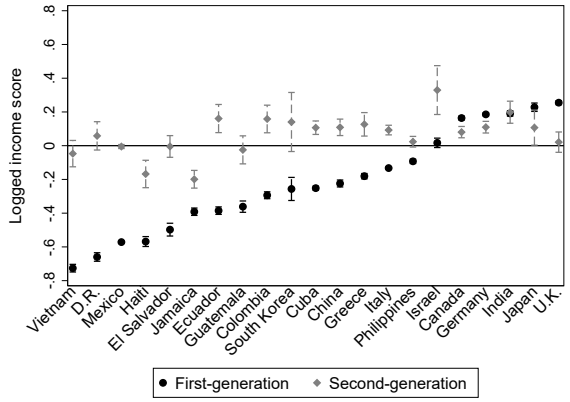
(a) 1880–1910 Cohort



(b) 1910–1940 Cohort



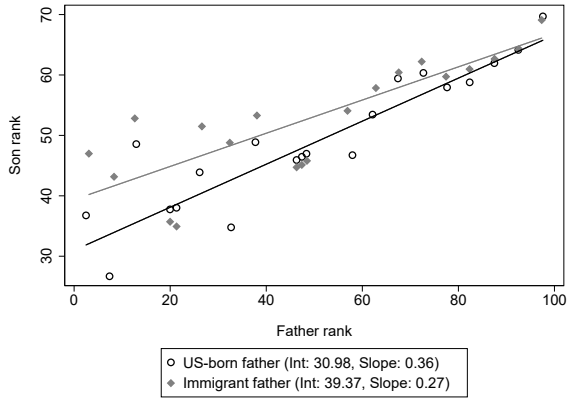
(c) 1980–2010 Cohort



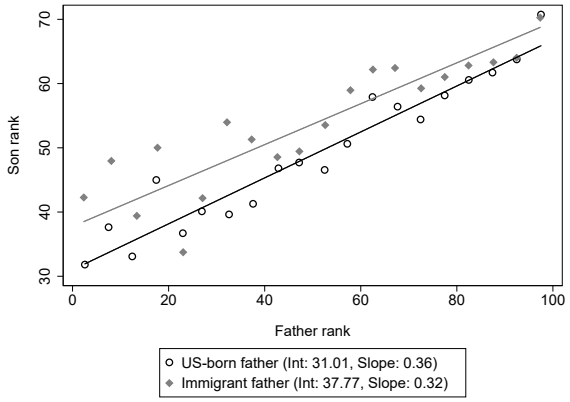
Note: The historical samples use the 1880, 1910, and 1940 Censuses. The historical samples of first-generation immigrants and the US-born are restricted to white men in the earlier Census who are aged 30–50, who have a child in the household, and who were either born in the US or in one of the 17 source countries. The second-generation samples are restricted to white men in the later Census who are aged 30–50 and whose fathers were either born in the US or in one of the 17 source countries. For the most recent cohort, we use the 1980 Census and the 2006–2015 Current Population Surveys (CPS). We restrict the 1980 sample of first-generation immigrants and their counterparts to men aged 30–50, who have a child present in the household, and who were born in the US or in one of the top sending countries identified in the Opportunity Insights sample. We restrict the CPS sample of second-generation immigrants and their counterparts to men aged 30–50 whose fathers were either born in the US or in one of the top sending countries identified in the Opportunity Insights sample. For all cohorts, we use predicted income (described in the main text) as the measure of income.

Figure 2: Intergenerational Mobility of Immigrants and the US-born, Rank-rank Correlations

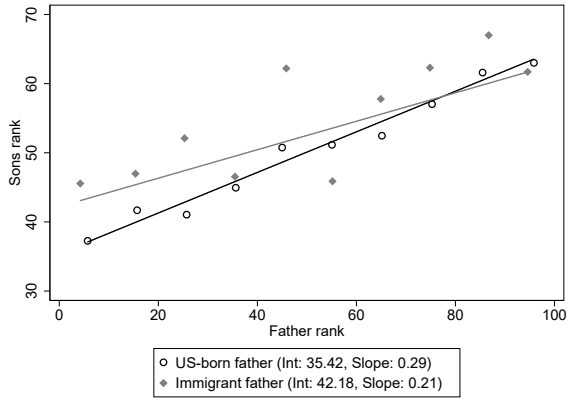
(a) 1880–1910 Cohort



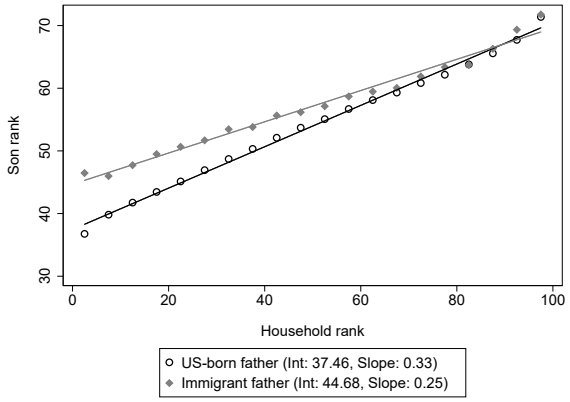
(b) 1910–to 1940 Cohort



(c) GSS: 1984–2006 Cohort



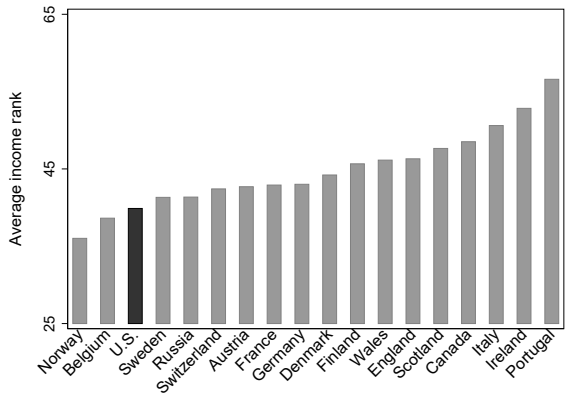
(d) Opportunity Insights: 1997–2015 Cohort



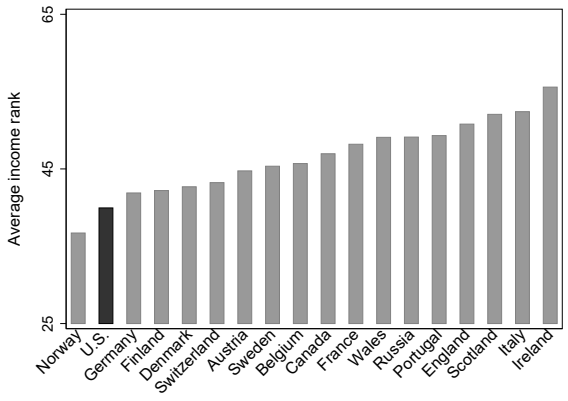
NOTE: Sons are assigned percentile ranks relative to all other sons born in the same birth year. Fathers are ranked relative to all fathers with children born in the same birth year. The figures plot the mean income rank of children by father's income rank, for sons with and without foreign-born fathers, as well as the corresponding regression lines using equation 2. In the GSS sample, the outcomes for sons are measured between 2000 and 2018 (with the median year being 2006). The outcomes for parents are measured via a retrospective question about parents' occupations when the respondent was 16 years old; the data therefore correspond to parental outcomes between 1966 and 2004 (with the median year being 1984). The Opportunity Insights Data come from Chetty et al. (2018a); son's income is measured in 2014–2015 and parental income is measured in 1994–2000.

Figure 3: Average Income Rank for Children Born to 25th Percentile, by Father's Birthplace

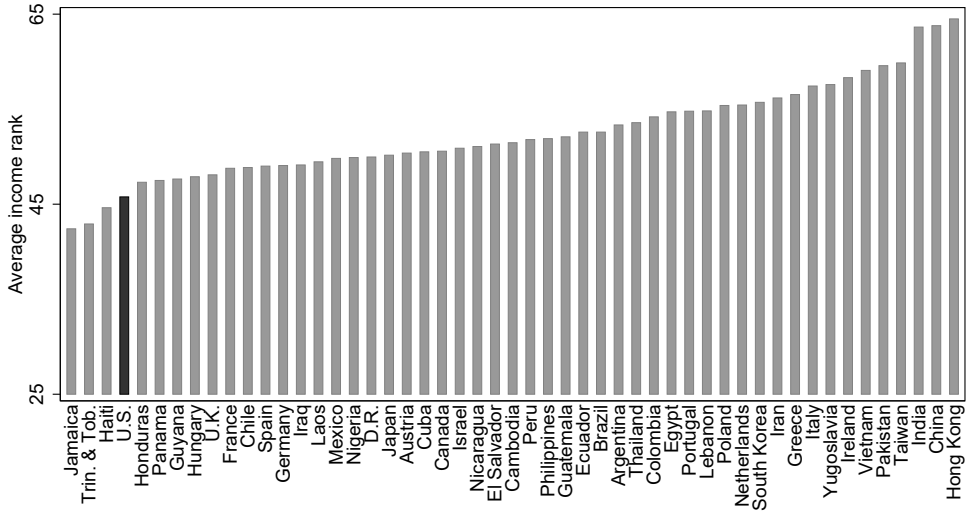
(a) 1880–1910 Cohort



(b) 1910–1940 Cohort



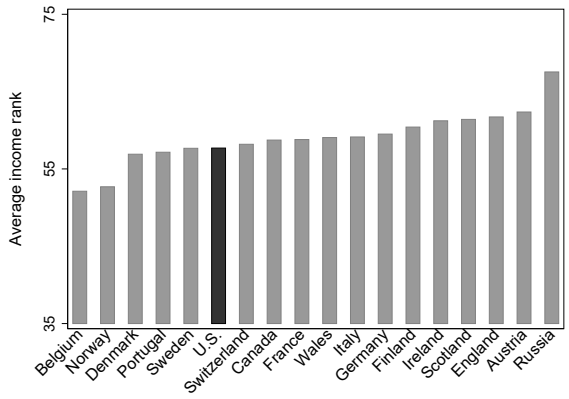
(c) Opportunity Insights: 1997–2015 Cohort



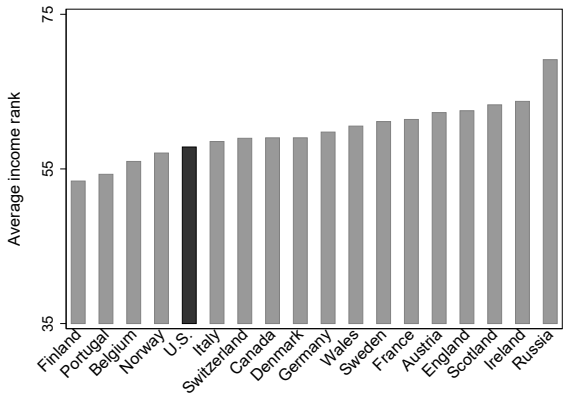
Note: This figure plots the average income rank for children born to the 25th percentile of the parental income distribution. Data for the Opportunity Insights cohort come from Chetty et al. (2018a); son's income is measured in 2014–2015 and parental income is measured in 1994–2000.

Figure 4: Average Income Rank for Children Born to 75th Percentile, by Father's Birthplace

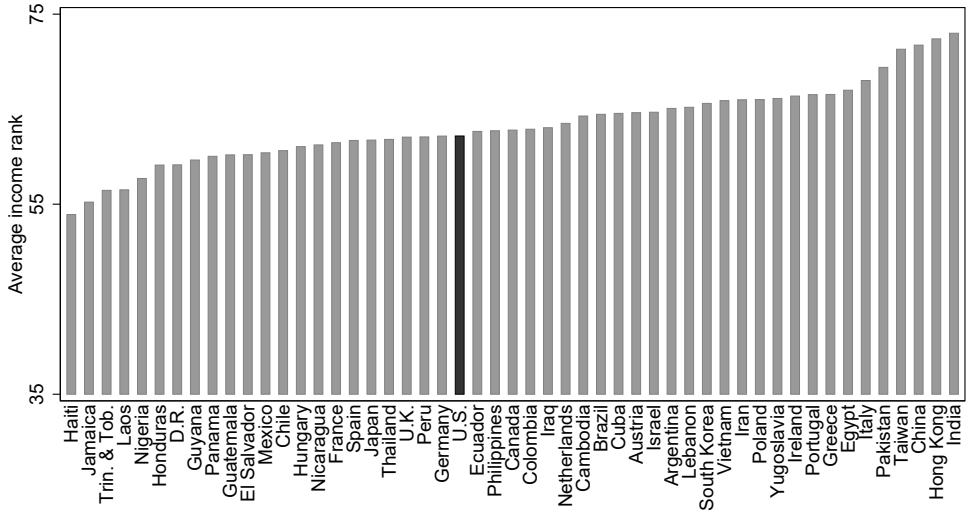
(a) 1880–1910 Cohort



(b) 1910–1940 Cohort



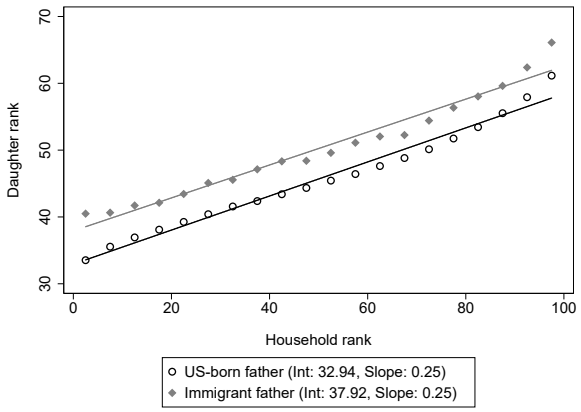
(c) Opportunity Insights: 1997–2015 Cohort



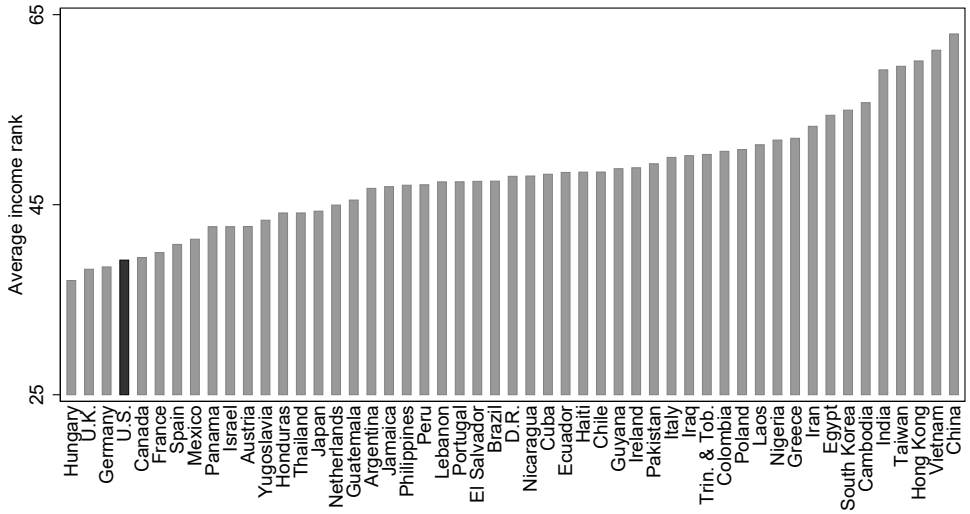
NOTE: This figure plots the average income rank for children born to the 75th percentile of the parental income distribution. Data for the Opportunity Insights cohort come from Chetty et al. (2018a); son's income is measured in 2014–2015 and parental income is measured in 1994–2000.

Figure 5: Intergenerational Mobility of Immigrants' Daughters, 1997–2015 Cohort

(a) Rank-rank Correlations



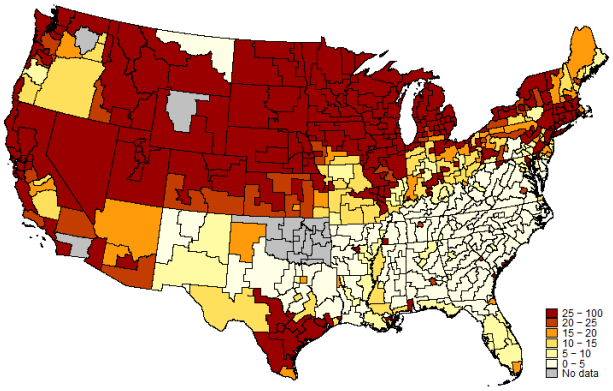
(b) Average Income Rank for Daughters Born to 25th Percentile, by Father's Birthplace



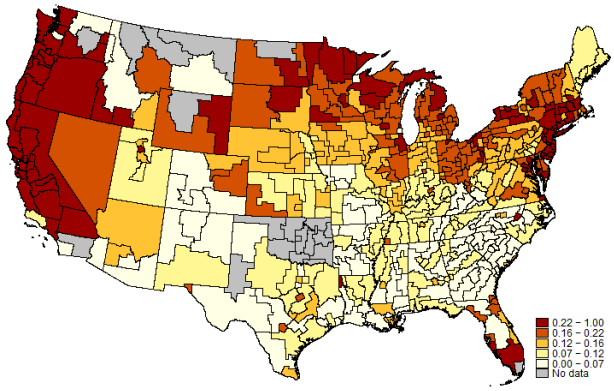
NOTE: Panel (a) plots the mean income rank of daughters by parental income rank, for daughters with and without foreign-born fathers, as well as the corresponding regression lines using equation 2. Panel (b) plots the average income rank for daughters born to the 25th percentile of the parental income distribution, by father's country of origin. Data are from Chetty et al. (2018a); daughter's income is measured in 2014–2015 and parental income is measured in 1994–2000.

Figure 6: Share of Immigrants and Upward Mobility, by State Economic Area or Commuting Zone

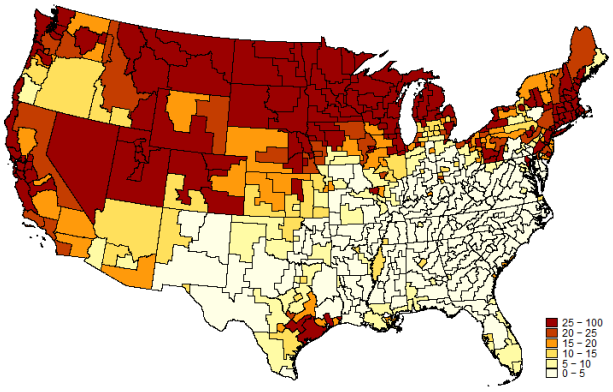
(a) Share of sons that are second-generation (1880–1910)



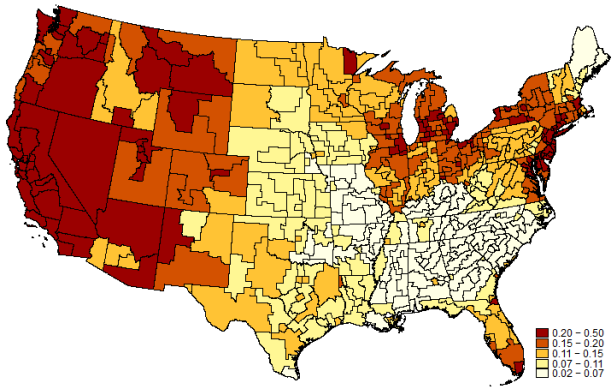
(b) Upward mobility of sons of US-born fathers (1880–1910)



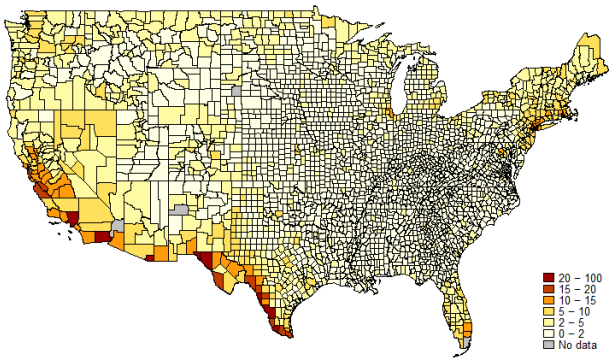
(c) Share of sons that are second-generation (1910–1940)



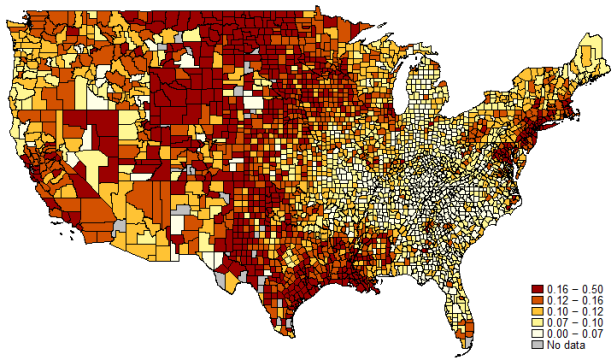
(d) Upward mobility of sons of US-born fathers (1910–1940)



(e) Share of population that is foreign-born, 1980



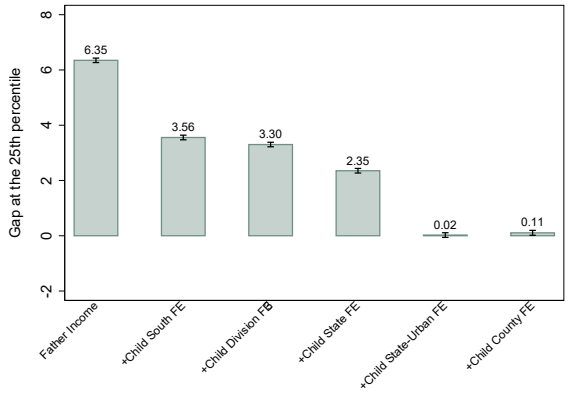
(f) Upward mobility of Opp. Insights cohort



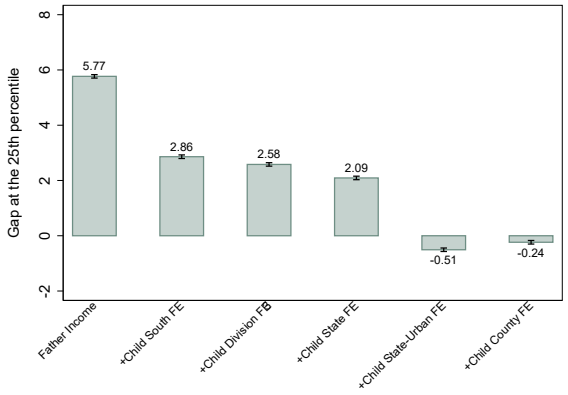
NOTE: Panels (a) and (c) show the share of the individuals in our linked sample that are second-generation immigrants in each 1880 or 1910 state economic area, respectively. Panel (e) shows the share of each county's population that is foreign-born in 1980, using data from Manson et al. (2019). In panels (b) and (d), upward mobility is measured as the share of sons of US-born fathers in each state economic area that reached the top fifth of the national income distribution, conditional on having had a father in the bottom fifth of the national income distribution. Panel (f) shows the share of non-Hispanic white sons who are upwardly mobile using data from Chetty et al. (2018b); upward mobility is measured as the share of sons in the top 20 percent of the income distribution, conditional on having had household income in the bottom fifth of the national income distribution.

Figure 7: Intergenerational Gap at the 25th percentile, Comparing Children in Similar Childhood Locations

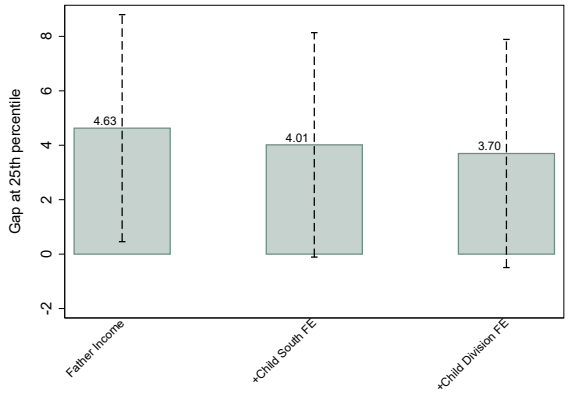
(a) 1880–1910 Cohort



(b) 1910–1940 Cohort



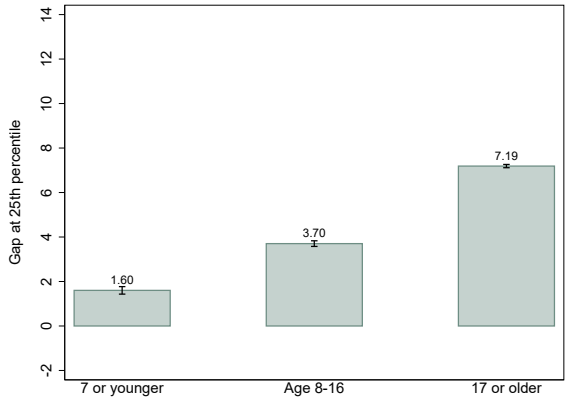
(c) GSS: 1984–2006 Cohort



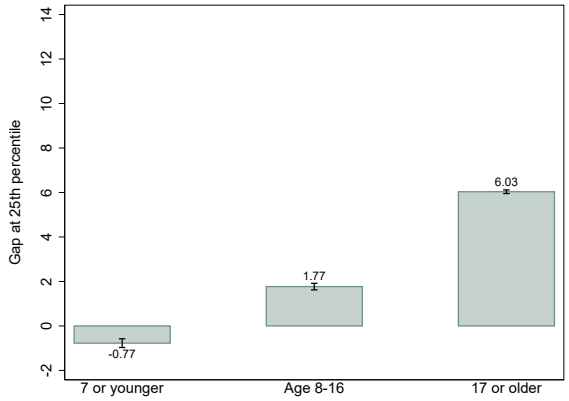
NOTE: This graph plots the intergenerational gap between children of immigrants and children of US-born men, for children born to the 25th percentile of the income distribution. The first bar for each cohort is estimated by regressing son's income ranks on their fathers' ranks, an indicator for having a foreign-born father, and the interaction of these two variables. The second and third bars plot the gap after including an indicator variable for growing up in the US South and childhood Census division fixed effects, respectively. In the historical cohorts, the fourth bar plots the gap after including childhood state fixed effects and the fifth bar plots the gap after including state × urban fixed effects. The sixth bar plots the gap after including childhood county fixed effects.

Figure 8: Intergenerational Gap at the 25th percentile for 1910–1940 Cohort, by Father's Age of Arrival and Language in Sending Country

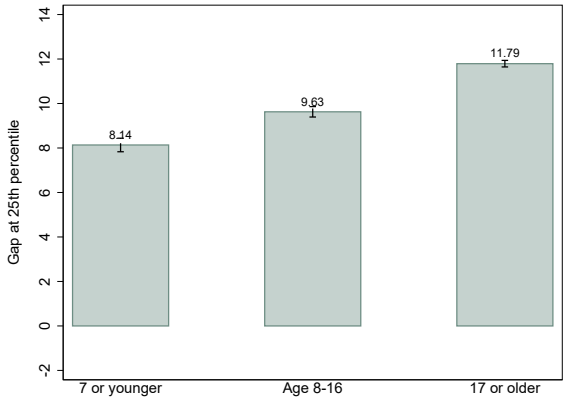
(a) All Countries



(b) Non-English-Speaking Countries



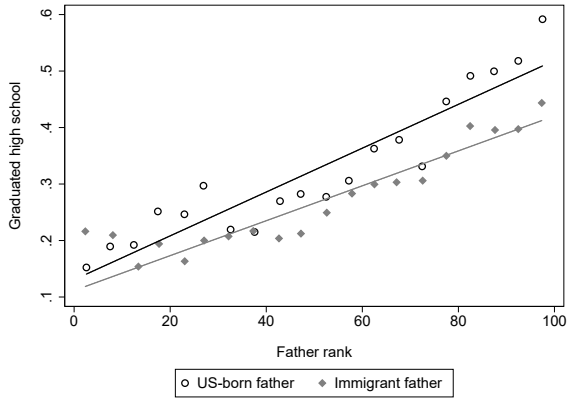
(c) English-Speaking Countries



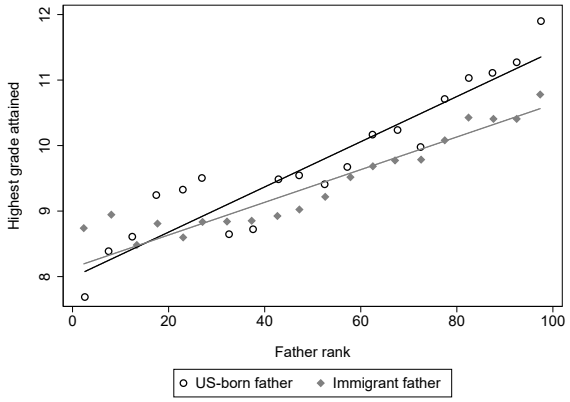
NOTE: This graph plots the intergenerational gap between children of immigrants and children of US-born men, for children born to the 25th percentile of the income distribution. Sons of immigrants are divided into groups based on their father's age of arrival to the United States. The latter two figures also separate sons based on the language spoken in their father's sending country (English-speaking countries are Canada, England, Ireland, Scotland, and Wales).

Figure 9: Educational Attainment, 1910–1940 cohort

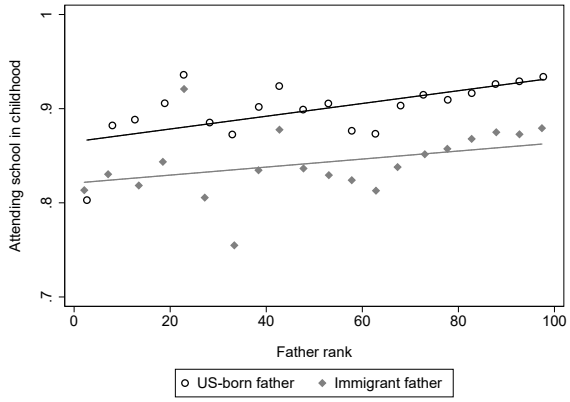
(a) Son graduated high school



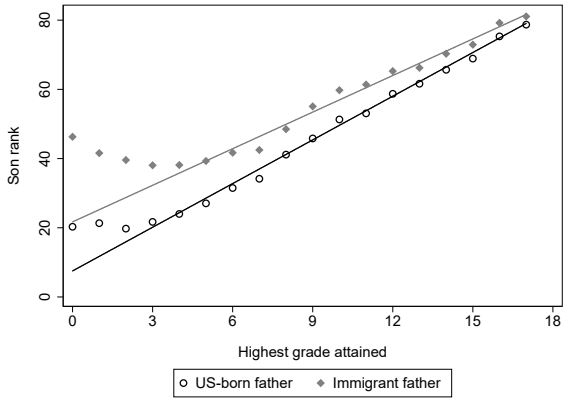
(b) Son's highest grade in school



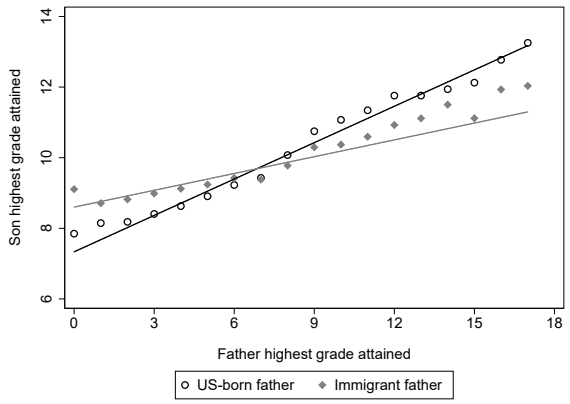
(c) Son's school attendance in 1910



(d) Son's income rank and educational attainment



(e) Son's and father's educational attainment



NOTE: Fathers are ranked relative to all fathers with children in the same birth year. Panels (a) and (b) plot the mean educational attainment of children by their father's income rank, for sons with and without foreign-born fathers, as well as the corresponding regression lines. Panel (c) considers linked sons aged 12–16 in the earlier Census and plots school attendance as a function of his father's income rank (after controlling for age fixed effects). Panel (d) plots a son's mean income rank as a function of completed years of schooling. Panel (e) plots a son's average years of completed schooling as a function of his father's years of schooling (for fathers who can be linked to the 1940 Census).

Table 1: Intergenerational Mobility Estimates, by Cohort

	Intercept (α)	Immigrant father (β_0)	Father's rank (β_1)	Immigrant father $\times$ rank (β_2)
1880–1910 Cohort	30.98 [0.06]	8.39 [0.12]	0.36 [0.00]	-0.08 [0.00]
1910–1940 Cohort	31.01 [0.04]	6.75 [0.09]	0.36 [0.00]	-0.04 [0.00]
GSS Cohort	35.42 [1.17]	6.76 [3.06]	0.29 [0.02]	-0.09 [0.06]
Opp. Insights Cohort	36.64 [0.40]	7.42 [0.60]	0.33 [0.01]	-0.08 [0.01]

NOTE: This table reports the estimated coefficients from equation 2 in the paper (standard errors in brackets below each estimate). Data for the Opportunity Insights cohort come from Chetty et al. (2018a).