NBER WORKING PAPER SERIES

FORTUNATE SONS: ELITE POLITICAL SELECTION IN AMERICAN HISTORY

James J. Feigenbaum
Andrew B. Hall
Daniel M. Thompson
Jesse Yoder

Working Paper 35569
http://www.nber.org/papers/w35569

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2026

This draft subsumes and substantially expands on NBER WP 26156. For research assistance, the authors thank Brittany Dutton, Qianmin Hu, Elise Kostial, Anna Nakai, Conor Orton, Josh Rose, and Anish Sundar, as well as Joseph Price's Record Linking Lab at BYU. For helpful discussion, the authors thank Ran Abramitzky, Avi Acharya, Alberto Alesina, Sam Bazzi, Matilde Bombardini, David Broockman, Jacob Brown, Nick Carnes, Dan Carpenter, Karen Clay, Cesi Cruz, Ernesto Dal Bó, Robert Erikson, Martin Fiszbein, Anthony Fowler, Matt Gentzkow, Liz Gerber, Rob Gillezeau, Ed Glaeser, Claudia Goldin, Josh Goodman, Justin Grimmer, Alisa Hall, Tarek Hassan, Seth Hill, Brian Knight, Shiro Kuriwaki, Eddie Lazear, Greg Martin, Jaakko Meriläinen, Christopher Muller, Suresh Naidu, Paul Novosad, Max Palmer, Daniele Paserman, Vincent Pons, Joseph Price, Hunter Rendleman, Dominic Rohner, Jesus Rojas Venzor, Wendy Schiller, Benjamin Schneer, Jesse Shapiro, Erik Snowberg, Jim Snyder, Marco Tabellini, Chris Tausanovitch, Danielle Thomsen, Francesco Trebbi, Jessica Trounstine, Jennifer Nicoll Victor, Zachary Ward, Ruthie Zussman, the students of Econ 220 in the Fall 2018–2019 quarter, seminar participants at BU, UBC, the Stanford Economics Brown Bag Lunch, the NBER Spring Political Economy Conference, the members of Stanford's Democracy Policy Lab, the Center for Effective Lawmaking Conference, the 2019 Congress and History conference at MIT and Harvard, the University of Toronto Economic History Conference, and the Harvard Political Economy Seminar. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Fortunate Sons: Elite Political Selection in American History
James J. Feigenbaum, Andrew B. Hall, Daniel M. Thompson, and Jesse Yoder
NBER Working Paper No. 35569
August 2026
JEL No. H10, H70, N31, N32

ABSTRACT

We trace the family origins of Members of Congress (MC) born between 1830 and 1950, linking each MC, their parents, and their brothers to the complete-count censuses. Future MCs have always been economic outliers. By 1940, close to 60 percent of MCs aged 18--40 held a college degree, against under 5 percent of comparable young men outside Congress. In the 19th-century censuses, MCs as adults held roughly three times the wealth of demographically matched controls. They also come from economically elite families: their fathers earn more, attain more education, and hold more wealth. The brothers of future MCs sit between the population and the MCs themselves. Across the socioeconomic measures we compare, we calculate that roughly half of an MC's adult premium is shared with his brother and the other half is his own. Even as the American economy and political system have changed, the family background gap has widened over the last century and a half. Wealth, household servants, and four separate occupation-based scores all show the gap holding or growing across cohorts. We study four Progressive-Era reforms in a triple-difference design---women's suffrage, the secret ballot, direct primaries, and the direct election of senators---and none produce detectable shifts in who reaches Congress. Our results document a Congress drawn from the most fortunate families in American society in every cohort we observe.

James J. Feigenbaum
Boston University
Department of Economics
and NBER
jamesf@bu.edu

Andrew B. Hall
Stanford University
Department of Political Science
andrewbhall@stanford.edu

Daniel M. Thompson
University of California, Los Angeles
Department of Political Science
dthompson@polisci.ucla.edu

Jesse Yoder
Stanford University
jesseyoder90@gmail.com

1 Introduction

Today, Members of the U.S. Congress are wealthier and more educated than the citizens they represent (Carnes 2018, 2013). Has this always been the case? The American myth of social mobility extends to politics: anyone born in a log cabin could grow up to serve in Congress or even as President like Lincoln or Garfield. The historical record says otherwise. In every cohort we observe, legislators come from high-status families.[1]

We ask three questions about elite political selection in American history. How economically unrepresentative are members of Congress? How much of that unrepresentativeness traces to the families legislators were born into, rather than to the individuals they grew up to be before entering Congress? And how has the gap changed across the long sweep of American political history, from the cohorts serving in the antebellum Congress to the cohorts serving at the end of the 20th century?

Future members of Congress are economically far above the median of the American population in every era we can observe. In 1940, the first census to record individual earnings and educational attainment, nearly 60% of future MCs aged 18–40 report 16 or more years of schooling, against fewer than 5% of similarly-aged men in the population. More than 12% report top-coded wage earnings above $5,000, compared to less than 1% of the population. They own homes three times as valuable and are nearly 12 times more likely to own homes in the top 1%. The gap survives tight matching on age, race, place, and birth state. Even among fellow lawyers, future MCs stand out, with more earnings and more education. In 19th century censuses, where the enumerators recorded wealth rather than wages, the gap looks much the same: future MCs hold roughly three to four times the real-estate wealth, and about three times the personal-property wealth, of their tightly matched controls. Across every measure of economic status, in every era we can study, future legislators are unusually successful and high-status men.

MCs also grow up in highly selected families. The fathers of future MCs in 1940 earn roughly twice the average father's wages and are more than five times as likely to hold a college degree. Their homes are far more valuable and they are 30 times more likely to have two or more live-in household servants. In the 19th century, MC fathers hold real-estate and personal-property wealth roughly twice that of demographically matched control fathers. To quantify how much of the MC's advantage is his own versus his family's, we compare future MCs to their own brothers. The eventual MC out-earns and out-educates the brother who stays out of politics, though the brother is still well ahead of the population. We calculate that roughly half of an MC's advantage over the population is shared with his brother, suggesting that family background and the MC himself play about equal roles in the MC advantage we document.

The unrepresentativeness of the future MCs' backgrounds has not narrowed across the century-plus our data cover. From the 40th Congress (1867) through the 115th (2017), MC fathers exceed the matched-control fathers by large margins on both the two direct measures of elite status we can

[1] An apocryphal story about the campaign of a different assassinated congressman-then-president may lie closer to the truth. Legend has it that, while bankrolling his son's run for office, Joe Kennedy boasted he could send his chauffeur to Washington with the sums he was spending.

track across the period—real-estate wealth and household servants—and the four occupation-based scores we study. The gap rises through the bulk of the sample. Despite dramatic changes in political institutions and the American economy—including extensions of the franchise to non-property-holders, to Black men, to women, and to younger voters; the High School Movement and then the expansion of college access—MCs continue to come from elite backgrounds, Congress after Congress.

We sharpen this finding by estimating the effects of specific Progressive-Era reforms on political selection. We analyze four reforms that all democratized American electoral politics—women's suffrage, the secret ballot, the direct primary, and the direct election of senators—in a triple-difference specification, comparing the backgrounds of MCs in treated states and years against their tightly matched population controls. None of the reforms produce a detectable or robust change in the family-background gap across our measures of economic status.

To establish these patterns, we build a dataset that follows both members of Congress and their families across censuses. We start with the Biographical Directory of the U.S. Congress and the universe of representatives and senators born between 1830 and 1950 who have died. We locate each on FamilySearch.org, a large open genealogical platform, and harvest the wiki-style person records and family trees that its users have built up over decades, supplementing them with our own records and linking and genealogical work where needed (Price et al. 2021). From each MC's FamilySearch record we follow a chain to the corresponding census record, which lets us pull every variable IPUMS records for that person from the complete-count censuses (Ruggles et al. 2024, 2025). The FamilySearch trees also give us each MC's parents and siblings, whom we link to the complete-count censuses the same way.

Three features of the resulting panel shape what this paper can say. First, the data are long. In Thompson et al. (2019), the predecessor to this paper, we only observed MCs in a single 1940 cross-section; we extend that to ten decennial censuses, allowing us to study MCs born across more than a century. The long panel is what lets us ask whether the family-background gap has moved over time, across more than a century of economic growth and political reform, including the Progressive-Era reforms we study directly. Snapshots of officeholders' economic backgrounds are common, but over-time data barely exist, a gap Carnes and Lupu (2023) identify and our panel's long-run view fills. Second, the FamilySearch trees include the brothers of future MCs, whom we link to the same complete-count census records. Comparing a future MC to his own brother allows us to net out the family component of his adult advantage and quantify the role of individuals versus their families in their pre-Congress success. Third, we can compare MCs to their matched controls across the full century, despite the lack of wealth data after 1870 and earnings data before 1940, by relying on recent improvements in occupation-based status measures.[2]

Our paper joins a comparative literature on the socioeconomic backgrounds of legislators (Dal Bó and Finan 2018; Carnes and Lupu 2023; Gulzar 2021). The closest predecessor in approach is Dal Bó et al. (2017), who use Swedish administrative registers to characterize political selection in

[2]We also make use of direct measures of status that we can observe over multiple censuses, including the presence of live-in household servants (Airoldi and Moser 2024; Kalsi and Ward 2025).

another long-lived democracy.[3] The Swedish picture is one of *inclusive meritocracy*: politicians are positively selected on competence even within their own families, but their parental backgrounds are roughly representative of the population. Recent work in Finland (Jokela et al. 2025) and Denmark (Dahlgaard and Pedersen 2025) reach similar conclusions about positive ability selection.[4] In the UK, Cagé and Dewitte (2024) build a long-run candidate-level dataset on UK general elections from 1918 to 2017 and document a persistent over-representation of high-SES candidates for Parliament.[5] Our results suggest that the American political system is an *exclusive meritocracy*: positive selection on labor-market success sits alongside a political class with substantially unrepresentative family backgrounds, and both patterns have held for more than a century.[6]

We also contribute to the broader literature on the historical selection into elite professions. The over-representation of high-status families is not unique to Congress, but no prior work has followed the family origins of its members across more than a century to ask whether that selection has tightened, loosened, or held steady. Inventors come from high-status backgrounds: Bell et al. (2019) document, using IRS tax records, that children from top-1% families are 10 times more likely than children from below-median families to become inventors, even conditional on early test scores; Akcigit et al. (2017) provide the historical counterpart, linking 1880–1940 patents to the complete-count censuses and finding strong father-income effects on inventive activity. Abramitzky et al. (2024) link U.S. university faculty to the censuses and show that academics are severely and persistently under-recruited from poorer families, with clear stability over the 20th century in selection. Airoldi and Moser (2024) use the 1921 *American Men of Science* biographical dictionary to show that scientists from high-SES families have roughly 40% higher odds of being recognized as leaders in their fields, conditional on publication. Novosad et al. (2024) extend the picture across the global century of Nobel laureates, finding that the average laureate grew up in roughly the 90th percentile of their country's distribution and that this has only very modestly opened over time.[7] Our findings for Congress fit this broader pattern: elite American institutions, despite their

[3] In concurrent work, Chyn et al. (2026) also study the family backgrounds of U.S. politicians in the late nineteenth and early twentieth century, identifying officeholders from their enumerated census occupations and characterizing family background with an imputed measure of wealth. Our project, which originates in Thompson et al. (2019), instead links the full roster of members of Congress to their census records, measures family background along several direct and occupation-based dimensions, and compares each future legislator to his own brothers.

[4] Folke and Rickne (2025) document an intra-party promotion bottleneck through which working-class candidates drop out of Swedish electoral careers at every step from candidacy to leadership.

[5] Reeves and Friedman (2024), drawing on a century of *Who's Who* records, document persistent reproduction across the broader British elite, including British MPs.

[6] Within American politics our paper also builds on Querubín and Snyder (2013), who use the 19th-century censuses to study wealth accumulation in office. Dal Bó et al. (2009) show that holding congressional office is self-perpetuating, raising the chance that a legislator's relatives later reach Congress. Recent work from Ho (2025) also links MCs to the census to study the effects of office-holding on MCs' families and neighbors. Studies of antebellum and postbellum political economy have also made use of links to and from the complete count censuses. Sehgal and Sehgal (2024) document a persistent wealth advantage among MCs descended from slaveholding families; Hall et al. (2019) link Civil War enlistment records to the census to study the effects of slaveholding on rebellion; Bellani et al. (2026) show increased political influence among former slaveowners after emancipation.

[7] Studies of elite educational gatekeepers tell a similar story: exclusive Harvard clubs of the 1920s and 1930s reproduced economic elite status through peer networks (Michelman et al. 2022), and Ivy-Plus admission today roughly doubles the probability of reaching the top 1% (Chetty et al. 2023).

differences, recruit disproportionately from already-advantaged families.

The century-plus of descriptive evidence we assemble speaks to mechanism, not to verdict. If economic success and educational attainment have nothing to do with whether a legislator serves her constituents well, as Carnes and Lupu (2015) argue, the selection patterns we document are a pure cost of unrepresentative government. If, on the other hand, earnings ability and education do correlate with political ability—as an important strand of research in political economy argues (Besley and Reynal-Querol 2011; Besley et al. 2011, 2017; Meriläinen 2022)—then the relevant question is how much we want our political system to prioritize these abilities at the cost of preventing those with less access to education, and therefore less ability to earn, from becoming politicians. The value of legislators with strong educational backgrounds and private-market success could be high, if these traits are useful for crafting better policy, but the costs of an unrepresentative legislature are also high: rich legislators may be out of touch with the desires of their constituents (Gilens 2012), a broadly representative legislature may add legitimacy to the democratic process (Mansbridge 1999), and economic representativeness may be a normative goal in and of itself. We do not adjudicate that question here. Our long panel establishes the empirical magnitudes within which any normative debate has to take place: how unrepresentative legislators have been, how much of that traces to family, and how stable the pattern has been across a century of institutional reform intended to flatten it.

The remainder of the paper is organized as follows. Section 2 describes the record-linking procedure that yields our panel of MCs, their parents, and their brothers across ten decennial censuses. Section 3 shows that future MCs sit far above their matched controls on every measure of adult economic status we observe, in every era we can study, and that the gap survives even when we restrict the comparison to fellow lawyers in the same county. Section 4 turns to family background and documents that MC fathers exceed matched-control fathers by large margins on wealth, earnings, education, and home value and that the gap rises through the bulk of our panel rather than falling. Section 5 compares future MCs to their non-politician brothers to separate family background from the future MC, and finds the two account for roughly equal shares of his premium. Section 6 estimates the effects of four Progressive-Era reforms on the family backgrounds of newly elected MCs, showing that none robustly moves selection. Section 7 asks why these family backgrounds may shape what Congress does. Section 8 concludes.

2 Data: Linking Members of Congress to the Census

To study the social and economic backgrounds of U.S. legislators, we link members of Congress to their records in the Federal Census. We draw on the recently digitized and transcribed complete count censuses from 1850 to 1950 (Ruggles et al. 2024, 2025).[8] We also link the parents and siblings of MCs to their own census records, allowing us to observe the family characteristics of future politicians. Our sample covers all deceased members of Congress born between 1830 and 1950, a

[8]The U.S. census releases personally identifiable records on a 72-year delay, so the 1850, 1860, 1870, 1880, 1900, 1910, 1920, 1930, 1940, and 1950 censuses are now publicly available in full.

group whose careers in Congress ran from the antebellum era to the last decades of the 20th century. The long period we study spans several fundamental transformations in American politics, including the Civil War and emancipation, the rise of mass democratic participation, the expansion of the franchise to women, and much more.

We build our data on FamilySearch.org, a genealogical platform with a wide assortment of digitized historical records, including every surviving Federal Census between 1850 and 1950.[9] Crucially, it also hosts open wiki-style user-contributed family trees connecting individuals to their kin across censuses and other records (Price et al. 2021).[10] Working from these trees, rather than matching records by name, age, and birthplace alone, gives us greater confidence in our links.

2.1 Linking Members of Congress to FamilySearch

We start with the Biographical Directory of the United States Congress, which lists every member of the House and Senate together with name, birth year, birth place, and a short biographical sketch. We restrict the sample to MCs born between 1830 and 1950 who were deceased as of May 2026, yielding 6077 legislators.[11] This is a substantial expansion of our earlier work (Thompson et al. 2019), which covered only MCs observable in the 1940 census. The current sample reaches back to legislators born just after the founding generation and forward to recent figures—Donald Rumsfeld, Robert Dole, John Lewis—who never appeared in the older data at all.[12]

For each MC we either locate or create a FamilySearch person record. Most legislators already have person records, built up over years by genealogists, historians, and descendants. For the rest we create new records from the Biographical Directory entry and other public sources. We then attach every available Federal Census record to the MC's profile, verifying links prior users have already made and adding any we find missing. We also attach non-census records to the MC's profile, such as birth, marriage, and death records, to enrich the data, further verify the links, and improve our ability to find their family members.

We then harvest the census links for each MC and connect them to the IPUMS complete-count census (Ruggles et al. 2024, 2025).[13] From the IPUMS complete-count census data, we have transcriptions and encoding of all the various census variables. We focus on the socio-economic status (SES) variables, which we use to measure the background and pre-congressional status of

[9] The 1890 Federal Census was destroyed by fire in 1921.

[10] FamilySearch.org is operated by the Church of Jesus Christ of Latter-day Saints and is freely accessible. The database contains over 8 billion historical records and links these records to over 1.4 billion person profiles.

[11] We impose the 1830-1950 birth year restrictions to ensure that each member of Congress could have appeared in at least one Federal Census during their lifetime. The 1850 Census is the first to enumerate all household members by name; the 1950 Census is the most recent census with publicly available microdata. We restrict to deceased members because FamilySearch person records are publicly visible only for deceased individuals.

[12] In Appendix B, we show that an alternative data construction strategy that starts with self-identified politicians or Members of Congress directly in the census would fail to capture a huge number of MCs who report different occupations during enumeration, even during censuses taken when they are serving in office.

[13] FamilySearch assigns each person a unique PID and each historical record a unique ARK ID. IPUMS, in turn, assigns its own identifier—the HISTID—to each enumerated individual in the complete-count censuses, and these HISTIDs map one-to-one onto ARK IDs. Following PID $\rightarrow$ ARK $\rightarrow$ HISTID lets us pull the IPUMS variables for each MC and each relative we have placed on a family tree.

MCs, recorded in the censuses before the MCs enter Congress.

2.2 Constructing Family Trees

Once we have located each MC on the FamilySearch tree and added additional historical records, including census records, we gather data on the MC's family. Many MC relatives already have records on FamilySearch, given the platform's broad coverage of American families. Where they do not, we create new FamilySearch records for each of their parents, siblings, spouses, children, paying particular attention to the MC's father and brother(s). When a future MC appears as a child living with his parents in, say, the 1880 census, the parents and siblings are sitting on the same enumeration page and we can read them off directly.[14] We then locate or create FamilySearch records for each and attach every record we can find to their profiles and verify the census links against the underlying census transcription and image. We then harvest the census data for parents and siblings and connect them to the IPUMS complete-count census, just as we do for MCs.

This differs from the automated record-linkage methods now standard in economic history (for a review, see Abramitzky et al. 2021). Whether the algorithm is exact matching (Ferrie 1996), supervised learning (Feigenbaum 2018), or expectation-maximization (Enamorado et al. 2019), the input is usually the same: names, ages, and birthplaces compared across censuses. These methods work well on average. But they generate both false positives—records belonging to different people that are wrongly joined—and false negatives—records that cannot be linked, often because the names are common (Bailey et al. 2020). Because MCs are many orders of magnitude rarer than the population from which their relatives are drawn, even a small false-positive rate could bias our comparisons.[15]

Working from the household sidesteps this problem. Rather than linking a father's 1880 record to his 1900 record on name similarity, we observe the father in the same household as his child—the future MC—in 1880, and then recover his 1900 record, and any others, from his own FamilySearch profile rather than by re-matching him across censuses. We never have to match fathers by name across censuses, the main source of error in standard approaches. Errors remain possible, of course: enumerators misrecorded names, households sometimes included non-relatives, and the genealogical data on FamilySearch is imperfect, despite the efforts of researchers and genealogists. But the scope for systematic bias is much smaller.[16]

[14] One advantage of working from family trees is that we observe the parents and siblings of future MCs directly in the genealogical record, rather than inferring kinship only by linking across decades or using census households.

[15] If, for example, we incorrectly linked a census record belonging to a typical member of the population as belonging to a future member of Congress's father, this would attenuate our estimates of the socioeconomic advantage of politicians' families. The asymmetry between the small treatment group and the large population makes false positives particularly damaging.

[16] Our results are robust to starting with only the FamilySearch records of MCs and then linking the MCs to their parents and siblings using more traditional algorithmic record linkage methods, in particular the CensusTree project (Buckles et al. 2025), the MLP links (Helgertz et al. 2021), and ABE methods as in the Census Linking Project (Abramitzky et al. 2021). However, doing so restricts our sample of fathers and brothers to those that are observed in the same household at some point with the MCs and who can be linked algorithmically. See Section A.2 for robustness to alternative linking pipelines.

2.3 Control Samples

For each MC observed in a census, we draw matched control sets from the universe of American enumerated in that same census year. We build these controls along a *telescope* of cumulative conditions: starting from all members of the MC's sex, then adding the same race and nativity, then age, then region, then state, county, urban/rural/farm status, and finally place of birth. The control set is drawn separately in each census wave, so the comparison group lines up with the MC at the right point in the life cycle—a child when the MC was a child, a young adult when the MC was a young adult. Our descriptive figures use the full telescope to show how MCs differ from the population once we hold more and more demographics fixed. In our main regressions, we restrict to the tightest set: controls who match the MC on every variable in the telescope.[17] From every control pool we exclude any census record linked to an MC, an MC's brother, or an MC's father, so that no MC's relative is ever another MC's control.

We use the same procedure to build matched control sets for MCs' fathers and brothers. Targets are the father and brother census records from the FamilySearch crosswalk; controls are drawn from the same census year using the same telescope. These are controls *for* the father (or brother) of an MC, not fathers (or brothers) *of* the MC's controls. Each control is a census record matched to the target (MCs fathers and brothers) on the telescope demographics, with no FamilySearch linkage of its own.

The control samples serve two purposes. First, they provide a benchmark for assessing how unrepresentative MCs or MC families are relative to comparable non-politicians, beyond raw comparisons to the full population. Second, in our analyses, the control sets provide a "placebo" group that should be unaffected by our various treatment measures.

2.4 Data Coverage

The resulting dataset covers more than a century of legislators and their families. Figure 1 plots the share of MCs we link to the census by Congress (panel a) and by year of birth (panel b). Linkage is above 90 percent in every cohort from the 1840s through the 1920s, falling off only at the endpoints, where the MC's life overlaps with fewer published census waves.

We locate fathers for 91.5% of MCs. Of those, 97.9% are linked to at least one census record reporting an occupation, and 92.1% are observed at prime working ages (25–55) in at least one census. We locate at least one brother for 75.0% of MCs, and those brothers appear in a census record before the MC took office for 99.3% of those MCs with a located brother.[18] There is no analogous location rate for the control sample because controls are not located through FamilySearch

[17] An MC with no control matching on every telescope variable drops from the tightest-set regressions. Such MCs still enter the descriptive figures, which show the full telescope and so retain them at the looser matching levels.

[18] We focus on brothers rather than sisters because for much of our period, female labor force participation in the formal labor market was low and highly selected. Brothers (and their occupations) carry a much clearer signal of economic status (Goldin 1990). Women also typically changed their surnames upon marriage, making it difficult to link them across censuses using name-based methods. While this linking limitation is common to all studies using linked historical census data, our use of the FamilySearch profiles mostly mitigates this issue.

but drawn directly from the census, matched to the target (MC, father, or brother) on the telescope demographics. Figure 2 reports the link counts by Congress for fathers (panel a) and brothers (panel b). Appendix Table A.6 shows the linked-brother subsample is balanced against the full MC universe on MC outcomes and control-matching coverage.

[Figure 1 about here.]

[Figure 2 about here.]

2.5 Measuring Socioeconomic Status

The Federal Census asked different questions in different years, and no single (perfect) measure of socioeconomic status spans the full century covered by our sample of MCs and their families. Instead, we draw on several different measures, of roughly two types. The first are direct: variables recorded for the individual or the household directly during enumeration. These include earnings, years of schooling, or wealth, in the censuses where such values were recorded. The second type are indirect scores constructed in the economic history literature, which try to assign a quantitative score to a person based on their occupation and (in some cases) additional census-recorded demographics. We apply these scores uniformly across every census wave from 1850 to 1950. We report results with both types of measures because they are complementary: while the direct measures are more intuitive, they exist only in the censuses they were recorded in and make long run comparisons difficult. But the occupation-based scores cover the full sample period, though at the cost of imputation.

Direct Measures

We are able to draw many of our most important measures of economic status directly from the census enumeration forms. These include wages, years of schooling, declared wealth, and live-in servants. Wage earnings and years of schooling appear only in 1940 (and in the 1-in-6 sample line in 1950).[19] The census asked respondents about their wealth in different forms in different decades: real estate value in 1850, 1860, and 1870; personal property in 1860 and 1870; and home value (among home owners) in 1930 and 1940. For home value we follow Kalsi and Ward (2025), attributing the household's reported value to the head and spouse of owner-occupied units and ranking within five-year age bins. We also count the live-in domestic servants in each household, using Kalsi and Ward's threshold of two or more as a marker of elite status that occupational scores miss. These measures of live-in domestics are available continuously from 1880 to 1950.[20]

[19]The 1940 Census top-coded wage earnings at $5,000. 12.6% of future members of Congress in our sample have top-coded earnings, compared to less than 1% of the general population. In 1950, the census recorded information on total and non-wage earnings, in addition to wage or labor earnings, as it had in 1940.

[20]Airoldi and Moser (2024) use 1+ live-in servant as their measure of elite household status. Our results are robust to this definition as well.

While none of the direct measures span our entire sample, from 1850 to 1950, each is useful and a clear signal of economic status. But to construct consistent measures of socio-economic status over our entire analysis period, we turn to occupation-based proxy measures.

Occupation-Based Measures

The simplest and most commonly used historical measure of occupation-based status is the IPUMS occupation score, which assigns each occupation the median earnings of its workers in the 1950 Census (Sobek 1995). It is useful but blunt: occscores vary only with occupation, not across cohorts, places, sex, race, or anything else, so it does a poor job distinguishing among elite occupations and fails to recover well-known disparities across race or sex or region (see Saavedra and Twinam 2020).

Recent work in economic history and related fields has produced finer occupation-based measures. The Song score (Song et al. 2020) ranks occupations by their members' literacy and educational attainment within birth cohort, scaled 0 to 100. Unlike the IPUMS occscore, the Song score varies over the long 19th and early 20th century. When occupations like clerk rise in status as the economy shifts, the Song scores mirror these changes. A further refinement is the adjusted Song score (Ward 2023), which layers race-by-region adjustments on top of the Song score. A lawyer in Mississippi and a lawyer in Massachusetts do not occupy the same rung of the SES hierarchies, nor would a Black and white lawyer in Mississippi be likely to have comparable social or economic status.

We also use three occupation-based measures of status that predict earnings or wealth from occupation and other census-recorded covariates. First, LIDO (Saavedra and Twinam 2020) is a LASSO-predicted income score trained on 1950 earnings, with industry-specific LASSO regressions on occupation, age, race, sex, and state of residence. The LIDO score is denominated in hundreds of 1950 dollars. Second, the Collins–Wanamaker 1940 income score (Collins and Wanamaker 2022) averages 1940 wage income within cells defined by occupation, race, and region (or division).[21] Because self-employed workers report no wages in 1940, Collins and Wanamaker (2022) impute the wage-worker mean in their (occupation, race, region) cell, scaled by a self-employed/wage-earner earnings ratio estimated from 1960. Farmers receive a special adjustment: the farm-laborer wage in the same race x division cell serves as the base, then is multiplied by a 1960 farmer/farm-laborer earnings ratio that varies by race, region (or southern division), and home ownership. Farmers, farm managers, and farm laborers have their scores adjusted upward to account for in-kind pay. Third, the Collins–Zimran 1870 wealth score (Collins and Zimran 2023) parallels the Collins–Wanamaker idea but for wealth and with simplifications. Specifically, Collins–Zimran use the wealth question on the 1870 census, and compute occupational-cell means of total wealth (real plus personal property) for white men aged 30–65 in the labor force, with cells defined by occupation x South-vs-non-South or occupation x state. We focus on the occupation x state variant for precision.

Together, we have six occupation-based measures and predictions of socio-economic status that complement the direct measures of economic status. With these measures, we can compare the

[21] For cells with few observations at the occupation by race by region level, Collins and Wanamaker (2022) fill by falling back from division to region to nation and from detailed to 1-digit occupation.

relative status of members of Congress and their families across cohorts, and compare them to the general population.

3 How Economically Unrepresentative are Members of Congress?

How do members of Congress compare to similarly aged Americans before they enter Congress? We focus in this section on future members of Congress aged 18 to 40, observed only in census years before the future MC was elected to Congress.[22] We study MCs in 1940, in the 19th century, and then across our entire sample, shifting between measures of SES available in different censuses and periods. We find that across all measures and eras, the future MCs we observe are high status. They earn more, finish more schooling, own more wealth, and work in higher-ranking occupations, both in general and compared to tightly matched control sets.

3.1 Earnings and Education in 1940

The 1940 census is the first federal census to record individual (labor) earnings and years of education, and it is the cleanest single cross-section for studying the labor-market position of future MCs.[23] Figure 3 compares the distributions of wage earnings (panel a) and years of education (panel b) for future MCs to those of similarly aged men in the general population.[24] Future MCs sit far to the right of the general distribution in both panels. More than 12% of future MCs are in the top-coded ($5000+) earnings bin, compared to less than 1% of the general population. Nearly 60% of future MCs report on the census enumeration that they have 16 or 17+ years of schooling, against just 5% of the general population.

[Figure 3 about here.]

Differences in raw distributions reflect real gaps between future MCs and the rest of the population. But they could also reflect differences in observables (age distribution within 18-40, race, and geography). Most of the gap, we find, is MC selection. We examine this in Figure 4 where we progressively tighten the comparison group, matching future MCs to non-MCs on an increasing set of covariates: all men, then age, then race and foreign-born status, then place (census region, then state, then county), then urban/rural/farm status, and then finally matching on birthplace. The future-MC–vs-control gap shrinks as we condition on more covariates but remains substantial in every specification. Even in our tightest comparison set, a group we will return to later for more

[22] Our approach gives us a clean view of the labor-market position and educational attainment of future MCs before they enter Congress, without any contamination from their time in office. We know from Ho (2025), Querubín and Snyder (2013), Palmer and Schneer (2019), and Palmer and Schneer (2016) that serving in Congress can have economic benefits for MCs and their families and we want to capture outcomes before that channel is possible.

[23] The results on MC selection based on 1940 data replicate many of the main findings from our previous paper, Thompson et al. (2019), but with a larger sample size and links made directly via FamilySearch.

[24] Specifically, all men aged 18-40 in 1940.

analysis, future MCs have several more years of schooling, nearly twice the earnings, and are vastly more likely to have top-coded earnings than the control set.[25]

[Figure 4 about here.]

Within an occupation, future MCs stand out. A career in law is a common pathway to political office, and 1940 was no different: roughly one in four future MCs is a lawyer at the time we first observe him in the census between the ages of 18 and 40. In Table 1, we restrict to lawyers (both among MCs and among their tightest matched set of controls). Even within this high-status occupation, future MCs earn more and are more educated than otherwise comparable controls. We estimate that future MC lawyers earn about $650 more annually and are 20 points more likely to be top-coded (among lawyers with positive wage earnings and no other income). They also report about half a year more of education and are 7 points more likely to have a college degree than matched non-MC lawyers.

[Table 1 about here.]

3.2 Wealth in the 19th Century

The pre-1940 censuses do not collect earnings or education, but during the 19th century, the census did ask respondents to report the value of real estate wealth (1850, 1860, 1870) and personal property wealth (1860, 1870). Figure 5 shows that future MCs in this era had substantially higher real estate (panel a) and personal-property (panel b) wealth than matched controls, at all levels of matching. The magnitudes are staggering.[26] Future MCs had more than $6000 in real estate wealth and nearly that much in personal property wealth, compared to just about $2000 for the tightest of matched controls.[27]

[Figure 5 about here.]

[25] Appendix Figure E.4 shows parallel telescopes on the binary education margins—college completion, college attendance, and high-school completion—with the same matched-control pattern.

[26] Wade Hampton III, born in 1818 just before our sample starts, is illustrative. His grandfather, Wade Hampton I, was among the wealthiest landholders in the antebellum South. By the eve of the war Hampton himself owned some 900 enslaved people and 12,000 acres across cotton plantations in South Carolina and Mississippi. He had inherited the fortune—along with roughly $400,000 of his father's debts—in 1858. The war undid most of it. Hampton's lands lay in the path of Sherman's march, and by the time he filed for bankruptcy in 1868 he owed more than $1 million against $235,500 in real estate; his household goods were auctioned in Columbia for just $118.65. Still, he went on to serve as governor of South Carolina and then in the U.S. Senate from 1879 to 1891 (Andrew 2008, pp. 315–16).

[27] All dollar figures are nominal, recorded by the enumerator in the census year and pooled across 1850, 1860, and 1870 without deflation. The modal census respondent in the 19th century had zero wealth; the share with zero wealth falls as we tighten our comparison groups in on the MCs. We report means including the zeros in Figure 5. In Appendix C, we present parallel figures for real estate and personal property wealth that plot the extensive margin (share with non-zero wealth) and the intensive margin (mean wealth excluding zeros). These show a similar pattern of large MC–control gaps.

3.3 Occupation-Based Measures Across All Censuses

Earnings, education, and wealth give a sharp picture of selection, but each appears in only a handful of census waves. To cover the full 1850–1950 period, we turn to occupation-based measures: occupation is the one economic variable recorded in every wave.

Figure 6 extends the matched comparison to four occupation-based measures across every census wave from 1850 to 1950. The adjusted Song score (panel a) ranks occupations on a 0–100 status scale within birth cohort and adjusts for race and region; in the tightest comparison set, we estimate that future MCs score about 20.6 points higher than their matched controls. LIDO (panel b), a LASSO-predicted income score trained on 1950 earnings, places future MCs roughly $750 (1950 $) above their matched controls. The Collins–Wanamaker 1940 occupational income score (panel c), built from mean 1940 wages in occupation-by-race-by-region cells, shows a future-MC–control gap of nearly $1000. The Collins–Zimran 1870 wealth score at the state level (panel d), built from mean total wealth in occupation-by-state cells, puts future MCs more than $5000 above controls. As in the 1940 and 19th-century results, the gap narrows under tighter matching but never closes, in any wave or on any measure.

[Figure 6 about here.]

Across the 1940 cross-section, the 19th-century wealth censuses, and the pooled occupation-based measures from 1850 to 1950, the picture is the same: future MCs sit well above their tightly matched controls on every dimension we can observe, in every wave. The pattern raises an obvious question: does the advantage come from the families future MCs were born into, or from the careers they built on their own before election to Congress? The next section turns to the family side of the story.

4 Family Background of Future MCs

Before they serve in Congress, MCs are high status and economically successful. In 1940, they are better educated and have higher earnings, compared to both the general population and a tight control group. In the 19th century, they are wealthier. Throughout our period, they are higher in occupational status, measured with a variety of occupation-based measures. MCs could be high-ability people we want selected into Congress, as in the Swedish case studied by Dal Bó et al. (2017). But their early-life status could just reflect their families' status. Do future MCs attain high status because of their own ability, or because they were born into the most fortunate families in America? To answer that question, we need to observe the families from which legislators are drawn.

In this section, we draw our data on MC fathers from the census and compare them to tightly matched control sets. Once again, we match on sex, then age, then race and foreign-born status, then place (census region, then state, then county), then urban/rural/farm status, and then finally on birthplace.[28] We document large gaps between the fathers of future MCs and their control

[28]We build the control sets as increasingly tight demographic matches to the MC fathers. An alternative approach

matches. These gaps appear across all measures of economic status—direct measures like education and wealth, and indirect measures based on occupation scores—and across every cohort.

Then, after documenting the large family-background selection of future MCs, we trace this selection across cohorts and across Congresses. Has family-background selection become more or less extreme over time? Has it moved with larger changes in inequality in the US economy? Has it shrunk as the political system has become more open? Has it increased with the costs of running for office? Our long panel of MC fathers linked to the census enables us to follow the patterns over more than a century of American political and economic history.

4.1 MC Fathers in 1940 and the 19th Century

We begin with the 1940 census, where we can measure father's earnings and education directly. Figure 7 compares the distributions of wage earnings (panel a) and years of education (panel b) for MC fathers to those of all working-age men in 1940. MC fathers are concentrated at the top of both distributions, just as their MC children were. More than 15% of MC fathers earned more than $5,000 in 1940, compared to just about 1% of the general population. About a quarter of MC fathers had a college degree, a large share compared to cohorts where only about 5% of the population had a college degree and one-third had only a common school education. These conclusions remain even when we compare MC fathers to a tight control group of fathers, matched on increasingly similar demographics (see Appendix Figures E.1 and E.2 for comparisons on earnings, top-coded earnings, and education).

[Figure 7 about here.]

In Table 2, we compare the MC fathers to all fathers in the 1940 census on a variety of outcomes. MC fathers earned roughly twice as much wage income as the average father (mean or median), were more than five times as likely to hold a college degree, and were nearly 10 times as likely to have top-coded earnings.[29] Some 3.4% of MC fathers had at least two household servants, compared to just 0.1% of the rest of the population, and they lived in homes more than twice as valuable.

[Table 2 about here.]

MC fathers in the 19th century were also a highly elite group. We use the wealth questions on the 1850, 1860, and 1870 censuses to measure this directly. Figure 8 shows that MC fathers had substantially higher real estate (panel a) and personal property (panel b) wealth than matched controls, at all levels of matching. MC fathers held more than $8,000 in real estate and personal

would be to start with the demographically matched controls for the MCs and link them backwards in the census to their fathers. Our results are robust to this, but we think attaching the controls to the MC fathers directly is more straightforward and takes better advantage of our FamilySearch linked data.

[29] Appendix Table E.3 reports occupation-based status scores (IPUMS Occupation Score, Song Score, Adjusted Song Score) for the same MC-fathers and all-fathers samples.

property wealth, compared to just about $4,000 for the tightest of matched controls.[30] Between a quarter and a third of MC fathers were in the top 10% of all fathers in wealth, depending on the exact measure. Moreover, about 1 in 5 were in the top 5% and 1 in 13 were in the top 1%.

[Figure 8 about here.]

4.2 Selection Across Cohorts

Has family-background selection into Congress grown or shrunk across U.S. history? Several forces should have flattened the gap. The franchise expanded over our period—to non-property-holders, to Black men during Reconstruction, to women in 1920, to all African Americans after the Voting Rights Act, to younger voters in 1971. Institutional reforms made the path to Congress more competitive and less controlled by party elites: the secret ballot spread across the states in the 1890s and 1900s, direct primaries replaced caucus and convention nomination, and the 17th Amendment moved Senate elections out of state legislatures and into the hands of voters. Education opened up too: the high school movement and then the postwar expansion of college access widened the pool of Americans with the credentials required by most political careers (Goldin and Katz 2009; Goldin 1998).

Other forces cut the other way. Twentieth-century income inequality is U-shaped, falling through midcentury and rising after 1980 (Goldin and Margo 1992; Piketty and Saez 2003); if a family's economic position at the top determines its political access, the family-background gap might track that U. Running for Congress has also become more expensive and more complex (Hall 2019), raising the floor on the resources—financial and social—a candidate's family must provide. These forces work against the democratizing ones. The net direction is an empirical question.

Across more than 130 years and more than 80 Congresses, we find that the gap between the fathers of future MCs and tightly matched control fathers has not shrunk; if anything, it has grown. Whatever has opened up about American politics over this period, it has not been the family backgrounds from which legislators are drawn.

No single variable lets us follow MC fathers across the full sample. The wealth questions disappear after 1870; the earnings and educational attainment questions arrive only in 1940. To trace selection continuously, we rely on two sets of measures. First, we use direct measures of elite status that we can piece together across our time period: real estate wealth (recorded in the 1850, 1860, and 1870 censuses) and household servants (recorded in the 1880, 1900-1940 censuses).[31] Second, we use occupation-based measures of economic status that we can construct in every census from 1850 forward: the adjusted Song score, the LIDO score, the Collins–Wanamaker earnings score, and the Collins–Zimran wealth score.

[30] Appendix C.2 decomposes these gaps into the extensive margin (any positive wealth) and the intensive margin (mean conditional on positive). For MC fathers the gap is almost entirely intensive: most fathers held some property, but MC fathers sat in the far right tail of the wealth distribution.

[31] These direct measures of elite status are constructed using the same methodology as in Kalsi and Ward (2025) which studies the persistence of wealth among the top 1% before, during, and after the Gilded Age.

4.2.1 Direct Measures of Elite Status

Real estate wealth and household servants are two direct measures of elite status that are both recorded in the multiple censuses covering our period. Figure 9 brings together these direct top-tail measures. We compare the outcomes on these measures of high-status for the fathers of MCs and their tightest controls, summarizing each cohort with a rate ratio to make clear the differences on these rare markers of extreme wealth. For every MC father we compute the gap between his own high-status indicator and the matched-control rate—the rate at which his tightest controls clear the same threshold—and regress that gap on MC year-of-birth dummies and father-age dummies. We focus on the coefficients on the MC year of birth; these are age-adjusted percentage-point differences. We then divide that difference by the control base rate in the cohort and recenter at one, turning the additive gap into a multiplicative one. A value of two means MC fathers are twice as likely as their controls to clear the threshold, a value of three means they are three times as likely. We plot these ratios on a log axis, so that proportional gaps above and below the null are equally spaced. Panel (a) overlays 1+ and 2+ household-servant counts from the 1880 and 1900–1940 censuses. Panels (b)–(d) present real-estate wealth at the top 1%, top 5%, and top 10% thresholds. The dashed line at one is the null of no difference in relative prevalence of these elite markers between MC fathers and their controls.

[Figure 9 about here.]

We start with the household-servant-based measure of high status because it covers the majority of the MCs in our sample. Multiple servants are a marker of highly elite status as only about 1% of matched-control fathers report two or more. We see in the figure that the MC-father-to-control ratio at the 2+ threshold starts around 3 and grows over time. The 1+ threshold, used by Airoldi and Moser (2024) to study elite family backgrounds among scientists, is more common but still a strong marker of elite status. The ratio between MC fathers and their controls on the 1+ threshold starts above 2 and grows over time. In both cases, the gap between MC fathers and their controls is large and persistent across cohorts, with no sign of narrowing over time.

Real estate wealth tells a very similar story, albeit for only the earlier part of our panel. Recorded in the 1850, 1860, and 1870 censuses, real estate wealth is the only other marker of high status we can observe in more than two censuses. Across all three thresholds, the MC-father rate is higher than the matched-control. At the top 1% (panel b), MC fathers are two to three times as likely as their tightest controls to be at the top of the wealth distribution. At the top 5% (panel c) the ratio is slightly smaller, and at the top 10% (panel d) it settles closer to 1.5–2. None of the three panels shows the gap closing over the half-century we have good coverage for MC birth cohorts.

Put together, the direct measures point in one direction. The family background gap is both large and shows no sign of narrowing across more than a century of MC cohorts. MCs are much more likely to come from the most elite and wealthiest families.

4.2.2 Occupation-Based Measures of Status

The direct measures of elite status are top markers and require us to stitch together two measures to cover our sample. Occupation, in contrast, was recorded in every census we study, from 1850 to 1950, and occupation-based scores translate it into economic status, assigning each occupation a score based on the earnings, wealth, or standing of the people who held it. But occupation scores, however carefully constructed, are not perfect proxies for economic status, so we draw on four: the adjusted Song score, the LIDO score, the Collins–Wanamaker occupational earnings score, and the Collins–Zimran occupational wealth score.

Only part of the true underlying MC-father gap is captured by these occupation-based scores. Much of the difference in economic status that we have seen so far is due to within-occupation differences—invisible to occupation-based measures of status—rather than across-occupation differences.[32] To quantify this precisely, in Table 3 we compare the gaps between the fathers of MCs and their controls, with and without controlling for detailed occupation fixed effects. We look at direct measures of economic status: 1940 earnings, 1930 and 1940 home values, household servants from 1900 to 1930, and the 19th-century wealth questions. For each outcome we regress it on an MC-father dummy with an MC-control pair fixed effect, first alone (column 1) and then with an occupation fixed effect added (column 2). The share of the gap that survives the within-pair occupation FE (column 3) reflects the part an occupation score cannot see, and the implied scaling factor (column 4) is an indication of how much our occupation-based time series will understate the true gap between MC fathers and controls. The implied scaling factor varies across outcomes. In 1940 it is about 2× for wage earnings and roughly 3× for home value. For 19th-century wealth, where most of the gap runs within occupations, the implied scaling runs from 3.9× on 1860 personal property to 6.0× on 1860 real property. Put simply, even carefully constructed occupation scores like the ones we use only capture part of MC-father selection. We suggest that the time-series figures that follow are best read as a lower bound on MC family selection.[33]

[Table 3 about here.]

Within versus between occupation variation aside, we still think that our four occupation score measures provide a valuable time series on MC selection over time. We show our main results for all four scores, focusing on the comparison between MC fathers and the tightest set of control observations, matching on sex, age, race, foreign-born status, county, urban/rural/farm status, and birthplace. For each score, we compute the gap between an MC's father and the mean of his matched control cell and regress that gap on a full set of MC birth-cohort dummies and a full set of father-age-in-census dummies. We restrict to MC fathers observed at ages 30–65 and cluster

[32] A simple example from our data: Willie Mae and Eddie Lewis were sharecroppers, and their third child was future MC and Civil Rights hero John Lewis. James Greer Bankhead was a planter and slaveowner with $45,425 personal wealth in 1860, and his son John Hollis Bankhead served 20 years in the House and another 13 years in the Senate from Alabama. Both Eddie Lewis (1950) and James Greer Bankhead (1860) were recorded as farmers in the census.

[33] Tables E.1 and E.2 run the same within-occupation exercise on future MCs and on their brothers, in place of MC fathers.

standard errors on MC bioguide.[34] The age dummies absorb a mechanical composition shift: across the censuses in which we observe MC fathers, the average father age in our analysis sample drifts with MC cohort even inside the 30–65 window, just based on the census years we can observe between 1850-1950.

[Figure 10 about here.]

Figure 10 plots, for each of our four occupation-based scores, the average gap between MC fathers and their matched control fathers by MC birth cohort. Point size scales with the number of MCs in each cohort. Two patterns hold across the occupation scores.

First, the gap is large and uniformly positive. MC fathers exceed their matched controls by 5 to 10 points on the adjusted Song scale, by \$100–\$300 (1950 \$) on LIDO, by \$200–\$500 on the CW 1940 occupational income score, and by \$1,000–\$2,500 on the CZ 1870 occupational wealth score.

Second, the gap grows over the bulk of the period. Three of the four measures (adjusted Song, LIDO, and CW) roughly double between the 1830s and 1900s birth cohorts; for CZ 1870, it rises from about \$1,250 to about \$2,250. The franchise expanded, education expanded, the political system opened up—and on every one of our four scores, MC fathers' economic position pulled away from that of comparable men, not toward it.[35]

Rather than plot against MC year of birth, we can also organize this graphical history of selection into federal office by the Congresses an MC served in.[36] The cuts are related but distinct because while a single MC comes from only one birth cohort, he may serve across many Congresses. The typical Congress contains MCs born across a wide number of cohorts. Plotting by Congress allows us to ask: at a given moment in American political history, how unrepresentative were the families of the men in the Capitol?

[Figure 11 about here.]

Figure 11 shows the same four scores by Congress, and matches the by-cohort version in spirit. From the 40th Congress (1867) to the 115th Congress (2017), MC fathers exceed their matched controls in every Congress and on every score. Again, we see that the gap rises over time, at least for the first 100 years of our sample. For three of the four scores, we see a peak in the Congresses of the 1960s and 1970s, right around the moment Congress was diversifying on other dimensions. The most recent Congresses show some attenuation on adjusted Song score and CW, while the LIDO and CZ scores show little or no decline. Of course there is selection into our sample for the later years as we can only collect data on MCs after they are deceased (and if they were born before

[34] Because we see some MC fathers multiple times between the ages of 30 and 65 in different censuses, we reweight each observation so that each MC has weights totalling up to 1.

[35] The trajectories are not strictly monotonic, but are also not the same across the four measures. We view this non-monotonicity as more speculative, given the difficulties in measuring economic status precisely. But no matter the flaws of our various measures of occupational status, the two broad patterns are clear. There is a gap that rises across most of the panel and never returns to its 19th-century levels.

[36] The by-Congress version of the figures replaces the cohort dummies with congress dummies and lets multi-term MCs contribute one row per term served.

the 1950 census), so we are cautious about the very end of the panel when there are many MCs in Congress who are not in our sample.

Does the persistent selection of MCs from high-status family backgrounds across cohorts and Congresses vary by party, region, or chamber? When we cut the data by political party, we see weak evidence that Republicans are more likely to come from high-status family backgrounds than Democrats (Table E.4, columns 2-3), at least according to three of our four measures of status, though the difference is only statistically significant for the Collins–Zimran wealth scores. The time trends between parties are not dramatic (Figure E.5). A similarly suggestive if weak pattern emerges by region: future MCs from the South appear to come from higher status families according to three of our four measures, but lag far behind on the CZ Wealth Score (Table E.4, columns 4-5; Figure E.7). The differences across chamber are the starkest: Members of the Senate appear to come from higher status families—though only marginally significant in one measure—and the gap is quite stable over our full period (Figure E.6). A fitting result for the chamber once known as the millionaires' club.

Taken together, we see that the gap between MC fathers and their tightly matched control sets is positive in every cohort and every Congress, larger in the middle of the 20th century than at any other point, and—even after the late-period attenuation—substantially larger today than at the start of our window. Our conclusions hold for both direct measures like wealth and household servants, but they also hold for occupation-based measures of economic status, which we can construct continuously across our full period. There is no evidence that, despite the many forces that could have flattened the gap, family-background selection into Congress has shrunk over time.

5 Comparing Future MCs to Their Brothers

MCs are clearly drawn disproportionately from economically elite households. We have also seen that MCs are high achievers themselves before they even reach Congress, in terms of earnings, education, wealth, and occupational status. How much of the MC's status before Congress simply reflects the privilege of his family?

There are two possibilities here. MCs could be the typical product of an elite household. But MCs are not the only members of those households we observe. By comparing the future MCs to their brothers (and then comparing both to their tightly matched control sets), we can examine selection into Congress and parse between- and within-family selection. If MCs are systematically more successful than their own brothers, there is a second margin: within an already-privileged family, the eventual MC is the most successful sibling. That still might suggest that politics and high office are only open to a select few, but at least there is some positive selection within the small eligible group.[37]

[37]Of course, there are other possibilities. One is that the non-MC is the higher ability brother, but given that the returns to ability outside of politics might be higher, he opts for a different career. Thus, the MC is the runt of the litter who chooses politics, perhaps because the family business is not open to him. We will not find evidence for this story. Another, related story, is that of the brother as financier (or father figure) to the MC brother. Before the Civil

We analyze brothers in two nested steps. First, following Dal Bó et al. (2017), we compare future MCs to their brothers directly. We find MCs outpace their siblings across nearly all outcomes before they reach Congress. Second, we compare both MCs and their brothers to their sets of tightly-matched control populations. This allows us to benchmark the MC's premium against his brother's and decompose the total premium into a family-shared piece and an individual MC piece.

5.1 MC vs. Brother

To estimate the within-family premium, we draw on our data on the brothers of future MCs. Recall that for each MC we located on FamilySearch, we have also collected his siblings (brothers) and added links for the brother to the Federal censuses.[38] This allows us to observe both the MC and his brother as young adults and to compare their economic status before the MC enters Congress.

For brother i in family j, observed in a census wave before j's future MC entered Congress, we estimate

$$Y_{ij} = \beta\,\mathrm{MC}_{ij} + \theta_{\text{age}} + \alpha_{\text{family}} + \delta_{\text{year}} + \varepsilon_{ij}, \tag{1}$$

where Y_{ij} is an outcome (earnings, education, wealth, or occupation score), MC_{ij} indicates that brother i goes on to serve in Congress, θ_{age} is a set of age fixed effects, α_{family} is a family (bioguide) fixed effect, and δ_{year} is a census-year fixed effect. The sample is restricted to men aged 18–40 with a reported occupation, observed before the future MC entered Congress. We cluster standard errors by family. The family fixed effect absorbs all sources of variation in Y that are shared within the household: parental earnings, education, occupation, location, race, ethnicity, and any unobserved cultural or genetic factors that families share. The coefficient β identifies positive selection on individual outcomes within family.

We start by comparing MCs to their brothers in the 1940 Census. Table 4 reports estimates of β in 1940 for our main outcomes. We see that future MCs out-earn their brothers by a substantial margin, have more years of schooling, and are more likely to have completed college.

[Table 4 about here.]

But we have many more measures of economic status than just those in 1940. Pooling brothers observed in every census wave from 1850 through 1950, Table 5 estimates Equation (1) on the long panel. Across decades, future MCs are positively selected on whatever outcomes we can measure: wealth in the 19th century, occupation score in every wave, and earnings and education in 1940. The brother gap is large and stable across cohorts.

[Table 5 about here.]

War, Jefferson Davis' political career was funded by his much older and wealthier brother Joseph. During the Gilded Age, Pennsylvania Senator Boies Penrose drew on his younger brother Spencer's mining fortune to bankroll his career and political machine, though the Penrose family was a prominent Old Philadelphia dynasty.

[38] Siblings are paired on FamilySearch when individual profiles are created, usually because two siblings co-occur in a historical record; these records are often census manuscripts from childhood, but could also be state censuses, birth records, or other documents. Thus, we are able to find many more siblings than just those linked via the complete-count censuses.

We can see clearly from Tables 4 and 5 that the within-family gap between the future MC and his brother is large. But it is also smaller than the raw gap between the future MC and his matched controls from the general population we estimated in Section 3. This suggests that Congress's unrepresentativeness reflects selection both within and across families. We will quantify the relative importance of these two margins in the next subsection.

5.2 How Much Is the MC, How Much Is the Family?

For the brothers, just as for the future MCs, we build a tightly-matched control group, again matching on sex, age, race, foreign-born status, county of residence, and birthplace in the same census year. Again, we restrict our sample of brothers to those aged 18 to 40 and observed in years before the MC brother's first Congress.[39] We also restrict to brothers who are observed in the same census year as the MC.

In Table 6, we compare MCs to their own controls (cols (1) and (2)) and MC brothers to their controls (cols (3) and (4)). We present three measures of socio-economic status: Panel A is adjusted Song score, Panel B is Collins–Wanamaker 1940 score, and Panel C is Collins–Zimran 1870 wealth score. We use all of our data in odd-numbered columns and restrict to the years we observe both MCs and their brothers in even columns. We then stack the MC- and brother-comparisons in cols (5) and (6). Column (6) is the sharpest within-family test. We include an MC family fixed effect, so the MC coefficient is identified entirely off variation between the MC and his brother in the same census year. Standard errors are clustered at the bioguide-level.

[Table 6 about here.]

We see that both MCs and their brothers are higher status than their tight control sets across our measures of socio-economic status. However, the MCs stand out relative to their control groups by much more than their brothers do. This is evident when we stack the data: we estimate a large within-family premium, with the MC-control gap roughly twice the brother-control gap once family fixed effects are included.

In Table 7, we present the column (6) specification results across all the different outcomes on which we can compare MCs and their brothers (and their controls), spanning the mid-19th to mid-20th centuries.[40] By dividing the coefficient on MC brothers by the coefficient on the MC, we can calculate the family share of the outcome. If this ratio is closer to zero (the MC coefficient much

[39] We only study MCs and their brothers before Congress, given the possibility that serving in Congress benefits an MC or their family, as documented by Ho (2025) for MC's children and neighbors and Querubín and Snyder (2013); Palmer and Schneer (2016, 2019) for MCs themselves.

[40] We feature four of the six full-period occupation-based proxies introduced in Section 2: the adjusted Song occupation score, the Saavedra–Twinam LIDO score, the Collins–Wanamaker 1940 score, and the state-based Collins–Zimran 1870 wealth score. We set aside the IPUMS occupation score, the unadjusted Song score, and the region-based Collins–Zimran score. Three direct measures appear only in 1940: wage income (level and log) and years of schooling. Three are wealth measures available only in the 1850, 1860, and 1870 censuses: real property, personal property, and total wealth. Finally, we include measures of elite status from Kalsi and Ward (2025) based on home value in 1930-1940 (coded in $1000s, with non-owners entering at $0 following Kalsi and Ward, in contrast to the owner-conditional mean home value reported in Table 2) and the number of servants in a household.

larger), that suggests that something about the future MC himself, and not his family, explains his selection. If the ratio is closer to one (the MC coefficient and the brother coefficient of similar size), that suggests family background plays a large role. We find in Table 7 that the family share is generally between roughly 0.4 and 0.5 for the continuous SES score proxies. This suggests that about half of the MC selection is really MC family selection.[41] These family shares are also quite stable over time, as we show in Figure E.3 where we split the sample by census year.

[Table 7 about here.]

5.3 Interpretation

Our analysis suggests that MC selection runs on two margins. The first is family-level: MCs are drawn disproportionately from families that are economically elite. But the second is individual-level: even within that elite, the eventual MC is much more economically successful than his own brother on most measures of status. The political system is able to select, at least in part, for candidates with successful records, and those records are not just a reflection of family background. But still, the family selection is large and the path to Congress is not open to all. The typical MC is the most successful son of a successful family.[42]

6 Effects of Institutional Change on Candidate Selection

We have shown that the family backgrounds of MCs have been stable over time, with a persistent overrepresentation of the children of elite fathers. Is this stability driven by features of the American political system? Have institutional reforms that change the rules of the game also changed who reaches Congress? In this section, we use our long panel of MCs, their parents, and their matched controls to test whether four major institutional reforms of the Progressive Era affected the family backgrounds of newly elected MCs.

The heart of our sample coincides with one of the most active periods of institutional reform in American political history. Between roughly 1880 and 1930, the United States adopted the direct election of senators, women's suffrage, the Australian (secret) ballot, and the direct primary. Each reform was designed, at least in part, to make politics more open and more responsive to ordinary voters.[43] Did these reforms also change who reached Congress?

[41] The decomposition is identified off MCs with at least one within-window linked brother. Appendix Table A.6 shows this subsample is well-balanced against the full MC universe on MC outcomes and tightest-control coverage, with father status only modestly lower (a few percent across the status measures).

[42] Family fixed effects absorb only what brothers share, so the within-family premium could in principle include birth-order, age-spacing, or differential parental-investment channels. However, existing estimates of family-composition effects on education, including in our historical US setting, are an order of magnitude smaller than the family shares we report (Tan 2019; Black et al. 2005).

[43] A parallel wave of Progressive-Era reforms reshaped municipal government; Carreri et al. (2023) document that adoption of the council-manager form and nonpartisan elections in U.S. cities raised local spending and shifted policy.

6.1 Institutional Reforms Background

The 19th Amendment: Women's Suffrage Women's suffrage roughly doubled the pool of eligible voters. Some states had granted it before 1920; the 19th Amendment nationalized it, barring states from denying the vote on the basis of sex. Formal enfranchisement was not the same as effective access, however: poll taxes, literacy tests, and intimidation continued to keep many women—Black women in the South above all—from the polls. Our treatment marks the year a state first removed sex as a legal barrier. Past work, mostly exploiting state and year variation in the expansion of the women's vote, has shown women's suffrage increased state government spending (Lott, Jr. and Kenny 1999), particularly on public health spending (Miller 2008), education spending (Carruthers and Wanamaker 2015; Kose et al. 2021), and other dimensions (Morgan-Collins 2021). Politicians clearly believed that women voters had distinct preferences from their male counterparts, at least initially (Moehling and Thomasson 2012). Newly enfranchised women voters may have brought distinct preferences over candidate backgrounds, including for or against elite families.[44]

The Secret Ballot Before the secret ballot reform, voters cast ballots printed and distributed by parties. Each party's ticket was often a distinctive color. This made vote choice observable to families, neighbors, bosses, and employers. Massachusetts was the first state to adopt the Australian (secret) ballot in 1888; the rest of the country followed over the next several decades, with South Carolina the last to adopt in 1950 (Lott, Jr. and Kenny 1999). The secret ballot reduced turnout, though scholars debate how much of the drop reflects fewer corrupt ballots versus the loss of intimidated but real voters (Cox and Kousser 1981; Heckelman 1995, 2000; Reed 2014; Fergusson et al. 2023). Earlier work argued that the secret ballot led to the emergence of candidate-centered politics; Moskowitz and Rogowski (2024) finds those effects on candidate or politician actions are mostly null with cleaner econometric specifications. We ask whether a reform intended to reduce corruption and eliminate the market for votes (Dal Bó 2007) also changed the kinds of candidates who won office.

The Direct Primary Historically, party nominations were controlled by party conventions dominated by bosses and delegates. The direct primary, adopted by most states by 1915, transferred nomination power to rank-and-file party voters.[45] The direct primary was designed to make politics more democratic and responsive to ordinary voters, and may have changed the family backgrounds of candidates who survived the nomination process to win office. Ansolabehere et al. (2010) show the direct primary did increase competition in American elections, while Hirano and Snyder Jr (2019) document positive effects on candidate quality, at least for open seats in areas with one dominant party.

[44] Even if female and male voters did not differ in their preferences over candidate family status or background, these aspects could vary systematically on dimensions the sex-based voting blocs did disagree about, like support for public spending, public health, or public education.

[45] In the South, the direct primary was paired with the white primary, which formally disenfranchised Black voters.

The 17th Amendment: Direct Election of Senators Our fourth reform is the 17th Amendment, ratified in 1913, which replaced state-legislature appointment of senators with direct popular elections. Gailmard and Jenkins (2009) show that senators from moderate states became less ideologically extreme after the direct election and that within-state delegations grew more ideologically diverse. The counterfactual House members (always directly elected) did not change in either direction. Reformers alleged that the appointment system was corrupt and favored monied special interests (Appendix Figure D.1 shows a contemporary cartoon from *Puck*). If the appointment system also favored candidates with elite ties, direct election might have opened the Senate to a more representative pool of candidates.

6.2 Effect of Reforms on MC Selection

To assess the effects of these Progressive-Era institutional reforms, we estimate a triple-difference specification on the panel of MCs and their matched controls. The outcomes are our measures of the MC's father's status. We present adjusted Song score, LIDO, Collins–Wanamaker income score, and Collins–Zimran wealth score in the baseline and the other measures in the appendix.

We study the three reforms with state and year variation (women's suffrage, secret ballot, and direct primary) first; Appendix Figure D.2 shows the state-level first-adoption year for each. In this case, the three differences are (1) future MC versus matched control, (2) first-election year before versus after the reform, and (3) state. Each matched set inherits an analogous first election year from his paired MC, which we use as a placebo timing. We only include MCs (and their controls) once in the panel, in the year of their first election. We cluster standard errors at the state level and report results for the years 1880–1930.

The reforms overlap in time (and often place), so we present a joint specification that includes all three state and time varying reforms together, controlling for the others when estimating the effect of each one. For person i in state s at first-election year t, we estimate

$$Y_{ist} = \beta_0 \,\mathrm{MC}_i + \sum_{r \in \mathcal{R}} \left[\beta_{1r} \,\mathrm{Post}_{rst} + \beta_{2r} (\mathrm{MC}_i \times \mathrm{Post}_{rst})\right] + \delta_s + \tau_t + X_i'\gamma + \varepsilon_{ist} \tag{2}$$

by OLS, where Y_{ist} is a measure of the MC father's socioeconomic status, MC_i equals one for future MC fathers and zero for their matched controls, and Post_{rst} equals one if state s had adopted reform r by year t. The set $\mathcal{R}$ contains the three state-timed reforms: women's suffrage, the secret ballot, and the direct primary. δ_s and τ_t are state and first-election-year fixed effects; X_i collects baseline controls (father's age and race). Standard errors are clustered by state. The coefficients of interest are $\hat{\beta}_{2r}$: each is a triple difference capturing how the MC–control gap in father status shifts across states that adopted reform r before versus after the MC's first election. Table 9 extends equation (2) to include the 17th Amendment by adding $\beta_{2,17}(\mathrm{MC}_i \times \mathrm{Senate}_i \times \mathrm{Post1914}_t)$; because the amendment was adopted nationally in 1913, its identifying variation is chamber-by-period rather than state-by-period.

[Table 8 about here.]

Table 8 reports estimates that are statistically indistinguishable from zero for all three state-timed reforms. MCs continue to come from higher-status families in every outcome. Women's suffrage point estimates are positive on Song and CZ wealth and negative on LIDO and CW1940, but all standard errors are several times the size of the coefficients. The secret ballot is negative on the score outcomes and positive on CZ wealth; only the CW1940 estimate in column 5 is marginally significant. The direct primary coefficients are a mix of positive and negative estimates, all statistically insignificant with point estimates near zero. MCs are highly selected, and we cannot detect movement in that gap from any of the three Progressive-Era reforms.

Bringing in the 17th Amendment, the joint specification in Table 9 tells a similar story. The 17th is identified differently from the other three reforms: we exploit chamber-by-period rather than state-by-period variation, so the three differences are (i) future MC versus matched control, (ii) first-election year before versus after the reform, and (iii) Senate versus House. The other three reforms retain their state-by-period structure. We cluster standard errors at the state level.

[Table 9 about here.]

Adding the 17th Amendment largely does not change the picture of MC selection. Again, the other three reforms do not appear to affect MC selection. For direct election, we see some weak evidence that selection is reduced: the interaction is negative on every outcome, but statistically significant on only two of the eight columns in Table 9, both when we measure the family-background gap with LIDO. This hints that the 17th Amendment may have lowered the background-status gap among newly elected senators, but the results are not robust enough to support a firm conclusion.[46]

These nulls are less surprising than they might first appear. Whatever their specific design, all four reforms acted at the ballot box, but voters, at least contemporaneously, are not where candidates from less elite backgrounds get screened out (Carnes and Lupu 2016; Carnes 2018; Carnes and Lupu 2023). The constraints that bind are earlier, in party gatekeeping, donor preferences, and the large costs of running (Hall 2019; Cagé and Dewitte 2021).

These four institutional reforms were designed in different ways to open the political system. None appear to have changed the family backgrounds of the men who reached Congress.

7 Consequences of Selection

We have documented persistent socioeconomic selection into Congress across a century of American history. The question is whether it matters. If the family circumstances and individual ability of legislators do not affect what they do once in office, then elite selection is a fact about who runs and who wins, but not about what Congress does. If background does shape behavior in office, then selection is a mechanism, and the social skew this paper documents translates into a skew in policy.

[46] Direct election is identified off the chamber-by-1914 contrast, so any other shock to senators around that date is a potential confounder we cannot rule out from this specification alone. While some states (Oregon, notably) did experiment with direct election before the 17th Amendment, the timing is very compressed and much of the identification even with state by year treatment would come from the change in 1914 with the 17th Amendment.

There are several channels through which legislators' backgrounds could affect what Congress does: they could vote differently on the floor, prioritize different bills, win at different rates, build different careers, and attend to different constituents. While we do not estimate the size of any of these channels directly in this paper, a substantial body of recent work has.

Across many dimensions on which legislators differ from one another, who they are has been shown to affect what they do (Krcmaric et al. 2020). Canon (1999) documents the consequences of race for representation, particularly in majority-minority districts drawn after the Voting Rights Act. Fridkin and Kenney (2014) show that female senators communicate with constituents in systematically different ways from their male counterparts. Working-class legislators cast different roll-call votes and sponsor different bills than their wealthier colleagues, a pattern documented across multiple congressional eras (Carnes and Lupu 2023; Carnes 2012, 2013, 2018). Carreri and Teso (2023) find that members who came of age during a recession take more conservative positions on redistribution than their co-partisans. Members who arrive in Congress with prior political experience vote more reliably with their party than those who do not (Keena and Knight-Finley 2018). Phillips et al. (2024) push this further on the candidate-pool side, showing that the ideological selection of state legislators who run for the U.S. House helps explain congressional polarization. McCrain and O'Connell (2023) show that state legislative service more than doubles a politician's chance of later contesting and winning a congressional seat. Even the composition of a legislator's own family can move policy: Washington (2008) shows that members with daughters cast more liberal votes on women's issues, and McGuirk et al. (2023) find that members with draft-age sons are less likely to vote in favor of conscription. These last two findings of personal circumstances moving a legislator's behavior on a contested national policy are striking examples of how legislator background matters. Taken together, this evidence establishes that who legislators are, and not only which districts they represent or which party they belong to, shapes what Congress does.

Why would a politician's background shape what he does in office at all? Scholars of political biography point to a handful of pathways: a politician's values, his competence, his material interests, and how others perceive him (Krcmaric et al. 2020). Wealth can run through any of these, and a few careers show it pulling in different directions. James Couzens made his fortune as Henry Ford's business manager, and when Ford bought out his stake for some $30 million in 1919 he became one of the richest men in the country. In the Senate he made himself a "one-man bloc," attaching himself to no faction and declaring that if being a Republican meant kowtowing to the president, he did not care to be a Republican (Barnard 1958). His wealth was part of the point: it was hard, his contemporaries conceded, to brand a man that rich a wild-eyed radical, even as he waged a "millionaires' war" against Andrew Mellon's tax cuts for the rich, suggesting that perhaps personal fortune does not dictate policy preferences in office. But the reverse story in Congress is just as easy to find. John C. Calhoun was born into the rich slaveholding planter class, and he spent four decades in Congress defending its interests: he pronounced slavery a "positive good" on the Senate floor in 1837 and devised the doctrine of nullification to spare the slaveholding South the burden of the federal tariff (Niven 1988). There are also cases where wealth buys only the appearance of

independence. John Heinz, heir to the food-processing fortune and one-time richest member of Congress, was widely assumed too rich to be bought, yet he had to raise campaign money like any other politician. During his 1976 Senate run, Heinz returned $6,000 in illegal Gulf Oil contributions he had accepted in the House (New York Times 1976).

Do legislators' family origins shape what they do in office? Studying recent Congresses, Grumbach (2015) uses Carnes's coding of legislators' class origins to show that upper-class Democrats raised by working-class parents vote more liberally than those raised by upper-class parents. But Carnes's coding exists only for recent Congresses, leaving open whether the link between family origins and legislative behavior reaches further back or has shifted across cohorts. Linking legislators directly to the complete-count census, as we do, enables the exploration of these relationships historically. Feigenbaum et al. (2025) link legislators to their parents and grandparents and find that MCs with more recent immigrant ancestry cast more pro-immigration roll-call votes and speak about immigration in more positive terms on the floor. The result holds conditional on district, party, and ideology, and is confirmed by a close-election regression-discontinuity design that holds district composition fixed by construction. Immigrant ancestry is only one slice of family-background variation. Chyn et al. (2026) apply the design from Feigenbaum et al. (2025)—name-based imputation of a family characteristic, close elections, and a comparison of the winners' roll-call votes—to family wealth, studying how legislators voted on ten major House tax bills between 1916 and 1940. Districts that narrowly elected the higher-status candidate were significantly less likely to support the higher-tax position. Yet none of this work draws on the continuous, century-long measures of fathers' earnings, education, wealth, and occupational status that we trace directly. The long panel we assemble here is what would let that work begin.

8 Conclusion

In every cohort we observe between 1830 and 1950, members of Congress have come from the most economically fortunate families in America. We asked three questions of this fact. How far above ordinary Americans do future MCs sit? How much of the gap reflects their families rather than themselves? And has the gap moved as American politics and the American economy have changed? The short answers are: very far, about half, and barely. Future MCs are economic outliers on every adult margin we can observe, including earnings and education in 1940 and real-estate and personal-property wealth in the 19th century. We estimate that future MCs' brothers share roughly half of that premium; the other half is the MC himself. Across more than 130 years and more than 80 Congresses, the gap between MC fathers and tightly matched control fathers has not narrowed; it has widened across most of our sample. And of four Progressive-Era reforms designed to open the political system, none robustly changed who reached Congress.

Our findings come with limitations. Our main design is observational: we compare future MCs and their families to tightly matched controls drawn from the complete-count census. The sample is also increasingly selected as we reach more recent Congresses because we only observe deceased

MCs. We do use triple-difference specifications that exploit the staggered state-level adoption of three of the reforms (women's suffrage, the secret ballot, and the direct primary), and identify the fourth, the direct election of senators, from the House–Senate contrast instead. None has a robust effect on selection. The one suggestive result, for the direct election of senators, is also the most weakly identified, resting on a single national change in 1913 rather than staggered state adoption. We are reassured by the consistency of the gap across both our two direct measures of elite family background and our four occupation-based scores. But we know that the occupation-based time series capture only the across-occupation part of MC family selection. The wealth and earnings outcomes in Table 3 show that within-occupation selection is also large, and the occupation scores miss it. Our time-series gaps are best read in tandem—real-estate wealth and servants along with occupation-based scores—and likely as a lower bound. How far below the true gap they sit depends on how much the occupation scores compress.

The long panel we have built opens questions beyond the scope of this paper. The most natural is whether the family-background variation we measure shapes what legislators do once they reach Congress. Feigenbaum et al. (2025) show that one slice of family history—immigrant ancestry—affects congressional actions. An existence proof. The other dimensions of family background we trace through the census are still largely untested as predictors of legislative behavior. A second open question is why the elite persistence we find has lasted. The Progressive-Era reforms we study did little to move it, but other economic or political shocks may have. Political campaigning and the media environment have both changed over the last two centuries, and those changes may have moved who runs and who wins. But the cost of running for office may simply have always been high enough to select candidates from elite backgrounds (Cagé and Dewitte 2021; Hall 2019). A third question is comparative. The U.S. picture we document sits between the inclusive meritocracy of Sweden, where parental backgrounds are roughly representative of the population (Dal Bó et al. 2017), and the persistent upper-class overrepresentation in the United Kingdom (Cagé and Dewitte 2024). The natural next step is to fill in more histories of elite political selection across other democracies.

The system through which Americans choose their legislators has changed over the past century and a half. The families from which those legislators come have not.

References

Abramitzky, Ran, Leah Boustan, Katherine Eriksson, James Feigenbaum, and Santiago Pérez, "Automated linking of historical data," *Journal of Economic Literature*, 2021, *59* (3), 865–918.

_ , Lena Greska, Santiago Pérez, Joseph Price, Carlo Schwarz, and Fabian Waldinger, "Climbing the ivory tower: How socio-economic background shapes academia," Technical Report, National Bureau of Economic Research 2024.

Airoldi, Anna and Petra Moser, "Inequality in Science: Who Becomes a Star?," Technical Report, National Bureau of Economic Research 2024.

Akcigit, Ufuk, John Grigsby, and Tom Nicholas, "The Rise of American Ingenuity: Innovation and Inventors of the Golden Age," 2017. NBER Working Paper 23047.

Ansolabehere, Stephen, John Mark Hansen, Shigeo Hirano, and James M Snyder Jr, "More democracy: The direct primary and competition in US elections," *Studies in American Political Development*, 2010, *24* (2), 190–205.

Bailey, Martha J, Connor Cole, Morgan Henderson, and Catherine Massey, "How well do automated linking methods perform? Lessons from US historical data," *Journal of economic literature*, 2020, *58* (4), 997–1044.

Barnard, Harry, *Independent Man: The Life of Senator James Couzens*, Charles Scribner's Sons, 1958.

Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen, "Who Becomes an Inventor in America? The Importance of Exposure to Innovation*," *The Quarterly Journal of Economics*, 2019, *134* (2), 647–713.

Bellani, Luna, Anselm Hager, and Stephan Maurer, "The Civil War Reduced Slave' Owners Economic Power but Increased Their Political Influence," Technical Report, IZA Network LISER 2026.

Besley, Timothy and Marta Reynal-Querol, "Do Democracies Select More Educated Leaders?," *American Political Science Review*, 2011, *105* (3), 552–566.

_ , Jose G Montalvo, and Marta Reynal-Querol, "Do Educated Leaders Matter?," *The Economic Journal*, 2011, *121* (554), F205–227.

_ , Olle Folke, Torsten Persson, and Johanna Rickne, "Gender Quotas and the Crisis of the Mediocre Man: Theory and Evidence from Sweden," *American Economic Review*, 2017, *107* (8), 2204–42.

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes, "The more the merrier? The effect of family size and birth order on children's education," *The Quarterly Journal of Economics*, 2005, *120* (2), 669–700.

Bó, Ernesto Dal, "Bribing voters," *American Journal of Political Science*, 2007, *51* (4), 789–803.

_ and Frederico Finan, "Progress and Perspectives in the Study of Political Selection," *Annual Review of Economics*, 2018, *10*, 541–575.

_ , Pedro Dal Bó, and Jason Snyder, "Political Dynasties," *The Review of Economic Studies*, 2009, *76* (1), 115–142.

Buckles, Kasey, Adrian Haws, Joseph Price, and Haley E.B. Wilbert, "Breakthroughs in historical record linking using genealogy data: The Census Tree project," *Explorations in Economic History*, 2025, *98*, 101717.

Cagé, Julia and Edgard Dewitte, "It takes money to make MPs: new evidence from 150 years of British campaign spending," 2021.

_ **and** _ , "The Rising Demand for Representation: Lessons from 100 Years of Political Selection in the UK," 2024. Working Paper, Sciences Po Paris.

Canon, David T, *Race, redistricting, and representation: The unintended consequences of black majority districts*, University of Chicago Press, 1999.

Carnes, Nicholas, "Does the numerical underrepresentation of the working class in Congress matter?," *Legislative studies quarterly*, 2012, *37* (1), 5–34.

_ , *White-collar Government: The Hidden Role of Class in Economic Policy Making*, University of Chicago Press, 2013.

_ , *The Cash Ceiling: Why Only the Rich Run for Office–and What We Can Do about It*, Princeton University Press, 2018.

_ **and Noam Lupu**, "What Good is a College Degree? Education and Leader Quality Reconsidered," *Journal of Politics*, 2015, *78* (1), 35–49.

_ **and** _ , "Do voters dislike working-class candidates? Voter biases and the descriptive underrepresentation of the working class," *American Political Science Review*, 2016, *110* (4), 832–844.

_ **and** _ , "The Economic Backgrounds of Politicians," *Annual Review of Political Science*, 2023, *26* (1), 253–270.

Carreri, Maria and Edoardo Teso, "Economic recessions and congressional preferences for redistribution," *Review of Economics and Statistics*, 2023, *105* (3), 723–732.

_ **, Julia Payson, and Daniel M Thompson**, "When progressives took power. The political and economic effects of municipal reform in US Cities," *Working paper*, 2023.

Carruthers, Celeste K and Marianne H Wanamaker, "Municipal housekeeping: The impact of women's suffrage on public education," *Journal of Human Resources*, 2015, *50* (4), 837–872.

Chetty, Raj, David Deming, and John Friedman, "Diversifying Society's Leaders? The Determinants and Causal Effects of Admission to Highly Selective Private Colleges," 2023. NBER Working Paper 31492.

Chyn, Eric, Katherine Cohen, Kareem Haggag, and Bryan A Stuart, "The Persistence of Power: How Family Origins Shape Political Representation and Policy," Working Paper 35180, National Bureau of Economic Research May 2026.

Collins, William J and Ariell Zimran, "Working their way up? US immigrants' changing labor market assimilation in the age of mass migration," *American Economic Journal: Applied Economics*, 2023, *15* (3), 238–269.

_ **and Marianne H Wanamaker**, "African American intergenerational economic mobility since 1880," *American Economic Journal: Applied Economics*, 2022, *14* (3), 84–117.

Cox, Gary W and J Morgan Kousser, "Turnout and rural corruption: New York as a test case," *American Journal of Political Science*, 1981, pp. 646–663.

Dahlgaard, Jens Olav and Rasmus T Pedersen, "Inclusive meritocracy: ability and descriptive representation among Danish politicians," *Political Science Research and Methods*, 2025, *13* (3), 567–590.

Dal Bó, Ernesto, Frederico Finan, Olle Folke, Torsten Persson, and Johanna Rickne, "Who Becomes A Politician?," *The Quarterly Journal of Economics*, 2017, *132* (4), 1877–1914.

Enamorado, Ted, Benjamin Fifield, and Kosuke Imai, "Using a probabilistic model to assist merging of large-scale administrative records," *American Political Science Review*, 2019, *113* (2), 353–371.

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

Feigenbaum, James, Maxwell Palmer, and Benjamin Schneer, ""Descended from Immigrants and Revolutionists": How Family History Shapes Immigration Policy Making," *The Quarterly Journal of Economics*, 2025, *140* (3), 2381–2457.

Fergusson, Leopoldo, Juan Felipe Riaño, and B. K. Song, "Media, Secret Ballot, and Democratization in the US," *Journal of Historical Political Economy*, November 2023, *3* (3), 165–297.

Ferrie, Joseph P., "A new sample of males linked from the public use microdata sample of the 1850 US Federal Census of Population to the 1860 US Federal Census Manuscript Schedules," *Historical Methods*, 1996, *29* (4), 141–156.

Folke, Olle and Johanna Rickne, "The class ceiling in politics," *American Political Science Review*, 2025, *119* (3), 1375–1392.

Fridkin, Kim and Patrick Kenney, *The changing face of representation: The gender of US senators and constituent communications*, University of Michigan Press, 2014.

Gailmard, Sean and Jeffery A. Jenkins, "Agency Problems, the 17th Amendment, and Representation in the Senate," *American Journal of Political Science*, 2009, *53* (2), 324–342.

Gilens, Martin, *Affluence and Influence: Economic Inequality and Political Power in America*, Princeton University Press, 2012.

Goldin, Claudia, *Understanding the gender gap: An economic history of American women*, Oxford University Press, 1990.

_ , "America's graduation from high school: The evolution and spread of secondary schooling in the twentieth century," *The Journal of Economic History*, 1998, *58* (2), 345–374.

_ **and Lawrence F Katz**, *The Race between Education and Technology*, Harvard University Press, 2009.

_ **and Robert A Margo**, "The Great Compression: The wage structure in the United States at mid-century," *The Quarterly Journal of Economics*, 1992, *107* (1), 1–34.

Grumbach, Jacob M, "Does the American dream matter for members of congress? Social-class backgrounds and roll-call votes," *Political Research Quarterly*, 2015, *68* (2), 306–323.

Gulzar, Saad, "Who Enters Politics and Why?," *Annual Review of Political Science*, 2021, *24*, 253–275.

Hall, Andrew B., *Who Wants to Run? How the Devaluing of Political Office Drives Polarization*, University of Chicago Press, 2019.

_ **, Connor Huff, and Shiro Kuriwaki**, "Wealth, Slave Ownership, and Fighting for the Confederacy: An Empirical Study of the American Civil War," *American Political Science Review*, 2019.

Heckelman, Jac C, "The effect of the secret ballot on voter turnout rates," *Public Choice*, 1995, *82* (1), 107–124.

_ , "Revisiting the relationship between secret ballots and turnout: A new test of two legal-institutional theories," *American Politics Quarterly*, 2000, *28* (2), 194–215.

Helgertz, Jonas, Joseph Price, Jacob Wellington, Kelly J Thompson, Steven Ruggles, and Catherine A. Fitch, "A new strategy for linking U.S. historical censuses: A case study for the IPUMS multigenerational longitudinal panel," *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 2021, *55* (1), 12–29.

Hirano, Shigeo and James M Snyder Jr, *Primary elections in the United States*, Cambridge University Press, 2019.

Ho, Madeleine, "The Role of Political Connections in Employment: Evidence from US Elections 1850–1930," 2025. Job Market Paper, University of California, Davis.

Jokela, Markus, Jaakko Meriläinen, Janne Tukiainen, and Åsa von Schoultz, "Personality traits and cognitive ability in political selection," *Journal of the European Economic Association*, 2025, *23* (5), 1908–1949.

Jr., Rod Andrew, *Wade Hampton: Confederate Warrior to Southern Redeemer* Civil War America, Chapel Hill: University of North Carolina Press, 2008.

Kalsi, Priti and Zachary Ward, "The Gilded Age and Beyond: The Persistence of Elite Wealth in American History," Technical Report, National Bureau of Economic Research 2025.

Keena, Alex and Misty Knight-Finley, "Governed by experience: Political careers and party loyalty in the senate," *Congress & the Presidency*, 2018, *45* (1), 20–40.

Kose, Esra, Elira Kuka, and Na'ama Shenhav, "Women's suffrage and children's education," *American Economic Journal: Economic Policy*, 2021, *13* (3), 374–405.

Krcmaric, Daniel, Stephen C Nelson, and Andrew Roberts, "Studying leaders and elites: The personal biography approach," *Annual Review of Political Science*, 2020, *23* (1), 133–151.

Lott, Jr., John R. and Lawrence W. Kenny, "Did Women's Suffrage Change the Size and Scope of Government?," *Journal of Political Economy*, 1999, *107* (6), 1163–1198.

Mansbridge, Jane, "Should Blacks Represent Blacks and Women Represent Women? A Contingent "yes"," *The Journal of Politics*, 1999, *61* (3), 628–657.

McCrain, Joshua and Stephen D O'Connell, "Experience, institutions, and candidate emergence: the political career returns to state legislative service," *Political Science Research and Methods*, 2023, *11* (4), 763–784.

McGuirk, Eoin F, Nathaniel Hilger, and Nicholas Miller, "No kin in the game: moral hazard and war in the us congress," *Journal of Political Economy*, 2023, *131* (9), 2370–2401.

Meriläinen, Jaakko, "Political selection and economic policy," *The Economic Journal*, 2022, *132* (648), 3020–3046.

Michelman, Valerie, Joseph Price, and Seth D Zimmerman, "Old Boys' Clubs and Upward Mobility Among the Educational Elite," *The Quarterly Journal of Economics*, 2022, *137* (2), 845–909.

Miller, Grant, "Women's Suffrage, Political Responsiveness, and Child Survival in American History," *Quarterly Journal of Economics*, 2008, *123* (3), 1287–1327.

Moehling, Carolyn M and Melissa A Thomasson, "The political economy of saving mothers and babies: The politics of state participation in the Sheppard-Towner program," *The Journal of Economic History*, 2012, *72* (1), 75–103.

Morgan-Collins, Mona, "The electoral impact of newly enfranchised groups: The case of women's suffrage in the United States," *The Journal of Politics*, 2021, *83* (1), 150–165.

Moskowitz, Daniel J and Jon C Rogowski, "Ballot Reform, the Personal Vote, and Political Representation in the United States," *British Journal of Political Science*, 2024, *54* (1), 22–39.

New York Times, "Rep. Heinz Says Illegal Gifts From Gulf Oil Were Accepted Unintentionally," *The New York Times*, February 1976.

Niven, John, *John C. Calhoun and the Price of Union: A Biography*, Louisiana State University Press, 1988.

Novosad, Paul, Sam Asher, Catriona Farquharson, and Eni Iljazi, "Access to opportunity in the sciences: Evidence from the nobel laureates," 2024.

Palmer, Maxwell and Benjamin Schneer, "Capitol Gains: The Returns to Elected Office from Corporate Board Directorships," *Journal of Politics*, 2016, *78* (1), 181–196.

_ and _, "Postpolitical careers: How politicians capitalize on public office," *The Journal of Politics*, 2019, *81* (2), 670–675.

Phillips, Connor Halloran, James M. Snyder Jr., and Andrew B. Hall, "Who runs for congress? a study of state legislators and congressional polarization," *Quarterly Journal of Political Science*, 2024, *19* (1), 1–25.

Piketty, Thomas and Emmanuel Saez, "Income Inequality in the United States, 1913–1998," *The Quarterly Journal of Economics*, 2003, *118* (1), 1–41.

Price, Joseph, Kasey Buckles, Jacob Van Leeuwen, and Isaac Riley, "Combining family history and machine learning to link historical records: The Census Tree data set," *Explorations in Economic History*, 2021, *80*, 101391.

Querubín, Pablo and James M. Snyder Jr., "The Control of Politicians in Normal Times and Times of Crisis: Wealth Accumulation by U.S. Congressmen, 1850-1880," *Quarterly Journal of Political Science*, 2013, *8*, 409–450.

Reed, Daniel C, "Reevaluating the vote market hypothesis: Effects of Australian ballot reform on voter turnout," *Social Science History*, 2014, *38* (3-4), 277–290.

Reeves, Aaron and Sam Friedman, *Born to Rule: The making and remaking of the British elite*, Harvard University Press, 2024.

Ruggles, Steven, Matt A. Nelson, Matthew Sobek, Catherine A. Fitch, Ronald Goeken, J. David Hacker, Evan Roberts, and J. Robert Warren, "IPUMS Ancestry Full Count Data: Version 4.0," 2024.

_ , Sarah Flood, Matthew Sobek, Daniel Backman, Grace Cooper, Julia A. Rivera Drew, Stephanie Richards, Renae Rodgers, Jonathan Schroeder, and Kari Williams, "IPUMS USA: Version 16.0," 2025.

Saavedra, Martin and Tate Twinam, "A machine learning approach to improving occupational income scores," *Explorations in Economic History*, 2020, *75*, 101304.

Sehgal, Neil KR and Ashwini R Sehgal, "Slaveholder ancestry and current net worth of members of the United States Congress," *Plos one*, 2024, *19* (8), e0308351.

Sobek, Matthew, "The comparability of occupations and the generation of income scores," *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 1995, *28* (1), 47–51.

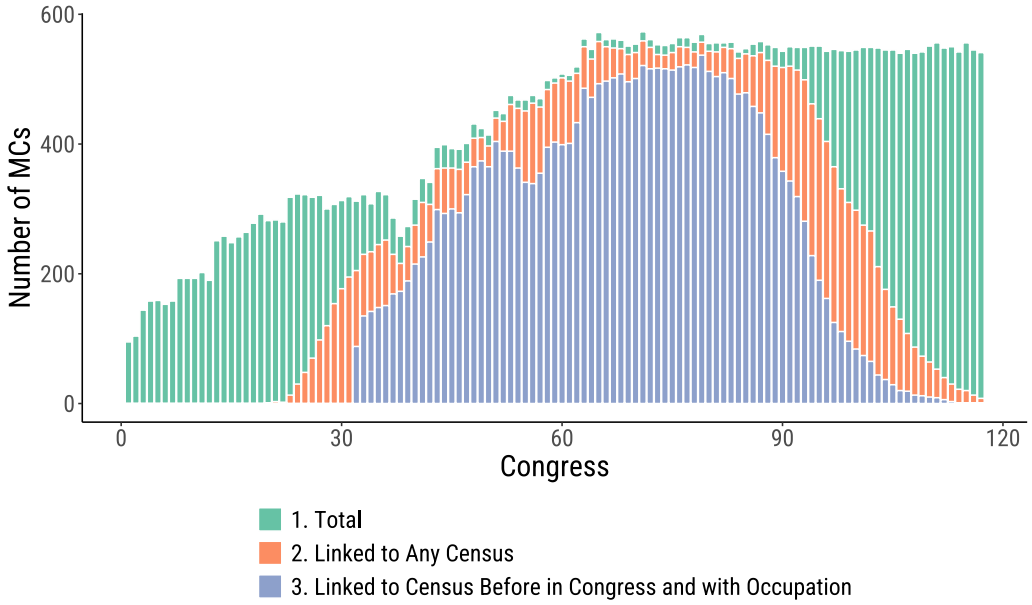
Song, Xi, Catherine G. Massey, Karen A. Rolf, Joseph P. Ferrie, Jonathan L. Rothbaum, and Yu Xie, "Long-term decline in intergenerational mobility in the United States since the 1850s," *Proceedings of the National Academy of Sciences*, 2020, *117* (1), 251–258.

Tan, Hui Ren, "More is less?: The impact of family size on education outcomes in the United States, 1850–1940," *Journal of Human Resources*, 2019, *54* (4), 1154–1181.

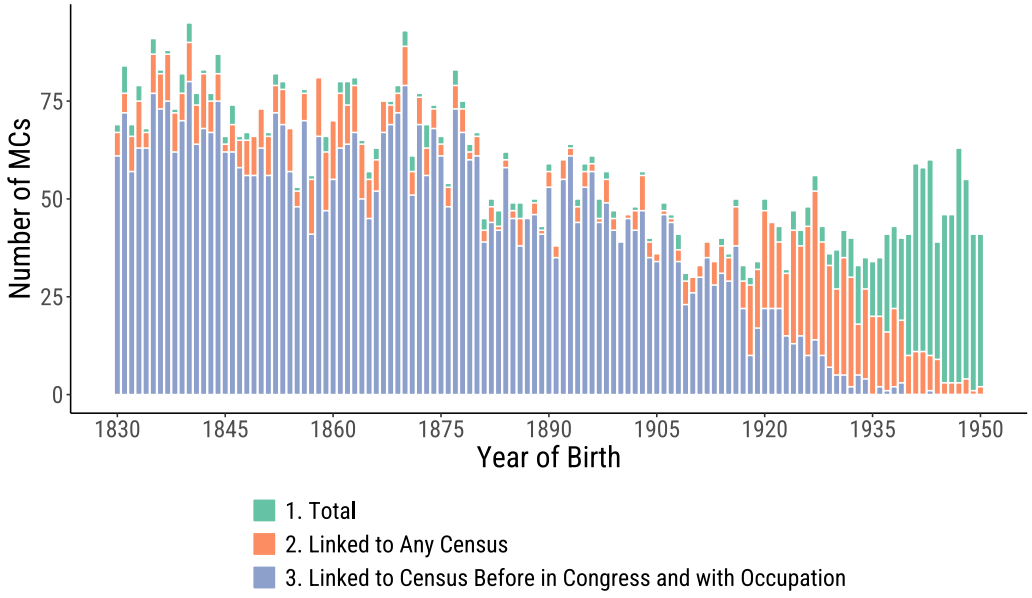
Thompson, Daniel M, James J Feigenbaum, Andrew B Hall, and Jesse Yoder, "Who becomes a member of Congress? Evidence from de-anonymized census data," Technical Report, National Bureau of Economic Research 2019.

Ward, Zachary, "Intergenerational mobility in American history: Accounting for race and measurement error," *American Economic Review*, 2023, *113* (12), 3213–3248.

Washington, Ebonya L, "Female socialization: How daughters affect their legislator fathers' voting on women's issues," *American Economic Review*, 2008, *98* (1), 311–332.

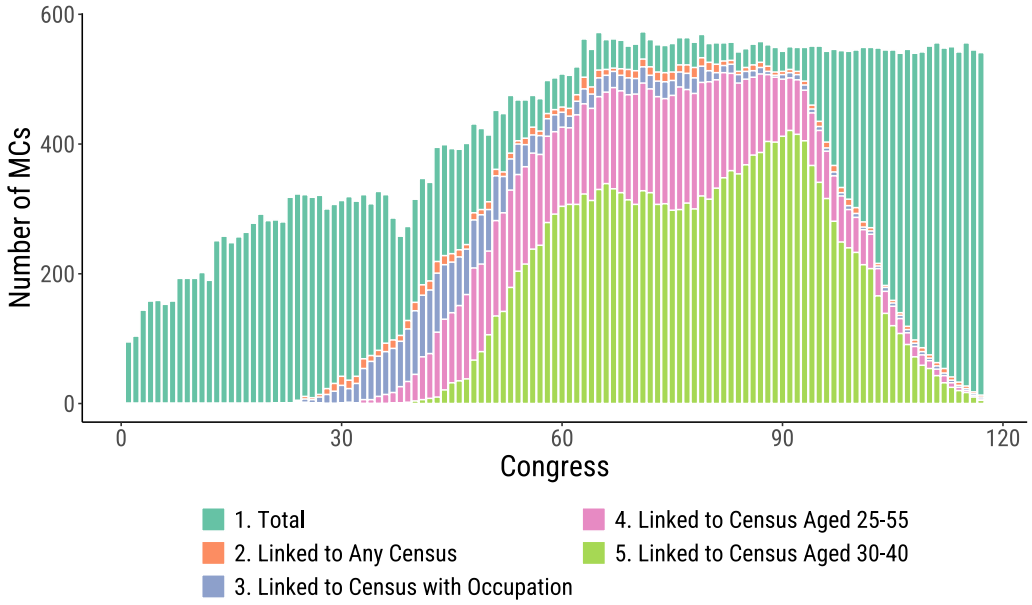


(a) By Congress

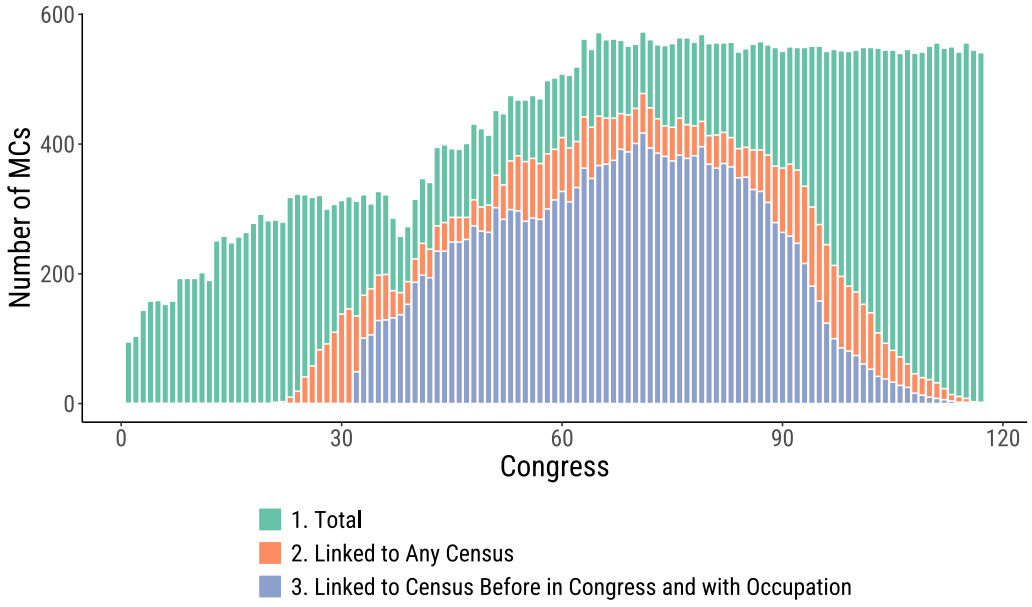


(b) By year of birth

Figure 1: Number of MCs linked via FamilySearch to the complete-count census, by Congress (a) and year of birth (b).

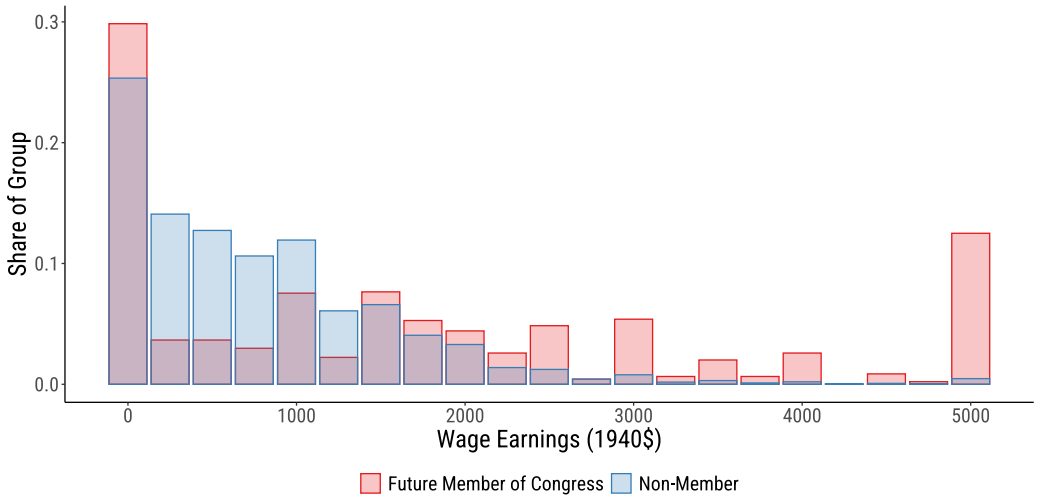


(a) Fathers

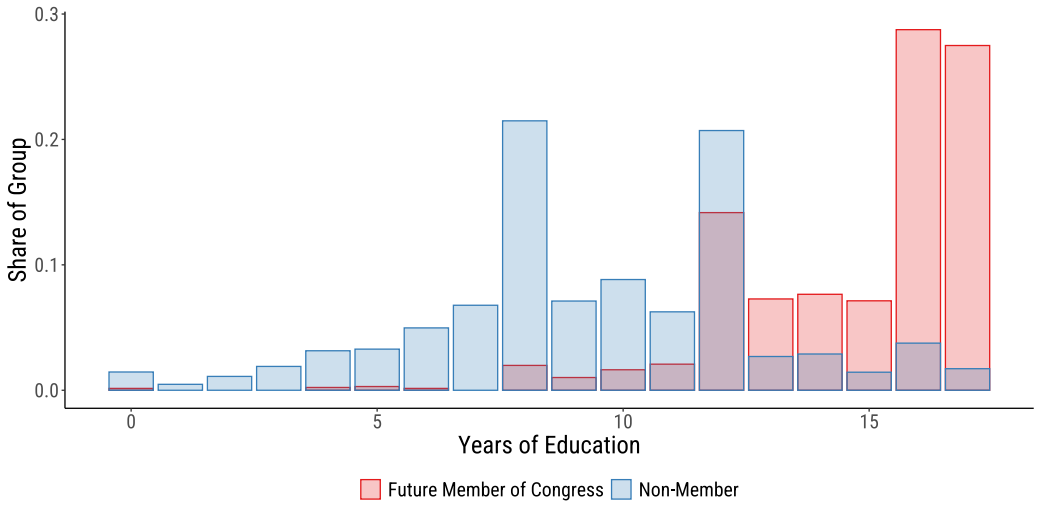


(b) Brothers (at least one)

Figure 2: Number of MC fathers (a) and at least one MC brother (b) linked via FamilySearch to the complete-count census, by Congress.

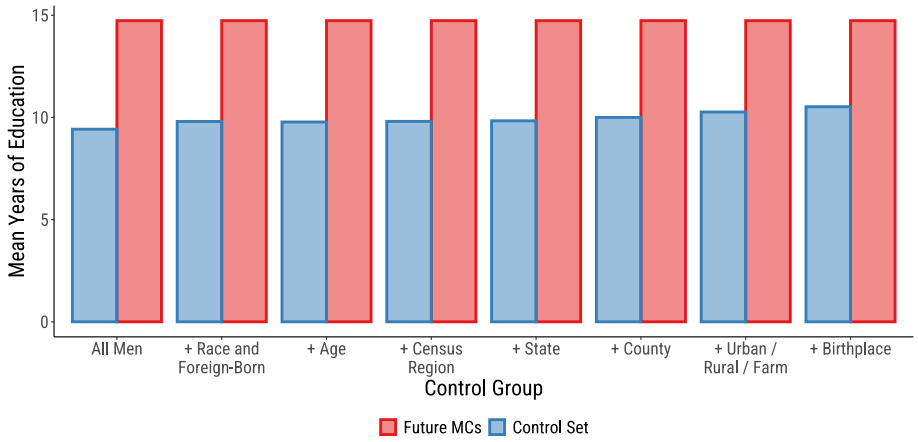


(a) Wage earnings

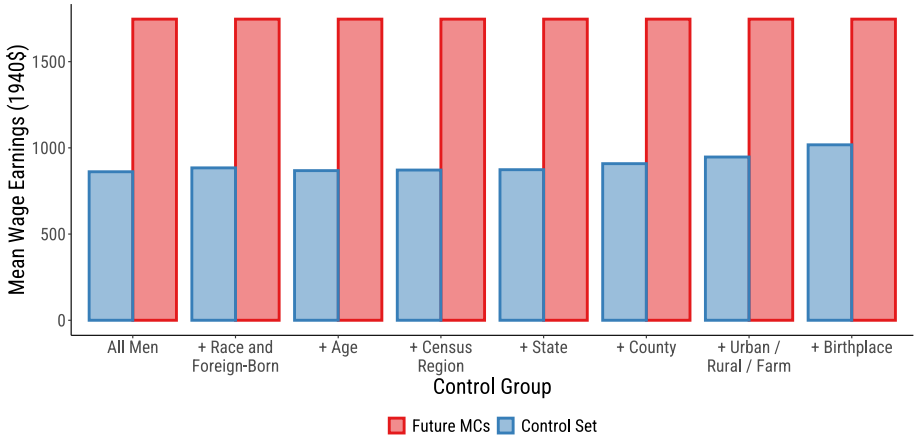


(b) Years of education

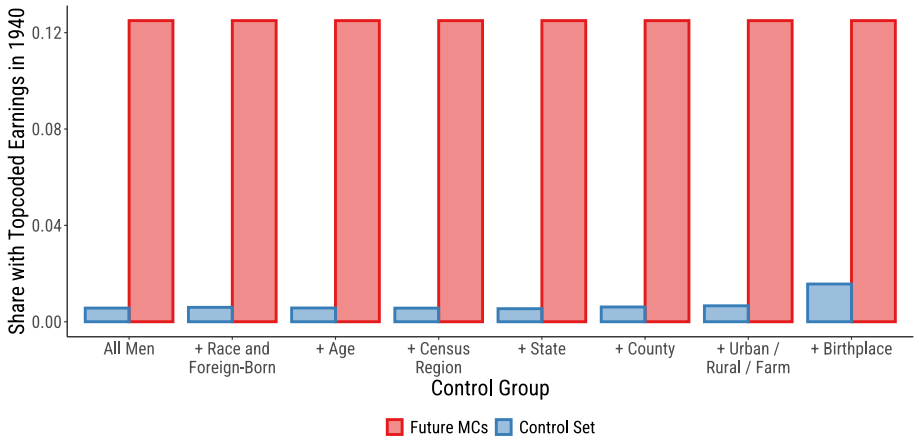
Figure 3: Distributions of 1940 wage earnings and years of education for future MCs and the general population. Future MCs are men who will become members of Congress observed in 1940 before they are elected and between the ages of 18 and 40. The 1940 census recorded wage and salary income only, so we set to missing the future MCs who reported non-wage income but no wages. The mass that remains at zero is mostly young men still in school. About two-thirds of the future MCs in the $0 bar were enrolled in school, and the majority were twenty or younger.



(a) Years of education

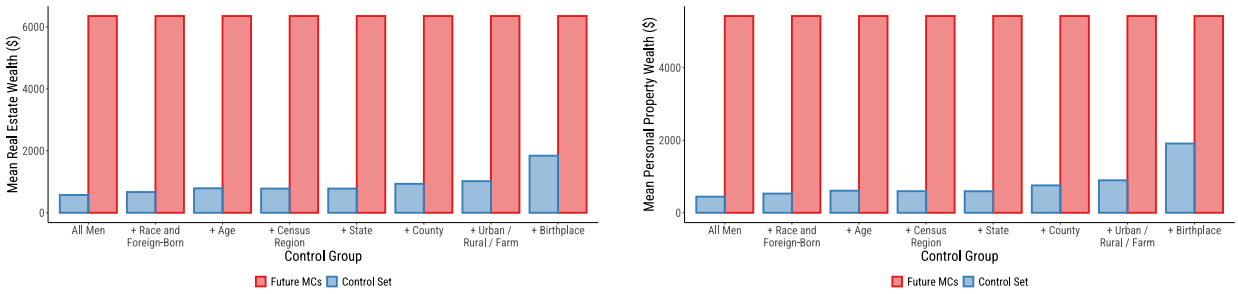


(b) Wage earnings



(c) Share with top-coded earnings

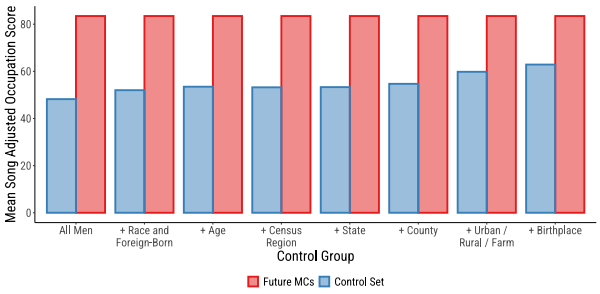
Figure 4: Future MCs versus matched controls in 1940, with progressively tighter matching, on years of education (a), wage earnings (b), and the share with top-coded earnings (c).



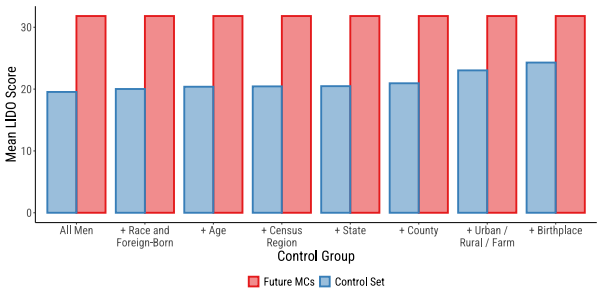
(a) Real estate wealth

(b) Personal property wealth

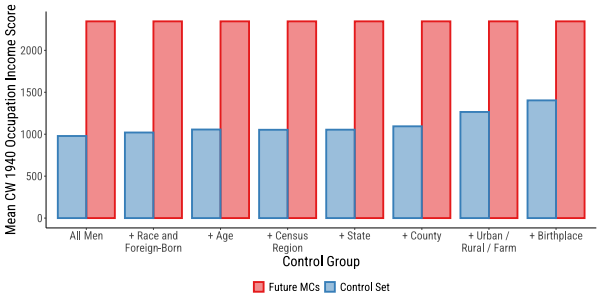
Figure 5: Future MCs versus matched controls, with progressively tighter matching, on real estate wealth in 1850–1870 (a) and personal property wealth in 1860–1870 (b).



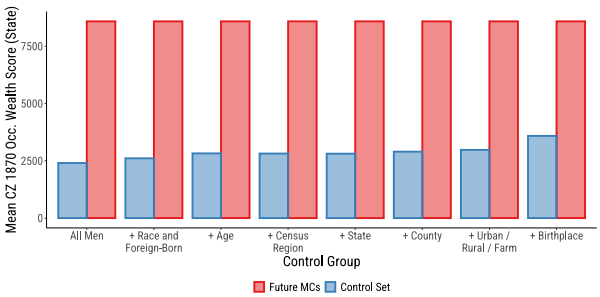
(a) Adjusted Song occupation score



(b) LIDO score

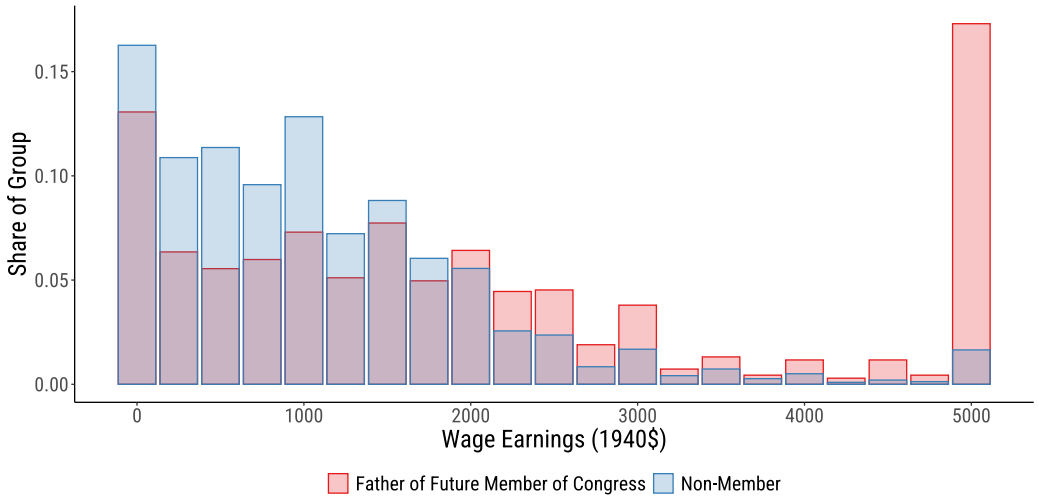


(c) CW 1940 occupation income score

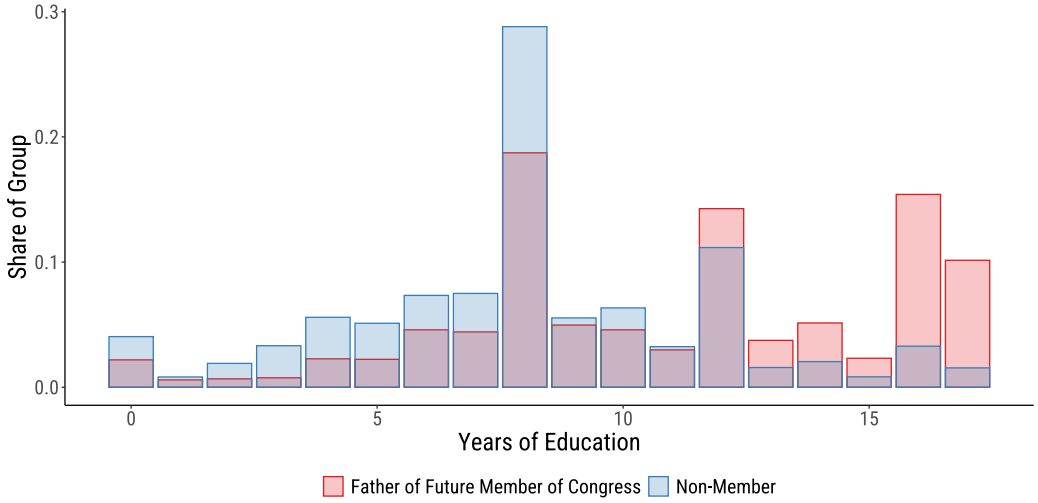


(d) CZ 1870 occupational wealth score (state)

Figure 6: Occupation-based measures of selection for future MCs versus matched controls, pooled across census waves 1850–1950, by control-group definition: adjusted Song occupation score (a), LIDO score (b), CW 1940 occupation income score (c), and CZ 1870 occupational wealth score at the state level (d).

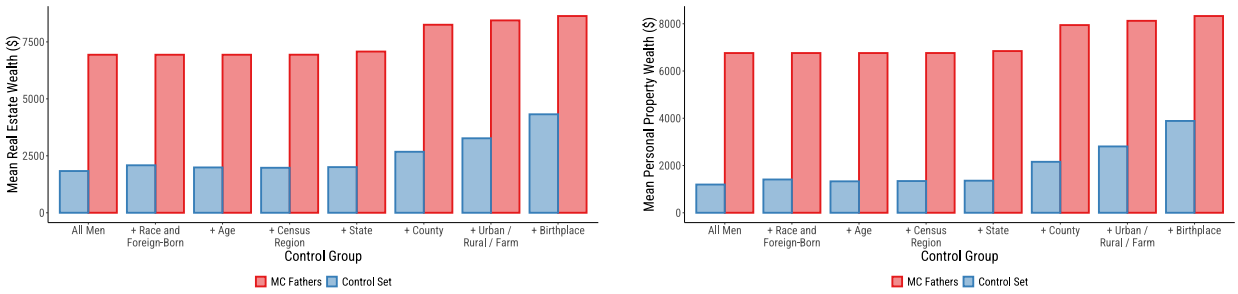


(a) Wage earnings



(b) Years of education

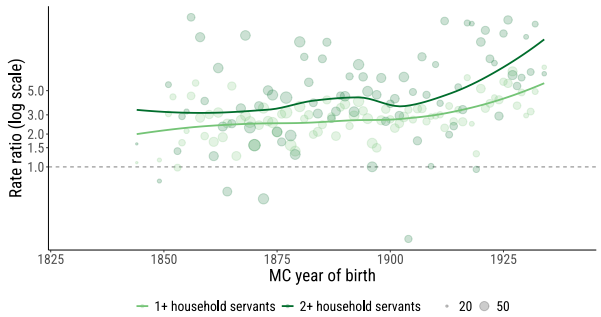
Figure 7: Distributions of 1940 wage earnings and years of education for MC fathers and the general population.



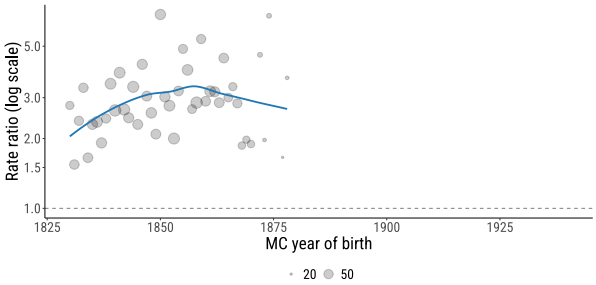
(a) Real-property wealth

(b) Personal-property wealth

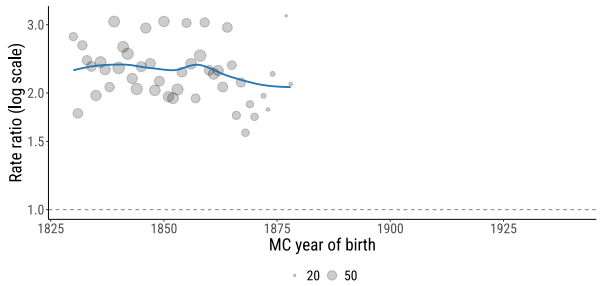
Figure 8: MC fathers versus matched controls, with progressively tighter matching, on real estate wealth in 1850–1870 (a) and personal property wealth in 1860–1870 (b).



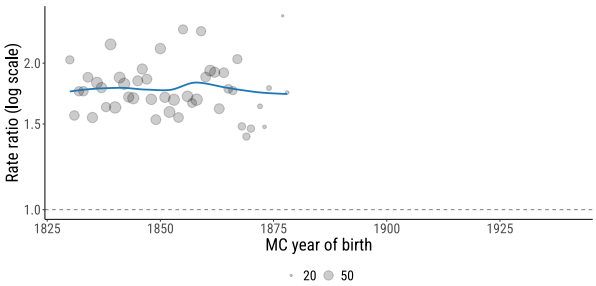
(a) Household servants (1+, 2+)



(b) Wealth: top 1%

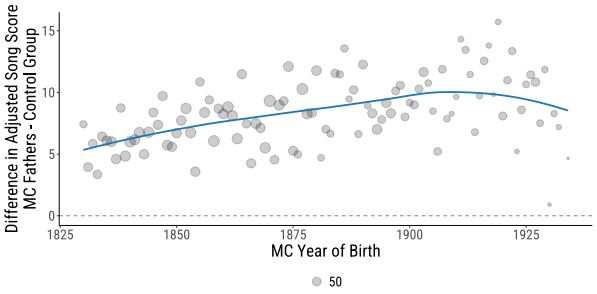


(c) Wealth: top 5%

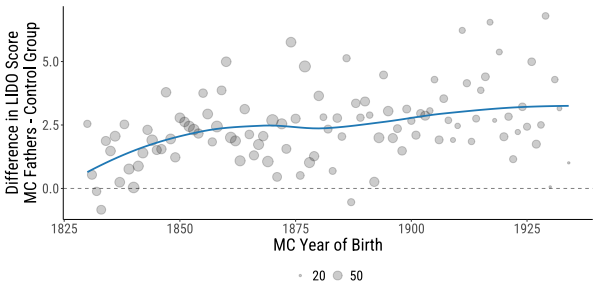


(d) Wealth: top 10%

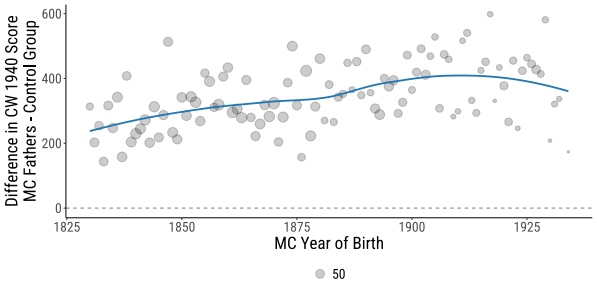
Figure 9: Family-background selection of future MCs across cohorts. MC fathers versus the tightest matched set of control fathers on direct high-status measures, by MC birth cohort.



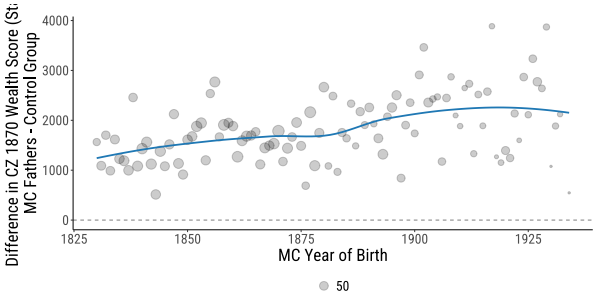
(a) Adjusted Song occupation score



(b) LIDO score

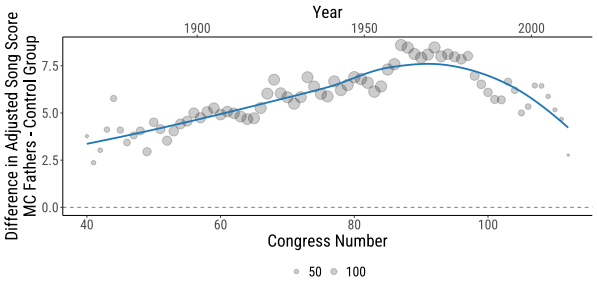


(c) CW 1940 occupation income score

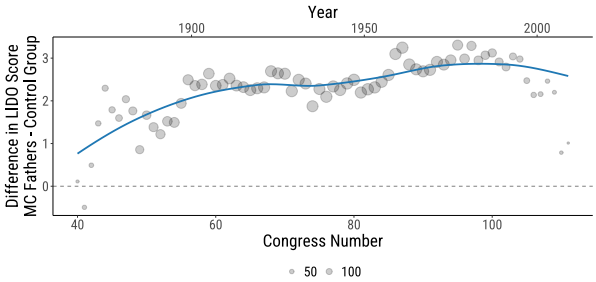


(d) CZ 1870 occupational wealth score (state)

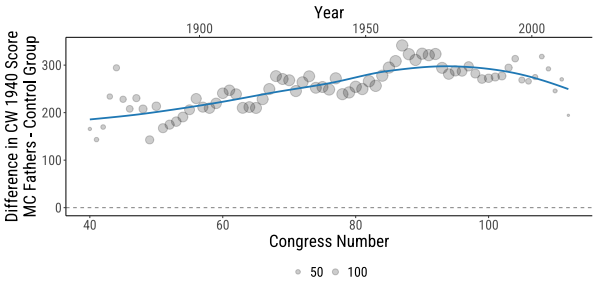
Figure 10: Family-background selection of future MCs across cohorts: MC fathers versus the tightest set of control fathers, by MC birth cohort.



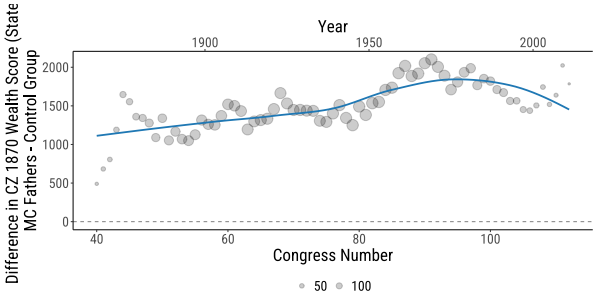
(a) Adjusted Song occupation score



(b) LIDO score



(c) CW 1940 occupation income score



(d) CZ 1870 occupational wealth score (state)

Figure 11: Family-background selection of future MCs across Congresses: MC fathers versus the tightest set of control fathers, by Congress served.

Table 1: Future MC Lawyers are Highly Selected, 1940

	Earnings			Education	
	Level	Log	Top-Coded	Years	College Grad.
	(1)	(2)	(3)	(4)	(5)
Member of Congress	651.25***	0.10	0.19**	0.41***	0.07**
	(234.47)	(0.24)	(0.08)	(0.11)	(0.03)
Demographic FE	Yes	Yes	Yes	Yes	Yes
Geographic FE	Yes	Yes	Yes	Yes	Yes
Observations	3,060	3,060	3,060	8,312	8,312
R^2	0.32	0.26	0.21	0.05	0.04
Y Mean	2,493.6	7.6	0.15	15.9	0.85

Note: This table reports regressions comparing future MCs against their tightest matched set of controls, restricting throughout to people (future MCs and controls) reporting a lawyer occupation in 1940 and to MCs who had not yet been seated in Congress. We use different lawyer subsamples to study earnings and education. The earnings columns (1)–(3) keep lawyers with positive wage earnings and no income from other sources (business, self-employment, rents, interest, or dividends). The education columns (4)–(5) instead keep all lawyers aged 30 and older, regardless of wage or non-wage earnings status, so that reported schooling reflects completed education. Demographic fixed effects are age, race, and birthplace; geographic fixed effects are state, county, urban, and farm. Standard errors clustered at the MC level are reported in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Comparing outcomes between MC fathers and the population of fathers, 1940

Outcome	MC Fathers	All Fathers
Earnings		
Wage earnings, 25th percentile	$1,020	$600
Wage earnings, median	$1,890	$1,100
Wage earnings, 75th percentile	$3,300	$1,768
Wage earnings, mean	$2,316	$1,297
Wage earnings top-coded	18.9%	2.0%
Education		
Years of education	10.98	8.13
High school completion	50.7%	20.5%
Some college	36.3%	9.3%
College completion	25.1%	4.8%
Household wealth		
Home value, mean	$8,828	$3,202
Home value, top 10% (within age bin)	37.2%	20.1%
2+ household servants	3.4%	0.1%
Demographics		
Age	51.2	44.4
Foreign-born	13.5%	16.9%
White	97.3%	92.5%
Black	2.5%	7.1%
Sample size		
N	1,248	20,664,640
N (with positive wage earnings)	625	13,529,304

Note: This table compares outcomes for MC fathers to the broader population of fathers in the 1940 U.S. complete-count census. The All Fathers sample comprises 1940 census men with at least one co-resident child, identified via the IPUMS POPLOC parent variable. The MC Fathers sample comprises father records linked to MCs through FamilySearch and matched into the 1940 census. Wage earnings percentiles, mean, and top-coded share are all calculated conditional on positive reported wage earnings. Mean home value is conditional on the father being an owner-occupied head or spouse, following Kalsi and Ward (2025). The Home Value, Top 10% indicator is based on the within-five-year-age-bin percentile rank of home value (VALUEH), restricted to household heads, again following Kalsi and Ward (2025). Appendix Table E.3 compares occupation-based status scores for the same samples.

Table 3: How much of the status gap between MC fathers and control fathers remains within occupations?

	$\hat{\beta}$ (MC)	$\hat{\beta}$ (MC | Occ)	% Missing	Scaling
1940 earnings				
1940 wage earnings	491.7***	290.3***	59.0%	2.44×
	(48.8)	(36.8)		
1940 log wage earnings	0.570***	0.293***	51.4%	2.06×
	(0.034)	(0.029)		
20th-century wealth				
1940 home value rank	0.083***	0.057***	68.6%	3.19×
	(0.0047)	(0.0045)		
1940 log home value	0.664***	0.453***	68.3%	3.15×
	(0.046)	(0.047)		
1930 home value rank	0.051***	0.034***	67.4%	3.06×
	(0.0029)	(0.0027)		
1930 log home value	0.661***	0.445***	67.3%	3.06×
	(0.038)	(0.036)		
1900 1{2+ servants}	0.052***	0.038***	73.1%	3.71×
	(0.0053)	(0.0047)		
1910 1{2+ servants}	0.054***	0.042***	77.4%	4.43×
	(0.0054)	(0.0048)		
1920 1{2+ servants}	0.037***	0.025***	67.0%	3.03×
	(0.0048)	(0.0042)		
1930 1{2+ servants}	0.044***	0.035***	81.3%	5.34×
	(0.0055)	(0.005)		
19th-century wealth				
1870 log real property	0.865***	0.692***	80.1%	5.02×
	(0.040)	(0.039)		
1870 log personal property	0.947***	0.706***	74.5%	3.92×
	(0.042)	(0.039)		
1870 log total wealth	1.13***	0.858***	75.6%	4.11×
	(0.043)	(0.040)		
1860 log real property	0.943***	0.789***	83.7%	6.12×
	(0.038)	(0.038)		
1860 log personal property	1.33***	0.989***	74.6%	3.94×
	(0.044)	(0.040)		
1860 log total wealth	1.43***	1.07***	75.1%	4.01×
	(0.043)	(0.040)		
1850 log real property	0.960***	0.811***	84.4%	6.42×
	(0.036)	(0.035)		

Note: This table reports how much of the status gap between MC fathers and matched control fathers survives within occupation (controlling for three-digit occupation fixed effects). Each row corresponds to a different outcome and reports a separate pair of regressions on the MC-father estimation sample, with controls weighted by $1/n_{\text{controls}}$ within each MC (bioguide) pair. Column (1) reports the coefficient on the MC-father indicator from a regression with bioguide-pair fixed effects. Column (2) adds occ1950 fixed effects. Column (3) is the ratio of column (2) to column (1), the share of the raw MC-father gap that remains within occupation. Column (4) is the ratio of column (1) to [column (1) − column (2)], the factor by which an occupation-score-based gap understates the gap in the direct outcome. Standard errors clustered at the bioguide level are reported in parentheses. Tables E.1 and E.2 repeat the exercise with future MCs and with MC brothers in place of MC fathers.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Differences between future MCs and their brothers, 1940

Outcome	$\hat{\beta}_{\text{MC}}$	SE	N
Education			
Years of schooling	2.10***	(0.20)	768
High school completion	0.19***	(0.03)	768
College attendance	0.27***	(0.03)	768
College completion	0.32***	(0.04)	768
Earnings			
Wage earnings	258.28**	(124.55)	720
Log earnings	0.36***	(0.08)	446
Earnings top-coded	0.10***	(0.02)	720
Capital earnings	0.19***	(0.04)	775
Capital earnings only	0.12***	(0.04)	709
Household wealth			
Home value ($1000s)	1.27***	(0.37)	791
Home value, top 10% (within age bin)	0.08**	(0.03)	791
2+ household servants	-0.00	(0.01)	791
Occupation scores			
Adjusted Song score	14.62***	(1.79)	787
LIDO	5.50***	(0.75)	622
CW score	663.56***	(59.67)	789
CZ wealth (state)	2842.92***	(381.70)	776
Occupation score	11.88***	(1.18)	791
Song score	20.27***	(2.08)	787
CZ wealth (region)	3809.49***	(372.59)	791

Note: This table reports differences in 1940 economic outcomes between future MCs and their brothers. The sample comprises MC records and their brothers, restricted to men aged 18–40 with a reported occupation and observed prior to the MC's first year in Congress. Each row estimates $y = \beta\,\text{MC} + \alpha_{\text{family}} + \theta_{\text{age}} + \varepsilon$, where y is the outcome named in the row and the regression includes family (bioguide) and age fixed effects. Home value follows the Kalsi and Ward (2025) convention: VALUEH for the head or spouse of an owner-occupied household, zero otherwise, reported in $1000s. Home Value, Top 10% is an indicator for the top decile of the within-five-year-age-bin percentile rank of VALUEH, with non-attributable observations (renters and non-head/spouse household members) entering at zero, also following Kalsi and Ward (2025). Standard errors clustered at the bioguide level are reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Differences between future MCs and their brothers across census waves, 1850–1950

Outcome	Years	$\hat{\beta}_{\text{MC}}$	SE	N
Education				
Years of schooling	1940-1950	2.00***	(0.19)	846
High school completion	1940-1950	0.18***	(0.03)	846
College attendance	1940-1950	0.25***	(0.03)	846
College completion	1940-1950	0.31***	(0.03)	846
Earnings				
Wage earnings	1940-1950	261.61**	(121.91)	797
Log earnings	1940-1950	0.32***	(0.08)	498
Earnings top-coded	1940-1950	0.09***	(0.02)	797
Capital earnings	1940	0.19***	(0.04)	775
Capital earnings only	1940	0.12***	(0.04)	709
19th-century wealth				
Real estate property	1850-1870	3131.94***	(473.43)	5848
Log real estate property	1850-1870	0.47***	(0.06)	2473
Personal property	1860-1870	2670.77***	(537.13)	3405
Log personal property	1860-1870	0.74***	(0.07)	1998
Household wealth				
Home value ($1000s)	1930-1940	1.18***	(0.39)	2222
Home value, top 10% (within age bin)	1930-1940	0.10***	(0.02)	2222
2+ household servants	1880-1940	0.01***	(0.00)	10731
Occupation scores				
Adjusted Song score	1850-1950	13.23***	(0.44)	17171
LIDO	1850-1950	6.32***	(0.24)	14566
CW score	1850-1950	713.24***	(17.45)	16936
CZ wealth (state)	1850-1950	3227.32***	(94.37)	16979
Occupation score	1850-1950	12.87***	(0.35)	17162
Song score	1850-1950	16.19***	(0.48)	17171
CZ wealth (region)	1850-1950	3532.38***	(86.65)	17175

Note: This table reports differences in economic outcomes between future MCs and their brothers, pooled across census waves from 1850 to 1950. The sample comprises MC records and their brothers, restricted to men aged 18–40 with a reported occupation and observed prior to the MC's first year in Congress. Each row estimates $y = \beta\,\text{MC} + \alpha_{\text{family}} + \delta_{\text{year}} + \theta_{\text{age}} + \varepsilon$, where y is the outcome named in the row and the regression includes family (bioguide), census-year, and age fixed effects. The Years column reports the census waves over which each outcome has non-missing data. Home Value, Home Value Top 10%, and 2+ Household Servants definitions all follow Kalsi and Ward (2025); see Table 4 note for definitions. Standard errors clustered at the bioguide level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Within-family MC vs. brother selection

	Panel A: Adj. Song Score					
	(1)	(2)	(3)	(4)	(5)	(6)
MC	18.77***	18.73***			19.61***	27.33***
	(0.43)	(0.53)			(0.53)	(0.41)
Brother			4.95***	6.38***	5.75***	13.70***
			(0.47)	(0.56)	(0.53)	(0.40)
Age Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Family FE	No	No	No	No	No	Yes
Observations	4,881,709	2,450,106	5,960,861	3,818,430	6,268,536	6,268,536
R^2	0.02	0.02	0.02	0.02	0.02	0.12
Y Mean	63.2	63.7	62.4	62.7	63.1	63.1
	Panel B: CW 1940 Score					
MC	933.85***	937.65***			968.25***	1,196.52***
	(14.07)	(18.53)			(18.92)	(17.11)
Brother			195.69***	227.56***	205.56***	453.39***
			(13.24)	(15.99)	(15.67)	(13.27)
Age Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Family FE	No	No	No	No	No	Yes
Observations	4,716,933	2,362,344	5,761,483	3,690,951	6,053,295	6,053,295
R^2	0.05	0.05	0.06	0.05	0.05	0.11
Y Mean	1,403.8	1,416.7	1,368.3	1,380.5	1,394.6	1,394.6
	Panel C: CZ Wealth (state)					
MC	5,377.91***	5,307.20***			5,314.14***	5,761.72***
	(85.04)	(110.16)			(109.82)	(106.20)
Brother			1,696.43***	1,832.09***	1,826.30***	2,264.43***
			(60.68)	(75.62)	(74.88)	(71.95)
Age Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Family FE	No	No	No	No	No	Yes
Observations	4,877,094	2,449,248	5,951,571	3,814,739	6,263,987	6,263,986
R^2	0.03	0.02	0.03	0.03	0.02	0.04
Y Mean	2,979.3	3,000.3	2,877.2	2,904.9	2,942.2	2,942.2

Note: This table compares future MCs and their brothers to matched controls drawn from the general population on three economic status measures, separately by panel. Panel A uses the adjusted Song score, Panel B uses the Collins–Wanamaker 1940 Score, and Panel C uses the Collins–Zimran wealth score constructed at the state level. All columns include census-year fixed effects and a quadratic in age. Columns (1)–(2) compare future MCs to their matched controls, first in the full MC sample and then in the intersection sample restricted to bioguide-years in which a brother is also observed. Columns (3)–(4) compare brothers of future MCs to their matched controls, using the same two samples. Columns (5)–(6) stack the MC and brother comparisons in the intersection sample with separate MC and brother indicators; column (6) adds family (bioguide) fixed effects. Standard errors clustered at the bioguide level are reported in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: The family component of MC selection

	Years	MC $\hat{\beta}_{\text{mc}}$	Brother $\hat{\beta}_{\text{bro}}$	Family share $\hat{\beta}_{\text{bro}}/\hat{\beta}_{\text{mc}}$	N
Occupation-based SES proxies (1850–1950)					
Adj. Song Score	1850–1950	27.33***	13.70***	0.50	6,268,536
		(0.41)	(0.40)		
LIDO	1850–1950	9.89***	3.81***	0.38	5,431,589
		(0.22)	(0.18)		
CW 1940 Score	1850–1950	1,196.52***	453.39***	0.38	6,053,295
		(17.11)	(13.27)		
CZ Wealth (state)	1850–1950	5,761.72***	2,264.43***	0.39	6,263,986
		(106.20)	(71.95)		
1940 direct measures					
Wage Income	1940	381.88***	170.08**	0.45	1,114,423
		(94.44)	(55.58)		
Log Wage Income	1940	0.67***	0.33***	0.50	870,499
		(0.07)	(0.05)		
Years of Schooling	1940	4.85***	2.93***	0.60	1,156,190
		(0.15)	(0.18)		
Antebellum wealth, levels (1850–1870)					
Real Property	1850–1870	4,334.08***	2,030.45***	0.47	384,942
		(370.53)	(261.43)		
Personal Property	1860–1870	4,355.11***	1,944.70***	0.45	251,250
		(536.32)	(316.11)		
Total Wealth	1850–1870	7,270.98***	3,309.94***	0.46	384,966
		(649.43)	(431.91)		
Family wealth indicators (1880–1940)					
Home Value ($1000s)	1930–1940	2.25***	1.29***	0.57	2,734,946
		(0.30)	(0.29)		
Home Value Top 10%	1930–1940	0.15***	0.07***	0.44	2,734,946
		(0.02)	(0.01)		
2+ Servants	1880–1940	0.05***	0.04***	0.79	6,461,841
		(0.00)	(0.00)		

Note: This table reports the share of the future-MC status premium that is also shared by his brother. Estimates come from the stacked within-family specification in column (6) of Table 6, with census-year and family (bioguide) fixed effects. $\hat{\beta}_{\text{mc}}$ is the premium of future MCs relative to their matched controls, and $\hat{\beta}_{\text{bro}}$ is the premium of their brothers relative to their matched controls. The family share is the ratio $\hat{\beta}_{\text{bro}}/\hat{\beta}_{\text{mc}}$, the fraction of the future-MC premium that is also observed in his brother. The sample is restricted to cases where the future MC and his brother are observed in the same census year. Standard errors clustered at the bioguide level are reported in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Effect of Women's Suffrage, Secret Ballot, and Direct Primary on Member of Congress Selection, 1880-1930

	Adj. Song Score		LIDO		CW 1940 Score		CZ Wealth (state)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Member of Congress	9.54***	11.70***	5.19***	4.94***	467.35***	465.99***	2,242.23***	2,283.66***
	(1.48)	(0.78)	(0.70)	(0.56)	(47.07)	(30.35)	(261.26)	(256.31)
Member of Congress × Women's Suffrage	2.40	1.11	−1.08	−1.02	−17.26	−19.02	259.31	283.21
	(1.83)	(1.50)	(0.74)	(0.70)	(52.03)	(45.69)	(219.72)	(237.37)
Member of Congress × Secret Ballot	−2.13	−1.54	−1.03	−0.37	−101.33*	−37.86	162.09	118.43
	(1.99)	(1.07)	(0.77)	(0.68)	(57.22)	(42.04)	(252.08)	(230.33)
Member of Congress × Partisan Primary	0.11	0.23	0.41	0.32	35.84	19.21	−315.57	−365.54
	(1.65)	(1.13)	(0.70)	(0.65)	(51.74)	(43.97)	(289.79)	(299.48)
Baseline FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Chamber FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-Decade FE	Yes	No	Yes	No	Yes	No	Yes	No
MC-matched-set FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	473,534	473,518	413,824	413,788	459,250	459,226	473,515	473,498
Y Mean	62.3	62.2	24.0	24.0	1,273.3	1,273.2	3,309.3	3,309.2

Note: This table reports estimates of the effect of three Progressive-Era reforms (women's suffrage, the secret ballot, and the direct primary) on the family background of MCs first elected between 1880 and 1930. The sample comprises fathers of these MCs and the fathers' tightest matched controls. Outcomes are father-level economic status measures: the adjusted Song occupation score, LIDO, the CW 1940 score, and CZ wealth. LIDO coefficients are in hundreds of 1950 dollars. All three reform indicators enter the same regression interacted with the MC indicator; each coefficient is the marginal effect of one reform, holding the others fixed. For each outcome, the first column includes state-by-decade fixed effects; the second subsumes those with MC-matched-set fixed effects. All columns include baseline fixed effects for father's age and race, MC's state, census year, first-election year, and chamber. Standard errors clustered at the state level are reported in parentheses. Appendix Table D.1 presents the remaining score and wealth outcomes.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Effect of Direct Election, Women's Suffrage, Secret Ballot, and Direct Primary on Member of Congress Selection, 1880-1930

	Adj. Song Score		LIDO		CW 1940 Score		CZ Wealth (state)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Member of Congress	11.89***	13.83***	7.11***	6.70***	580.04***	542.87***	2,332.33***	2,352.76***
	(2.76)	(1.80)	(1.27)	(1.09)	(98.33)	(81.68)	(486.42)	(471.02)
Member of Congress × Women's Suffrage	2.47	1.17	−1.03	−0.97	−13.79	−16.67	262.08	285.34
	(1.84)	(1.52)	(0.73)	(0.70)	(52.41)	(46.15)	(225.41)	(242.11)
Member of Congress × Secret Ballot	−2.12	−1.52	−1.01	−0.34	−100.60*	−37.07	162.58	119.04
	(1.99)	(1.07)	(0.77)	(0.68)	(57.12)	(41.90)	(251.80)	(229.72)
Member of Congress × Partisan Primary	0.18	0.29	0.46	0.37	38.94	21.24	−312.98	−363.60
	(1.66)	(1.12)	(0.68)	(0.65)	(51.12)	(43.83)	(288.57)	(298.26)
Member of Congress × Direct Election	−2.53	−2.30	−2.08**	−1.91*	−121.65	−83.14	−97.28	−74.74
	(2.10)	(1.65)	(1.00)	(0.97)	(85.17)	(74.22)	(483.56)	(463.50)
Baseline FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Chamber FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-Decade FE	Yes	No	Yes	No	Yes	No	Yes	No
MC-matched-set FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	473,534	473,518	413,824	413,788	459,250	459,226	473,515	473,498
Y Mean	62.3	62.2	24.0	24.0	1,273.3	1,273.2	3,309.3	3,309.2

Note: This table reports estimates of the effect of four Progressive-Era reforms (the direct election of senators, women's suffrage, the secret ballot, and the direct primary) on the family background of MCs first elected between 1880 and 1930. The sample comprises fathers of these MCs and the fathers' tightest matched controls. Outcomes are father-level economic status measures: the adjusted Song occupation score, LIDO, the CW 1940 score, and CZ wealth. LIDO coefficients are in hundreds of 1950 dollars. All four reform indicators enter the same regression interacted with the MC indicator; each coefficient is the marginal effect of one reform, holding the others fixed. The direct election effect is identified off chamber-by-period variation, while the other three are identified off state-by-period variation. For each outcome, the first column includes state-by-decade fixed effects; the second subsumes those with MC-matched-set fixed effects. All columns include baseline fixed effects for father's age and race, MC's state, census year, first-election year, and chamber. Standard errors clustered at the state level are reported in parentheses. Appendix Table D.2 presents the remaining score and wealth outcomes.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A Linking

A.1 MCs on FamilySearch

We illustrate our manual record-linking workflow for the most prominent economist turned politician in our sample, Paul Douglas. Douglas, the UChicago labor economist, served as a Senator from Illinois from 1949 to 1967.

We start with Douglas's entry from the Congressional Bioguide entry in Figure A.1. From here, we collect his name, place of birth, year of birth, and other relevant biographical information.

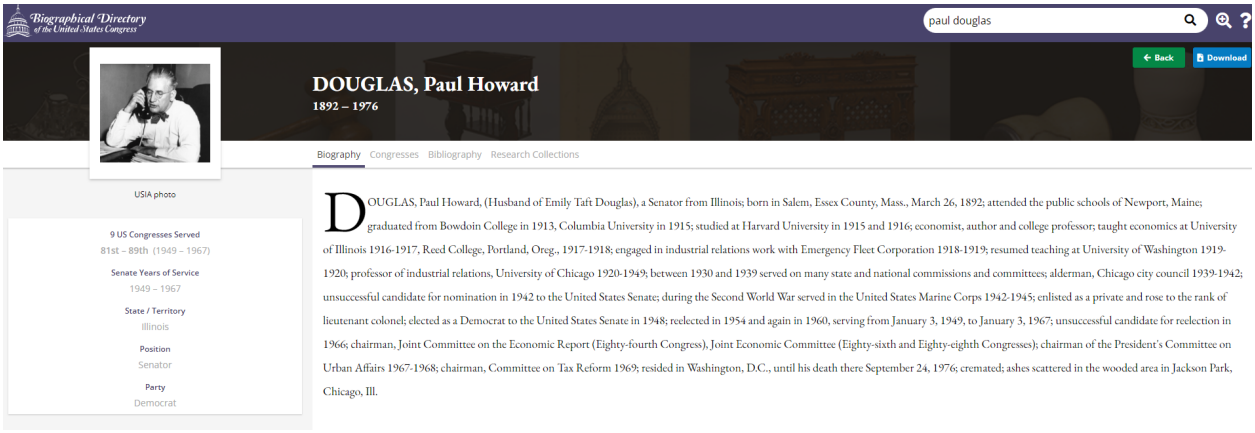


Figure A.1: Paul Douglas in the Congressional Bioguide.

We then search for Douglas on FamilySearch. His record, which already existed on the platform, is in Figure A.2. We can see that his biographical details are recorded (linked to specific sources with information on his name, birth, death, etc). We also see that there are 24 total sources connected to his record.

When we inspect those sources (Figure A.3), we see that several of them are relevant census records from 1900, 1920, 1930, and 1940. His 1910 record is not connected.

We can also see that FamilySearch has created a family tree for Douglas, showing his parents and siblings (Figure A.4). We will use the information on his father (who is himself linked to several censuses) and his brother John and connect them to the IPUMS complete count data as well.

A.2 Linking Robustness

Our main analysis sample recovers the census data for each MC, his father, and his brothers via the FamilySearch tree. To ensure that our main results are robust to alternative methods of reconstructing linked records, we rebuild both our father and brother samples using three algorithmic linking pipelines:

CLP (Census Linking Project, ABE-conservative). From the Census Linking Project, we draw the conservative, exact-match links (Abramitzky et al. 2021).

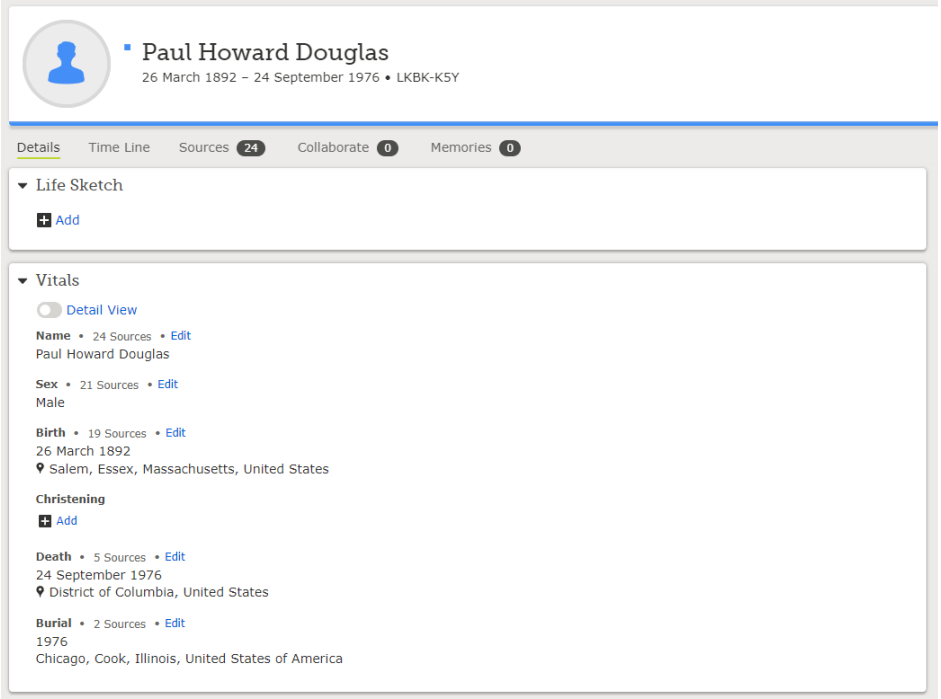


Figure A.2: Paul Douglas on FamilySearch.

Figure A.3: Paul Douglas on FamilySearch: attached sources (`https://www.familysearch.org/tree/person/sources/LKBK-K5Y`).

MLP (IPUMS Multigenerational Longitudinal Panel). Directly from IPUMS, we draw on their probabilistic record linking method which uses both individual-level and family-level and neighborhood-level characteristics in linking (Helgertz et al. 2021).

CensusTree. Finally, we use the family-history-augmented machine-learning linking method from Price et al. (2021) and Buckles et al. (2025). At its heart, CensusTree includes seven match indicators (CLP, direct-hint, family-tree, implied, MLP, profile-hint, and XGBoost) and we include census links made with any of the methods.

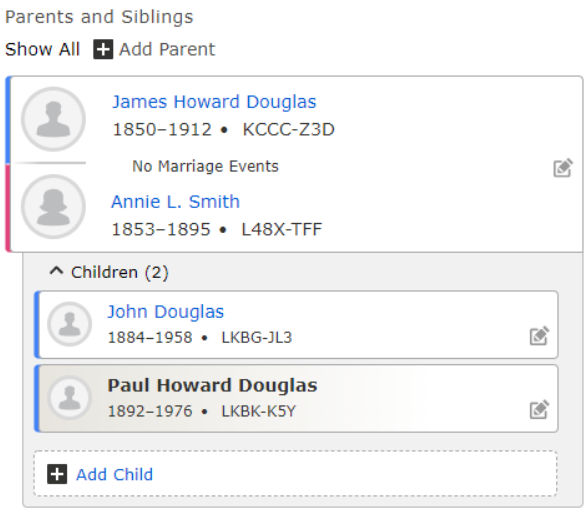


Figure A.4: Paul Douglas' family on FamilySearch.

Fathers pipeline. We hold the MC side of the comparison fixed, starting with each MC from the FamilySearch crosswalk used in the main analysis. We then connect MCs to their tightest control set from the complete count census. We then look for fathers of both MCs and the control set, linking each index person (MC or control) into every census year from 1850 to 1950 using the chosen algorithmic method and reading off the `POPLOC` pointer in the linked household to identify a co-resident father. Once we have found co-resident fathers for both MCs and controls, we link those fathers forward and backward to locate those fathers across all census years using the same algorithmic linking method. We keep only candidate fathers who are male and at least 15 years older than the index person.

Brothers pipeline. Again, we anchor on the FamilySearch-identified MC record. Within the MC's enumeration household we look for other male persons who share at least one of `POPLOC` or `MOMLOC` with the MC; we exclude any household member who is himself an MC. Each candidate brother is then linked across census years using the chosen algorithmic census linking method.

The remainder of this section is organized in four pieces. Section A.2.1 reports MC-level sample coverage by linking method. Section A.2.2 asks how often two methods agree on the identity of the linked father or brother. Section A.2.3 re-estimates the father-side selection gap with each method. Section A.2.4 re-estimates the brother-side within-family family-share decomposition.

A.2.1 Sample overlap by linking method

Showing the same regression under four pipelines tells us whether the estimates move with the linker. It does not tell us whether the four pipelines are running on the same MCs. So before the regressions we report sample overlap.

Figures A.5 and A.6 are upset plots of MC-level coverage. Each MC is one element. Horizontal

bars on the left give the four method-specific set sizes (number of MCs for whom the method links at least one father or brother to a census record). The dot-and-line columns name the intersection on display; the vertical bars on top count the MCs in that intersection. Intersections are sorted by count, so the most common coverage patterns sit on the left. The universe for these plots is the same restricted sample we analyze in the main text—MCs born 1830–1950 in our search universe—so the counts here never exceed that sample and line up with the family-link rates reported there.

For fathers, FamilySearch finds a father for 5,559 MCs, with CensusTree, MLP, and CLP at 5,517, 5,296, and 5,220. About one in twenty MCs in the union is found by some algorithmic method but missed by the FamilySearch data we use in our main analysis. For brothers, FamilySearch covers 4,557 MCs while the three algorithmic methods each cover exactly 4,444. The identical coverage is structural rather than coincidental: brother discovery uses `POPLOC` and `MOMLOC` inside the MC's anchor record and does not depend on the cross-census linker, so the algorithmic pipelines coincide at the MC level by construction. They diverge only at the brother-observation level (the subject of the next subsection).

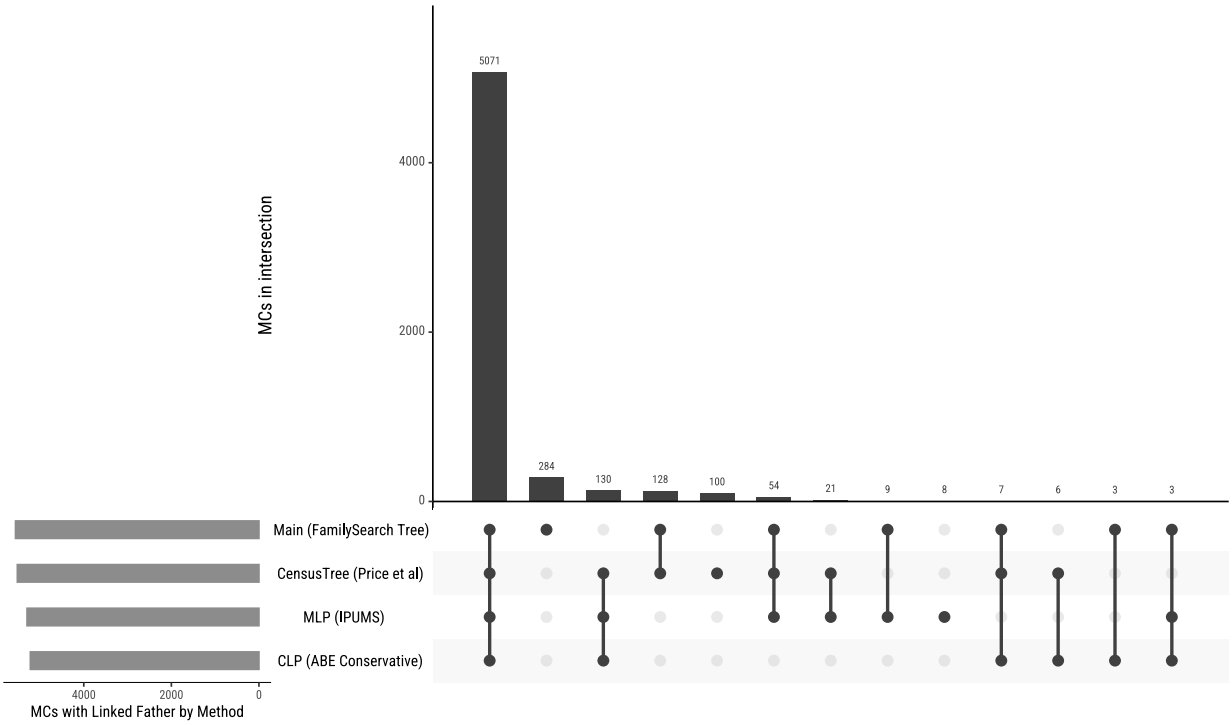


Figure A.5: MC coverage by linking method: fathers.

Note: Each MC is one element. Set membership in method M means M links the MC to at least one census record for the MC's father in any year, 1850–1950. Horizontal bars on the left are the four set sizes; the dot-and-line matrix encodes which methods are in each intersection (columns sorted by intersection size, descending).

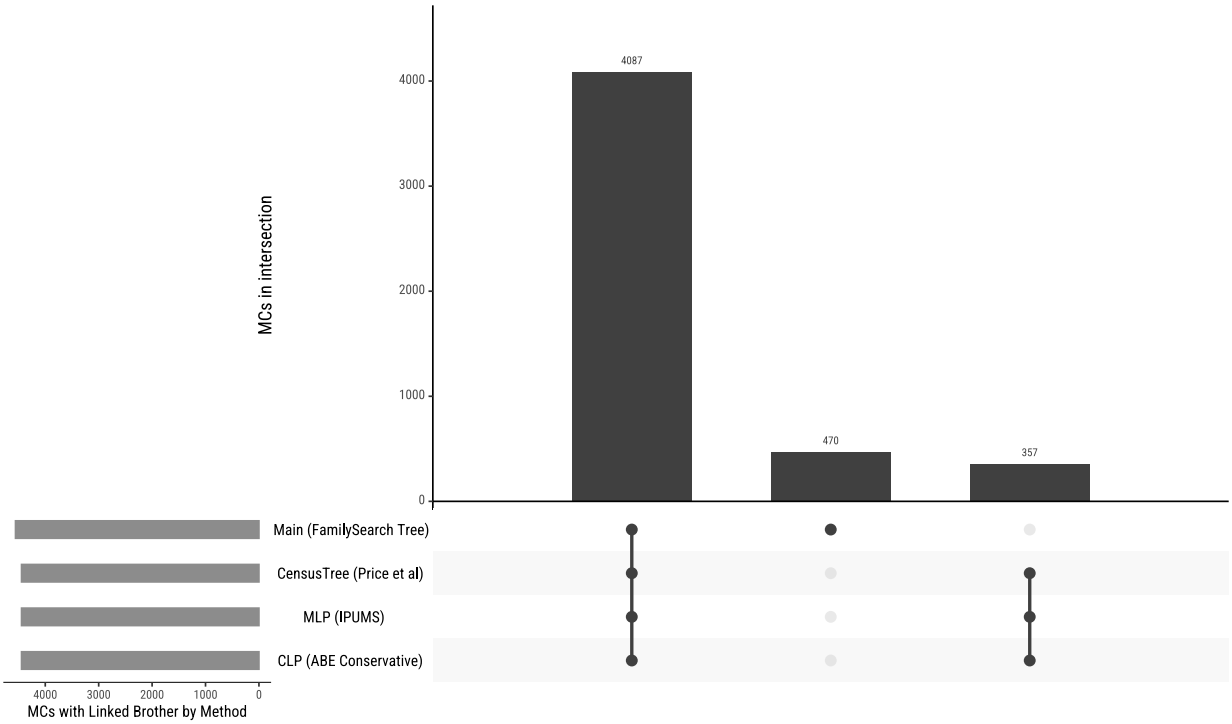


Figure A.6: MC coverage by linking method: brothers.

Note: Each MC is one element; set membership requires at least one linked brother for that MC in any year, 1850–1950. The three algorithmic methods (CLP, MLP, CensusTree) discover candidate brothers via shared `POPLOC` or `MOMLOC` in the MC's anchor census record and therefore coincide at the MC level by construction. See Figure A.5 note for layout.

A.2.2 Linked-person agreement across methods

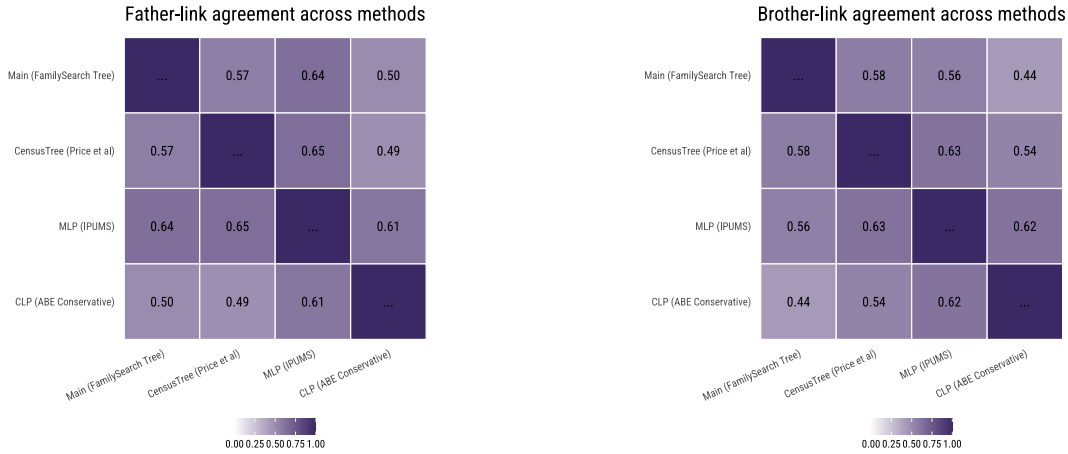
Conditional on two methods linking an MC m, do they link to the same person? Two pipelines could find a father for the same MC, yet tag different census records as that father. Figure A.7 reports the agreement rate across methods.

For each MC and each pair of methods, we compute the Jaccard agreement between the methods' linked tuples for that MC, where a tuple is a (census year, `HISTID`) pair. Each MC contributes weight one to the heatmap cell regardless of how many years his father or brothers are observed. A Jaccard of 1 means the two methods link the exact same set of census records for that relationship; 0 means disjoint sets.

For fathers, MLP and CLP agree most strongly (0.61) and FamilySearch–CLP least (0.50). For brothers the same MLP–CLP cell is highest (0.62) and FamilySearch–CLP again the lowest (0.44). The pattern across both panels is the same: the FamilySearch data–with links that come from the family tree rather than from cross-census record linking–disagrees most with the stricter algorithmic linkers.

Brother agreement runs systematically below father agreement, even between the same pair of methods. The reason is mechanical: each MC has at most one father per year but typically several

brothers per year, so a method that finds one more or one fewer brother-record per cell pushes the Jaccard denominator without necessarily changing the identity of the people linked.



(a) Fathers

(b) Brothers

Figure A.7: Per-MC link agreement between linking methods.

Note: Each cell is the mean per-MC Jaccard agreement on linked (year, `HISTID`) tuples between two methods, averaged over MCs where both methods produce at least one tuple. Per-MC Jaccard is $|A_m \cap B_m|/|A_m \cup B_m|$ where A_m, B_m are the sets of linked tuples for MC m under methods A and B (1 = identical link sets; 0 = disjoint). Diagonals are 1 by construction.

A.2.3 Fathers: selection by cohort

Figure A.8 re-estimates the headline father-side selection gap (the MC-father vs. control-father difference in various occupation-based scores) under each of the four linking pipelines, plotted by MC birth-year cohort. We use the same age-adjusted cohort-dummy specification as Figure 10, with weights split evenly across an MC's linked father histids and standard errors clustered on bioguide. The four series sit on top of each other across the panel: the MC–control father gap is large, positive, and stable in birth cohort, and the levels do not move materially when we swap FamilySearch for any of the three algorithmic methods.

A.2.4 Brothers: family-share decomposition

Tables A.1–A.5 re-estimate the within-family family-share decomposition (column (6) of Table 6) under three alternative brother-linking pipelines. In each, the MC side (FS-identified MCs and their tightest controls) is held fixed; what rotates across columns is the source of the brother. The Main (FS) column reproduces the headline specification, recovering brothers via FamilySearch sibling links. The remaining three columns discover brothers from complete-count census households via shared `POPLOC` or `MOMLOC` variables in any of the MC's enumeration years, then link each brother across census waves using one of three algorithmic linkers: CLP/ABE (exact, conservative), MLP, or the CensusTree.

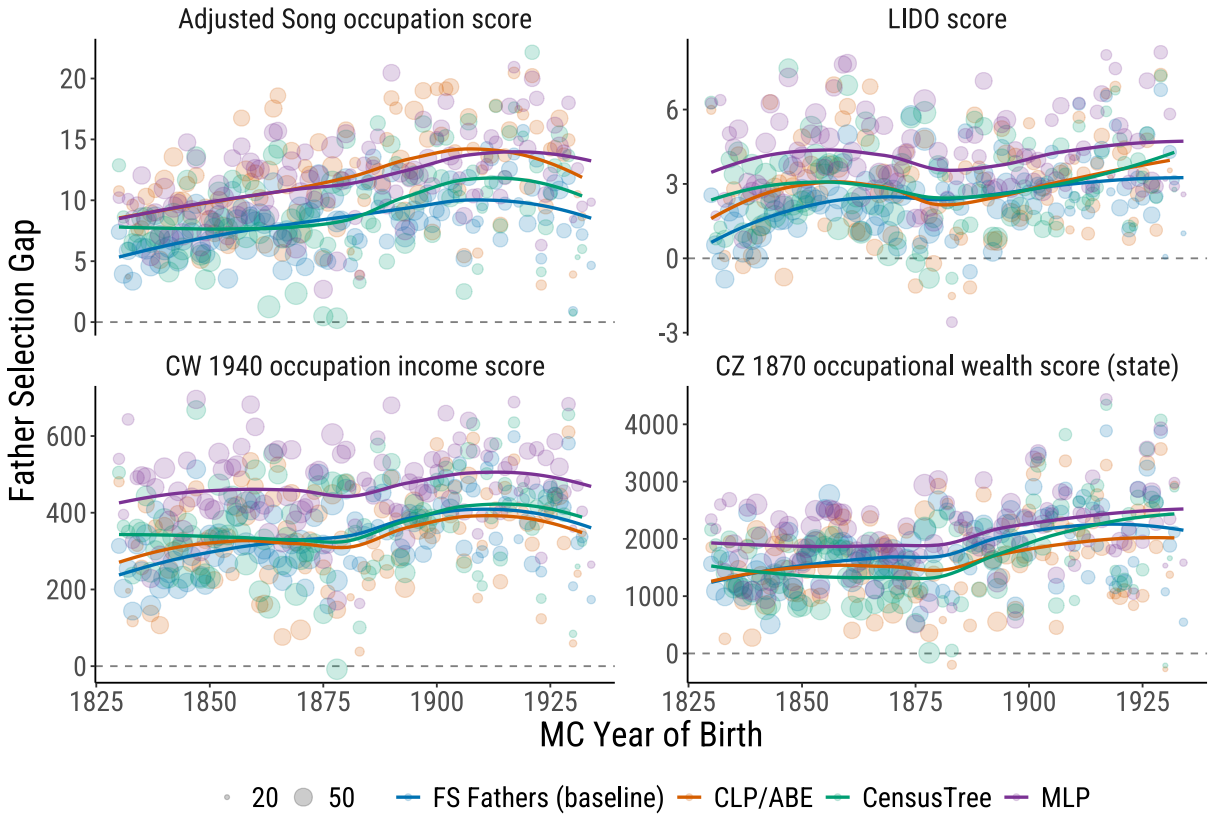


Figure A.8: Father occupation-score selection gap by MC birth cohort and linking method.

Note: Each series plots the MC-father vs. control-father gap in occupation score by MC birth-year cohort under one linking pipeline. The four pipelines are FamilySearch (Main), CLP (ABE-conservative), MLP, and CensusTree, with fathers discovered via the within-household `POPLOC` pointer as described in Section A.2.

A.3 Brothers: balance of the linked subsample

B Member of Congress Occupations in the Census

We build our dataset of members of Congress (MCs) in the census starting with the Congressional Biographical Directory. We then hunt down the MCs in the census via FamilySearch, looking for MCs before, during, and after their terms in Congress.

One could imagine building the dataset in the opposite direction or without the Congressional Biographical Directory at all. It would be convenient if we could pull members of Congress—and perhaps even a wider universe of American politicians—straight off the census enumeration page. But how occupations are reported, enumerated, and coded in the historical censuses makes this hard. For lower-level offices the difficulty is structural: city councilors, mayors, school board members, and even many state legislators historically held part-time positions and, then as is likely the case now, would not have described themselves to the enumerator as politicians or legislators. Though

Table A.1: Brother linking robustness: main outcomes

	Main (FS)	CLP/ABE	MLP	CensusTree
Adj. Song Score				
MC	27.33***	25.79***	26.76***	27.76***
	(0.41)	(0.66)	(0.47)	(0.45)
Brother	13.70***	13.82***	13.30***	12.48***
	(0.40)	(0.63)	(0.46)	(0.42)
Family share	0.50	0.54	0.50	0.45
LIDO				
MC	9.89***	7.26***	8.80***	9.75***
	(0.22)	(0.33)	(0.24)	(0.23)
Brother	3.81***	2.47***	2.89***	3.04***
	(0.18)	(0.26)	(0.19)	(0.18)
Family share	0.38	0.34	0.33	0.31
CW 1940 Score				
MC	1,196.52***	1,018.25***	1,124.01***	1,182.53***
	(17.11)	(25.82)	(19.14)	(18.39)
Brother	453.39***	404.49***	405.43***	376.32***
	(13.27)	(19.21)	(14.75)	(13.44)
Family share	0.38	0.40	0.36	0.32
CZ Wealth (state)				
MC	5,761.72***	5,038.44***	5,469.20***	5,703.46***
	(106.20)	(166.93)	(120.35)	(114.93)
Brother	2,264.43***	1,830.36***	2,075.52***	1,882.98***
	(71.95)	(113.47)	(85.77)	(73.52)
Family share	0.39	0.36	0.38	0.33
N (obs)	6,158,640	4,153,358	6,176,611	5,929,809
N (MCs)	3,179	1,697	2,583	2,820

Note: This table reports how robust the within-family MC vs. brother selection estimates of Table 6 are to the choice of brother-linking method. Each column re-estimates the stacked within-family specification of Table 6, column (6), $y = \beta_{\text{mc}}\,\text{MC} + \beta_{\text{bro}}\,\text{Brother} + \alpha_{\text{family}} + \delta_{\text{year}} + \theta_{\text{age}} + \varepsilon$, on the intersection sample of bioguide–years in which both an MC and a brother are observed. The MC side of the comparison (FamilySearch-identified MCs and their tightest matched controls) is held fixed across columns; only the brother-linking method rotates. Each row corresponds to a different occupation-based status outcome. Family share is the ratio $\hat{\beta}_{\text{bro}}/\hat{\beta}_{\text{mc}}$, the fraction of the future-MC premium that is also observed in his brother. N (obs) is the median intersection-sample size across the outcomes in the panel; N (MCs) is the count of distinct bioguides in that sample. Standard errors clustered at the bioguide level are reported in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

member of Congress (in the House or in the Senate) was very much a full-time job, we can still use the census occupations to gauge how serious this problem might be.

Restricting attention to MCs who were officially sitting members of Congress on census day and whose positions are about as full-time and visible as elective office gets, Figure B.1 shows that even among the MCs we link to the census, the majority typically do not report themselves as members of

Table A.2: Brother linking robustness: additional occupation-based outcomes

	Main (FS)	CLP/ABE	MLP	CensusTree
Song Score				
MC	33.26***	32.28***	33.28***	33.54***
	(0.45)	(0.75)	(0.53)	(0.49)
Brother	16.56***	17.06***	16.11***	14.69***
	(0.46)	(0.73)	(0.53)	(0.49)
Family share	0.50	0.53	0.48	0.44
Occscore				
MC	20.30***	16.61***	18.87***	20.03***
	(0.33)	(0.49)	(0.37)	(0.36)
Brother	6.93***	5.67***	5.89***	5.64***
	(0.25)	(0.35)	(0.28)	(0.26)
Family share	0.34	0.34	0.31	0.28
CZ Wealth (region)				
MC	6,031.65***	5,155.31***	5,759.36***	5,978.18***
	(91.05)	(145.10)	(104.49)	(98.01)
Brother	2,292.28***	1,783.78***	2,052.66***	1,894.84***
	(68.00)	(108.56)	(80.15)	(68.49)
Family share	0.38	0.35	0.36	0.32
N (obs)	6,268,536	4,227,467	6,287,709	6,041,117
N (MCs)	3,179	1,697	2,583	2,820

Note: See Table A.1 note.

Congress. Through much of the nineteenth and early twentieth century the share giving an occupation of "politician"—which we define as IPUMS `OCC1950` code 250 ("Officials and administrators (n.e.c.), public administration") or `IND1950` code 916 ("Federal public administration")—hovered between eighteen and forty-five percent of found MCs, never reaching a majority. Many reported themselves as lawyers (`OCC1950`=55), a few as farmers, the rest scattered across other occupations. Only in 1940 and 1950 do a clear majority of the MCs we link to the census actually report themselves as politicians.

A few examples from our data give a sense of the patterns in the data:

Kingsley S. Bingham (MI), 1850 and 1860. Bingham appears in the 1850 census as a forty-two-year-old farmer with $8,000 in real-estate wealth, sitting in the middle of his four-year run as a U.S. Representative from Michigan. He turns up again in 1860, age fifty-one, *still* reporting his occupation as farmer, this time as a sitting U.S. Senator. In between, he had been the governor of Michigan (1855–1859).

Amos Tuck (NH), 1850. Tuck was a thirty-nine year old New Hampshire Representative in his second term in 1850, with $6,000 in wealth. The manuscript reports lawyer, but the IPUMS coding lands in `OCC1950`=674 rather than 55, so he ends up in the "other occupation" bucket. This is probably because lawyer was misread as sawyer by the IPUMS transcription. Tuck sat in Congress

Table A.3: Brother linking robustness: 1940 wage and schooling

	Main (FS)	CLP/ABE	MLP	CensusTree
Wage Income				
MC	381.88***	268.67**	312.80***	317.53***
	(94.44)	(104.82)	(90.00)	(119.06)
Brother	170.08***	178.08**	130.73**	202.53***
	(55.58)	(69.19)	(53.89)	(77.23)
Family share	0.45	0.66	0.42	0.64
Log Wage Income				
MC	0.67***	0.58***	0.58***	0.77***
	(0.07)	(0.09)	(0.08)	(0.09)
Brother	0.33***	0.30***	0.28***	0.30***
	(0.05)	(0.07)	(0.05)	(0.07)
Family share	0.50	0.52	0.48	0.40
Years of Schooling				
MC	4.85***	4.34***	4.66***	4.73***
	(0.15)	(0.18)	(0.15)	(0.19)
Brother	2.93***	2.86***	2.91***	2.89***
	(0.18)	(0.20)	(0.17)	(0.22)
Family share	0.60	0.66	0.62	0.61
N (obs)	1,114,423	905,452	1,267,541	820,753
N (MCs)	292	212	303	204

Note: See Table A.1 note.

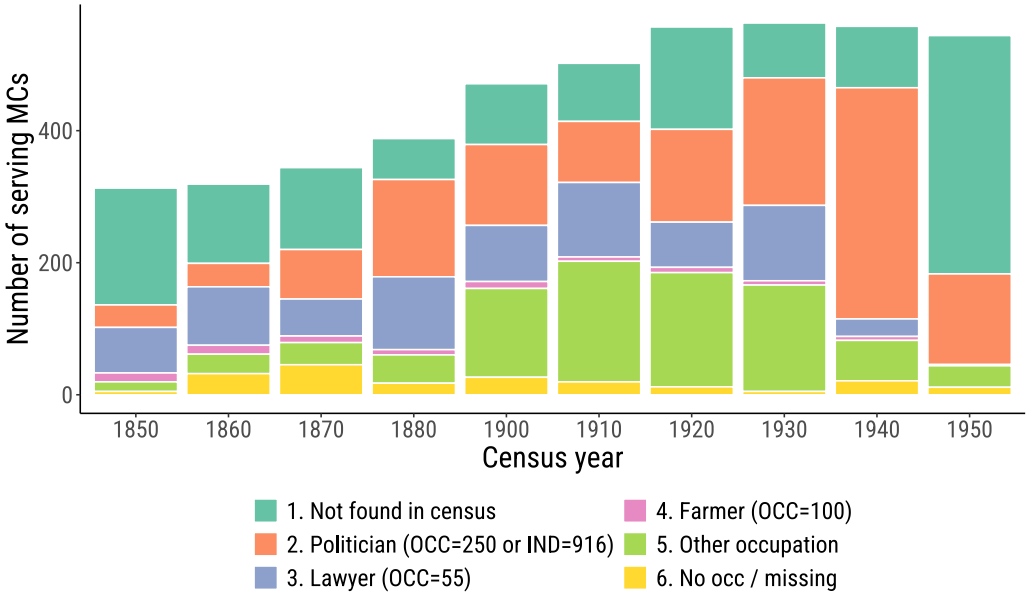


Figure B.1: Reported occupations of sitting MCs, by census year. Each bar conditions on MCs whose terms spanned the enumeration day for that census; categories are based on IPUMS `OCC1950` and `IND1950`.

Table A.4: Brother linking robustness: antebellum wealth, levels

	Main (FS)	CLP/ABE	MLP	CensusTree
Real Property				
MC	4,334.08***	1,775.64***	3,290.67***	4,067.11***
	(370.53)	(455.67)	(486.99)	(773.33)
Brother	2,030.45***	583.57**	1,267.37***	1,212.15***
	(261.43)	(228.43)	(298.79)	(253.93)
Family share	0.47	0.33	0.39	0.30
Personal Property				
MC	4,355.11***	2,737.55***	3,802.47***	3,532.55***
	(536.32)	(1,008.15)	(709.58)	(563.10)
Brother	1,944.70***	1,313.59**	1,420.36***	1,088.02***
	(316.11)	(611.14)	(382.80)	(258.58)
Family share	0.45	0.48	0.37	0.31
Total Wealth				
MC	7,270.98***	3,720.84***	6,409.13***	7,081.71***
	(649.43)	(1,076.68)	(973.21)	(1,038.47)
Brother	3,309.94***	1,519.12**	2,440.93***	2,167.16***
	(431.91)	(629.27)	(578.81)	(444.49)
Family share	0.46	0.41	0.38	0.31
N (obs)	384,942	220,032	289,980	366,145
N (MCs)	1,155	479	670	818

Note: See Table A.1 note.

as both a Free Soiler and a Whig during the party realignments of the 1840s and 1850s, and lent his name to the Tuck School at Dartmouth.

Alexander Stephens (GA), 1850 and 1880. Stephens—the future Vice President of the Confederacy—was a U.S. Representative from Georgia in 1850, and the census enumerator recorded his occupation as "atty at law." Three decades later, after leading an armed rebellion against the United States, he was back in the House representing Georgia. The 1880 census enumerated him at his D.C. boarding house as a member of Congress.

William Claflin (MA), 1880. Claflin appears in 1880 as a sixty-three year old in his second term representing Massachusetts's 4th district, having previously served three terms as Governor of the commonwealth (1869–1872). In addition to his political activities, Claflin had joined his father's boot-and-shoe manufacturing firm. The census reports him as a leather merchant, so Claflin's occupation is technically accurate, if misleading.

Ben Johnson (KY), 1920. Johnson served twenty years in the House of Representatives from Kentucky. In 1920 he is reported as a farmer, in the middle of his thirteenth year in the House. When he retired from Congress in 1927, he told a reporter, "I am away behind on hunting and fishing and must catch up." In reality, he was trained as a lawyer and he was the son of a Kentucky

Table A.5: Brother linking robustness: gilded-age family-wealth indicators

	Main (FS)	CLP/ABE	MLP	CensusTree
Home Value ($1000s)				
MC	2.25***	1.81***	2.01***	2.34***
	(0.30)	(0.27)	(0.27)	(0.30)
Brother	1.29***	1.12***	1.08***	1.02***
	(0.29)	(0.37)	(0.26)	(0.30)
Family share	0.57	0.62	0.54	0.44
Home Value Top 10%				
MC	0.15***	0.13***	0.14***	0.17***
	(0.02)	(0.02)	(0.02)	(0.02)
Brother	0.07***	0.05***	0.05***	0.04***
	(0.01)	(0.01)	(0.01)	(0.01)
Family share	0.44	0.35	0.39	0.25
2+ Servants				
MC	0.05***	0.05***	0.05***	0.05***
	(0.00)	(0.01)	(0.01)	(0.00)
Brother	0.04***	0.05***	0.04***	0.03***
	(0.00)	(0.01)	(0.01)	(0.00)
Family share	0.79	1.02	0.78	0.71
N (obs)	2,734,946	2,242,320	2,966,887	2,665,063
N (MCs)	566	432	565	523

Note: See Table A.1 note.

state senator and lieutenant governor.

Richard B. Wigglesworth (MA), 1940. Wigglesworth was a Massachusetts Representative serving in the House in 1940. On the manuscript, his occupation is recorded as lawyer. He did give an industry of "Member of Congress," but IPUMS coded the entry into legal services (`IND1950`=879) with `OCC1950`=55 (lawyer), so a "politician" rule would not catch him either.

C 19th-Century Wealth: Extensive and Intensive Margins

C.1 Future MCs

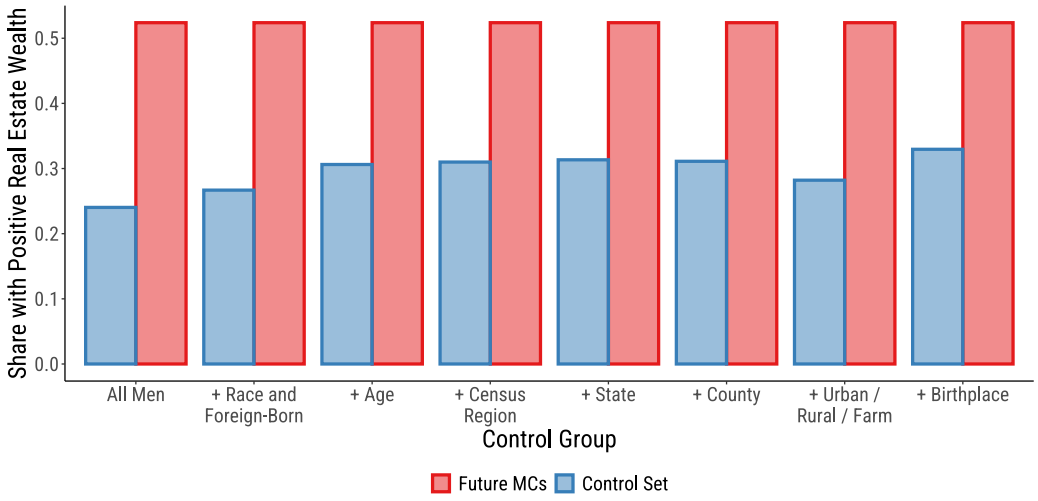
Both margins contribute. About 52 percent of future MCs report positive real estate between 1850 and 1870, against 24 percent of all matched controls and 33 percent at the tightest cell, which matches on race, age, region, state, county, urban/rural status, and birthplace; 64 percent of future MCs report positive personal property, against 41 percent of all controls and 50 percent at the tightest cell (Figure C.1). Conditional on holding any wealth, future MCs are far ahead. Their mean real estate is roughly $12,000 in every cell of the telescope. Among All Men controls the conditional mean is just $2,400; even after the full set of matches it rises only to $5,800. The personal-property gap is wider in ratio terms: future MCs hold $8,500 on average, against $1,100 among All Men and $3,700 at the tightest cell (Figure C.2). Future MCs were more likely to hold any wealth at all and,

Table A.6: Balance of the linked-brother subsample against the full MC universe

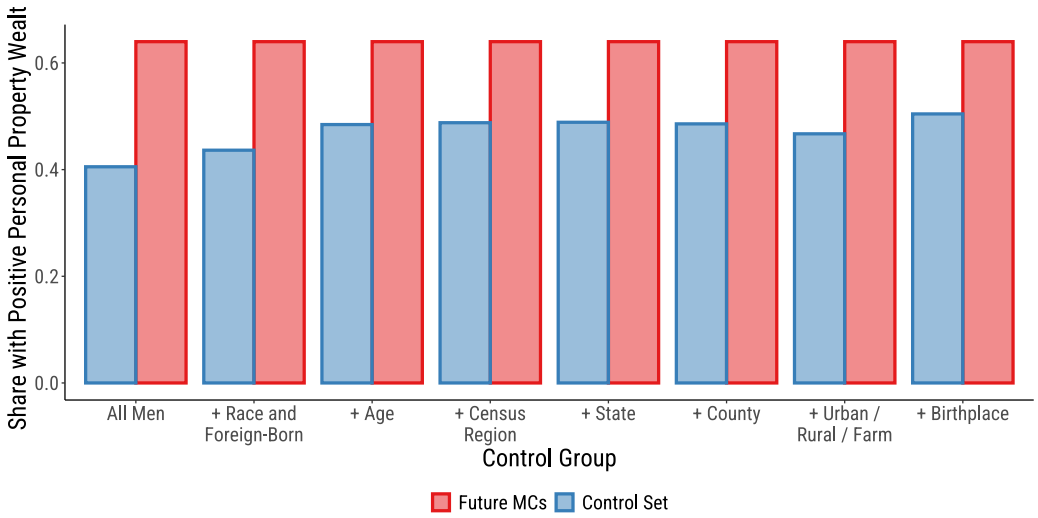
	MCs			Difference	
	All	With Linked Brother	With Brother in Window	(2)−(1)	(3)−(1)
	(1)	(2)	(3)	(4)	(5)
Demographics					
Birth year (mean)	1875.00	1872.51	1870.50	-2.49***	-4.50***
Region: Northeast	0.31	0.30	0.30	-0.01***	-0.01
Region: Midwest	0.33	0.34	0.34	0.00	0.01**
Region: South	0.29	0.31	0.30	0.01***	0.01**
Region: West	0.06	0.06	0.05	-0.01***	-0.01***
Family link coverage					
Share with linked father	0.83	0.87	0.87	0.04***	0.04***
Father status					
Adjusted Song score	67.87	66.65	65.78	-1.22***	-2.09***
LIDO	26.43	25.73	25.11	-0.70***	-1.32***
CW score	1576.27	1523.03	1470.02	-53.25***	-106.25***
CZ wealth (state)	5575.59	5463.32	5389.13	-112.27***	-186.46***
MC adult outcomes					
Adjusted Song score	86.38	86.53	86.47	0.14	0.08
LIDO	35.78	35.85	35.99	0.07	0.21
CW score	2564.47	2574.62	2574.86	10.15	10.39
CZ wealth (state)	9605.73	9572.09	9539.21	-33.64	-66.52
Years of schooling	14.48	14.38	14.42	-0.11	-0.07
Control-matching coverage					
Share with ≥ 1 tightest control	0.99	0.99	0.99	0.00**	0.00
Sample size					
MCs	6,073	4,590	3,532		

Note: This table reports balance of MC-side characteristics across three nested samples of MCs. Column (1) is the full MC sample we study (MCs born 1830–1950 and now deceased). Column (2) restricts to MCs with at least one brother located in any census. Column (3) further restricts to MCs with at least one brother whose census appearance falls in the within-family analysis window: brother enumerated before the MC entered Congress, age 18–40, with a reported occupation; column (3) is the sample underlying Table 7. Columns (4) and (5) report the difference in means between each subsample and the full sample—that is, column (2) or (3) minus column (1). Father-status rows are conditional on the MC having a linked father and average the score across his father's qualifying census appearances at age 30–55. MC adult outcomes are taken from the MC's most recent pre-Congress census observation at age 30–50. Birth-regions use the Census Region at the MC's earliest appearance in the census as a proxy for region of upbringing. We cannot separate MCs with no brother in the FamilySearch tree from MCs with a brother in the tree whose census record could not be located; both cases make up the change from column (1) to column (2). Stars test that same difference, from a regression of each row variable on a subsample-membership indicator with heteroskedasticity-robust standard errors.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

conditional on holding any, several times wealthier than their neighbors.



(a) Real estate wealth

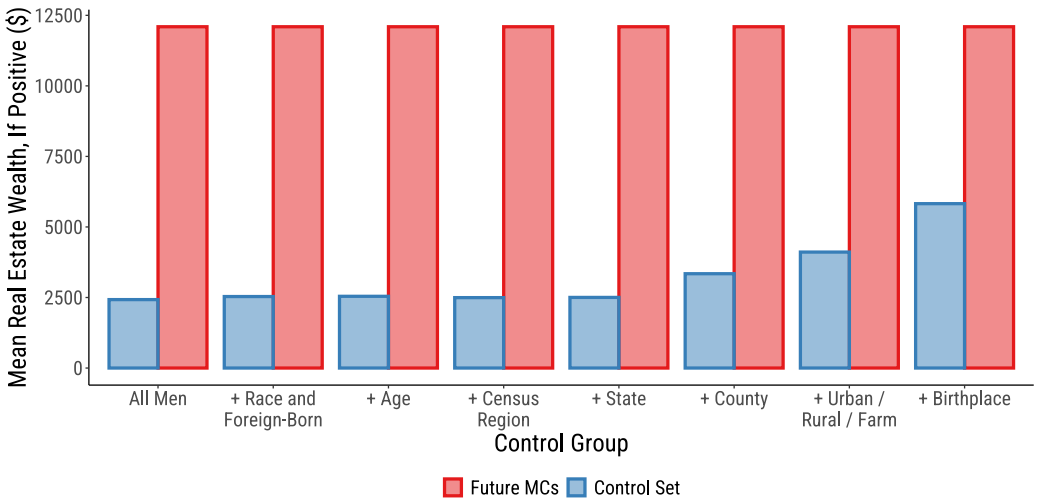


(b) Personal property wealth

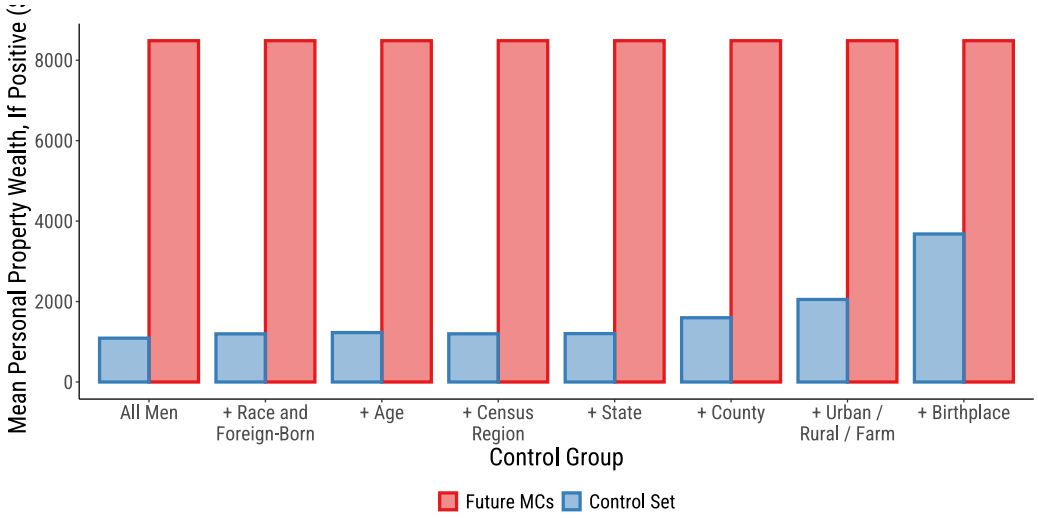
Figure C.1: Future MCs versus matched controls, with progressively tighter matching, on the share with positive real estate wealth in 1850–1870 (a) and positive personal property wealth in 1860–1870 (b).

C.2 MC Fathers

For fathers the decomposition looks different. About 57 percent of MC fathers hold positive real estate in 1850–1870, against 50 to 61 percent of matched control fathers, and the small gap reverses sign once we match on county and birthplace (Figure C.3, panel a). For personal property, 66 percent of MC fathers hold positive amounts versus 67 to 79 percent of controls; if anything, control fathers are slightly more likely to hold any (panel b). What distinguishes MC fathers is how much.



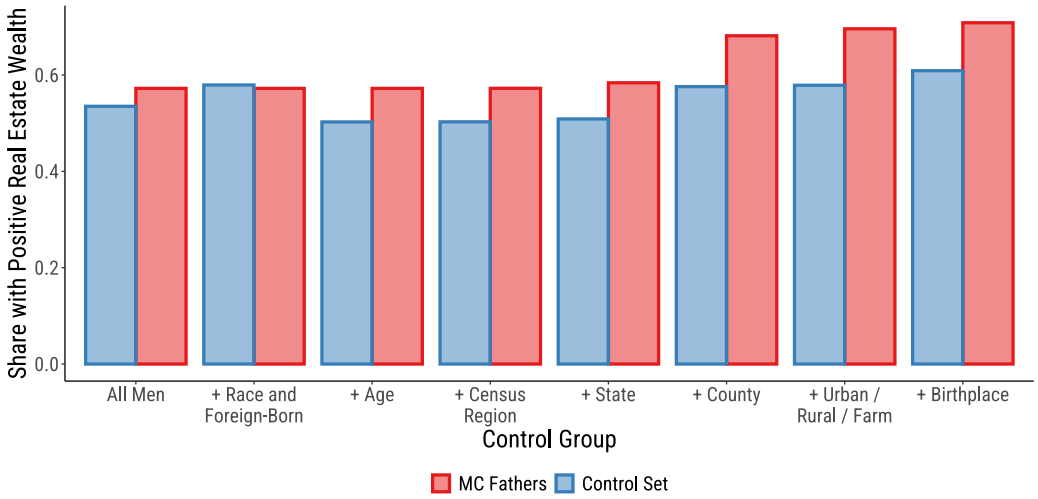
(a) Real estate wealth, conditional on being positive



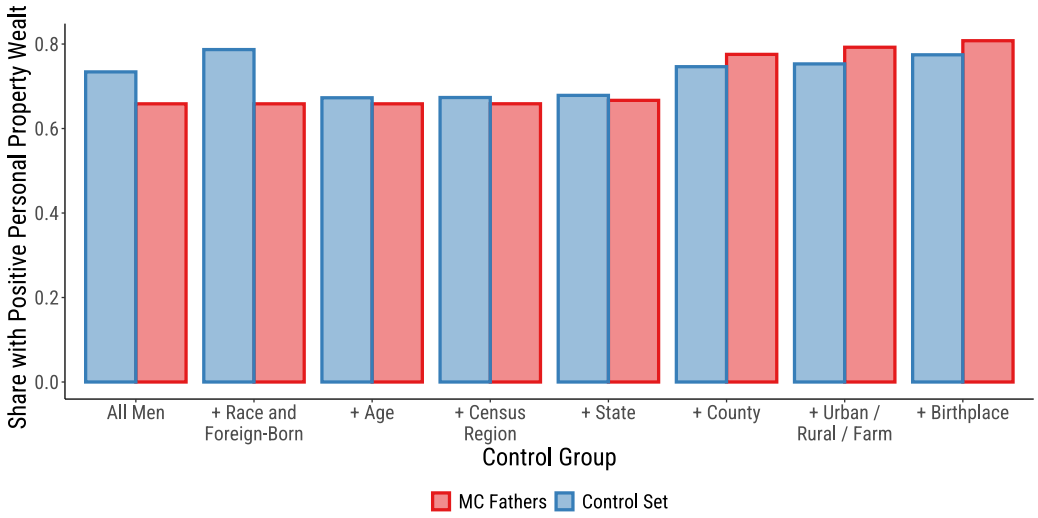
(b) Personal property wealth, conditional on being positive

Figure C.2: Future MCs versus matched controls, with progressively tighter matching, on mean real estate wealth among men with positive real estate wealth in 1850–1870 (a) and mean personal property wealth among men with positive personal property wealth in 1860–1870 (b).

Conditional on positive holdings, MC fathers report mean real estate of about $12,200 in every cell. Among All Men control fathers the conditional mean is $3,500, rising to $7,400 at the tightest cell. The personal-property gap follows the same pattern: MC-father means of $10,200 against $1,700 unmatched and $5,400 fully matched (Figure C.4). The unconditional gaps in MC-father wealth in Figure 8 are therefore almost entirely intensive: most fathers held some property, but those whose sons would go to Congress sat firmly in the right tail.



(a) Real estate wealth



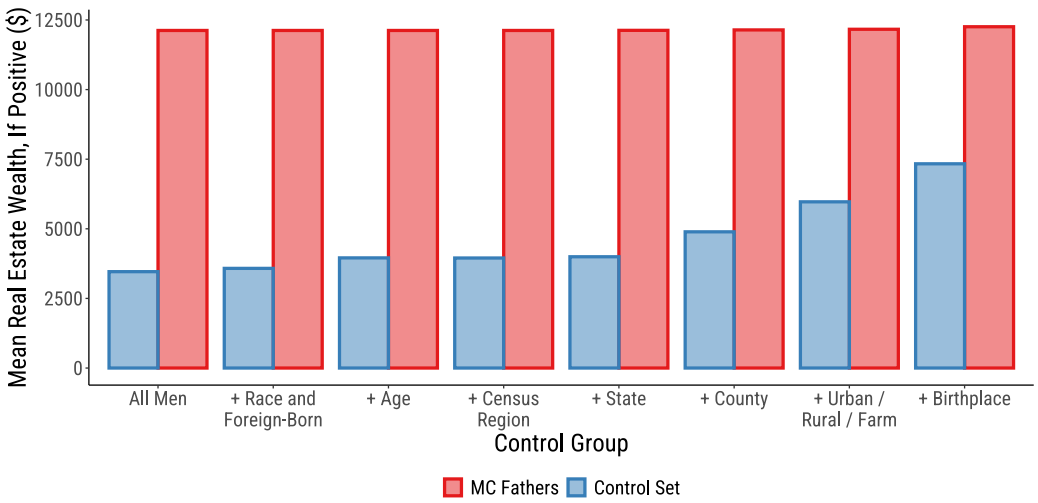
(b) Personal property wealth

Figure C.3: MC fathers versus matched control fathers, with progressively tighter matching, on the share with positive real estate wealth in 1850–1870 (a) and positive personal property wealth in 1860–1870 (b).

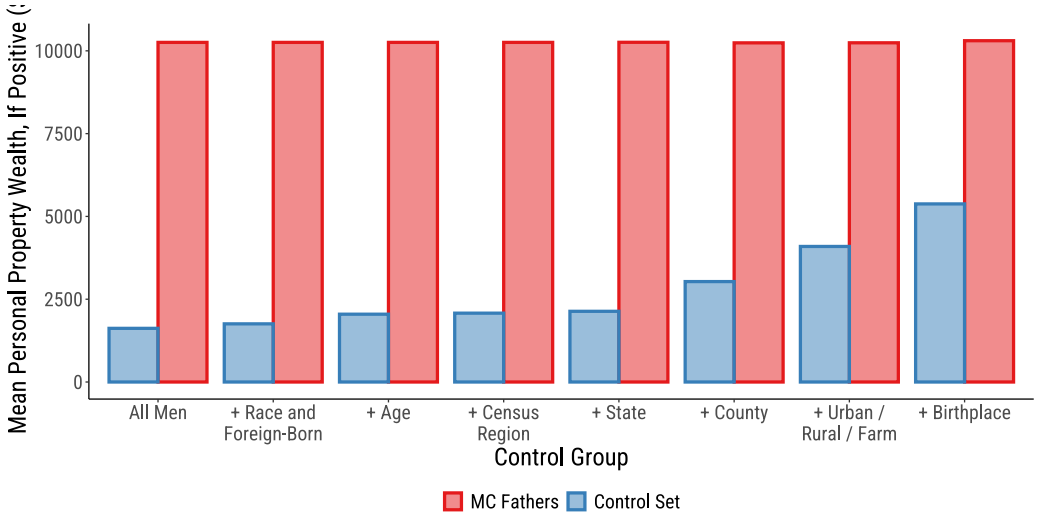
D Institutional Reforms: Additional Outcomes

This appendix presents the joint DDD specifications from Section 6 for the remaining occupational-score outcomes. Table D.1 reports the 3-reform joint specification (women's suffrage, secret ballot, direct primary) for Song score, Occscore, and the Collins–Zimran region-based wealth score; Table D.2 adds the 17th Amendment.[47] As in the main text, each pair of columns within an outcome reports

[47] As in the main table, direct election is negative across all six specifications in Table D.2 and marginally significant (at the 10% level) in two of the six: the state-by-decade-FE columns for the Song score and Occscore. The estimates do not survive the tighter MC-matched-set fixed effects.



(a) Real estate wealth, conditional on being positive



(b) Personal property wealth, conditional on being positive

Figure C.4: MC fathers versus matched control fathers, with progressively tighter matching, on mean real estate wealth among fathers with positive real estate wealth in 1850–1870 (a) and mean personal property wealth among fathers with positive personal property wealth in 1860–1870 (b).

the state-by-decade FE specification (col. 1) and the MC-matched-set FE specification (col. 2); baseline FE and standard errors are as in Table 8.

E Additional Figures and Robustness

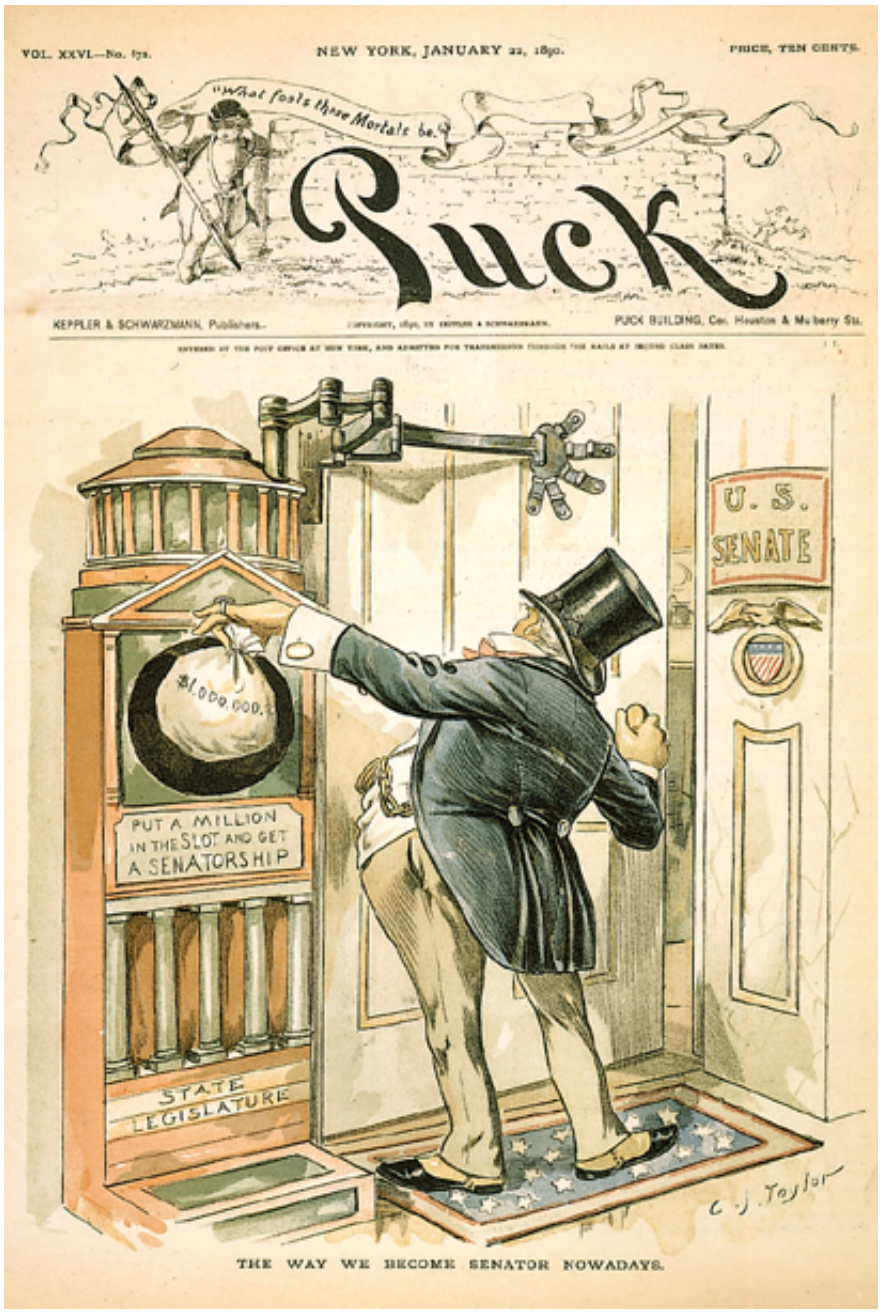
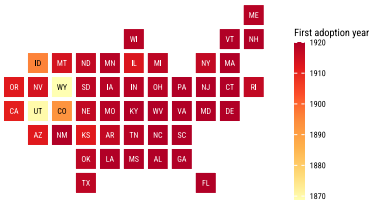
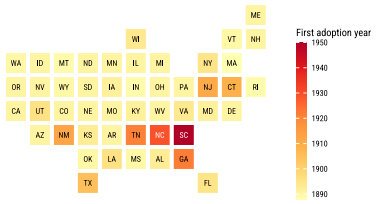


Figure D.1: "The Way We Become Senator Nowadays."

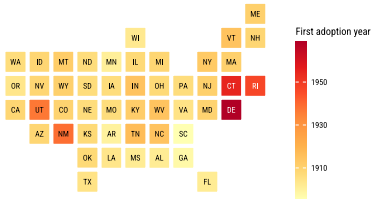
Note: Charles J. Taylor, *Puck*, January 22, 1890. U.S. Senate Collection, item 38_00598. https://www.senate.gov/art-artifacts/historical-images/political-cartoons-caricatures/38_00598.htm.



(a) Women's suffrage



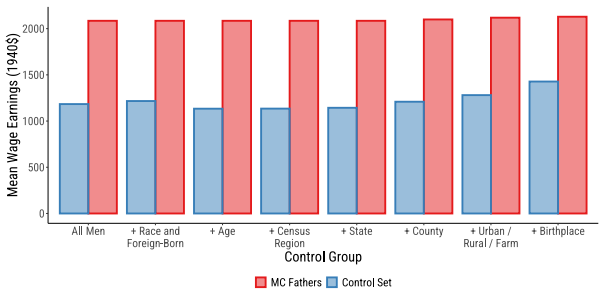
(b) Australian (secret) ballot



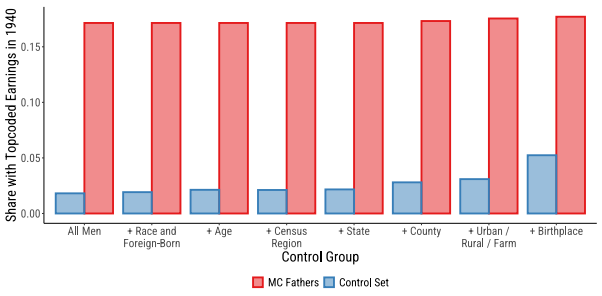
(c) Direct partisan primary

Figure D.2: State-level first-adoption year for the three state-varying Progressive-Era reforms.

Note: This figure plots the first-adoption year for each state on three Progressive-Era reforms: women's suffrage (a), the Australian/secret ballot (b), and the direct partisan primary (c). Color encodes adoption year. Women's-suffrage and secret-ballot adoption years are from Lott, Jr. and Kenny (1999), Table 1 ("Changes in Voting Laws, 1870–1940"); the direct-primary adoption year is from Hirano and Snyder Jr (2019). The suffrage column records the first enfranchisement, so the earliest adopters are the western states (Wyoming 1869, Utah 1870, Colorado 1893, Idaho 1896) and several states carry pre-Nineteenth-Amendment presidential-suffrage years rather than 1920. Two corrections to the published Lott, Jr. and Kenny (1999) table: Kentucky's secret-ballot year is set to 1892 rather than their 1882, which predates the Australian ballot (first adopted by Massachusetts in 1888) and Kentucky's 1891 abolition of *viva voce* voting; and Utah's suffrage year of 1870 reflects its territorial grant, revoked in 1887 and restored at statehood in 1896. The 17th Amendment is not shown: it is national rather than state-varying, ratified in 1913, and identified off chamber-by-period variation.



(a) Wage earnings



(b) Share with top-coded earnings

Figure E.1: MC fathers versus matched control fathers in 1940, across progressively tighter control sets: wage earnings (a) and share with top-coded earnings (b).

Table D.1: Effect of Women's Suffrage, Secret Ballot, and Direct Primary on Member of Congress Selection, 1880-1930 (Remaining Outcomes)

	Song Score		Occscore		CZ Wealth (region)	
	(1)	(2)	(3)	(4)	(5)	(6)
Member of Congress	12.00***	12.71***	7.49***	7.04***	2,131.27***	2,182.17***
	(1.17)	(0.72)	(0.94)	(0.66)	(179.09)	(183.17)
Member of Congress × Women's Suffrage	2.40	2.31	−0.87	−0.93	224.34	227.86
	(1.61)	(1.44)	(0.92)	(0.76)	(200.24)	(221.04)
Member of Congress × Secret Ballot	−1.83	−0.93	−1.49	−0.23	215.17	98.82
	(1.65)	(1.04)	(1.06)	(0.85)	(203.49)	(184.70)
Member of Congress × Partisan Primary	−0.49	−0.77	0.68	0.60	−169.84	−209.12
	(1.49)	(1.21)	(0.90)	(0.73)	(239.62)	(261.74)
Baseline FE	Yes	Yes	Yes	Yes	Yes	Yes
Chamber FE	Yes	Yes	Yes	Yes	Yes	Yes
State-Decade FE	Yes	No	Yes	No	Yes	No
MC-matched-set FE	No	Yes	No	Yes	No	Yes
Observations	473,534	473,518	473,425	473,409	473,534	473,518
Y Mean	59.3	59.3	25.9	25.9	3,169.6	3,169.5

Note: This table reports estimates of the effect of three Progressive Era reforms (women's suffrage, the secret ballot, and the direct primary) on the family background of MCs first elected between 1880 and 1930; it is the companion of Table 8 for the remaining occupation-based status outcomes. The sample comprises fathers of these MCs and the fathers' tightest matched controls. Outcomes are father-level economic status measures: the Song occupation score, IPUMS `OCCSCORE`, and the CZ wealth score aggregated to the Census region (vs. state in the main table). All three reform indicators enter the same regression interacted with the MC indicator; each reported coefficient is the marginal effect of one reform on the MC father–control gap, conditional on the other two. For each outcome, the first column includes state-by-decade fixed effects; the second subsumes those with MC-matched-set fixed effects. All columns include baseline fixed effects for father's age and race, MC's state, census year, first-election year, and chamber. Standard errors clustered at the state level are reported in parentheses.

Table D.2: Effect of Direct Election, Women's Suffrage, Secret Ballot, and Direct Primary on Member of Congress Selection, 1880-1930 (Remaining Outcomes)

	Song Score		Occscore		CZ Wealth (region)	
	(1)	(2)	(3)	(4)	(5)	(6)
Member of Congress	15.16***	14.85***	10.06***	8.81***	2,240.45***	2,295.97***
	(2.06)	(2.00)	(1.73)	(1.48)	(400.44)	(383.99)
Member of Congress × Women's Suffrage	2.50	2.37	−0.79	−0.87	227.71	231.37
	(1.63)	(1.46)	(0.92)	(0.76)	(204.36)	(224.51)
Member of Congress × Secret Ballot	−1.81	−0.91	−1.47	−0.22	215.76	99.82
	(1.66)	(1.03)	(1.06)	(0.85)	(203.38)	(184.27)
Member of Congress × Partisan Primary	−0.39	−0.71	0.75	0.65	−166.69	−205.93
	(1.48)	(1.20)	(0.88)	(0.73)	(240.46)	(262.50)
Member of Congress × Direct Election	−3.41*	−2.31	−2.78*	−1.92	−117.87	−123.09
	(1.85)	(1.86)	(1.44)	(1.33)	(416.21)	(395.82)
Baseline FE	Yes	Yes	Yes	Yes	Yes	Yes
Chamber FE	Yes	Yes	Yes	Yes	Yes	Yes
State-Decade FE	Yes	No	Yes	No	Yes	No
MC-matched-set FE	No	Yes	No	Yes	No	Yes
Observations	473,534	473,518	473,425	473,409	473,534	473,518
Y Mean	59.3	59.3	25.9	25.9	3,169.6	3,169.5

Note: This table reports estimates of the effect of four Progressive Era reforms (the direct election of senators, women's suffrage, the secret ballot, and the direct primary) on the family background of MCs first elected between 1880 and 1930; it is the companion of Table 9 for the remaining occupation-based status outcomes. The sample comprises fathers of these MCs and the fathers' tightest matched controls. Outcomes are father-level economic status measures: the Song occupation score, IPUMS `OCCSCORE`, and the CZ wealth score aggregated to the Census region (vs. state in the main table). All four reform indicators enter the same regression interacted with the MC indicator; each reported coefficient is the marginal effect of one reform on the MC father–control gap, conditional on the other three. The direct election effect is identified off chamber-by-period variation, while the other three are identified off state-by-period variation. For each outcome, the first column includes state-by-decade fixed effects; the second subsumes those with MC-matched-set fixed effects. All columns include baseline fixed effects for father's age and race, MC's state, census year, first-election year, and chamber. Standard errors clustered at the state level are reported in parentheses.

Table E.1: How much of the status gap between future MCs and control men remains within occupations?

	$\hat{\beta}$ (MC)	$\hat{\beta}$ (MC | Occ)	% Missing	Scaling
1940 earnings				
1940 wage earnings	1,201***	585.6***	48.7%	1.95×
	(46.6)	(45.7)		
1940 log wage earnings	1.02***	0.475***	46.8%	1.88×
	(0.026)	(0.033)		
20th-century wealth				
1940 home value rank	0.073***	0.034***	46.0%	1.85×
	(0.0031)	(0.0031)		
1940 log home value	1.13***	0.852***	75.2%	4.03×
	(0.036)	(0.056)		
1930 home value rank	0.064***	0.032***	50.7%	2.03×
	(0.0022)	(0.0021)		
1930 log home value	1.07***	0.835***	78.3%	4.60×
	(0.027)	(0.041)		
1900 1{2+ servants}	0.124***	0.105***	84.7%	6.52×
	(0.0059)	(0.0064)		
1910 1{2+ servants}	0.108***	0.090***	83.8%	6.16×
	(0.0056)	(0.0061)		
1920 1{2+ servants}	0.077***	0.067***	86.7%	7.50×
	(0.0049)	(0.0054)		
1930 1{2+ servants}	0.072***	0.072***	100.4%	-267.14×
	(0.0048)	(0.0058)		
19th-century wealth				
1870 log real property	1.67***	1.44***	86.0%	7.12×
	(0.043)	(0.059)		
1870 log personal property	2.06***	1.72***	83.7%	6.14×
	(0.044)	(0.060)		
1870 log total wealth	2.23***	1.77***	79.2%	4.81×
	(0.041)	(0.058)		
1860 log real property	1.71***	1.44***	84.5%	6.44×
	(0.041)	(0.060)		
1860 log personal property	2.27***	1.73***	76.1%	4.19×
	(0.044)	(0.060)		
1860 log total wealth	2.45***	1.82***	74.2%	3.87×
	(0.044)	(0.058)		
1850 log real property	1.57***	1.43***	91.1%	11.24×
	(0.053)	(0.074)		

Note: This table reports how much of the status gap between future MCs and matched control men survives within occupation (controlling for three-digit occupation fixed effects); it is the companion of Table 3 for future MCs in place of MC fathers. Each row corresponds to a different outcome and reports a separate pair of regressions on the future-MC estimation sample, with controls weighted by $1/n_{\text{controls}}$ within each MC–(bioguide) pair. Column (1) reports the coefficient on the future-MC indicator from a regression with bioguide-pair fixed effects. Column (2) adds occ1950 fixed effects. Column (3) is the ratio of column (2) to column (1), the share of the raw future-MC gap that remains within occupation. Column (4) is the ratio of column (1) to [column

Table E.2: How much of the status gap between MC brothers and control men remains within occupations?

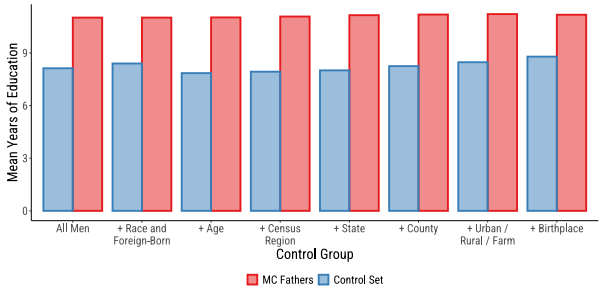
	$\hat{\beta}$ (Brother)	$\hat{\beta}$ (Brother | Occ)	% Missing	Scaling
1940 earnings				
1940 wage earnings	387.4***	216.5***	55.9%	2.27×
	(29.0)	(22.8)		
1940 log wage earnings	0.505***	0.222***	44.1%	1.79×
	(0.025)	(0.021)		
20th-century wealth				
1940 home value rank	0.051***	0.030***	59.4%	2.46×
	(0.0028)	(0.0025)		
1940 log home value	0.687***	0.469***	68.3%	3.16×
	(0.036)	(0.035)		
1930 home value rank	0.031***	0.017***	54.7%	2.21×
	(0.0017)	(0.0015)		
1930 log home value	0.572***	0.340***	59.4%	2.46×
	(0.030)	(0.030)		
1900 1{2+ servants}	0.058***	0.048***	81.7%	5.46×
	(0.0047)	(0.0044)		
1910 1{2+ servants}	0.059***	0.045***	77.4%	4.42×
	(0.0048)	(0.0044)		
1920 1{2+ servants}	0.043***	0.038***	87.1%	7.75×
	(0.0045)	(0.0046)		
1930 1{2+ servants}	0.041***	0.035***	85.3%	6.79×
	(0.0045)	(0.0042)		
19th-century wealth				
1870 log real property	0.930***	0.763***	82.0%	5.56×
	(0.037)	(0.036)		
1870 log personal property	1.03***	0.739***	71.9%	3.56×
	(0.039)	(0.035)		
1870 log total wealth	1.17***	0.855***	73.2%	3.73×
	(0.041)	(0.036)		
1860 log real property	0.966***	0.794***	82.2%	5.62×
	(0.041)	(0.041)		
1860 log personal property	1.22***	0.853***	70.0%	3.33×
	(0.044)	(0.042)		
1860 log total wealth	1.34***	0.928***	69.4%	3.27×
	(0.046)	(0.044)		
1850 log real property	0.824***	0.680***	82.6%	5.74×
	(0.043)	(0.044)		

Note: This table reports how much of the status gap between MC brothers and matched control men survives within occupation (controlling for three-digit occupation fixed effects); it is the companion of Table 3 for MC brothers in place of MC fathers. Each row corresponds to a different outcome and reports a separate pair of regressions on the MC-brother estimation sample, with controls weighted by $1/n_{\text{controls}}$ within each MC (bioguide) pair. Column (1) reports the coefficient on the MC-brother indicator from a regression with brother–control pair fixed effects. Column (2) adds occ1950 fixed effects. Column (3) is the ratio of column (2) to column (1), the share of the raw MC-brother gap that remains within occupation. Column (4) is the ratio of

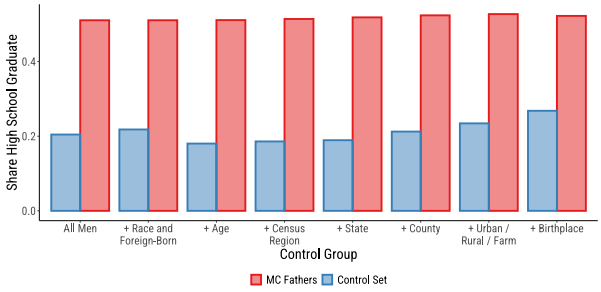
Table E.3: Comparing occupation-based status scores between MC fathers and the population of fathers, 1940

Outcome	MC Fathers	All Fathers
Occupation scores		
Adjusted Song score	71.3	51.5
LIDO	30.2	23.0
CW score	$1,822	$1,227
CZ wealth (state)	$5,548	$2,851
Occupation score	33.8	25.0
Song score	66.4	43.1
CZ wealth (region)	$5,993	$2,971
Sample size		
N	1,248	20,664,640

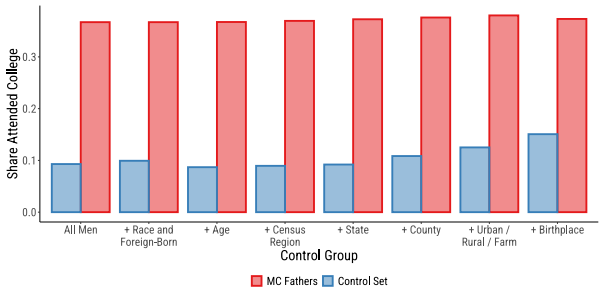
Note: This table compares occupation-based status scores for MC fathers to the broader population of fathers in the 1940 U.S. complete-count census; it is the companion of Table 2 for occupation scores in place of the wage and home-value outcomes reported there. The All Fathers sample comprises 1940 census men with at least one co-resident child, identified via the IPUMS POPLOC parent variable. The MC Fathers sample comprises father records linked to MCs through FamilySearch and matched into the 1940 census. The Occupation Score is IPUMS `OCCSCORE` (Sobek 1995); the Song Score (Song et al. 2020) and its Adjusted variant (Ward 2023) are 0–100 occupational rank scales; LIDO (Saavedra and Twinam 2020) is a LASSO-predicted income score in hundreds of 1950 dollars; the CW score (Collins and Wanamaker 2022) is the Collins–Wanamaker 1940 occupational income score in 1940 dollars; and the CZ wealth scores (Collins and Zimran 2023) are Collins–Zimran 1870 occupational wealth scores (state and region variants) in 1870 dollars. See Section 2 for full definitions.



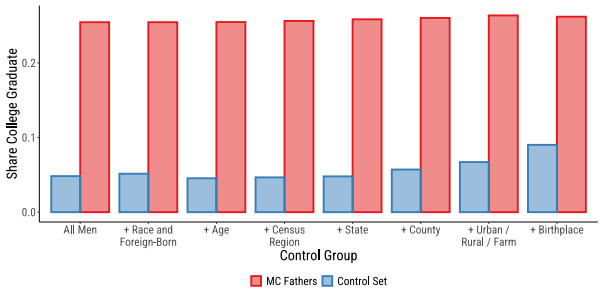
(a) Years of education



(b) High-school completion

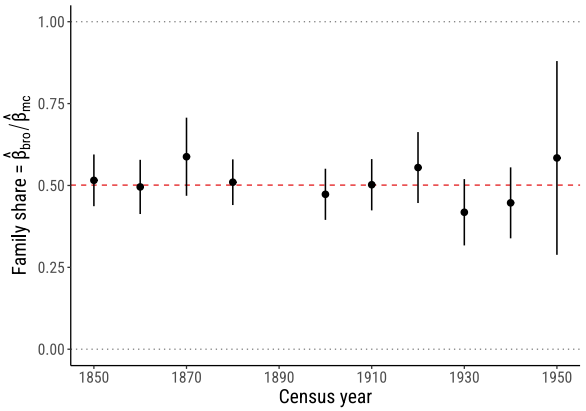


(c) College attendance

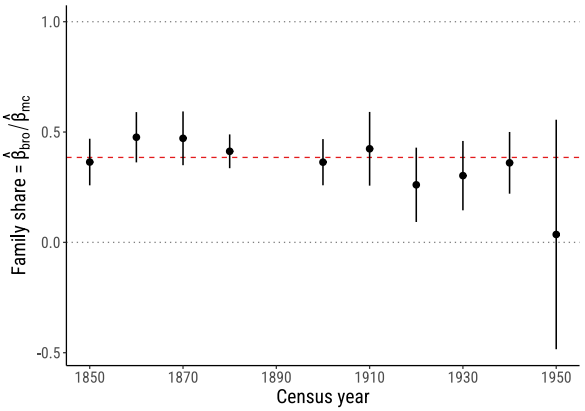


(d) College completion

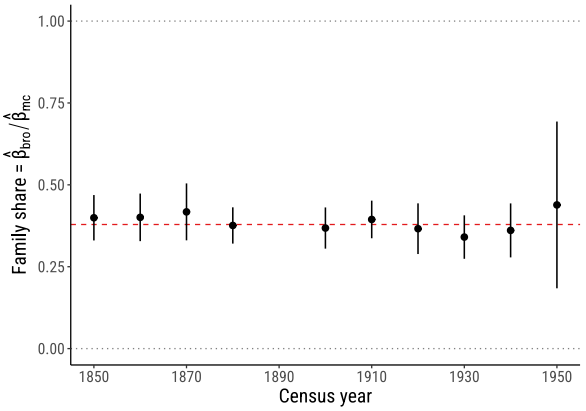
Figure E.2: MC fathers versus matched control fathers in 1940, across progressively tighter control sets: years of education (a), high-school completion (b), college attendance (c), and college completion (d).



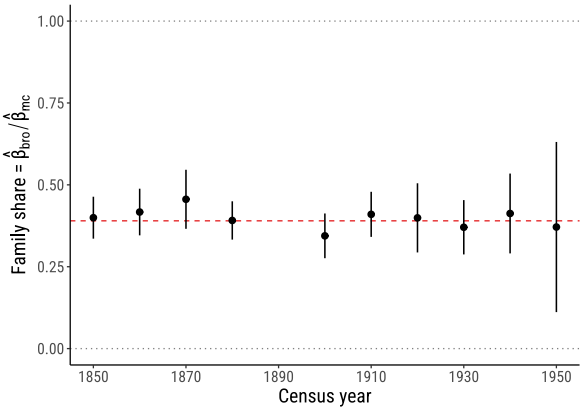
(a) Adjusted Song occupation score



(b) LIDO score

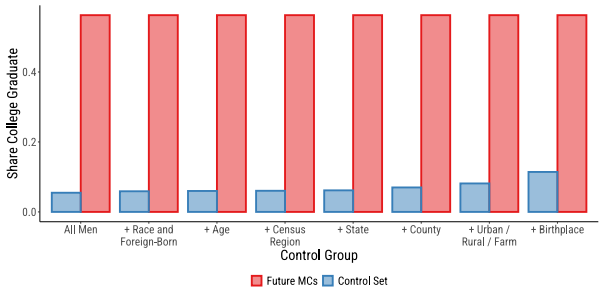


(c) CW 1940 occupation income score

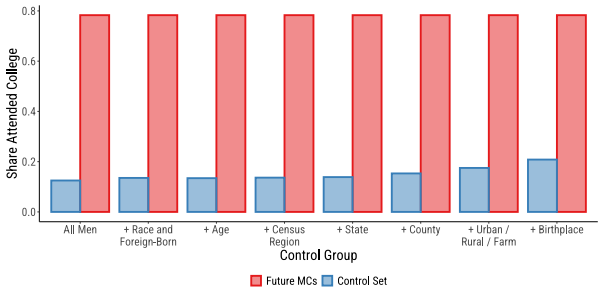


(d) CZ 1870 occupational wealth score (state)

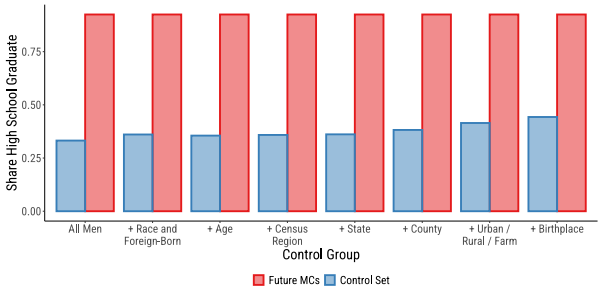
Figure E.3: Within-family share $\hat{\beta}_{\text{bro}}/\hat{\beta}_{\text{mc}}$ by census year, for four occupation-based scores. Dashed red line is the pooled (1850–1950) family share from Table 7.



(a) College completion

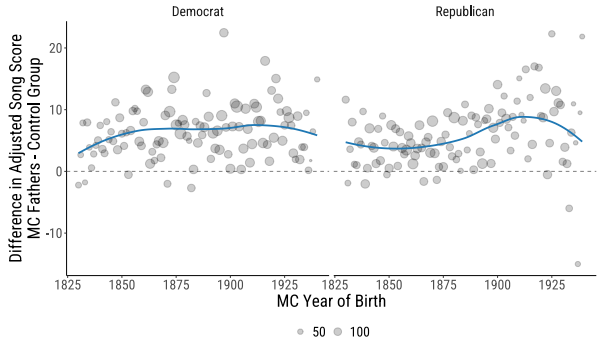


(b) College attendance

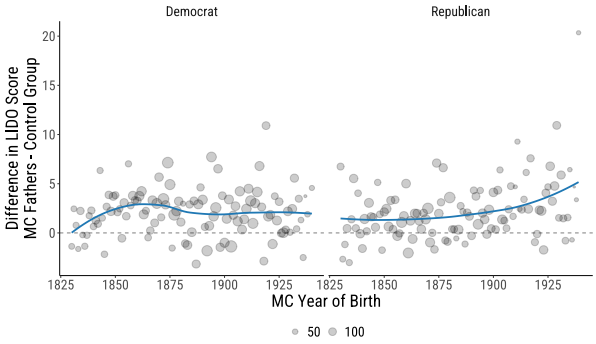


(c) High-school completion

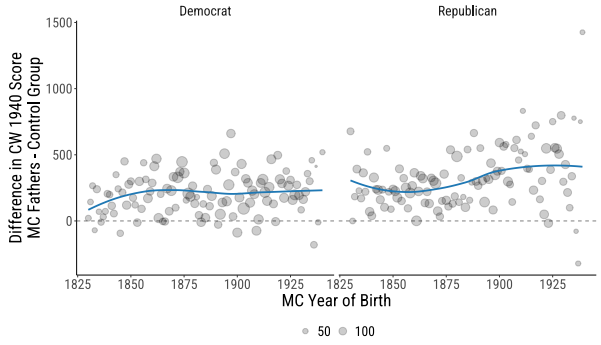
Figure E.4: Future MCs versus matched controls in 1940, across progressively tighter control sets: college completion (a), college attendance (b), and high-school completion (c).



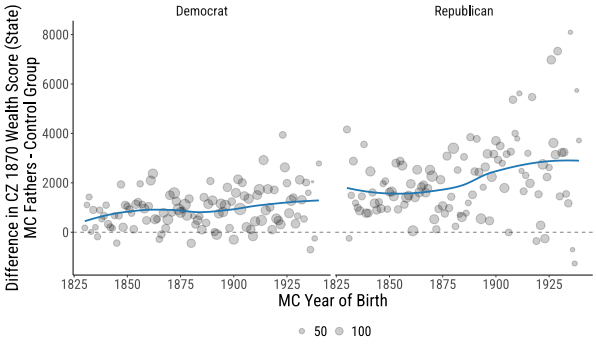
(a) Adjusted Song occupation score



(b) LIDO score

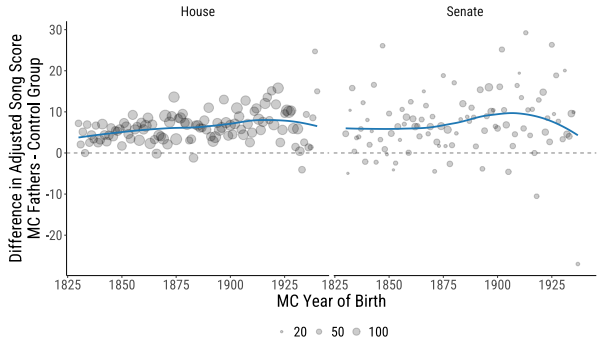


(c) CW 1940 occupation income score

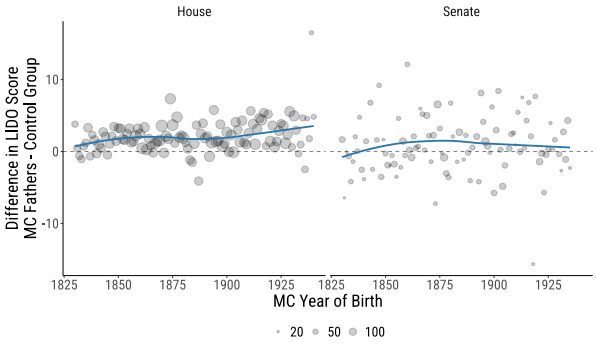


(d) CZ 1870 occupational wealth score (state)

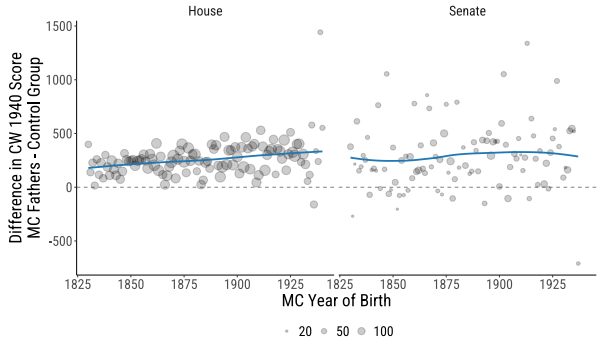
Figure E.5: Family-background selection by MC party: MC fathers versus the tightest set of control fathers, by MC birth cohort.



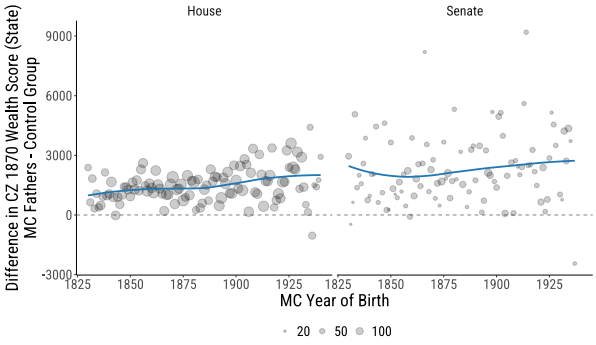
(a) Adjusted Song occupation score



(b) LIDO score

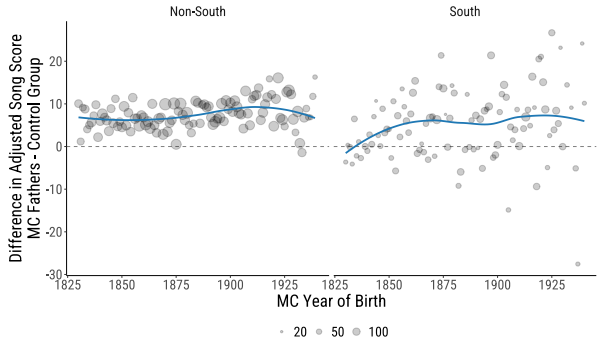


(c) CW 1940 occupation income score

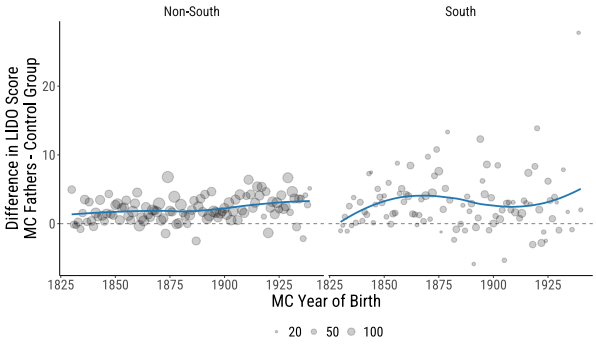


(d) CZ 1870 occupational wealth score (state)

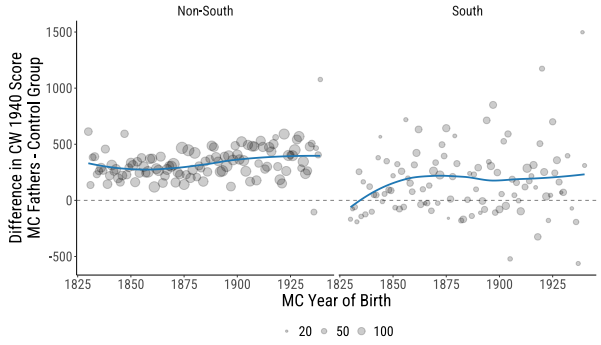
Figure E.6: Family-background selection by chamber: MC fathers versus the tightest set of control fathers, by MC birth cohort.



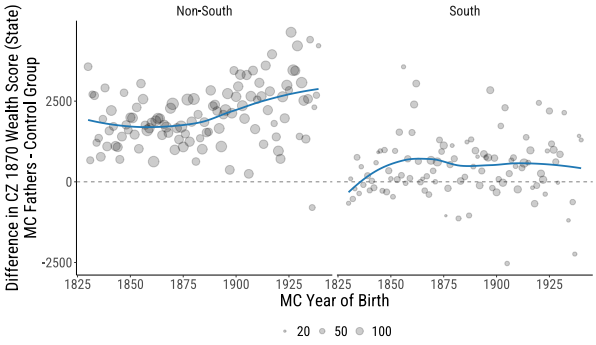
(a) Adjusted Song occupation score



(b) LIDO score



(c) CW 1940 occupation income score



(d) CZ 1870 occupational wealth score (state)

Figure E.7: Family-background selection by region (South vs. non-South): MC fathers versus the tightest set of control fathers, by MC birth cohort.

Table E.4: Family-background selection of future MCs by subsample: levels

	Panel A: Adj. Song Score						
		Party		Region		Chamber	
	All	Dem	Rep	South	Non-South	Senate	House
Adj. Song Score	8.41***	8.29***	8.62***	10.08***	7.85***	9.12***	8.24***
	(0.31)	(0.47)	(0.40)	(0.74)	(0.33)	(0.76)	(0.32)
Diff		0.33		-2.23**		-0.88	
		(0.61)		(0.81)		(0.80)	
Observations	51,975	27,342	24,302	13,212	38,763	9,974	42,001
Y Mean	8.6	8.6	8.7	9.9	8.2	9.6	8.4
	Panel B: LIDO Score						
LIDO	3.30***	3.36***	3.27***	3.95***	3.08***	3.90***	3.15***
	(0.16)	(0.23)	(0.22)	(0.36)	(0.17)	(0.40)	(0.16)
Diff		-0.09		-0.88*		-0.75	
		(0.32)		(0.40)		(0.41)	
Observations	48,465	25,473	22,683	12,257	36,208	9,298	39,167
Y Mean	3.2	3.3	3.0	3.6	3.0	3.8	3.0
	Panel C: CW 1940 Score						
CW 1940 Score	330.48***	311.41***	355.05***	346.63***	325.10***	383.47***	317.69***
	(11.74)	(16.66)	(16.48)	(26.47)	(12.93)	(29.99)	(11.96)
Diff		43.64		-21.53		-65.78*	
		(23.39)		(29.45)		(31.22)	
Observations	51,459	27,081	24,048	13,117	38,342	9,901	41,558
Y Mean	336.0	314.3	363.4	343.7	333.3	391.0	322.9
	Panel D: CZ 1870 Wealth Score (state)						
CZ 1870 (state)	1,730.21***	1,354.11***	2,167.98***	1,020.63***	1,967.25***	1,861.53***	1,698.66***
	(69.66)	(85.59)	(111.96)	(86.20)	(87.65)	(156.76)	(75.24)
Diff		813.87***		946.62***		-162.87	
		(140.75)		(122.92)		(169.14)	
Observations	51,503	26,979	24,193	13,205	38,298	9,864	41,639
Y Mean	1,775.7	1,408.7	2,198.6	998.0	2,043.8	1,924.8	1,740.3

Note: This table reports the average gap between MC fathers and their tightest matched controls on four occupation-based status measures, splitting the samples by MC party, region, and chamber. The sample restricts to MC fathers observed at ages 30–65 with MCs born 1830–1940. Each panel uses a different outcome: Panel A is the Adjusted Song Score, Panel B is the LIDO Score, Panel C is the Collins–Wanamaker 1940 Score, and Panel D is the Collins–Zimran 1870 wealth score at the state level. Each cell reports an MC-father estimation-sample regression of the MC-father minus control-father gap on a constant, weighted by $1/n_{\text{controls}}$ within each MC (bioguide) pair. Multi-term MCs contribute one observation per congress term and so enter multiple subsamples weighted by terms served; the "All" column uses the same per-term expansion and so matches the cuts. The "Diff" row reports the right-minus-left difference in means within each pair (Rep − Dem, Non-South − South, House − Senate), estimated by pooling the two columns and regressing the gap on a right-group indicator. Standard errors clustered at the bioguide level are reported in parentheses.