Three Essays on the Substance and Methods of Economic History

by

Morgan Henderson

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2017

Doctoral Committee:

Associate Professor Martha J. Bailey, Chair
Assistant Professor Achyuta Adhvaryu
Assistant Professor Mark Dincecco
Professor Paul W. Rhode

Morgan Henderson

morghend@umich.edu

ORCID iD: 0000-0002-0869-5738

Acknowledgements

I would like to thank my advisor, Martha Bailey, for her unfailing support and enthusiasm; the rest of my dissertation committee –Paul Rhode, Mark Dincecco, and Ach Adhvaryu – for their patience and generosity; Elyce Rotella, for her data and friendship; Laura Flak and the rest of the administrative staff for their help with matters great and small; Warren Whatley, for his mentoring early in my graduate career; all faculty in the economic history and labor groups; and my family – Helen, Ian, Susanna, and Andre. Two especially large debts are to Nick Henderson and Morgane Mouslim, without whom I would not be where I am today.

Table of Contents

List of Tables

List of Figures

Abstract

This dissertation explores questions on the substance and methods of economic history. Chapter one studies a little-known policy change in the late 19th and early 20th centuries to explore the causal effects of political exclusion on the economic wellbeing of immigrants. Starting in the mid-19th century, twenty-four states and territories expanded their electorates to allow non-citizen immigrants the right to vote; from 1864-1926, however, these same jurisdictions reversed this policy, creating a mass disenfranchisement for which the timing varied across states. Using this variation as well as a discontinuity in nationalization proceedings of the era, I find that political exclusion led to a 25-60% reduction in the likelihood that affected immigrants obtained public sector employment. I also document significant negative intergenerational effects: individuals of immigrant parentage born around the time of disenfranchisement earned 5-9% less as adults than comparable individuals of native parentage. I am able to rule out as mechanisms for this intergenerational effect a variety of policy and spending channels, but find evidence for a reduction in English-language proficiency among disenfranchised immigrants, which may have adversely affected the human capital of their children.

Chapter two explores the *causes* of the adoption and repeal of alien voting in the United States. This policy shift offers a valuable opportunity to understand the forces determining political inclusion and exclusion in a formative period of American democracy, and contributes to the broader literature on theories of democratization. I use qualitative evidence from the historical record to outline competing theories of both

adoption and repeal of alien voting, and then rationalize these hypotheses within the context of a median voting model. Using a discrete time hazard specification, I find evidence consistent with the hypothesis that states used alien voting as a locational amenity, with the objective of inducing immigrant in-migration in order to foster agricultural development. The results indicate that the timing of repeal was driven by social costs, rather than economic or political factors, although there is evidence for heterogeneity in correlates of support for repeal across states. Finally, the costs of constitutional change were salient for both adoption and repeal: states for which it was less costly to re-write or amend the constitution were more likely to adopt and repeal alien voting.

Chapter three is a co-authored methodological study intended to assess the efficacy of commonly used techniques to create name-linked historical datasets. The recent digitization of historical microdata has led to a proliferation of research using linked data, in which researchers use various methods to match individuals across datasets by observable characteristics; less is known, however, about the quality of the data produced using those different methods. Using two hand-linked ground-truth samples, we assess the performance of four automated linking methods and two commonly used name-cleaning algorithms. Results indicate that automated methods result in high rates of false matches – ranging from 17 to over 60 percent – and the use of phonetic name cleaning increases false match rate by 60-100 percent across methods. We conclude by exploring the implications of erroneous matches for inference, and estimate intergenerational income elasticities for father-son pairs in the 1940 Census using samples generated by each method. We find that estimates vary with linking method, suggesting that caution must be used when interpreting parameters estimated from linked data.

Chapter 1

The Economic Consequences of Immigrant Disenfranchisement: Evidence from the United States

What are the effects of disenfranchisement on immigrants? This paper studies a little-known episode in United States history in which twenty-four states and territories disenfranchised non-citizen immigrants from 1864 to 1926. This represented a significant shock to the political equilibrium of the time: mayoral and gubernatorial voting activity fell by approximately 10 and 5 percentage points, respectively. Moreover, disenfranchisement had real economic effects: the likelihood of having public sector employment fell by 25-60% for affected immigrants. There is no measured effect on other labor market outcomes, but it is possible this reflects the coarseness of the available data, rather than an absence of true impact. Consistent with this, I document significant negative intergenerational effects. Individuals of immigrant parentage born around the time of disenfranchisement earned 5-9% less as adults than comparable individuals of native parentage. I am able to rule out as mechanisms a variety of policy and spending outcomes, but find suggestive evidence for a reduction in the human capital of the children of immigrants as mediated through parental English-language proficiency.

1.1 Introduction

In 2013, the Supreme Court struck down a key provision of the 1965 Voting Rights Act requiring certain jurisdictions to obtain federal permission – "pre-clearance" - in order to amend their voting laws.[1] This decision occurred in the midst of a broad movement towards restricting the franchise: since 2010, twenty states have implemented measures that have increased the costs of voting along several dimensions.[2] Given the vast body of research linking political institutions and economic development (La Porta et al. 1997, Acemoglu et al. 2001), it is possible that this

[1] *Shelby County v Holder*, 70 U.S. (2013)

[2] Measures include requiring photo ID in order to vote, reduction of early voting periods, restrictions on voter registration drives, felon disenfranchisement, and mandating documentation of citizenship (Brennan Center 2016).

movement could have adverse economic consequences for affected individuals or for the economy more broadly.

However, in contrast to the large literature on the effects of *en*franchisement, there has been less work on the consequences of *dis*enfranchisement.[3] This is likely a function of the progressive march of history: while the disenfranchisement of African-Americans in the Jim Crow South provides a notable exception, developed nations are more characterized by extensions, rather than contractions, of the franchise. This study addresses this gap using a little-studied series of state-level constitutional changes that, from 1864-1926, repealed the right to vote for non-citizen immigrants in twenty-four states and territories (Keyssar 2000, Hayduk 2006).[4] These changes directly disenfranchised recent immigrants who were not yet eligible for citizenship and, among other immigrants who were not citizens, increased the cost of becoming eligible to vote by requiring naturalization. I exploit the timing of these repeals to estimate the effects of disenfranchisement on the economic wellbeing of immigrants and their children.

There are several theoretical channels through which losing the right to vote may affect the economic lives of immigrants. Vote-maximizing public officials may direct expenditure away from disenfranchised immigrants, toward more electorally valuable constituencies. To the extent that public spending is responsive, the adverse effects may be borne by either the immigrants themselves (if, for example, public employment was redirected) or their children

[3] Papers that have studied the effects of enfranchisement include Miller 2008, Lott and Kenny 1999, Fujiwara 2015, Hoffman et al. 2016, Carruthers and Wannamaker 2015, Funk and Gathman 2006, Aidt et al. 2006, Aidt and Dallal 2008, Cascio and Washington 2014, Hinnerich and Petterson-Lidbom 2014, Hodler et al. 2015, Kose et al. 2015. Naidu (2012) addresses the impact of disenfranchisement in the Jim Crow south, and is the sole recent work I have been able to locate on the topic.

[4] One legal scholar declared the “history of alien suffrage” as being “largely unwritten” (Raskin 1993). While I know of no economic research that has explored the effects of alien (dis)enfranchisement in the United States, Vernby (2013) explores the effects of Swedish alien suffrage on public spending.

(through, for example, reductions in education spending). Direct effects on labor markets are possible: employers may react to the reduced political voice of the disenfranchised group and offer lower wages to affected employees, and the disenfranchised may out-migrate if the right to vote is a strong enough locational amenity. Finally, formal civic exclusion may have reduced the benefits of social and cultural assimilation, leading to less integration into labor markets or reduced human capital accumulation of the children of immigrants.

Econometrically, the challenge of estimating the consequences of disenfranchisement is to distinguish the effects from the underlying causes. If unobserved forces drove the repeal of alien voting, then it is possible that any estimated effects could be a result of these forces rather than disenfranchisement *per se*. In this setting, this concern is mitigated by the fact that these disenfranchisements were constitutional, rather than statutory. States vary significantly in the process of amending their constitutions, and I present evidence that heterogeneity in the amending procedures generated an element of exogeneity in the timing of the disenfranchisements (Tarr 1996),

I test for the relevance of the disenfranchisements in two ways. First, I use newly assembled data on elections from 1870-1940 to assess the effect of disenfranchisement on political participation. Using an event-study specification that leverages the staggered timing of these disenfranchisements, I find that disenfranchisement led to significant reductions in voting activity: effect sizes are 2-4, 5-10, and 10-20 percentage points for presidential, gubernatorial, and mayoral elections, respectively. These are robust to multiple specifications, and the timing of effects suggests a causal role for disenfranchisement. Moreover, there is evidence of substantial spillovers in voting: the estimated impact on vote totals in mayoral and gubernatorial elections exceeded the magnitude of the disenfranchised population by a factor of at least two.

Second, using Census samples from 1900-1930, I estimate the effect of disenfranchisement on citizenship status. Whereas, prior to disenfranchisement, immigrants that had only *begun* citizenship proceedings – that is, filed “first papers” - had been allowed to vote, these disenfranchisements restricted the electorate to immigrants that had *completed* citizenship proceedings. Without the benefit of voting eligibility, immigrants that never intended to complete citizenship proceedings may have had less incentive to initiate those proceedings; therefore, disenfranchisement should have led to a reduction in the average incidence of filing first papers. However, if monitoring was costly in elections from this era, or if immigrants were not politically engaged, then the disenfranchisements may have had little actual impact (Allen and Allen 1981).

I test for this using a specification that exploits a discontinuity in the naturalization procedure of the era. Immigrants were required to maintain five years of continuous residence in the U.S. in order to qualify for citizenship, but could file first papers at any time after arrival (Department of Commerce and Labor 1909; Muir 1898). This delineates a pool of immigrants – those with fewer than five years of residence – that were rendered ineligible to vote following disenfranchisement, and allows for a triple-difference specification that compares outcomes for recent and less recent immigrants around this cutoff, before and after repeal, across states that did and did not disenfranchise immigrants. This specification nets out unobserved shocks common to *all* immigrants, and I find evidence that disenfranchisement affected the decision to initiate citizenship: the incidence of filing first papers fell by a statistically significant 50-54%. Taken together with results of disenfranchisement on aggregate voting activity, this suggests that the repeal of alien voting represented a significant shock to the political equilibrium of the era.

Using this same triple-difference specification, I estimate the short-run effect of disenfranchisement on labor market outcomes from 1900-1930. I find a 25-60% reduction in the likelihood of holding public sector employment for recent immigrants relative to the baseline mean, and this is robust to alternative specifications and definitions of public employment. Patronage employment - in which government employment was traded for electoral support - was legal for most states and cities in this era, and there is evidence that public-sector jobs paid significantly higher wages than comparable private sector positions (Rauch 1995, Troesken 1998). These results imply that politicians responded to disenfranchisement by shifting targetable public spending away from politically expendable constituencies.

I find null effects on other labor market outcomes, but this may reflect data availability rather than a lack of true impact. Census microdata samples with immigration covariates are only available from 1900-1930, which spans only part of the time period in which states disenfranchised immigrants, and these samples contain only relatively coarse measures of labor market outcomes.[5] Unobserved effects, however, may be manifested in the next generation; therefore, as a check, I estimate *intergenerational* effects of disenfranchisement. Using variation in paternal nativity, birthplace, and birth year in the 1940 full count Census and Census samples from 1950 and 1960, I estimate cohort-based event study specifications and find evidence for the intergenerational impacts of disenfranchisement. Adult males of immigrant parentage who were exposed to disenfranchisement at a young age obtained less education, and earned 5-9% less as adults, than comparable individuals of native parentage. These latter results are statistically significant, and are robust to alternative specifications and placebo checks.

[5] Wage and salary income, in particular, was not collected until the 1940 Census.

Controlling for educational attainment explains only approximately ¼ of the observed intergenerational effect on income. What else might drive this result? I test for environmental mechanisms by estimating the effect of disenfranchisement on a variety of aggregate demographic and political outcomes. I find no evidence for effects on migration, state-level political control, state- or city-level public spending, legislative preferences as captured by the Poole-Rosenthal index, or state-level Congressional support for federal immigration restrictions from 1897-1917. I do find suggestive evidence consistent with reduced social assimilation as a result of disenfranchisement: following repeal, immigrants were 7.6-11.0% less likely to speak English. Given the evidence on the importance of parental language skills on outcomes for children, this result suggests that reduced cultural assimilation and a consequent reduction in human capital for children may have been a factor driving the intergenerational effect (Bleakly and Chin 2008; Casey and Dustmann 2008).

This paper contributes to three literatures. First, it augments what is known about the effects of changes in electorate on policies and outcomes. As noted, there is a voluminous body of literature on the effects of *en*franchisement of various groups on voter turnout, public spending, infant mortality, and long-run educational outcomes, but much less on the effects of *dis*enfranchisement (Naidu 2012, Kousser 1974). Second, it contributes to the literature on immigrant assimilation in the early 20th century. The United States experienced a massive wave of immigration from 1850-1930, and while scholars are beginning to understand the processes of economic and social assimilation of these immigrants, relatively little is known about the *political* assimilation of immigrants in this era (Abramitzky and Boustan 2016; Shertzer 2016; Bandiera et al. 2016). Moreover, that disenfranchisement has adverse impacts on the children of immigrants suggests that assimilation may be a more complex process than previously thought.

Finally, it adds to the literature on the political economy of public finance in the late 19th and early 20th centuries. The research is mixed on the extent to which public spending was politically motivated in this era (Eli and Salisbury 2015; Wallis 1998), and this paper provides new evidence that government employment was highly responsive to voting eligibility.

1.2 Historical Context and Causes of Disenfranchisement

1.2.1 *Historical Context*

Starting with Wisconsin in 1848, twenty-three other states and territories from 1848-1890 adopted provisions in their state or territorial constitutions that allowed non-citizen immigrants the right to vote in local, state, and federal elections (Hayduk 2006, Keyssar 2000). The pattern of adoption was regional, and proceeded largely in three waves: Midwestern states and territories enacted these provisions in the late 1840s and 1850s; Western territories during the Civil War, and Southern states following the Civil War. Qualitative and quantitative evidence suggests a development motivation: labor-scarce states with relatively undeveloped economies used alien voting as a locational amenity, intended to stimulate in-migration (Henderson 2017; Neumann 1992; Raskin 1993).

An important feature of these provisions is that they were intended as a "pathway to citizenship": they enfranchised those non-citizen immigrants who had initiated citizenship proceedings by filing a "Declaration of Intent". During this era, naturalization was a two-part process in which first, the would-be citizen filed a Declaration of Intent – also known as "first papers" – that stated his intention to become a citizen of the United States. This declaration was not binding, and could be filed at any point following arrival at any "competent court", of which there were approximately 5,000 nation-wide (Rosberg 1977). No less than two years after filing first papers, and after a total period of residence of no less than five years, the applicant could file a petition for citizenship. Obtaining citizenship was not a requirement in this era, and many immigrants never even initiated the citizenship process: among adult male immigrants who had been in the United States for at least five years, the fraction who were *not* citizens was 30% in 1900, 34% in 1910, 48% in 1920, and 35% in 1930 (Ruggles 2015).[6]

States and territories began to repeal alien voting provisions starting with Nevada Territory in 1864, when it drafted a state constitution and was admitted to the Union (see figure 1.1 for a map of the dates of repeal by state). South Carolina followed in 1868, with its reconstruction constitution, and then Idaho Territory in 1874 and Georgia in 1877; ten states and territories repealed their provisions in either 1889 or the 1890s; three in the first decade of the 20th century; and eight after the outbreak of World War I. I approximate the magnitude of the shock by estimating the fraction of the adult male population that was non-citizen at the time of disenfranchisement using Census samples from 1870-1930 (see table 1.1). The fraction of the voting population that was affected varied by state, ranging from under 1% in Alabama, Arkansas, and Georgia to over 20% in North Dakota and Nevada. Dates of repeal are from Hayduk (2006), Keyssar (2000), the NBER/Maryland state constitution database, state blue books, legislative manuals, and other administrative reports.

1.2.2 *Causes of Disenfranchisement*

Legal scholars generally agree that these non-citizen voting provisions were repealed largely due to rising anti-immigrant sentiment in the late 19th and early 20th century (Rosberg 1977, Raskin 1993, Tienda 2002). Nativism was growing strongly during this period due to

[6] Although the fraction of *all* immigrants who were not citizens may be a more complete representation of immigrant behavior, the Census only collected citizenship status for adult males in 1900 and 1910.

factors such as the changing composition of immigrants from Northwest Europe to Southeast Europe, the Panic of 1893 and ensuing depression, the assassination of President McKinley in 1901 by a second-generation immigrant, and the United States' involvement in World War One (Jaret 1999; Van Nuys 2002). Quantitative evidence is consistent with this: Henderson (2017), finds that the fraction of the population that was immigrants was a significant predictor of repeal, with, in some specifications, a one standard deviation in the fraction foreign born in a given state nearly tripling the hazard of disenfranchisement.

It is important to note that legal basis for non-citizen voting was *constitutional*, rather than statutory. Because states vary in the costs of amending their constitutions, this raises the possibility that these costs were also a factor in the timing of disenfranchisement (Tarr 1996). Constitutional amendment is a "highly deliberative process… more difficult than that used for normal legislation" and during this era, constitutional amendments were primarily legislatively referred: they originated as statutes in the state legislature, and then had to win a popular vote for ratification at a general election (Lutz 1994). This multi-step process meant that, in some cases, it took states decades from when the state legislature *would* have disenfranchised non-citizens, to when non-citizens finally were disenfranchised.[7]

Furthermore, states varied along three dimensions of the constitutional amending process: the majority required for initial legislative passage (1/2 – 2/3), the number of successive passages in the state legislature (once or twice), and the fraction required for ratification by popular vote.

[7] There are examples of this in the historical record. Alien disenfranchisement reached the point of popular referendum in Missouri in 1912, Oregon in 1894, and Texas in 1919 implying that it had first passed in both houses of the state legislature, but did not achieve majorities of the popular vote (Official Manual of State of Missouri 1913-1914; *Sunday Oregonian* 1894; *Dallas Morning News* 1919). Missouri would go on to repeal in 1924, Oregon in 1914, and Texas in 1921.

A summary table of state amending procedures is presented in table 1.2. Crucially, certain states required a majority of *all* voters in a given election for passage; other states only required a majority of voters *voting on that particular issue* in a given election. This generates variation in the effective standard for passage because of the practice of "roll-off voting", in which voters vote only on a subset of questions on a ballot, with abstentions increasing in the number of questions (Bullock and Dunn 1996, Augenblick and Nicholson 2016, Stephens and Charles 2013). Given a certain amount of roll-off voting, states that required a majority of *all* electors effectively imposed a higher passage rate for the amendment than states that required only a majority of voters voting thereon in the referendum.[8,9]

There is qualitative evidence that variation in the amending procedures introduced an element of randomness in the timing of non-citizen voting.[10] In 1920, voters passed a proposed amendment in Arkansas that would have restricted the franchise to citizens, 87,237-49,757. Given that 190,113 total votes were cast in the election, however, the votes for the proposed amendment did not constitute a majority of all voters at the election (which would have been 190,113/2 = 95,057), and so the measure failed (Alysworth 1931). Similarly, Nebraska in 1910 attempted to disenfranchise non-citizen immigrants and the referendum passed, 100,450-74,878,

[8] Consider two states, A and B, each with 100 voters. If there is 20% roll-off voting, then only 80 voters in each state vote on the proposed amendment. If state A mandates a majority of *all* voters for passage, then passage requires 51 voters, or an effective majority of 100*(51/80) = 64%. If state B requires only a majority of the voters voting *thereon*, then only 41 votes are needed for passage.

[9] A measure of the importance of this provision is implicitly provided by Section 16 of the Nevada state constitution of 1864, which discusses the procedures for amending the state constitution. It specifies that "in determining what is a majority of the electors voting at such election, reference shall be had to the highest number of votes cast at such election for the candidates for any office or on any question."

[10] The issue of roll-off voting in contests for constitutional amendment is explicitly addressed in the Massachusetts state constitution of 1780, as amended in 1918. Constitutional amendments originating via initiative process – a rarity at the time – were required to obtain the support of at least 30% of all voters at a "state election", and a majority of voters "voting on such amendment". This suggests that roll-off voting was recognized as a salient factor in elections for constitutional amendments (Constitution of the Commonwealth of Massachusetts, 1918).

but failed to achieve a majority of the 243,390 electors voting in the election (Nebraska Blue Book and Historical Register, 1918). Only in 1918 did voters in Nebraska pass the constitutional amendment with a majority of *all* voters voting at the election.[11] The determinants of repeal are addressed systematically in Henderson (2017), who finds that states requiring on a majority of voters voting thereon were eight times more likely to repeal alien voting in a given decade than those requiring a majority of *all* voters.

It is possible that structural change in the demand for immigrant labor was a salient factor determining the timing of disenfranchisement. Given that the adoption of alien voting in the mid-19th century was in part due to increased demand for labor in order to spur economic development, declining demand may have driven the repeal of these provisions (although Henderson (2017) finds no evidence of economic determinants of repeal). While this would be problematic for estimating effects of disenfranchisement, this is mitigated by the variation in timing of repeal introduced by nuances of constitutional amending procedures. Furthermore, I am able to control for shocks common to all immigrants by leveraging a discontinuity in the naturalization proceedings of the era, which I explain in more detail in section 1.3.2.

1.3 Relevance of Alien Disenfranchisement

Were these disenfranchisements binding? There is a body of literature arguing that election fraud was pervasive during this era (Allen and Allen 1981), and there are examples of urban political machines conducting widespread immigrant naturalization for political purposes (Orth 1914).[12] Such behavior would argue against the importance of immigrant

[11] 123,292 for, 51,600 against, with 225,717 total votes cast (Nebraska State Canvassing Board, 1918)

[12] About the New York City political machine, the authors write “Most observers were convinced that Tammany’s grip on New York City government was based in large part upon wholesale election fraud in the foreign-born, working-class districts.” (Allen and Allen, pp. 158)

disenfranchisement: if it were possible to generate immigrant votes via "naturalization mills", for example, then the repeal of alien voting may not have been binding. Similarly, if monitoring in elections was costly in this era, then provisions delineating voter eligibility – like alien disenfranchisement – may not have affected actual voting behavior. I test for the relevance of disenfranchisement in two ways: first, I assess the effect on aggregate vote totals; second, I use Census samples from 1900-1930 to test for the effect on initiating naturalization proceedings.

1.3.1 *Effect of Disenfranchisement on Aggregate Vote Totals*

I explore the effect of disenfranchisement on political participation using data on mayoral elections from 1880-1924 and gubernatorial and presidential elections from 1870-1940. The mayoral data is a newly assembled dataset of vote totals for a panel of seventeen large cities.[13] Leveraging the staggered timing of alien disenfranchisement, I estimate event-study models at the state (or city)-year level to assess the relevance of the disenfranchisement on political activity. These specifications identify the effect of disenfranchisement based on sharp changes in voting activity in years relative to the year of repeal, net of changes in states (or cities) that had never enacted alien voting. The intuition is similar to that of differences-in-differences, but flexibly allowing the effect to differ over time relative to the date of disenfranchisement. Specifically, I estimate:

$$y_{jt} = a_j + c_t + t * a_s + \sum_{f=-h}^{h} D_j \pi_f * 1(t - R_j \in f) + \beta X_{jt} + \varepsilon_{jt} \quad (1)$$

[13] I was unable to find data for any cities other than those in my sample. The cities in the sample are Baltimore, Boston, Buffalo, Chicago, Cincinnati, Cleveland, Detroit, Los Angeles, Milwaukee, New Orleans, New York, Philadelphia, Pittsburgh, St. Louis, San Diego, San Francisco, and Seattle. All cities except San Diego and Seattle are from Holli and Jones (1982). Historical San Diego mayoral results are available from https://www.sandiego.gov/sites/default/files/legacy/city-clerk/pdf/mayorresults.pdf; Results from Seattle are available from http://www.seattle.gov/cityarchives/seattle-facts/historical-election-results

Y_{jt} is the ratio of total votes to the estimated population of males age 21 and up in area j (city or state) in year t. This is my preferred measure of aggregate political activity because it captures political engagement in a jurisdiction while implicitly controlling for a migration response.[14] Because frequency of mayoral and gubernatorial elections varied across cities and states in this era – ranging from one to four years in my sample – I use four-year event-time bins in the event-study specification. D_j is a 0/1 indicator for whether area j was in an alien disenfranchisement state.[15] R_j is the year of disenfranchisement in area j; h denotes the number of event-time bins used; and f indicates that voting outcome in year t, relative to year of repeal, is grouped into a particular event-time bin. The coefficients of interest are π_f, which trace out the effect of disenfranchisement over time, net of changes in states or cities that had never enacted alien voting. As formulated in specification (1), π_f are not identified; therefore, I omit the event-time bin immediately prior to repeal, and so estimates of π_f should be interpreted as relative to the omitted category. I estimate three pre- and post-disenfranchisement bins for mayoral contests, and, because of greater data availability, four for gubernatorial and presidential.

X_{jt} is a vector of state or city level controls. Because the disenfranchised population varied significantly across states, I control for intercensal estimates of the fraction of adult males that were non-citizens in area j and year t.[16] Furthermore, to account for the possibility that

[14] If migration was a primary response to disenfranchisement, then aggregate vote totals would fall sharply as voters left the state, spuriously implying a large disenfranchisement effect on current residents. I address the possibility of migration in greater detail in section 6.

[15] I prefer to use this indicator for ease of exposition and interpretation of coefficient estimates, but results are consistent using, instead of a 0/1 indicator, a measure of the "dose" of disenfranchisement as captured by the fraction of adults that were not citizens at the year of disenfranchisement. Results are available upon request.

[16] Because of data availability, I am unable to do this at the city-level; instead, I control for the fraction of the population that is foreign-born. This is highly correlated with the fraction of adult males that are non-citizens: in states from 1870-1940, the correlation coefficient is .96.

restrictions on immigrant voting coincided with other significant changes in voting policy in this era, I include a indicators for the presence of female suffrage, poll taxes, and literary tests (from Lott and Kenny, 1999). In my baseline specification, I include year and state or city fixed effects; in Model 2, I also include state- or city-specific linear trends in order to account for smoothly evolving unobservables. Coefficients are estimated using OLS. For mayoral elections, standard errors are calculated using wild bootstrap to account for the small number of clusters (Cameron et al. 2008); for all other specification, standard errors are clustered at the state level.

Results are presented in figures 1.2 (mayoral), 1.3 (gubernatorial), and 1.4 (presidential), where coefficients from the event-time indicators are plotted for the two models. Underlying regression results are in appendix tables 1.1.1, 1.1.2, and 1.1.3. The mayoral estimates are measured with some noise – which is not unexpected, given the small sample – but both models indicate that the share voting fell significantly upon immigrant disenfranchisement: at face value, the estimates suggest that the share voting fell by 10-12 percentage points initially, and then fell further to approximately 20 percentage points. I am unable to statistically differentiate the post-disenfranchisement coefficients in either specification ($F = .43$, $p = .75$ in baseline; $F = .23$, $p = .87$ in Model 2) but I am able to strongly reject that the post-disenfranchisement coefficients jointly equal zero ($F = 5.45$, $p < .001$ in the baseline; $F = 3.74$, $p = .006$ in Model 2). That the results are similar for both the baseline specification and Model 2 suggests that unobserved trends in political participation were not correlated with the timing of disenfranchisements. Moreover, consistent with a causal effect of disenfranchisement, there is no evidence of a pre-trend influencing the results: I cannot reject the hypothesis that the pre-repeal coefficients are jointly zero ($F = .80$, $p = .50$ in the baseline; $F = .77$, $p = .52$ in Model 2).

I find a similar pattern of results for gubernatorial elections in figure 1.3. There is an immediate and statistically significant drop in voting activity of 5-7 percentage points for gubernatorial elections in Model 2, with these effects remaining approximately constant for sixteen years. As above, the similarity of the estimates across the two models suggests that the econometric model is properly specified; however, controlling for state-specific trends adds to the precision of the estimates. There is no evidence of a pre-trend in either specification: pre-repeal coefficients do not jointly statistically differ from zero (F = .77, p = .55 in the baseline; F = .93, p = .43 in model 2). Effects are similar, albeit smaller, for presidential elections (figure 1.4). I document a reduction in political participation of 2 - 4 percentage points persisting for at least 16 years. Again, the time-path of effects suggests a causal role for disenfranchisement, as pre-repeal coefficients do not jointly differ from zero in either specification (F = .08, p = .97 in baseline; F = 1.08, p = .37 in Model 2). While the individual coefficients are not generally statistically precise, the time path of effects suggests that disenfranchisement served as a negative shock to voting activity.

Are these effect sizes reasonable? Immigrants constituted 44.3% of the adult male population in my sample cities, and 24.7% in my sample states (Ruggles et al., 2015). These provisions directly disenfranchised those immigrants who had filed first papers and were not yet citizens, and from 1900-1930, this constituted 11.6% of the adult male immigrant population. Assuming all of these individuals *would* have voted had they not been disenfranchised, the expected effects of disenfranchisement on aggregate voting activity range from 5.1 percentage points in cities to 2.9 percentage points in states. My estimates are approximately double that in both mayoral and gubernatorial elections, but are in line with the effect on presidential voting.

This can be rationalized by the fact that the act of voting is, itself, an equilibrium decision. The benefits of voting include any direct utility from participating in the democratic process *and* the expected value of casting the decisive vote in the election (Downs 1957). A reduction in the perceived competitiveness of an election –caused, for example, by exclusion of non-citizens from the electorate – would tend to reduce the likelihood of casting a decisive vote, and therefore lower the overall utility gain from voting. The competition-turnout link has been documented in certain electoral settings (Settle et al., 2014; Grofman et al., 1998) and implies that the magnitude of the aggregate voting response may exceed the scale of the direct disenfranchisement. This phenomenon can also explain why there is no evidence of spillover reductions in presidential elections: the level of public attention is high, so the direct utility gain from voting is large compared to that from the perceived likelihood of casting the decisive vote.

1.3.2 Effect of Disenfranchisement on Citizenship Status

As a second check for the relevance of the disenfranchisements, I estimate their effect on the incidence of filing first papers for adult male immigrants. Immigrants that intended to become citizens would have been relatively unaffected by the repeal of alien voting provisions: the repeal simply deferred eligibility for voting until after the five-year residency period. However, individuals that did *not* intend to become citizens – possibly because they did not intend to remain in the United States long enough to become eligible – would, following disenfranchisement, have had little incentive to initiate naturalization proceedings.[17] Therefore, if

[17] Citizenship conferred few benefits for immigrants in this era. Williams (1912) writes: “To the foreigners in Nebraska, practically the only advantages which accrue from American citizenship are protection abroad, certain homestead rights, and the inheritance of citizenship to minor children.” Moreover, I have been able to find no evidence that changes in these few benefits of citizenship were correlated with the timing of disenfranchisement.

the disenfranchisement was binding, there should be a *negative* effect on the average incidence of filing first papers (that is, initiating naturalization proceedings) for immigrants.[18]

Identification is threatened by the underlying causes of the disenfranchisements. While there is evidence that nuances of procedures to amend state constitutions determined the timing of disenfranchisement, it is possible that structural shifts in labor demand for immigrants were also a factor. If so, then any estimated effects of disenfranchisement may be attributable to these underlying shifts, rather than disenfranchisement *per se*. I account for this by exploiting institutional details of the naturalization process in this era. As set forth in the Immigration Act of 1798, immigrants were required to maintain a residence of five continuous years in the United States in order to obtain citizenship; however, immigrants could declare their intent to file for citizenship at any time after arrival. Thus, this five-year discontinuity identifies a pool of individuals – that is, immigrants with fewer than five years of residence - that *were* eligible voters prior to the disenfranchisement, and that were rendered *ineligible* following disenfranchisement. This motivates a specification in which outcomes are compared between recent immigrants and less recent immigrants, before and after repeal, between states that did and did not have alien voting. To the extent that these recent and non-recent immigrants are similar on unobservable characteristics, this effectively nets out all shocks affecting immigrants, including those generated by shifts in underlying demand for immigrants.

[18] Contemporary observers noted that alien suffrage – or lack thereon – strongly impacted citizenship behavior of immigrants. Williams (1912) cites the report of the Commission-General of immigration which notes that “in Indiana [which has alien suffrage] there were filed last year about 5000 declarations, while but 276 petitions were made for citizenship… in Ohio, on the other hand [which has citizen suffrage], where about the same number declared their intention to become citizens, 1676 petitions were filed…”. Chaney (1894) notes that “in the [Michigan] convention of 1867, delegate Blackman said that in nine cases out of ten the alien who had sworn his intention to become naturalized do declared simply for the purpose of becoming a voter, and that he left the matter there five, ten, fifteen, or twenty years, without fulfilling his declared intentions.”

For estimation, I use Census samples from 1900-1930 (Ruggles 2015). These Censuses recorded detailed citizenship information, and, importantly, differentiate non-citizens between those that had filed first papers, and those that had not. I estimate the following specification:

$$Y_{ist} = a_s + c_t + \rho New_{ist} + \beta Repeal_{st} + \gamma New_{ist} * Repeal_{st} + \delta X_{ist} + G_{st} + \varepsilon_{ist} \quad (2)$$

Y_{ist} is an indicator for having filed first papers for individual *i* in state *s* and year *t*. X_{ist} is a vector of individual controls comprised of age, age squared, literacy status, and an indicator for urban residence. New_{ist} is an indicator variable that takes the value of 1 for immigrants with fewer than five years of residence in the U.S., and $Repeal_{st}$ is an indicator variable that takes the value of 1 following repeal, in states that had repealed the provisions. The coefficient of interest is γ: this is an estimate of the differential effect of residing in an alien voting state following repeal between recent and less recent immigrants, net of changes in non-alien voting states. The sample is restricted to non-citizen foreign-born males aged 21 and over that had resided in the United States for fewer than fifteen years. This implicitly defines a "control" group as those immigrants that had resided in the U.S. for five to fifteen years (results are robust to alternative definitions of the control group: five to ten years and five to twenty years).

State fixed effects, a_s, and year fixed effects, c_t, are included to account for fixed factors within states and common shocks across time. G_{st} is a state-level 0/1 variable presence of laws barring aliens from public employment (Holmes 2003; Fishback et al. 2009).[19] I also include a state-level 0/1 variable indicating the decade *prior* to repeal of the voting provisions intended to capture, in the repeated cross-section, unobserved forces leading to repeal. I add two sets of controls in order to account for unobserved trends potentially driving results. First, in order to

[19] 17 states had enacted this type of law as of 1924 (Holmes 2003; Fishback et al. 2009). There is little overlap between this group and the states that had repealed, alien voting: only Indiana and Oregon are common to both.

account for the changing sending regions of immigrants over time, I include a linear trends in year of immigration interacted with birth region. This allows for the effect of origin region (categorized into North, West, East, Southern Europe, and non-Europe) to vary over time. I label this “Model 2”. Second, in order to control for potentially changing demand for immigrant labor, I also include state-specific census year linear trends. I label this “Model 3”. Coefficients are estimated using WLS (utilizing IPUMS person weights), and standard errors are clustered at the state level (Bertrand et al., 2004).

Results are presented in table 1.3. I find stable, statistically significant results indicating that immigrants responded to disenfranchisement by reducing the incidence of filing first papers: a reduction of 6.7 – 7.3 percentage points, which scales to 50-54% relative to the baseline mean for pre-repeal recent immigrants (13.5 percentage points). Interestingly, there is evidence of significant spillover effects: following disenfranchisement, *all* non-citizen immigrants, not only recent immigrants, were less significantly likely to file first papers. I now turn to estimating the effects of disenfranchisement on short-run labor market outcomes.

1.4 Immediate Effects of Disenfranchisement on Labor Market Outcomes

Economic theory provides multiple channels through which disenfranchisement may affect short-run labor market outcomes. No longer accountable to the disenfranchised population, elected officials may redirect public spending toward more valuable constituencies. Public sector employment was not subject to civil service regulations in most states and cities in this era, suggesting re-allocation of this employment as a potentially direct labor market consequence of disenfranchisement (Ujhelyi 2014, Folke et al. 2011).[20] Disenfranchisement may spur

[20] While data on public employment levels is scarce before 1930, Libecap (2007) estimates that federal civilian employment grew from 240,000 in 1901 to almost 400,000 in 1913; adjusting this by state and local government

outmigration of the disenfranchised, which, depending on the selection into migration, could have positive or negative effects on remaining immigrants. Finally, employers may respond to the reduction in political voice by discriminating against affected immigrants.

I test for the effects of disenfranchisement on short-run labor market outcomes by re-estimating specification (2) but using as outcomes labor force participation, occupational standing, and an indicator for public sector employment.[21] The Census Bureau did not specifically record public sector employment until 1940, so I construct three alternative measures based on IPUMS occupation codes (see appendix 1.2 for details). The estimation sample is non-farm foreign-born males aged 21-64 with fewer than 15 years of residence in the U.S.

Results for labor force participation and occupational standing are in table 1.4. Columns (1) and (2) present results from the baseline model; columns (3) and (4) from model 2; and columns (5) and (6) from model 3. Consistent with other cross-sectional evidence on immigrant assimilation into U.S. labor markets in this era, I find that recent immigrants experienced a penalty in both labor force participation and occupational standing, although this varies in statistical precision (Abramitzky et al., 2014). There is no evidence of an effect of disenfranchisement on immigrants (as captured by the interaction between $Repeal_{st}$ and New_{ist}), and the estimated effects are sufficiently small as to rule out even modest effects: sample means are approximately .95 for labor force participation and 25.0 for occupational standing.

---

share of total spending, I estimate that 600,000 to 1.2 million individuals were employed in some form of public service starting in 1900 (excluding military).

[21] Occupational standing is an IPUMS-constructed variable that assigns to each occupational category the median wage from the 1950 Census. This is provided in lieu of wages and income, which were not measured in the Census until 1940. The limitations of this measure are well known, but it is useful as an indicator of labor market outcomes in this era. See Abramitzky et al. (2014) for further discussion.

There is some evidence of negative spillovers for the wider population of immigrants. In both the baseline and Model 2, disenfranchisement has a negative effect on the labor force participation of all immigrants in the sample, although the effect size is small, and this is not statistically precise across specifications. As a further check for spillover effects, I re-estimate specification (2) on a sample that ought to have been directly unaffected by repeals: native adult males. Results are in table 1.5. I find weak evidence for *positive* spillover effects: adult native males in states following alien disenfranchisement were more likely to participate in the labor force, and work in a higher-status occupation.

Results for public sector employment are presented in table 1.6. For brevity's sake, I omit results from the baseline specification, and report results only from Model 2 and Model 3 (baseline results are not substantively different; full results available upon request). I find that disenfranchised immigrants were significantly less likely to be employed in the public sector: depending on the definition of public sector employment, the effect size ranges from .002-.004. While small in magnitude, these effect sizes are large relative to baseline means for recent immigrants prior to repeal: .009 (public 1), .008 (public 2), and .005 (public 3). The estimates scale to effect sizes of 25% - 60%, and the effects are statistically significant for two of the three definitions of public employment. There is little evidence of spillovers, either among the remaining immigrants in the sample, or in a placebo check using native men as an alternative treatment group (results reported in table 1.7). Given that public sector employment in this era was well compensated relative to private sector positions – Troesken (1998) estimates a wage premium of 25% - these results suggest that politicians directed lucrative positions away from immigrants that, because of disenfranchisement, had lost their political value.

1.5 Intergenerational Effects of Disenfranchisement

That disenfranchisement led to a strong reduction in public sector employment of affected immigrants suggests that voting power affected the economic wellbeing of immigrants; however, data availability from this era limits the extent for analysis of short-run labor market impacts. Census microdata samples with immigration covariates are only available from 1900-1930, which spans only part of the time period in which states disenfranchised immigrants. Immigrant-heavy states Michigan, Minnesota, and North Dakota, repealed in the 1890s; if true effects are larger in more immigrant-heavy areas, then estimated effects, based on the sample of states with smaller immigrant populations, will be biased downward. Moreover, the Census samples from 1900-1930 contain only relatively coarse measures of labor market outcomes. Income, a key measure of labor market success, was not collected by the Census until 1940; labor force participation varied little across the population of adult males; and measures of occupational standing do not allow for effects within occupations.

The true impact of disenfranchisement, however, may be reflected in the outcomes of the *children* of the disenfranchised.[22] Therefore, I estimate the intergenerational effects of disenfranchisement on the outcomes of adults from 1940-1960, who, as children, were exposed to disenfranchisement. In order to capture effects net of potential spillovers, I employ an econometric intergenerational design in which, instead of comparing children of disenfranchised immigrants to non-disenfranchised immigrants, outcomes are compared between children of all immigrants and children of natives. I restrict attention, however, to *native-born* individuals,

[22] A very large literature documents intergenerational correlations of income; see Lee and Solon (2009). Moreover, a growing literature has established that shocks to parental wealth or wellbeing have effects in the outcome of children (Leininger and Kalil 2014; Black et al. 2014; Currie and Rossin-Slater 2013).

thereby imposing that any disenfranchisement effects are indirect: either via parental transmission or changes to the environment.

Using variation in paternal birthplace, birth state, and birth cohort in the 1940 full-count Census and IPMS 1950-1960 Census samples, I leverage the staggered timing of disenfranchisement across states to estimate the following event-study model:

$$y_{ibcst} = b_b + c_c + g_t + j_s + \sum_{f=-h}^{h} D_b \pi_f 1(c \in f) + \sum_{f=-h}^{h} P_{ibc} D_b \alpha_f 1(c \in f) + \beta X_{ibc} + \varepsilon_{ibc} \quad (3)$$

Y_{ibct} is an outcome – either the natural log of real income, or years of completed education - for individual *i* born in birth state *b*, in birth cohort *c*, observed in census year *t* in state of residence s (for brevity, I omit the t and s subscripts on all other individual-level covariates in the specification). The sample is restricted to sample-line native-born males aged 25 and over, so that the individuals likely would have completed their education at the time of observation.[23] When income is used as the outcome, I further restrict the sample to individuals under age 65 in order to avoid conflating effects of age with effects of disenfranchisement. I use only variation in paternal nativity, rather than paternal and maternal, because of the institutional structure of the era: women in most states could not vote until 1920, and, until 1922, foreign-born women were assigned the citizenship status of their husbands (Smith 1998).

The event-time indicators are grouped into bins of five years, with membership of individual born in cohort c in bin f denoted by $1(c \in f)$. These are a measure of exposure to the repeal of alien voting; for example, an individual born in 1906 in Wisconsin (which had a repeal year of 1908) will receive a value of 1 for the indicator denoting birth year 0-5 years prior to repeal, and a value of 0 for all other years. Individuals not born in alien voting states will receive

[23] Parental nativity in 1940 and 1950, and income in 1950, were census sample-line variables (Ruggles et al. 2015).

a value of 0 for all exposure indicators. P_{ibc} is an indicator for having a foreign born father. The coefficients of interest are α_f: these trace the time path of the father-nativity differential for birth cohorts in relation to the year of repeal of alien voting. In order to identify α_f, an event-time bin must be omitted. In order to facilitate interpretation of the coefficient estimates, the omitted bin is for individuals born 30 years or more, prior to disenfranchisement. This implies that α_f should be interpreted relative to individuals who would have been adults, and therefore unaffected by childhood exposure, at the time of disenfranchisement.

As in specification (2), I include two sets of interactions intended to control for the changing composition of the immigrant population over time, as well as unobserved non-linear state-specific trends in the demand for immigrant labor. “Model 2” includes a set of birth-region specific trends in birth year for each parent; “Model 3” includes a set of birth-state specific trends in birth year. I control for fixed factors within birthplaces and factors common across birth years by including birthplace and birth cohort fixed effects. Individual controls are race indicators, age, age squared, father and mother birth region indicators and an indicator for father’s nativity (either foreign born or native). Census year fixed effects are included to control for the aging of the population. Coefficients are estimated using WLS. I employ IPUMS-provided individual-level weights in estimation (sample-line weights for 1940 and 1950 samples, person-weights for 1960). Standard errors are clustered at the birth state level.

It is important to note that exposure indicators imperfectly capture true exposure. Individuals born in a repeal state may move out of state prior to repeal, thus avoiding exposure to the post-repeal environment, and individuals not born in a repeal state may do the opposite. If such migration does not systematically vary with labor market potential, then it should only serve to attenuate my estimates. If, however, disenfranchisement spurred out-migration among positively selected individuals, then the composition of the remaining individuals, and their children, may generate a spurious negative intergenerational effect of disenfranchisement. Such a migration response would only be consistent with a very high implicit valuation of the right to vote; therefore, this possibility is unlikely. I return to this question in section 1.6, in which I discuss potential mechanisms for observed results.

See figures 1.5 and 1.6 for the plot of α_f from specification (2) for outcomes, respectively, years of completed education and the natural log of income due to wages and salary. I report the underlying coefficients in appendix tables 1.3.1 and 1.3.2.[24] Each point in the figures represents the coefficient from the interaction term between period of birth relative to year of repeal and an indicator for whether the father is foreign-born, and each figure plots the evolution of the differential between outcomes for native adult males with foreign-born fathers, and native males with native fathers. This specification allows for evaluation of a natural placebo group: those individuals that would have been adults (that is, aged 20 or older) at the time of disenfranchisement, and who would, therefore, have not been exposed directly to the effects of disenfranchisement as children.

The results from figure 5 indicate a negative effect of immigrant disenfranchisement on the educational attainment of children born around the time of disenfranchisement: the educational attainment differential – not statistically different from zero for older cohorts - widens sharply for cohorts born 5-10 years before disenfranchisement. These individuals would have been in prime schooling age at the time of disenfranchisement, and the effects for cohorts born 5-10 years prior to disenfranchisement, and cohorts both 10-15 years prior, are statistically

[24] For the sake of brevity, I report only the α_f, and not the event-time estimates for native males (that is, π_f). Full results available upon request.

distinguishable from one another (F = 3.47, p = .07 in Model 3). The effects are stable across specification: a reduction in educational attainment of .24.- .28 years for cohorts born 5-10 years prior to disenfranchisement, relative to those born 10-15 years prior. Results are robust to sequentially dropping birth regions (see appendix table 1.4.1). It is reassuring to note that the magnitude of these effects is roughly in line with previous research (Kose et al. 2015).[25]

I also find a negative effect on real adult income for exposed cohorts. Specifically, I find that individuals of immigrant parentage that are exposed to disenfranchisement at an early age – those born fewer than five years before repeal of alien voting – earn significantly less in income as adults relative to individuals of native parentage born in the same state and year. Point estimates of the income penalty for these cohorts are individually statistically significant and range from 5 - 9% across specifications. Results are substantively robust to sequentially dropping birth regions (see appendix table 1.4.2). In order to assess the extent to which the reduction in income is driven by the reduction in education, I re-estimate specification (4) for the natural log of income but include, as a covariate, completed education. These regression results are reported in appendix table 1.5 with, and without, the control for educational attainment. The inclusion of education can explain approximately ¼ of the observed effect on income.

The narrative evidence presented in section 1.2 suggests that there was an element of exogeneity in the timing of disenfranchisements due to nuances in the amending processes of state constitutions. It is still possible, however, that anti-immigrant attitudes may have partially driven disenfranchisement. If so, and if unobserved anti-immigrant sentiment peaked concurrently with the timing of the disenfranchisements, then the estimated intergenerational

[25] Kose et al (2015) estimate that women's suffrage led to up to a year of additional education for exposed children.

effect may be attributable to anti-immigrant attitudes, rather than disenfranchisement *per se*. I address this by leveraging the fact that women were assigned the citizenship status of their husbands until 1922 (Smith 1998). This provision implies that foreign-born women who were married to native men would not have been *directly* affected by the disenfranchisement, as they would have been assigned citizenship at the time of marriage, but, because of their nativity, *would* have been exposed to any unobserved prevailing anti-immigrant sentiment. Therefore, I repeat specification (3), but further restricting the sample only to individuals of immigrant parentage: those with a foreign-born father comprise the "treatment" group, and those with a foreign-born mother and native father comprise the "control" group. As this only applied until 1922, I include only those alien-voting states that repealed their provisions prior to 1910, and restrict attention only to the 1940 Census. Results are in appendix figures 1.1 and 1.2, and, while somewhat noisy, are substantively similar to results presented in figures 1.5 and 1.6. This suggests that intergenerational effects are not driven by unobserved anti-immigrant attitudes.

1.6 Mechanisms for Intergenerational Effects

What drives the intergenerational effects? Unfortunately, as noted above, data availability precludes more detailed analysis of the immediate effects of disenfranchisement on economic wellbeing. However, using estimates of intergenerational elasticity of income, it is possible to back out the effect on wages that *would* have generated the observed intergenerational effects. Feigenbaum (2015) estimates the intergenerational elasticity of income from 1915-1940 to be .249; given that educational attainment explains ¼ of the observed 5-9% effect, a reduction of 15-27% in wage income for immigrants relative to natives could explain the remaining intergenerational effect.

While such a penalty would be substantial, it is not unrealistic: public sector jobs in the late 19th century - from which recent immigrants were largely excluded following disenfranchisement - enjoyed a wage premium of 25% over the private sector (Troesken 1998). Moreover, Naidu (2012) estimates that the disenfranchisement of African-Americans in the South led to a loss in black income of at least 15%, which is in line with the magnitude of the effect that could rationalize the intergenerational effects. However, alternative mechanisms are possible. Below, I test for three channels for the intergenerational effect: migration, shifts in policy and spending, and cultural assimilation.

1.6.1 Migration Response

Scholars have shown that labor was relatively mobile in this era (Fishback 1998), which raises the possibility of a migration response to disenfranchisement: recent immigrants may have "voted with their feet" and re-located to states with immigrant voting provisions still in place or, failing that, returned to their home countries. While mass out-migration could effectively improve labor market outcomes for the immigrants that remain, positively selected out-migration could generate a spurious compositional effect: the estimates based on the remaining immigrants will, mechanically, appear to have worsened.

There are two factors, however, arguing against such a migration response: first, for immigrants who had resided in the United States for at least five years, applying for citizenship was a means of regaining the vote, thus mitigating the need for migration. While this did entail certain costs – waiting the appropriate residency period, filing paperwork, and paying small fees – they were almost certainly far less than costs of relocation. Second, anecdotal evidence from this time suggests that the price of an individual vote relatively low: a single vote, according to one observer in Indiana in 1896, cost "sandwiches and liquor and $5 a head" (Allen and Allen

1981). Theoretically, the value of the right to vote at any moment is the present discounted value of the future stream of benefits of voting. At this modest period valuation of a vote, even permanent disenfranchisement would generate little financial incentive for an individual to move to a jurisdiction in which voting was permitted; temporary disenfranchisement generated by the repeal of alien voting would have provided even less. Consistent with this, Miller (2008) finds no evidence of a migration response to women's suffrage, and contemporary observers documented no migration effects of alien voting. [26]

However, I empirically test for a migration response of disenfranchisement. Using IPUMS samples and published Census aggregates, I create a state-decade panel from 1870-1940 and test for effects of disenfranchisement on overall immigrant population, the fraction of the population that is immigrant, and, in order to assess the extent of selection into migration, the fraction of immigrants in the population that are illiterate. I repeat Model 2 from specification (1) and estimate an event-study model controlling for state fixed effects, year fixed effects, state-specific trends, and indicators for women's suffrage, literacy tests, and poll taxes. The data – occurring only at decadal intervals – is not fine enough to estimate precise treatment effects, but allows for general assessment of the extent to which migration may be driving the estimated results (either through its effects on labor supply, or the composition of remaining workers).

Results are in table 1.8. I find little evidence of a systematic migration response. Both total population and foreign-born population were falling in alien voting states over this period –

[26] Chaney (1894) writes that "immigrants flocked into Illinois, which required citizenship, and in doing so passed through Indiana which was satisfied with a declaration of intention." This is echoed in Williams (1912): "Alien suffrage was a drawing card used by western states to attract foreign immigration… the failure of alien suffrage to accomplish its original purpose appears from the fact that foreigners today do not hesitate to leave Nebraska [in which alien suffrage still held] for Colorado, Montana, Wyoming and other states in which they cannot vote, provided they can better their economic condition."

foreign-born slightly faster than total – but the effect is largely statistically insignificant. More importantly, it does not appear that the out-migration was driven by the foreign-born: the fraction of the overall population that is foreign born falls slightly in alien voting states following disenfranchisement, but the coefficients are not jointly statistically distinguishable from zero ($F = .94$, $p = .48$). Finally, there appears to be a reduction in the incidence of illiteracy among the foreign-born following repeal – suggesting, if anything, negative selection out of the state – but, again, these coefficients are not jointly distinguishable from zero ($F = .28$, $p = .96$). If anything, this selection effect should serve to bias the negative effects of disenfranchisement toward zero.

1.6.2 Policy Responses

A growing body of research has established the importance of political regime and individual policies for labor market outcomes (Beland, forthcoming; Beland and Unel 2015). Scholars have also demonstrated the link between enfranchisement and allocation of public resources (Miller 2008; Lott and Kenny 1999) and the pernicious effects of childhood exposure to political unrest (Smythe et al. 2004). I test for the salience of policy channels by repeating Model 3 from event-study specification (1) on the following outcomes: share voting democratic in gubernatorial elections; Poole-Rosenthal index (which measures Congressional political ideology, with larger index values denoting more conservative positions); the share of upper and lower houses of the state legislature that are held by democrats (presented in figure 1.7.1); and the log of real per capita state-level total spending and budget shares devoted to education, transportation, and social services (all presented in figure 1.7.2).

In order to test for the effect on local spending, I entered city-level total government spending and spending on educational services from odd years ranging from 1899-1929, collected from the *Financial Statistics of Cities Having a Population of over 100,000* and Bureau of Labor Statistics Bulletins 20 and 36. To test for shifts in sentiment against immigrants, I leverage the fact that Congress attempted to impose a literacy requirement for immigration on thirteen separate occasions from 1897-1917, and calculate the share of congressmen in a given state and year voting in favor of immigration restriction (Goldin 1994; Swift et al. 2009).[27] These votes are all either affirmative votes on enacting restrictions or attempted overrides of presidential vetoes, implying that a higher share of the congressmen in a particular state and year voting for this legislation indicated strong anti-immigrant sentiment. For both outcomes, I collapse the event-time indicators into a pre- and post-repeal indicator. Results from city spending and legislative voting are presented in table 1.9.

The results almost uniformly indicate a lack of effect of immigrant disenfranchisement on these outcomes. There is no appreciable effect on measures of democratic vote share or democratic presence in either chamber of the state legislature. I do find a modest increase in the Poole-Rosenthal index following disenfranchisement indicating a conservative shift in congressional ideology, but this is not statistically precise. Moreover, any shift in ideology is not reflected in specific attitudes toward immigration: legislators were no more likely to vote for immigration restrictions. There is no effect on public expenditures at either the state- or city-level. While these results certainly do not preclude the possibility of governmental response, they suggest that alien disenfranchisement was not accompanied by large swings in state or municipal spending or policy.

1.6.3 Social Assimilation

[27] Goldin (1994) documents that, from 1897-1917, the literacy test was passed by the House of Representatives four times, and by the Senate five times. The House voted to override Presidential vetoes twice, and the Senate once, in 1917, when the literacy test finally went into effect.

Evidence on immigration during this period has established that greater cultural and social assimilation of immigrants led to improved contemporary and future economic outcomes (Abramitzky et al. 2016, Goldstein and Stecklov 2015). Using more recent data, scholars have shown that the English-language skills of parents are salient for outcomes for offspring: greater language fluency leads to improved educational attainment and labor force performance for children (Bleakly and Chin 2008; Casey and Dustmann 2008). If immigrant disenfranchisement led to a reduction in the benefits of social assimilation, then it may have reduced incentives for immigrants to acquire English-language skills. This, in turn, could have led to lower human capital for children of immigrants, thus rationalizing the observed intergenerational effect.

To test this, I repeat specification (2) on adult male immigrants in Census samples from 1900-1930, but use as an outcome variable an indicator for English language proficiency. I include foreign-born males living on farms in this sample, although the results are robust to restricting to non-farm males. Results are presented in table 1.10. While I find no differential negative effect on English language proficiency for recent immigrants relative to less-recent immigrants, *all* immigrants in this sample are significantly less likely to be proficient in English following disenfranchisement. The magnitudes of the effect are large, ranging from 4.9 to 7.1 percentage points. Relative to a pre-repeal baseline mean of 64.3%, this reduction scales to 7.6-11.0%. The fact that this result is robust to the inclusion of birth region-specific linear trends in immigration year suggests that it is not driven solely by an increasingly negatively selected immigrant pool.

While this suggests that the cultural assimilation of immigrants as measured by English language skill may have driven, in part, the adverse intergenerational effects of disenfranchisement, I interpret this evidence as only suggestive. The fact that this result applies to *all* immigrants in the sample, rather than recent immigrants, implies that I cannot distinguish between the effects of disenfranchisement per se, and the unobserved effects of anti-immigrant sentiment that may have led to the timing of disenfranchisement (after accounting for the exogeneity introduced by variation in constitutional amending procedures). I leave the task of untangling these two effects to future research.

1.7 Conclusion.

This paper presents novel evidence on the effects of immigrant disenfranchisement on labor market outcomes using a little-studied set of state-level constitutional changes that disenfranchised non-citizen immigrants in twenty-four states and territories from 1864-1926. I document that these disenfranchisements had significant effects on political participation at the time: vote totals as a share of men aged 21 and over fell by approximately 3, 5, and 10 percentage points in presidential, gubernatorial, and mayoral elections, respectively. A first-order consequence of disenfranchisement appears to be an overall reduction in electoral competitiveness, with significant spillovers: the magnitude of the effect on political participation in gubernatorial and mayoral elections is double the magnitude of the disenfranchised population. Given recent research on the beneficial effects of political competition, this suggests the possibility of lasting negative consequences on economic growth (Besley et al. 2010).

Using difference-in-difference-in-difference specifications, I find that disenfranchisement significantly reduced public sector employed for affected immigrants. This is robust to alternative specifications and definitions of public employment in this era, and implies that politicians responded to disenfranchisement by shifting targetable spending away from electorally expendable populations. There is little evidence of an effect on other labor market outcomes, but it is possible that this masks unobserved effects due to the coarseness of the data.

As an implicit check for this, I estimate the effects of immigrant disenfranchisement on outcomes for exposed children. I find statistically significant evidence of a lasting shock: educational attainment fell sharply for children of immigrants, relative to children of natives, who were aged 5-10 at the time of disenfranchisement. Moreover, individuals of immigrant parentage exposed to the disenfranchisement at an early age earn approximately 5-9% less wage income as adults relative to comparable individuals of native parentage. There is no evidence that this intergenerational effect was driven by migration or shifts in the policy or spending environment; instead, I find suggestive evidence that disenfranchisement led to a reduction in English-language proficiency for all immigrants, which may have adversely affected the human capital accumulation of the next generation.

Taken together, these results suggest that alien disenfranchisement had widespread, negative, lasting effects. Therefore, despite the potential short-run electoral advantages to raising the cost of voting, current policy-makers considering measures that restrict the franchise would be wise to consider the lessons of history.

1.8 Works Cited

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

Abramitzky, Ran, Leah Platt Boustan and Katherine Eriksson. 2016. "Cultural Assimilation during the Age of Mass Migration." NBER Working Paper 22381.

Abramitzky, Ran and Leah Platt Boustan. 2016. "Immigration in American History." *Journal of Economic Literature.*

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2000. "The Colonial Origins of Comparative Development." *American Economic Review,* 91(5): 1369-1401.

Aidt, Toke S. and Bianca Dallal. 2008. "Female voting power: The contribution of women's suffrage to the growth of social spending in Western Europe (1869- 1960)", *Public Choice,* 134(3-4): 391-417.

Allen, Howard W. and Kate Warren Allen, "Vote Fraud and Data Validity," in Jerome Clubb, William Flanigan, Nancy Zingale (eds.), *Analyzing Electoral History: A Guide to the Study of American Voter Behavior* Beverly Hills: Sage Publications, 1981): 153

Augenblick, Ned and Scott Nicholson. 2016. "Ballot Position, Choice Fatigue, and Voter Behavior," *Review of Economic Studies,* 83(2): 460-480.

Aylsworth, Leon E. 1931. "The Passing of Alien Suffrage," *The American Political Science Review,* 25(1):114-116.

Bandiera, Oriana, Myra Mohnen, Imran Rasul, and Martina Viarengo. 2016. "Nation-Building Through Compulsory Schooling During the Age of Mass Migration." Manuscript.

Beland, Louis-Philippe. 2015. "Political Parties and Labor Market Outcomes. Evidence from US States." *American Economic Journal: Applied Economics*, 7(4): 198-220.

Beland, Louis-Philippe and Bulent Unel. 2015. "The Impact of Party Affiliation of U.S. Governors on Immigrant' Labor Market Outcomes." Louisiana State University Department of Economics Working Paper 2015-01.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-In-Differences Estimates?" *The Quarterly Journal of Economics* 119 (1): 249-275.

Besley, Timothy, Torsten Persson, and Daniel M. Strum. 2010. "Political Competition, Policy and Growth: Theory and Evidence from the US". *The Review of Economic Studies,* 77: 1329-1352.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2016. "Does Grief Transfer across Generations? Bereavements during Pregnancy and Child Outcomes." *American Economic Journal: Applied Economics*. 8(1): 193-223.

Bleakly, Hoyt and Aimee Chin. 2008. "What Holds Back the Second Generation? The Intergenerational Transmission of Language Human Capital among Immigrants." *Journal of Human Resources*. 43(2): 267-298.

Brennan Center. "New Voting Restrictions in Place for 2016 Presidential Election." September 12, 2016. Accessed September 30, 2016.

Bullock, Charles S. III and Richard E. Dunn. 1996. "Election Roll-Off: A Test of Three Explanations." *Urban Affairs Review*, 32(1): 71-86.

Burnham, W. Dean. Partisan Division of American State Governments, 1834-1985. Conducted by Massachusetts Institute of Technology. ICPSR ed. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor], 1989. http://doi.org/10.3886/ICPSR00016.v1

Burnham, W. Dean, Jerome M. Clubb, and William Flanigan. State-level Congressional, Gubernatorial, and Senatorial Election Data for the United States, 1824-1972. ICPSR ed. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor], 1991. http://doi.org/10.3886/ICPSR00075.v1

Cain, Louis P. and Elyce J. Rotella. 2001. "Death and Spending: Urban Mortality and Municipal Expenditure on Sanitation." *Annales De Demographie Historique* 1: 139-54

Cameron, Colin A, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-based Improvements for Inference with Clustered Standard Errors." *Review of Economics and Statistics*, 90(3): 414-427

Carruthers, Celeste and Marianne Wanamaker. 2015. "Municipal Housekeeping: The Impact of Women's Suffrage on Public Education." *Journal of Human Resources*, 50(4): 837-872.

Cascio, Elizabeth and Ebonya Washington. 2014. "Valuing the Vote: The Redistribution of Voting Rights and State Funds Following the Voting Rights Act of 1965", *Quarterly Journal of Economics*: 379-433

Casey, Teresa, and Christian Dustmann. 2008. "Intergenerational Transmission of Language Capital and Economic Outcomes." *Journal of Human Capital*, 43(3): 660-687.

Chaney, Henry A. 1894. "Alien Suffrage." Publications of the Michigan Political Science Association

Currie, Janet and Maya Rossin-Slater, 2013. "Weathering the storm: Hurricanes and birth outcomes," *Journal of Health Economics*, 32(3): 487-503

Cutler, David M. and Grant Miller. 2005. "The Role of Public Health Improvements in Health Advances: The 20th Century United States." *Demography* 42(1): 1-22

Daily Oregonian, September 2, 1984 (America's Historical Newspaper Database)

Department of Commerce and Labor, 1909, "Naturalization Laws and Regulations"

Feigenbaum, James J. 2015. "A New Old Measure of Intergenerational Mobility." Manuscript.

Fishback, Price. 1998. “Operations of ‘Unfettered’ Labor Markets: Exit and Voice in American Labor Markets at the Turn of the Century,” *Journal of Economic Literature* 36: 722-765.

Fishback, Price, Rebecca Holmes, and Samuel Allen. 2009. “Lifting the Curse of Dimensionality: Measures of the Labor Legislation Climate in the States During the Progressive Era” *Labor History*.

Folke, Olle, Shigeo Hirano, and James M. Snyder. 2011. Patronage and Elections in U.S. States. *American Political Science Review*. 105 (3):567-585.

Goldin, Claudia. 1994. “The Political Economy of Immigration Restriction in the United States.” *The Regulated Economy: A Historical Approach to Political Economy*. University of Chicago Press.

Goldstein, Joshua R. and Guy Stecklov. 2016. “From Patrick to John F.: Ethnic Names and Occupational Success in the Last Era of Mass Migration.” *American Sociological Review*, 81(1): 85-106.

Grofman, Bernard, Christian Collet, and Robert Griffin. 1998. “Analyzing the turnout-competition link with aggregate cross-sectional data.” *Public Choice*, 95: 233-246.

Haines, Michael R., and Inter-university Consortium for Political and Social Research. Historical, Demographic, Economic, and Social Data: The United States, 1790-2002. ICPSR02896-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2010-05-21. http://doi.org/10.3886/ICPSR02896.v3

Harper-Ho, Virginia. 2000. “Noncitizen Voting Rights: The History, the Law and Current Prospect for Change*,” Immigration and Nationality Law Review*, 21:477

Hayduk, Ron. *Democracy for all: Restoring Immigrant Voting Rights in the United States.* Taylor and Francis Group, 2006

Henderson, Morgan. 2017. “What Drives Expansions and Contractions of the Electorate? Evidence from Alien Voting in the United States.” Manuscript.

Holli, Melvin G. and Peter D’a Jones. 1981. *Biographical Dictionary of American Mayors, 1820-1980*. Greenword Press: Westport, Connecticut.

Holmes, Rebecca. 2003.“The Impact of State Labor Regulations on Manufacturing Input Demand During the Progressive Era.” Unpublished Ph.D. dissertation, University of Arizona, 2003

Jaret, Charles.1999. “Troubled by Newcomers: Anti-Immigrant Attitudes and Actions During Two Eras of Mass Immigration to the United States.” *Journal of American Ethnic History*, 9-39.

Keyssar, Alexander. 2000. *The Right to Vote: The Contested History of democracy in the United States*, Basic Books.

Kose, Esra, Elira Kuka, and Na’ama Shenav. 2015. “Women’s Enfranchisement and Children’s Education: the Long-Run Impact of the U.S. Suffrage Movement.” Manuscript.

Kousser, J.M. 1974. *The Shaping of Southern Politics: Suffrage Restriction and the Establishment of the One-party South, 1880-1910*. Yale University Press.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert W. Vishny. 1998. "Law and Finance," *Journal of Political Economy*, 106, 1113-1155

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

Leininger, Lindsey Jeanne, and Ariel Kalil. 2014. "Economic Strain and Children's Behavior in the Aftermath of the Great Recession." *Journal of Marriage and Family*, 76(5): 998–1010.

Libecap, Gary. 2007. "The Federal Bureaucracy: From Patronage to Civil Service", in *Government and the American Economy*

Lott, John and Lawrence Kenney. 1999. "Did Women's Suffrage Change the Size and Scope of government?" *Journal of Political Economy*, 107(6)

Lutz, Donald S. 1994. "Toward a Theory of Constitutional Amendment". *The American Political Science Review*, Vol. 88, No. 2 pp. 355-370

Meltzer, A. H., and S. F. Richard. 1981. "A rational theory of the size of government." *Journal of Political Economy*, 89: 914-927

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

Morlan, R. L. 1984. "Municipal versus national election voter turnout: Europe and the United States." *Political Science Quarterly*, 99:457-70.

Muir, C. H. 1898. "Naturalization." U.S. Infantry and Cavalry School Lectures.

Naidu, Suresh. 2012. "Suffrage, Schooling, and Sorting in the Post-Bellum South", NBER Working Paper 18129

Nebraska Blue Book and Historical Register, 1918

Neuman, Gerald. 1992. "We are the People: Alien Suffrage in German and American Perspective", *Michigan journal of International Law*

Official Report of the Nebraska State Canvassing Board, 1918

Official Manual of the State of Missouri for the years 1913-1914

Pritchett, J.B. 1989. "The Burden of Negro Schooling: Tax Incidence and Racial Redistribution in Postbellum North Carolina." *Journal of Economic History*, 966–973.

Raskin, James. 1993. "Legal Aliens, Local Citizens: the Historical, Constitutional, and Theoretical Meanings of Alien Suffrage." *University of Pennsylvania Law Review*

Rauch, James. 1995. "Bureaucracy, Infrastructure, and Economic Growth: Evidence from US Cities During the Progressive Era", *American Economic Review*, 85(4): 968-979

Rosberg, Gerald M. 1977. "Aliens and Equal Protections: Why Not the Right to Vote?" Michigan Law Review.

Rosenbloom, Joshua. 1994. “Looking for Work, Searching for Workers: U.S. Labor Markets after the Civil War,” *Social Science History,* 18, 377-403

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. *The Integrated Public Use Microdata Series: Version 6.0* [Machine-readable database]. Minneapolis: University of Minnesota, 2015.

Settle, Jamie E., Robert Bond, Lorenzo Coviello, Christopher J. Fariss, James Fowler, Jason Jones, Adam D. I. Kramer, Cameron Marlow. 2016. “From Posting to Voting: The Effects of Political Competition on Online Political Engagement”, *Political Science Research and Methods*, 4(2): 361-378

Shertzer, Allison. 2013. “Immigrant Group Size and Political Mobilization: Evidence from European Migration to the Unites States.” *Journal of Public Economics*, 139: 1-12.

Smith, Marian L. 1998. *Prologue Magazine*. National Archives and Records Administration.

Stephens, Mel ad Dou-Yan Yang. 2014. “Compulsory Education and the Benefits of Schooling.” *American Economic Review*, 104(6):1777-1792.

Stephens, Mel and Kerwin Charles. 2013. “Employment, Wages and Voter Turnout.” *American Economic Journal: Applied Economics*, 5(4):111-143

Swift, Elaine K., Robert G. Brookshire, David T. Canon, Evelyn C. Fink, John R. Hibbing, Brian D. Humes, Michael J. Malbin, and Kenneth C. Martis. Database of [United States] Congressional Historical Statistics, 1789-1989. ICPSR03371-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2009-02-03.

Tarr, Alan G. 1996. “State Constitutional Politics: an Historical Perspective.” *Constitutional Politics in the States: Contemporary Controversies and Historical Patterns.* Greenwood Press.

Tienda, M., 2002. “Demography and the social contract.” *Demography,* 39(4): 587-616.

Troeksen, Werner. 1998. “Patronage and Public-Sector Wages in 1896.” Journal of Economic History, 59(2): 424-446.

Ujhelyi, Gergely. 2014 “Civil Service Rules and Policy Choices: Evidence from US State Governments”, *American Economic Journal: Economic Policy*, 6(2): 338-380

Van Nuys, Frank. 2002. *Americanizing the West: Race, Immigrants, and Citizenship, 1890-1930.* University Press of Kansas.

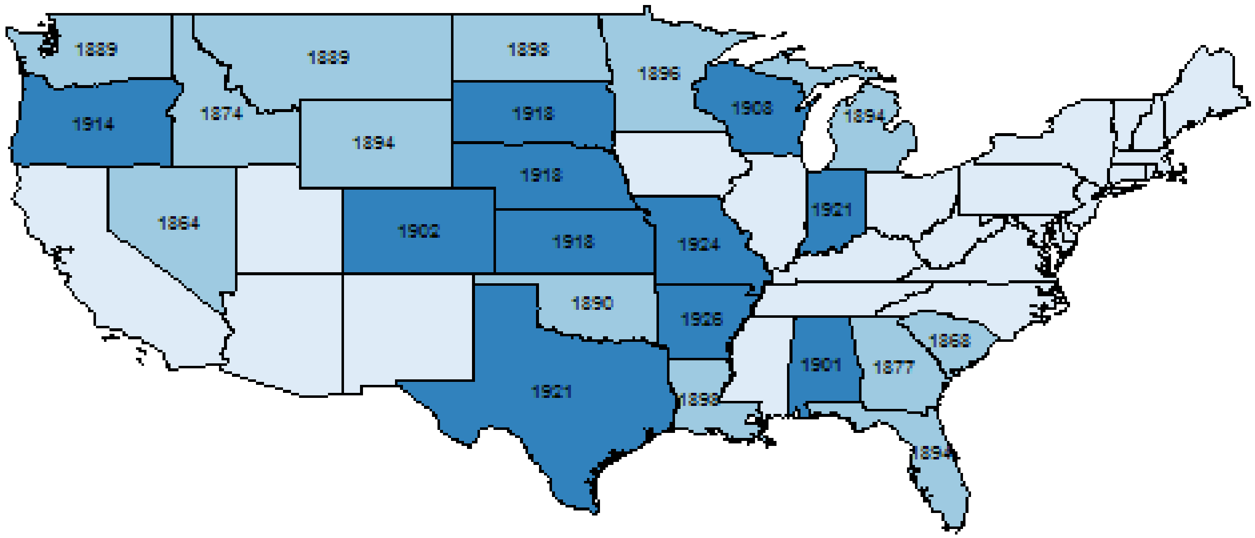
Vernby, Kare. 2013. “Inclusion and Public Policy: Evidence from Sweden’s Introduction of Noncitizen suffrage.” American Journal of Political Science, 57(1): 15-29.

Wallis, John. 1998. “The Political Economy of New Deal Spending Revisited, Again: With and without Nevada.” *Explorations in Economic History*, 35: 140-170

Wallis, John Joseph, “NBER/University of Maryland State Constitution Project”, www.stateconstitutions.umd.edu

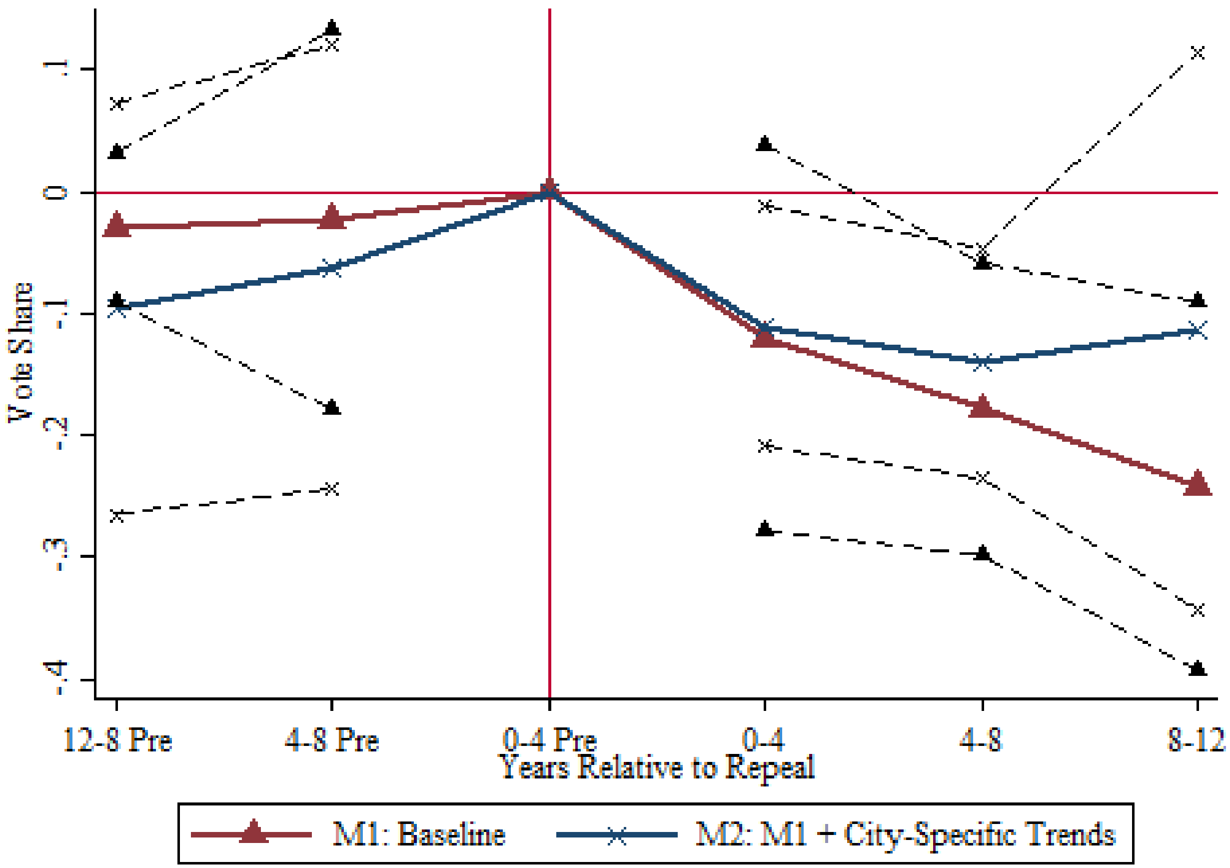
Williams, Hattie Plum. 1912. “The Road to Citizenship: A study of Naturalization in a Nebraska County.” *Political Science Quarterly.*

Figure 1.1: Dates of Repeal of Alien Voting



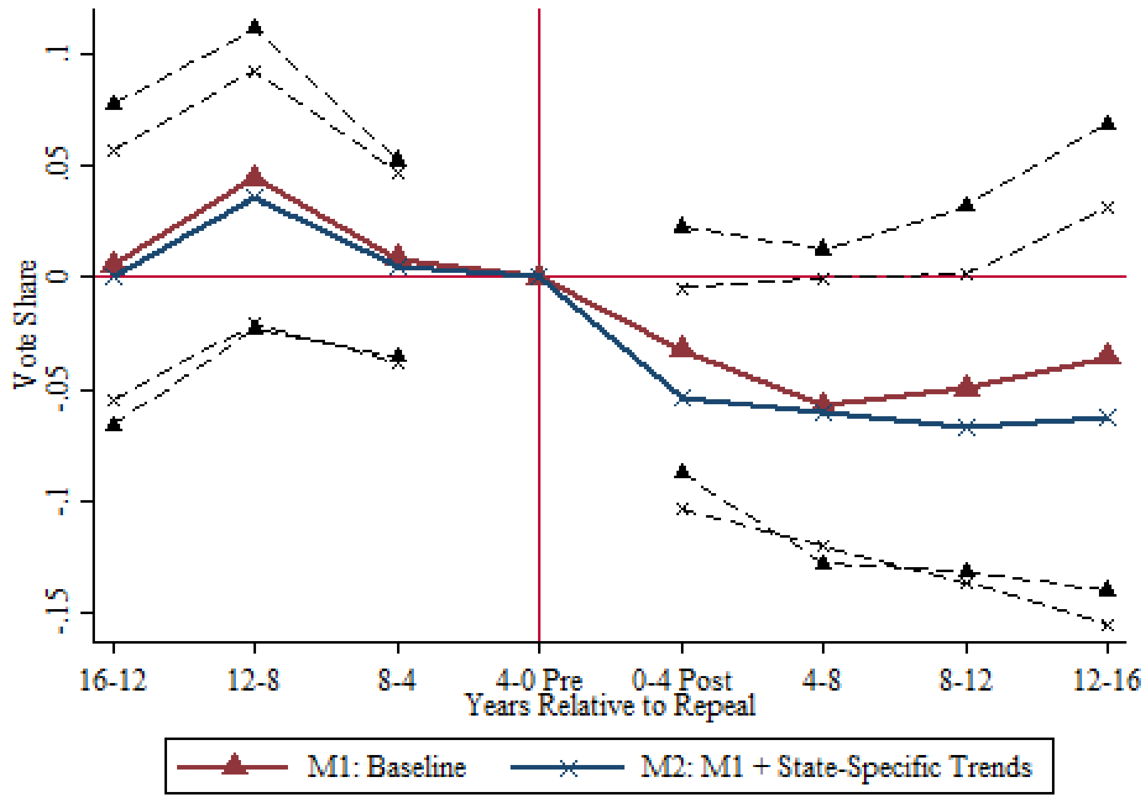
Notes: dates of repeal are based on Hayduk (2006), Keyssar (2000), and state constitutions.

Figure 1.2: Effect of Disenfranchisement on Mayoral Vote Totals



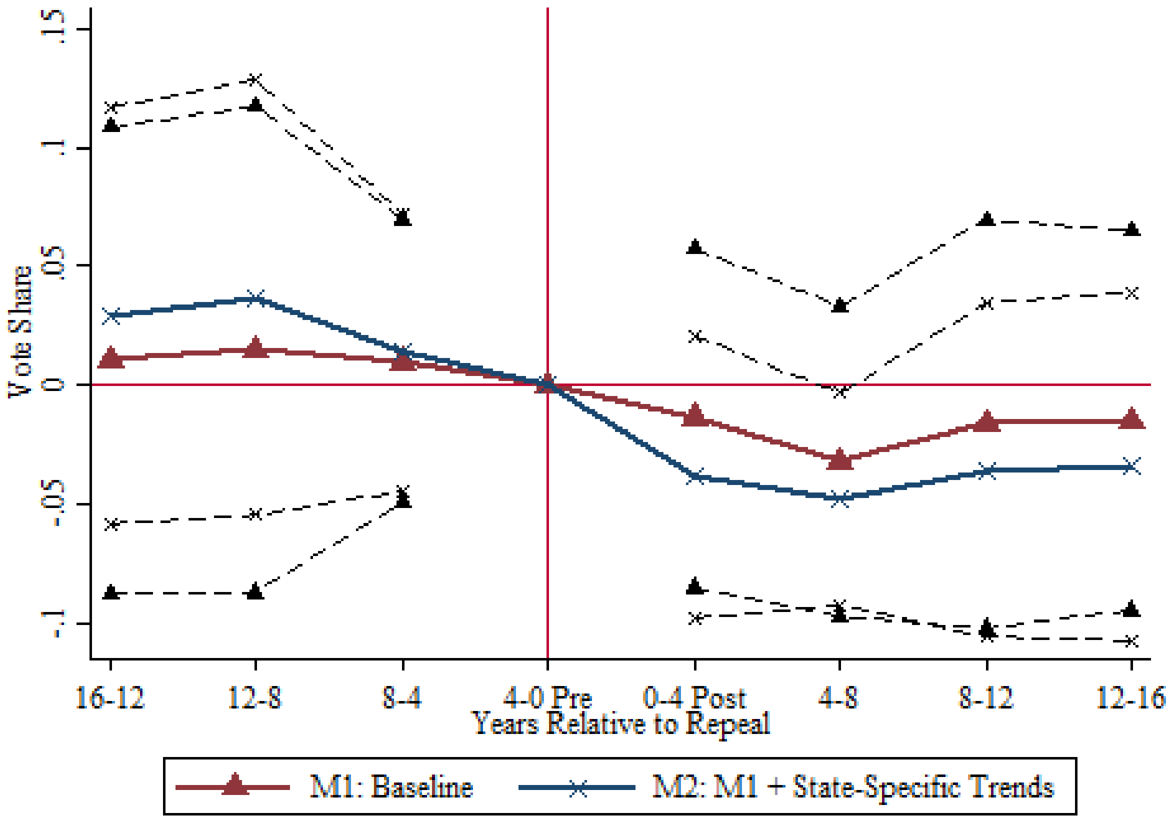
Notes: data on mayoral votes totals is from Holli and Jones (1981) and the city of San Diego. The cities in the sample are Baltimore, Boston, Buffalo, Chicago, Cincinnati, Cleveland, Detroit, Los Angeles, Milwaukee, New Orleans, New York, Philadelphia, Pittsburgh, St. Louis, San Diego, San Francisco. Controls are fraction foreign born and an indicator for the presence of women's suffrage, poll taxes, and literacy tests. The specification includes city and year fixed effects and city-specific linear trends. 95% confidence intervals are displayed. Standard errors are calculated via wild bootstrap.

Figure 1.3: Effect of Immigrant Disenfranchisement on Gubernatorial Vote Totals



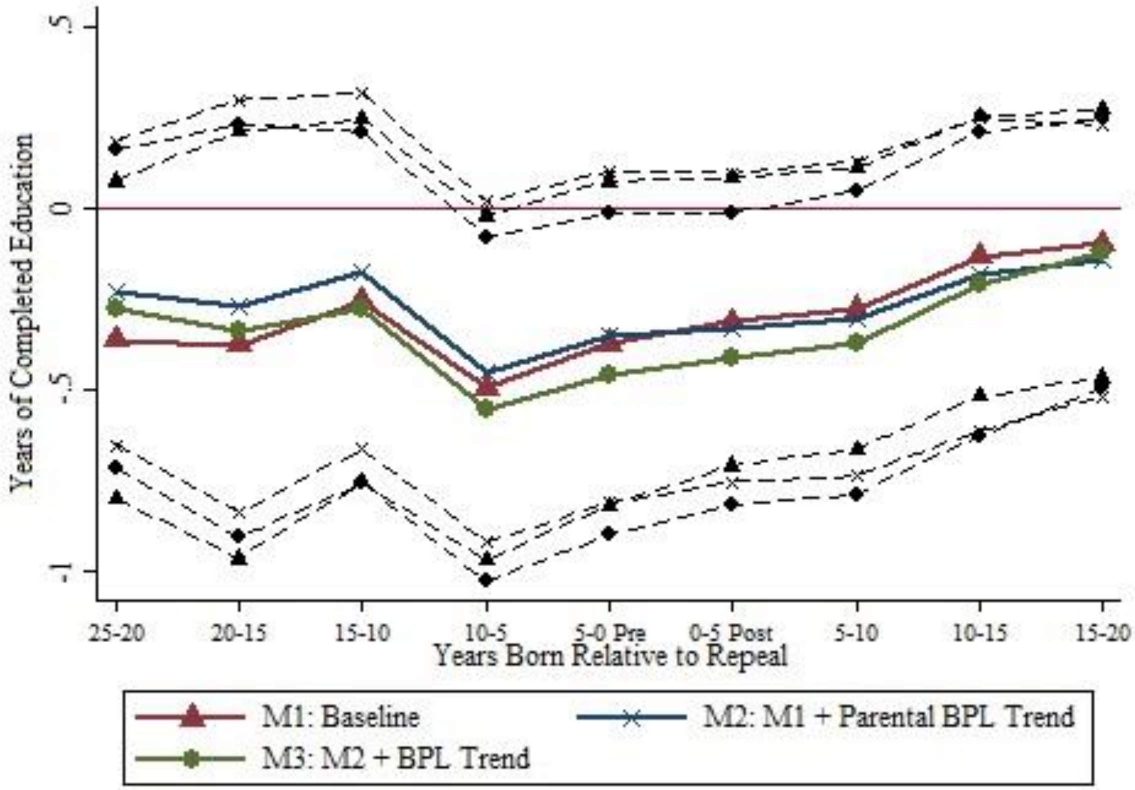
Notes: gubernatorial election data are from Burnham et al. (1991). State level controls are fraction of adult males that are non-citizen and indicators for the presence of women's suffrage, the poll tax, and presence of literacy tests. WA, ID, OK, and NV are excluded from the sample because they repealed alien voting upon entry to the U.S., and so have no pre-repeal observations in the sample. 95% confidence intervals are shown. Standard errors are clustered at the state level.

Figure 1.4: Effect of Immigrant Disenfranchisement on Presidential Vote Totals



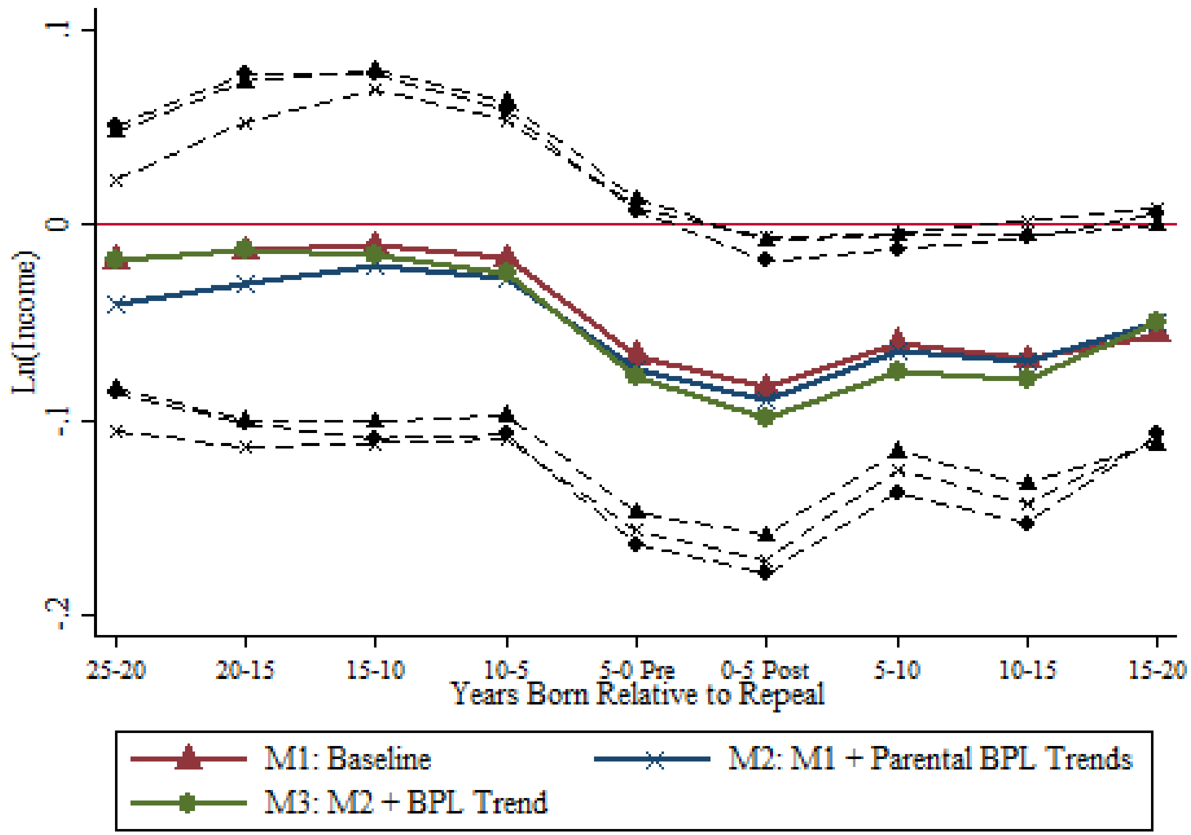
Notes: presidential election data are from Clubb et al. (2006). State level controls are the fraction of adult men that are not citizens and indicators for the presence of women's suffrage, the poll tax, and presence of literacy tests. WA, ID, OK, and NV are excluded from the sample because they repealed alien voting upon entry to the U.S., and so have no pre-repeal observations in the sample. 95% confidence intervals are shown. Standard errors are clustered at the state level.

Figure 1.5: Intergenerational Effect of Parental Disenfranchisement on Years of Completed Education



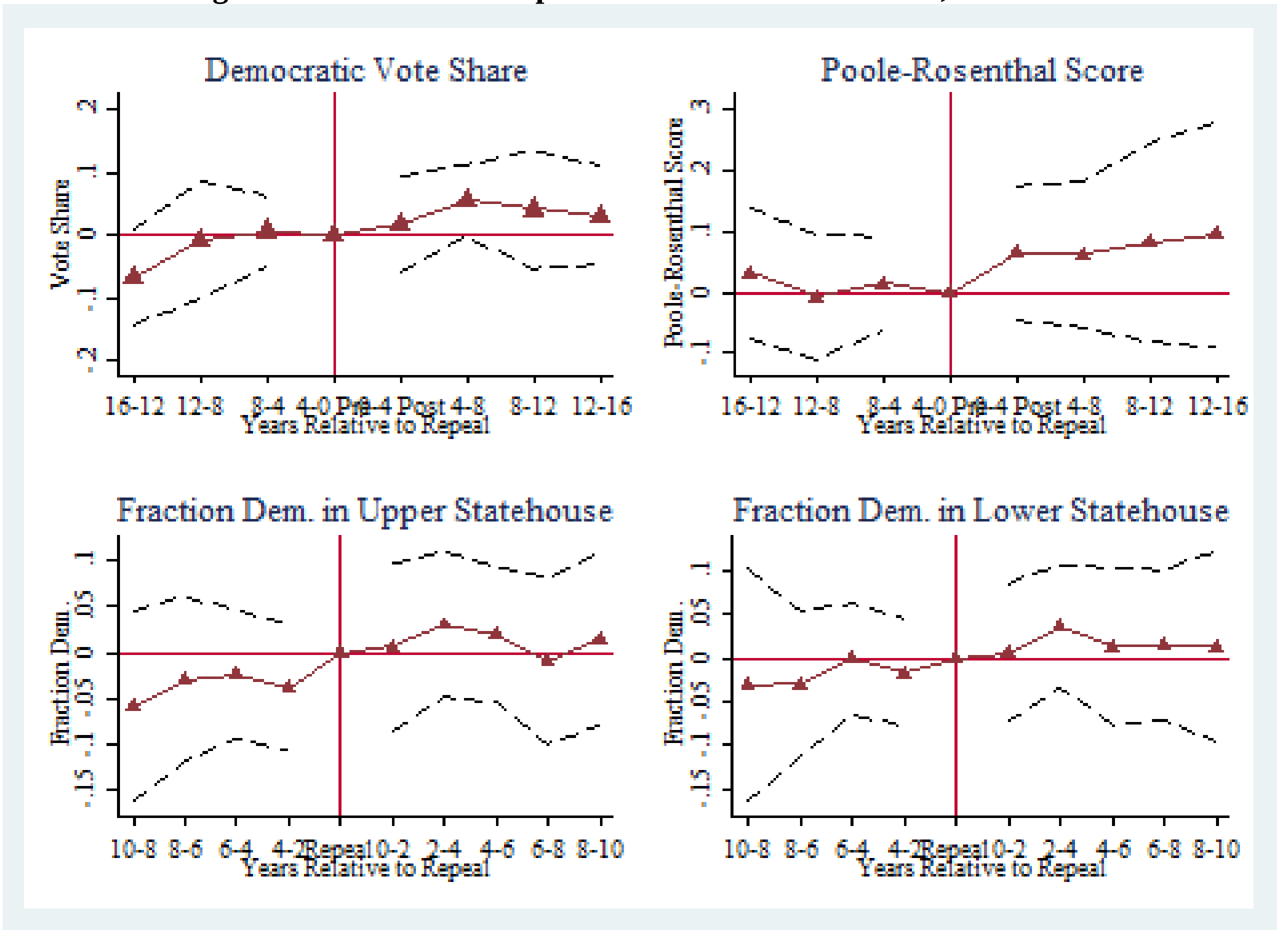
Notes: Data are from the full count 1940 Census and IPUMS 1950 and 1960 Census samples. Sample is restricted to sample line native males aged 25 and older. Individual level controls include age, age squared, educational attainment, race indicators, father nativity, and father and mother indicators. All specifications include birthplace fixed effects, birth cohort fixed effects, state of residence indicators, and survey year fixed effects. “Model 2” includes a set of birth-region specific trends in birth year for each parent; “Model 3” includes a set of birth-state specific trends in birth year. IPUMS sample line weights are used in estimation. 95% confidence intervals are displayed. Standard errors are clustered at the birth state level.

Figure 1.6: Intergenerational Effect of Parental Disenfranchisement on Ln(Income)



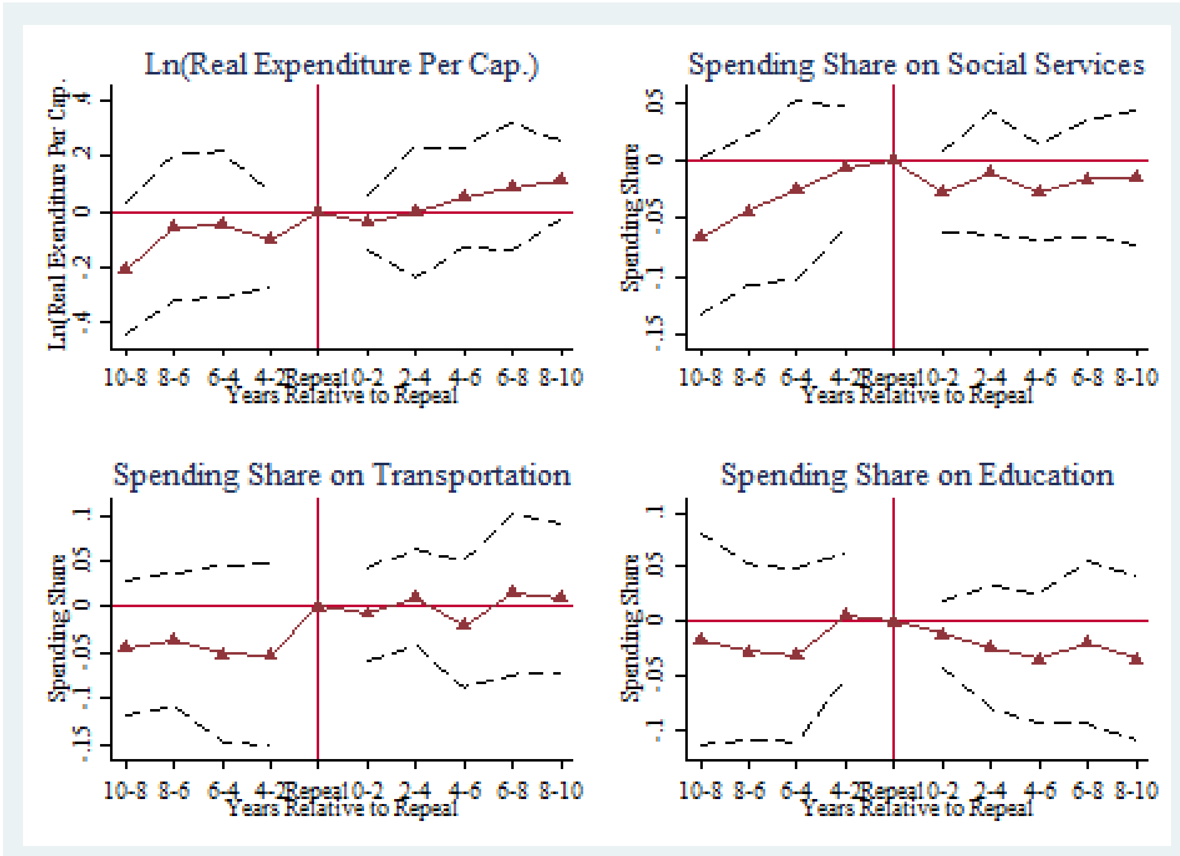
Notes: Data are from the full count 1940 Census and IPUMS 1950 and 1960 Census samples. Sample is restricted to sample line native males aged 25-64 (inclusive). Individual level controls include age, age squared, educational attainment, race indicators, father nativity, and father and mother indicators. All specifications include birthplace fixed effects, birth cohort fixed effects, state of residence indicators, and survey year fixed effects. “Model 2” includes a set of birth-region specific trends in birth year for each parent; “Model 3” includes a set of birth-state specific trends in birth year. IPUMS sample line weights are used in estimation. 95% confidence intervals are displayed. Standard errors are clustered at the birth state level.

Figure 1.7.1: Political Responses to Disenfranchisement, 1870-1940



Notes: gubernatorial election data are from Burnham et al. (1991). Data on compsition of state legislatures are from Burham et al. (1989). "DWNOM1" was used from the Poole-Rosenthal index. State level controls are the fraction of adult males that are non-citizen, and indicators for the presence of women's suffrage, poll taxes, and literacy tests. All specifications include state-specific linear trends. WA, ID, OK, and NV are excluded from the sample because of data availability. 95% confidence intervals are shown. Standard errors are clustered at the state level.

Figure 1.7.2 State Expenditure Responses to Disenfranchisement, 1870-1940



Notes: state spending data are from Lott and Kenny (1999). State level controls are the fraction of adult males that are non-citizen, and indicators for the presence of women's suffrage, poll taxes, and literacy tests. All specifications include state-specific linear trends. WA, ID, OK, and NV are excluded from the sample because of data availability. 95% confidence intervals are shown. Standard errors are clustered at the state level.

Table 1.1: Fraction of Males age 21 and over that were Non-Citizens in year of Disenfranchisement

State	Fraction Non-citizen
AL	.008
AR	.004
CO	.110
FL	.047
GA	.004
ID	.142
IN	.042
KS	.038
LA	.035
MI	.130
MN	.158
MO	.028
MT	.181
ND	.226
NE	.059
NV	.300
OK	.019
OR	.143
SD	.061
TX	.090
WA	.166
WI	.173
WY	.135

Notes: the fraction affected is defined as the population of non-citizen foreign-born males over age 21 divided by the entire population of males over age 21. For 1870 and 1900-1930, IPUMS census samples contain citizenship information, and were used in construction. For 1890, Census aggregates from Haines (2010) and table 71 from the 1890 Census were used. Numerator and denominator were linearly interpolated (separately) between decades. Nevada and Georgia, which repealed their provisions in 1864 and 1877, respectively, are assigned values from the 1870 Census.

Table 1.2 – Details of State Constitution Amending Process

State	Year Repeal	Year Enter Union	Relevant Constitutional Amending Article	Legislative Majority	# Passages Needed	Election Majority Standard
AL	1901	1819	17	2/3	One	All voters
AR	1926	1836	19	½	One	All voters
CO	1902	1876	19	2/3	One	Voting thereon
FL	1894	1845	17	3/5	One	Voting thereon
GA	1877	1788	12	2/3	Two	
ID	1874	1890	20	2/3	One	All voters
IN	1921	1816	16	½	Two	All voters
KS	1918	1861	14	2/3	One	Voting thereon
LA	1898	1812	256	2/3	One	Voting thereon
MI	1894	1837	20	2/3	One	Voting thereon
MN	1896	1859	14	½	One	Voting thereon*
MO	1924	1821	15	½	One	All voters
MT	*1894*	*1889*	*14*	*2/3*	*One*	*Voting thereon*
ND	1898	1889	15	½	Two	Voting thereon
NE	1918	1867	15	3/5	One	All voters
NV	*1864*	*1864*	*16*	*½*	*Two*	*Voting thereon*
OK	1890	1907	24	½	One	Voting thereon
OR	1914	1859	17	½	One	Voting thereon
SD	1918	1889	23	½	One	Voting thereon
TX	1921	1845	17	2/3	One	Voting thereon
WA	*1894*	*1889*	*23*	*2/3*	*One*	*Voting thereon*
WI	1908	1848	12	½	Two	Voting thereon
WY	*1894*	*1890*	*20*	*2/3*	*One*	*All voters*

Notes: amending information taken from state constitution in year of repeal (Wallis). Italics indicate that the state repealed alien voting upon entry into the Union.
*This was amended in 1898 to "majority of all voters having voted at said election"

Table 1.3 - Effect of Disenfranchisement on Incidence of Filing First Papers

	Filed First Papers?	Filed First Papers?	Filed First Papers?
Repeal	-0.152	-0.133	-0.068
	(0.033)***	(0.034)***	(0.037)
New	-0.137	-0.063	-0.061
	(0.010)***	(0.011)***	(0.010)***
New * Repeal	-0.073	-0.067	-0.072
	(0.019)***	(0.021)***	(0.020)***
Specification	Baseline	Model 2	Model 3
R^2	0.13	0.17	0.18
N	161,821	161,170	161,170

Notes: Data are from IPUMS Census samples, 1900-1930. Sample is restricted to non-citizen, foreign-born males over age 21 with less than 15 years of residence in the U.S. All specifications include individual level controls (age, age squared, literacy status, urban status), state and year fixed effects, and state-level controls: an indicator for the pre-repeal period in alien voting states, and an indicator for the presence of state regulation barring aliens from public employment (Fishback et al. 2009). Model 2 includes birth-region specific linear trends in year of immigration. Model 3 includes state-specific census-year linear trends. IPUMS person weights are used in estimation. * p<0.1; ** p<0.05; *** p<0.01

Table 1.4 - Short-run effects of Disenfranchisement on Labor Market Outcomes for Immigrant Males

	In Labor Force	Occupational Standing	In Labor Force	Occupational Standing	In Labor Force	Occupational Standing
	(1)	(2)	(3)	(4)	(5)	(6)
Repeal	-0.011	-0.183	-0.011	-0.052	-0.003	0.375
	(0.006)*	(0.486)	(0.006)*	(0.421)	(0.007)	(0.312)
New	-0.003	-1.227	-0.000	-0.175	0.000	-0.171
	(0.002)	(0.108)***	(0.003)	(0.159)	(0.003)	(0.160)
New * Repeal	-0.000	0.083	0.000	0.132	-0.001	0.144
	(0.005)	(0.293)	(0.005)	(0.226)	(0.005)	(0.242)
Specification	Baseline	Baseline	Model 2	Model 2	Model 3	Model 3
R^2	0.02	0.07	0.02	0.08	0.02	0.08
N	205,309	205,309	204,675	204,675	204,675	204,675

Notes: Data are from IPUMS Census samples, 1900-1930. Sample is restricted to non-farm males age 21-64 (inclusive), with less than 15 years of residence in the U.S. All specifications include individual level controls (age, age squared, literacy status, urban status), state and year fixed effects, and state-level controls: an indicator for the pre-repeal period in alien voting states, and an indicator for the presence of state regulation barring aliens from public employment (Fishback et al. 2009). Model 2 includes birth-region specific linear trends in year of immigration. Model 3 includes state-specific census-year linear trends. IPUMS person weights are used in estimation. * p<0.1; ** p<0.05; *** p<0.01

Table 1.5 – Short-run effects of Disenfranchisement on Labor Market Outcomes of Natives for Native Males

	In Labor Force (1)	Occupational Standing (2)	In Labor Force (3)	Occupational Standing (4)
Repeal	0.002	0.093	0.002	0.256
	(0.004)	(0.164)	(0.006)	(0.212)
Specification	Baseline	Baseline	Model 3	Model 3
R^2	0.02	0.07	0.02	0.07
N	1,757,589	1,757,589	1,757,589	1,757,589

Notes: Data are from IPUMS Census samples, 1900-1930. Sample is restricted to native non-farm males age 21-64 (inclusive). All specifications include individual level controls (age, age squared, literacy status, urban status), state and year fixed effects, and state-level controls: an indicator for the pre-repeal period in alien voting states, and an indicator for the presence of state regulation barring aliens from public employment (Fishback et al. 2009). Model 3 includes state-specific census-year linear trends. IPUMS person weights are used in estimation. * p<0.1; ** p<0.05; *** p<0.01

Table 1.6 – Effect of Disenfranchisement on Probability of Obtaining Public-Sector Employment for Immigrant Males

	Public 1	Public 2	Public 3	Public 1	Public 2	Public 3
	(1)	(2)	(3)	(4)	(5)	(6)
Repeal	-0.003	-0.002	-0.001	-0.006	-0.007	-0.001
	(0.003)	(0.004)	(0.001)	(0.003)*	(0.006)	(0.001)
New	0.002	0.001	0.001	0.001	0.001	0.001
	(0.001)*	(0.001)	(0.001)	(0.001)*	(0.001)	(0.001)
New * Repeal	-0.004	-0.002	-0.003	-0.004	-0.002	-0.002
	(0.001)***	(0.001)	(0.001)**	(0.001)***	(0.001)	(0.001)**
Specification	Model 2	Model 2	Model 2	Model 3	Model 3	Model 3
R^2	0.00	0.01	0.00	0.00	0.01	0.00
N	204,675	204,675	204,675	204,675	204,675	204,675

Notes: Data are from IPUMS Census samples, 1900-1930. Sample is restricted to non-farm males age 21-64 (inclusive), with less than 15 years of residence in the U.S. All specifications include individual level controls (age, age squared, literacy status, urban status), state and year fixed effects, and state-level controls: an indicator for the pre-repeal period in alien voting states, and an indicator for the presence of state regulation barring aliens from public employment (Fishback et al. 2009). Model 2 includes birth-region specific linear trends in year of immigration. Model 3 includes state-specific census-year linear trends. IPUMS person weights are used in estimation. * p<0.1; ** p<0.05; *** p<0.01

Table 1.7 –Effect of Disenfranchisement on Probability of Obtaining Public-Sector Employment for Native Males

	Public 1	Public 2	Public 3	Public 1	Public 2	Public 3
	(1)	(2)	(3)	(4)	(5)	(6)
Repeal	-0.001	-0.002	-0.001	0.002	-0.002	0.000
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)
Specification	Baseline	Baseline	Baseline	Model 3	Model 3	Model 3
R^2	0.01	0.01	0.00	0.01	0.01	0.00
N	1,757,589	1,757,589	1,757,589	1,757,589	1,757,589	1,757,589

Notes: Data are from IPUMS Census samples, 1900-1930. Sample is restricted to native non-farm males age 21-64 (inclusive). All specifications include individual level controls (age, age squared, literacy status, urban status), state and year fixed effects, and state-level controls: an indicator for the pre-repeal period in alien voting states, and an indicator for the presence of state regulation barring aliens from public employment (Fishback et al. 2009). Model 3 includes state-specific census-year linear trends. IPUMS person weights are used in estimation. * p<0.1; ** p<0.05; *** p<0.01

Table 1.8 - Migration Response to Disenfranchisement, 1870-1940

	Ln(Total Pop) (1)	Ln(Total Foreign Born Pop) (2)	Fraction Foreign Born (3)	Fraction Foreign Born Illiterate (4)
More than 30 years pre repeal	0.025	-0.061	-0.049	0.015
	(0.276)	(0.431)	(0.050)	(0.030)
20-30 years pre-repeal	0.175	0.157	-0.010	0.026
	(0.251)	(0.318)	(0.019)	(0.026)
10-20 years pre-repeal	0.031	0.045	-0.009	0.017
	(0.130)	(0.176)	(0.014)	(0.027)
0-10 years post-repeal	-0.066	-0.084	0.011	-0.007
	(0.140)	(0.180)	(0.010)	(0.010)
10-20 years post-repeal	-0.303	-0.342	0.022	-0.011
	(0.226)	(0.287)	(0.017)	(0.020)
20-30 years post-repeal	-0.638	-0.738	0.028	-0.014
	(0.342)*	(0.423)*	(0.025)	(0.021)
30 +	-1.200	-1.266	0.045	-0.008
	(0.447)***	(0.560)**	(0.029)	(0.032)
R^2	0.97	0.95	0.91	0.76
N	479	479	479	479

Data are from IPUMS samples, published Census aggregates from 1890 (Haines 2010), and table 37 from the 1890 Census. The dependent variable in column 4 is the ratio of the number of foreign-born individuals over age 10 that is illiterate to the entire population of foreign-born. All specifications include state fixed effects, decade fixed effects, and state-specific linear trends, and indicators for the presence of women's suffrage, poll taxes, and literacy tests (Lott and Kenny 1999). Standard errors are clustered at the state level.* p<0.1; ** p<0.05; *** p<0.01

Table 1.9 - Effect of Disenfranchisement on Legislator Voting Behavior and Municipal Public Spending

	Votes for imm. Restriction		Municipal Spending Outcomes			
	Fraction Yes	Fraction Yes	Log Real spending per capita	Log Real spending per capita	Share on Education	Share on Education
Repeal	-0.130	-0.043	0.010	0.024	-0.003	-0.008
	(0.135)	(0.074)	(0.070)	(0.090)	(0.006)	(0.010)
Fraction Foreign Born	4.779	2.051	-0.207	-0.543	-0.018	-0.017
	(3.882)	(7.184)	(0.351)	(0.288)*	(0.041)	(0.049)
Fraction Foreign Born ^2	-28.162	-14.429				
	(12.230)**	(19.770)				
State (city) fixed effects	Y	Y	Y	Y	Y	Y
State (city)- specific linear trends	N	Y	N	Y	N	Y
R^2	0.39	0.62	0.81	0.87	0.72	0.79
N	565	565	3,124	3,124	3,114	3,114

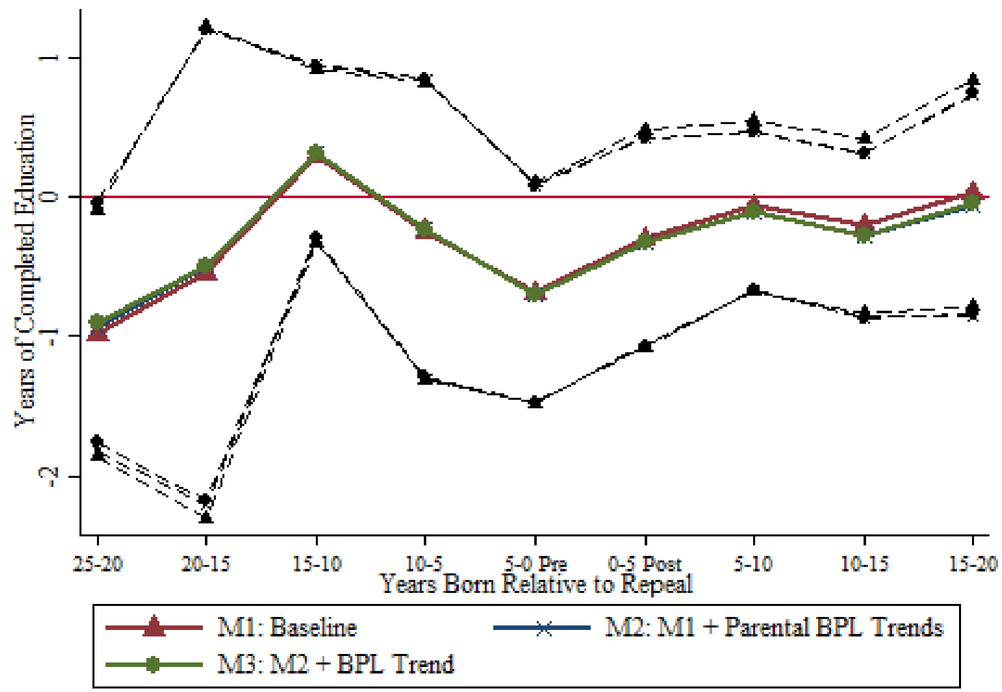
Dates of votes are from Goldin (1994). I use all votes listed in table 7.1, except those on 12/14/98, 5/27/2, 6/25/6, and 4/19/12. Roll-call data are from Swift et al. (2009). City spending data are from the *Financial Statistics of Cities and Bureau* of Labor Statistics Bulletins 20 and 36. I use odd-numbered years from 1899-1929 in my sample except for 1913: data for that year is missing, so 1912 is used instead. All specifications include year fixed effects, indicators for the presence of women's suffrage, poll taxes, and literacy tests, and the fraction foreign-born. Standard errors have been clustered at the state (or city) level. * p<0.1; ** p<0.05; *** p<0.01

Table 1.10 – Effect of Disenfranchisement on English Language Proficiency of Immigrants

	Speak English	Speak English	Speak English
Repeal	-0.071	-0.049	-0.050
	(0.029)**	(0.025)*	(0.015)***
New	-0.230	-0.307	-0.456
	(0.017)***	(0.019)***	(0.026)***
Repeal * New	0.019	0.018	0.018
	(0.030)	(0.024)	(0.022)
Specification	Baseline	Model 2	Model 3
R^2	0.28	0.34	0.35
N	225,453	224,775	224,775

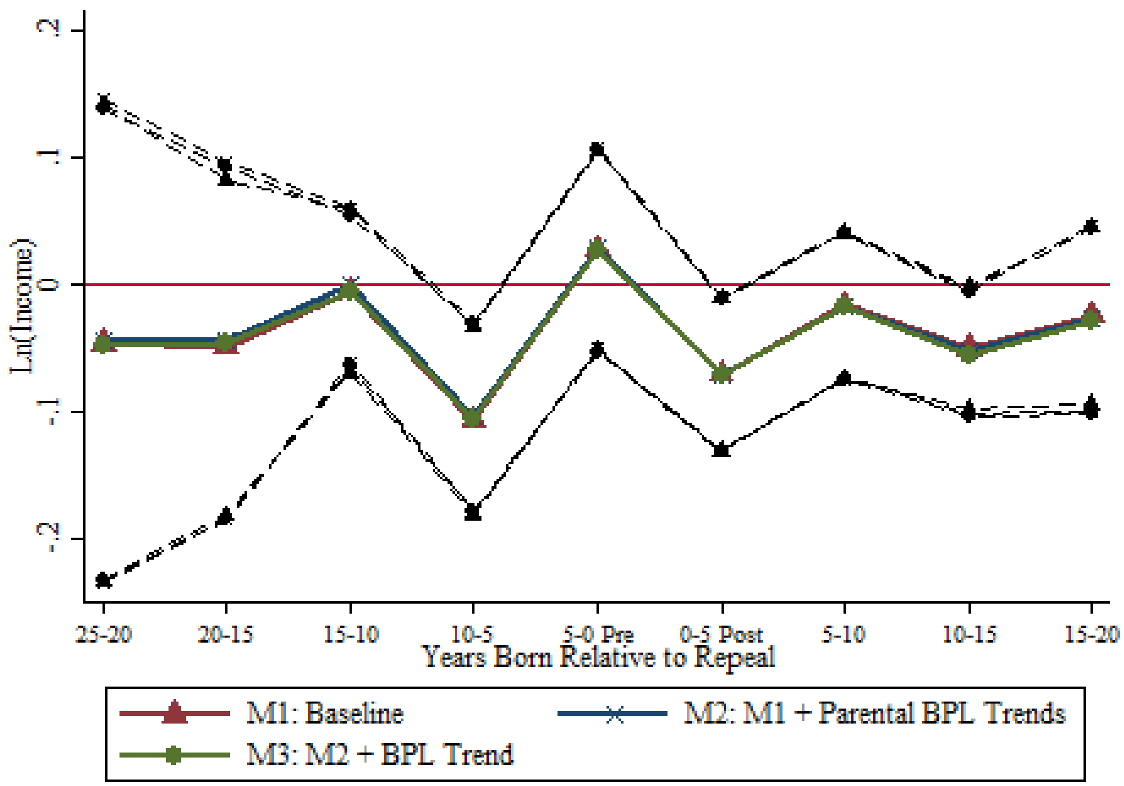
Notes: Data are from IPUMS Census samples, 1900-1930. Sample is restricted to foreign-born males age 21 and over with less than fifteen years' residence in the U.S. All specifications include individual level controls (age, age squared, literacy status, urban status), state and year fixed effects, and state-level controls: an indicator for the pre-repeal period in alien voting states, and an indicator for the presence of state regulation barring aliens from public employment (Fishback et al. 2009). Model 2 includes birth-region specific linear trends in year of immigration. Model 3 includes state-specific census-year linear trends. IPUMS person weights are used in estimation. * p<0.1; ** p<0.05; *** p<0.01

Appendix Figure 1.1 – Alternative Specification for Intergenerational Effect of Parental Disenfranchisement on Years of Completed Education



Notes: Data are from the full count 1940 Census. Sample is restricted to sample line native males aged 25 and older, born in states that either never had alien voting, or that repealed alien voting prior to 1910. Individual level controls include age, age squared, educational attainment, race indicators, father nativity, and father and mother indicators. All specifications include birthplace fixed effects, birth cohort fixed effects, state of residence indicators, and survey year fixed effects. "Model 2" includes a set of birth-region specific trends in birth year for each parent; "Model 3" includes a set of birth-state specific trends in birth year. IPUMS sample line weights are used in estimation. 95% confidence intervals are displayed. Standard errors are clustered at the birth state level.

Appendix Figure 1.2 – Alternative Specification for Intergenerational Effect of Parental Disenfranchisement on Ln(Income)



Notes: Data are from the full count 1940 Census. Sample is restricted to sample line native males aged 25-64 (inclusive). Individual level controls include age, age squared, educational attainment, race indicators, father nativity, and father and mother indicators. All specifications include birthplace fixed effects, birth cohort fixed effects, state of residence indicators, and survey year fixed effects. “Model 2” includes a set of birth-region specific trends in birth year for each parent; “Model 3” includes a set of birth-state specific trends in birth year. IPUMS sample line weights are used in estimation. 95% confidence intervals are displayed. Standard errors are clustered at the birth state level.

Appendix Table 1.1.1 - Effect of Disenfranchisement on Mayoral Voter Activity, 1880-1924

	Voting Activity	Voting Activity
12+ Years Pre	-0.071	-0.131
	(0.059)	(0.167)
8-12 Years Pre	-0.029	-0.096
	(0.031)	(0.086)
4-8 Years Pre	-0.023	-0.062
	(0.079)	(0.093)
0-4 Years Post	-0.120	-0.110
	(0.081)	(0.050)**
4-8 Years Post	-0.178	-0.140
	(0.061)***	(0.048)***
8-12 Years Post	-0.242	-0.114
	(0.077)***	(0.117)
12+ Years Post	-0.249	-0.065
	(0.223)	(0.079)
Specification	Baseline	Model 2
R^2	0.66	0.73
N	254	254

Notes: data on mayoral votes totals is from Holli and Jones (1981) and the cities of San Diego and Seattle. The cities in the sample are Baltimore, Boston, Buffalo, Chicago, Cincinnati, Cleveland, Detroit, Los Angeles, Milwaukee, New Orleans, New York, Philadelphia, Pittsburgh, St. Louis, San Diego, San Francisco, and Seattle. City level controls an indicator for the presence of women's suffrage, poll taxes, literacy tests, and the fraction of the population that is foreign born. Baseline specification includes city and year fixed effects. Model 2 includes city-specific trends. Standard errors are calculated via wild bootstrap.
* p<0.1; ** p<0.05; *** p<0.01

Appendix Table 1.1.2 - Effect of Disenfranchisement on Gubernatorial Voting Activity, 1870-1940

	Voting Activity	Voting Activity
16+ Years Pre	0.037	0.021
	(0.051)	(0.041)
12-16 Years Pre	0.006	0.001
	(0.037)	(0.029)
8-12 Years Pre	0.044	0.036
	(0.034)	(0.029)
4-8 Years Pre	0.009	0.005
	(0.022)	(0.022)
0-4 Years Post	-0.032	-0.054
	(0.028)	(0.025)**
4-8 Years Post	-0.058	-0.060
	(0.036)	(0.031)*
8-12 Years Post	-0.050	-0.067
	(0.042)	(0.035)*
12-16 Years Post	-0.036	-0.062
	(0.053)	(0.047)
16+ Years Post	-0.087	-0.056
	(0.060)	(0.054)
Specification	Baseline	Model 2
R^2	0.83	0.93
N	1,169	1,169

Notes: gubernatorial election data are from Burnham et al. (1991). State level controls an indicator for the presence of women's suffrage, poll taxes, literacy tests, and the fraction of the adult male population that is non-citizen. Baseline specification includes state and year fixed effects. Model 2 includes state-specific trends. WA, ID, OK, and NV are excluded from the sample because there are no pre-disenfranchisement observations. 95% confidence intervals are shown. Standard errors are clustered at the state level. * p<0.1; ** p<0.05; *** p<0.01

Appendix Table 1.1.3 - Effect of Disenfranchisement on Presidential Voting Activity, 1870-1940

	Voting Activity	Voting Activity
16+ Years Pre	0.021	0.066
	(0.064)	(0.056)
12-16 Years Pre	0.010	0.029
	(0.050)	(0.045)
8-12 Years Pre	0.015	0.037
	(0.052)	(0.047)
4-8 Years Pre	0.010	0.014
	(0.030)	(0.030)
0-4 Years Post	-0.014	-0.039
	(0.036)	(0.030)
4-8 Years Post	-0.032	-0.048
	(0.033)	(0.023)**
8-12 Years Post	-0.016	-0.036
	(0.044)	(0.036)
12-16 Years Post	-0.015	-0.034
	(0.041)	(0.037)
16+ Years Post	-0.077	-0.039
	(0.054)	(0.038)
Specification	Baseline	Model 2
R^2	0.86	0.95
N	744	744

Notes: gubernatorial election data are from Clubb et al. (2006). State level controls an indicator for the presence of women's suffrage, poll taxes, literacy tests, and the fraction of the adult male population that is non-citizen. Baseline specification includes state and year fixed effects. Model 2 includes state-specific trends. WA, ID, OK, and NV are excluded from the sample because there are no pre-disenfranchisement observations. 95% confidence intervals are shown. Standard errors are clustered at the state level. * p<0.1; ** p<0.05; *** p<0.01

Appendix Table 1.2 - Public Sector Employment Codes

Coding 1 ("Public 1"). Government job indicator equals 1 if occupation equals:

1900: Clerks in government offices; employees of government (not clerks); Officials of Government

1910 and 1920: Firemen – fire department; guards, watchmen, and doorkeepers; garbagemen and scavengers; other laborers (public service); detectives; marshals and constables; probation and truant offices; sheriffs; Officials and inspectors – city; officials and inspectors – county; official and inspectors – state; postmasters; other United States officials; policemen; life-savers; lighthouse keepers; other occupations (public service)

1930: Firemen – fire department; Guards, watchmen, and doorkeepers; garbagemen and scavengers; detectives; marshals and constables; probation and truant offices; sheriffs; Officials and inspectors – city; officials and inspectors – county; official and inspectors – state; officials and inspectors – United States; policemen; other public service pursuits

Coding 2 ("Public 2"): Government Job indicator equals 1 if IND1950 ("Industry, 1950 basis") equals:
"Postal Service", "Federal Public Administration", "State public administration", "local public administration", and "Public Administration, level not specified"
All instances when the occupation string indicated a military position (for example, "Soldier", "Sailor", "U S Navy") were excluded from this definition.

Coding 3 ("Public 3") Government Job indicator equals 1 if IND1950 ("Industry, 1950 basis") equals:
"Local Public Administration"

Appendix Table 1.3.1 - Coefficients for Intergenerational Effect of Parental Disenfranchisement on Educational Attainment

	Highest Grade	Highest Grade	Highest Grade
Age at Repeal 25-30	-0.002	0.142	0.117
	(0.250)	(0.238)	(0.251)
Age at Repeal 20-25	-0.363	-0.235	-0.281
	(0.223)	(0.213)	(0.224)
Age at Repeal 15-20	-0.378	-0.270	-0.340
	(0.301)	(0.290)	(0.291)
Age at Repeal 10-15	-0.256	-0.175	-0.276
	(0.256)	(0.251)	(0.247)
Age at Repeal 5-10	-0.497	-0.451	-0.556
	(0.243)**	(0.239)*	(0.243)**
Age at Repeal 0-5	-0.373	-0.356	-0.459
	(0.228)	(0.234)	(0.226)**
Age at Repeal -5-0	-0.314	-0.330	-0.417
	(0.202)	(0.217)	(0.206)**
Age at Repeal -5--10	-0.278	-0.305	-0.372
	(0.198)	(0.222)	(0.213)*
Age at Repeal -10--15	-0.135	-0.184	-0.212
	(0.196)	(0.220)	(0.213)
Age at Repeal -15--20	-0.098	-0.147	-0.124
	(0.188)	(0.191)	(0.189)
Age at Repeal <-20	0.127	0.013	0.116
	(0.157)	(0.174)	(0.176)
Specification	Baseline	Model 2	Model 3
R^2	0.19	0.19	0.19
N	3,751,233	3,751,233	3,751,233

Notes: Data are from the full count 1940 Census and IPUMS 1950 and 1960 Census samples. Sample is restricted to sample line native males aged 25 and older. Individual level controls include age, age squared, educational attainment, race indicators, father nativity, and father and mother indicators. All specifications include birthplace fixed effects, birth cohort fixed effects, state of residence indicators, and survey year fixed effects. “Model 2” includes a set of birth-region specific trends in birth year for each parent; “Model 3” includes a set of birth-state specific trends in birth year. IPUMS sample line weights are used in estimation. 95% confidence intervals are displayed. Standard errors are clustered at the birth state level.

Appendix Table 1.3.2 - Coefficients for Intergenerational Effect of Parental Disenfranchisement on Ln(Income)

	Ln(Income)	Ln(Income)	Ln(Income)
Age at Repeal 25-30	0.084	0.062	0.084
	(0.105)	(0.106)	(0.103)
Age at Repeal 20-25	-0.018	-0.041	-0.018
	(0.033)	(0.033)	(0.035)
Age at Repeal 15-20	-0.013	-0.031	-0.012
	(0.044)	(0.042)	(0.046)
Age at Repeal 10-15	-0.011	-0.021	-0.016
	(0.046)	(0.046)	(0.047)
Age at Repeal 5-10	-0.017	-0.028	-0.024
	(0.041)	(0.042)	(0.042)
Age at Repeal 0-5	-0.067	-0.074	-0.078
	(0.041)	(0.042)*	(0.044)*
Age at Repeal -5-0	-0.084	-0.089	-0.098
	(0.038)**	(0.042)**	(0.041)**
Age at Repeal -5--10	-0.061	-0.065	-0.075
	(0.028)**	(0.031)**	(0.032)**
Age at Repeal -10--15	-0.069	-0.070	-0.080
	(0.032)**	(0.037)*	(0.037)**
Age at Repeal -15--20	-0.056	-0.051	-0.051
	(0.029)*	(0.030)	(0.029)*
Age at Repeal <-20	-0.017	0.001	0.002
	(0.021)	(0.020)	(0.020)
Specification	Baseline	Model 2	Model 3
R^2	0.18	0.18	0.18
N	2,479,828	2,479,828	2,479,828

Notes: Data are from the full count 1940 Census and IPUMS 1950 and 1960 Census samples. Sample is restricted to sample line native males aged 25-64 (inclusive). Individual level controls include age, age squared, educational attainment, race indicators, father nativity, and father and mother indicators. All specifications include birthplace fixed effects, birth cohort fixed effects, state of residence indicators, and survey year fixed effects. “Model 2” includes a set of birth-region specific trends in birth year for each parent; “Model 3” includes a set of birth-state specific trends in birth year. IPUMS sample line weights are used in estimation. 95% confidence intervals are displayed. Standard errors are clustered at the birth state level.

Appendix Table 1.4.1 – Intergenerational Effect on Education, sequentially dropping Regions (Model 3)

	Highest Grade (no South)	Highest Grade (no West)	Highest Grade (no Midwest)	Highest Grade (no Northeast)
Age at Repeal 25-30	0.270	0.107	-0.041	0.015
	(0.291)	(0.254)	(0.336)	(0.238)
Age at Repeal 20-25	-0.100	-0.319	-0.371	-0.351
	(0.233)	(0.225)	(0.406)	(0.206)*
Age at Repeal 15-20	-0.097	-0.360	-0.599	-0.429
	(0.278)	(0.294)	(0.574)	(0.266)
Age at Repeal 10-15	-0.046	-0.351	-0.512	-0.356
	(0.241)	(0.239)	(0.560)	(0.216)
Age at Repeal 5-10	-0.272	-0.641	-0.927	-0.633
	(0.160)*	(0.240)**	(0.593)	(0.204)***
Age at Repeal 0-5	-0.274	-0.546	-0.519	-0.545
	(0.156)*	(0.223)**	(0.585)	(0.188)***
Age at Repeal -5-0	-0.235	-0.507	-0.580	-0.486
	(0.137)*	(0.212)**	(0.480)	(0.160)***
Age at Repeal -5--10	-0.267	-0.426	-0.500	-0.464
	(0.129)**	(0.223)*	(0.514)	(0.176)**
Age at Repeal -10--15	-0.195	-0.285	-0.121	-0.302
	(0.150)	(0.226)	(0.466)	(0.186)
Age at Repeal -15--20	-0.329	-0.183	0.575	-0.280
	(0.134)**	(0.206)	(0.250)**	(0.194)
Age at Repeal <-20	-0.083	0.173	0.471	-0.020
	(0.135)	(0.224)	(0.321)	(0.188)
R^2	0.13	0.18	0.21	0.21
N	2,421,948	3,524,605	2,507,488	2,799,658

Notes: Data are from the 1940 full count Census and 1950 and 1960 Census samples. Sample is restricted to sample line native males aged 25 and older. Individual level controls include age, age squared, race indicators, birth place indicators, birth year indicators, father birth region indicators, and mother birth region indicators. Model 3 includes linear time trends in birth year for each mother and father birth region and linear time trends for each birthstate. All specifications include state of residence and census year fixed effects. IPUMS sample line weights are used in estimation. $p<0.1$; ** $p<0.05$; *** $p<0.01$

Appendix Table 1.4.2 – Intergenerational Effect on Ln(Income), sequentially dropping Regions (Model 3)

	Ln(Income) (no South)	Ln(Income) (no West)	Ln(Income) (no Midwest)	Ln(Income) (no Northeast)
Age at Repeal 25-30	0.169	0.083	-0.069	0.053
	(0.111)	(0.104)	(0.037)*	(0.101)
Age at Repeal 20-25	0.029	-0.024	-0.065	-0.044
	(0.042)	(0.035)	(0.041)	(0.038)
Age at Repeal 15-20	0.051	-0.024	-0.112	-0.040
	(0.034)	(0.044)	(0.093)	(0.047)
Age at Repeal 10-15	0.042	-0.013	-0.135	-0.045
	(0.043)	(0.048)	(0.024)***	(0.045)
Age at Repeal 5-10	0.003	-0.037	-0.053	-0.056
	(0.045)	(0.043)	(0.076)	(0.039)
Age at Repeal 0-5	-0.029	-0.089	-0.136	-0.110
	(0.039)	(0.045)*	(0.067)**	(0.041)**
Age at Repeal -5-0	-0.087	-0.120	-0.048	-0.130
	(0.049)*	(0.044)***	(0.047)	(0.040)***
Age at Repeal -5--10	-0.032	-0.073	-0.169	-0.109
	(0.027)	(0.033)**	(0.047)***	(0.030)***
Age at Repeal -10--15	-0.057	-0.098	-0.050	-0.114
	(0.037)	(0.039)**	(0.073)	(0.035)***
Age at Repeal -15--20	-0.058	-0.048	-0.009	-0.087
	(0.025)**	(0.033)	(0.053)	(0.029)***
Age at Repeal <-20	-0.013	0.001	0.038	-0.036
	(0.021)	(0.024)	(0.029)	(0.021)*
R^2	0.13	0.18	0.20	0.19
N	1,628,456	2,315,560	1,698,762	1,796,706

Notes: Data are from the 1940 full count Census and 1950 and 1960 IPUMS Census samples. Sample is restricted to sample line native males aged 25-64 (inclusive). Individual level controls include age, age squared, race indicators, birth place indicators, birth year indicators, father birth region indicators, and mother birth region indicators. Model 3 includes linear time trends in birth year for each mother and father birth region and linear time trends for each birth state. All specifications include state of residence and census year fixed effects. IPUMS sample line weights are used in estimation. $p<0.1$; ** $p<0.05$; *** $p<0.01$

Appendix Table 1.5 - Intergenerational Effect on Ln(Income), controlling for Educational Attainment (Model 3)

	Ln(Income)	Ln(Income)
Age at Repeal 25-30	0.084	0.052
	(0.103)	(0.078)
Age at Repeal 20-25	-0.018	-0.022
	(0.035)	(0.036)
Age at Repeal 15-20	-0.012	0.004
	(0.046)	(0.037)
Age at Repeal 10-15	-0.016	-0.004
	(0.047)	(0.034)
Age at Repeal 5-10	-0.024	0.001
	(0.042)	(0.034)
Age at Repeal 0-5	-0.078	-0.057
	(0.044)*	(0.032)*
Age at Repeal -5-0	-0.098	-0.072
	(0.041)**	(0.036)*
Age at Repeal -5--10	-0.075	-0.053
	(0.032)**	(0.024)**
Age at Repeal -10--15	-0.080	-0.068
	(0.037)**	(0.030)**
Age at Repeal -15--20	-0.051	-0.047
	(0.029)*	(0.025)*
Age at Repeal <-20	0.002	-0.005
	(0.020)	(0.016)
Education Control?	N	Y
R^2	0.18	0.23
N	2,479,828	2,479,828

Notes: Data are from the 1940 full count Census and 1950 and 1960 IPUMS Census samples. Sample is restricted to sample line native males aged 25-64 (inclusive). Individual level controls include age, age squared, race indicators, birth place indicators, birth year indicators, father birth region indicators, mother birth region indicators. Model 3 includes linear time trends in birth year for each mother and father birth region and linear time trends for each birth state. Column 1 repeats results for Model 3 from Appendix table 3B. Column 2 controls for completed education. All specifications include state of residence fixed effects. IPUMS sample line weights are used in estimation. $p<0.1$; ** $p<0.05$; *** $p<0.01$

Chapter 2

What Drives Expansions and Contractions of the Electorate? Evidence from Alien Voting in the United States

In a little-known episode of United States history, 24 states and territories expanded the franchise from 1848-1890 to include non-citizen immigrants. From 1864-1926, however, these same states and territories repealed this right to vote, creating a mass disenfranchisement that affected up to 20% of the voting population in certain states. This policy shift offers a valuable opportunity to understand the forces determining political inclusion and exclusion in a formative period of American democracy, and contributes to the broader literature on theories of democratization. I use qualitative evidence from the historical record to outline competing theories of both adoption and repeal of alien voting, and then rationalize these hypotheses within the context of a median voting model. Using a discrete time hazard specification, I find evidence consistent with the hypothesis that states used alien voting as a locational amenity, with the objective of inducing immigrant in-migration in order to foster agricultural development. The results indicate that the timing of repeal was driven by social costs, rather than economic or political factors, although there is evidence for heterogeneity in correlates of support for repeal across states. Finally, the costs of constitutional change were salient for both adoption and repeal: states for which it was less costly to re-write or amend the constitution were more likely both to adopt and repeal alien voting.

2.1 Introduction

In a little-known chapter in American history, twenty-four states and territories extended the right to vote to non-citizen immigrants from 1848-1890. This constituted a significant departure from existing voting policy, which formerly had restricted the franchise only to *citizens* of the United States. From 1864-1926, however, these same states and territories repealed those provisions, returning to the standard of citizen voting. This latter shift affected up to 20% of the voting population, and had adverse consequences for the political engagement and economic wellbeing of immigrants and their offspring (Henderson 2017).

Theoretically, an expansion of the franchise to include individuals lower on the income distribution results in increased demands for redistribution via public expenditure (Meltzer and

Richard 1981). This is supported by a considerable – although not universal – body of empirical evidence, from a variety of settings, finding that expanding the right to vote results in increased and more widely distributed public spending (Miller 2008; Lott and Kenny 1999; Aidt and Dallal 2008; Cascio and Washington 2014). Given that increasing the electorate has predictable effects upon redistribution, a deeper economic question is: why would a group *voluntarily* dilute its own stake in the determination of resource allocation, thereby possibly sacrificing a portion of those resources? And, once included in the electorate, why would the included group allow itself to be voted out?

In this paper, I explore the determinants of expansion, and then contraction, in state-level immigrant voting rights from the mid-19th to early-20th centuries. Scholars acknowledge that relatively little work exists on alien voting in the United States, but have advanced several hypotheses for the causes of adoption and repeal.[1] Using qualitative and narrative evidence, I first present the competing hypotheses – economic, political, and social – and then rationalize these in the context of a median voting model, in which the pivotal voter decides whether or not to expand (or contract) the electorate based on the costs and benefits of doing so.

To test these hypotheses, I pursue two complementary empirical strategies: first, I assemble a state-decade dataset from 1850 – 1930 containing information on the timing of adoption and repeal of alien voting, economic development, political control, demographics, and the prevailing methods of constitutional amendment. I use a discrete-time proportional hazard specification in order to quantify the factors that affected the hazard of adoption or repeal of alien voting. I find that labor scarcity was a predictor of adoption: a one standard deviation

[1] Raskin (1993) notes that "legal observers who have ventured into this field have correctly noted the dramatic absence of professional historical accounts of alien suffrage." (Page 1393)

increase in the fraction of the population that was adult men was associated with a 76.2% reduction in the hazard of adoption. Moreover, adoption was less likely to occur in more economically developed states: one standard deviation increases in the fraction of acreage in the state that was agriculturally improved and the real value of manufacturing output per capita were associated with, respectively, reductions in the hazard of adoption of 80%, and 49%. This is consistent with qualitative evidence suggesting that alien voting was used as a locational amenity, intended to induce in-migration into labor-scarce states in order to accelerate economic development. I find no statistical evidence of political or social forces affecting the hazard of adoption of alien voting.

There is also no evidence for systematic economic or political determinants of *repeal* of alien voting; instead, the results suggest that repeal was driven by the social costs of immigration during this era. I find a positive relationship between the proportion of the population that is foreign born and the hazard of repeal, and this strengthens in both magnitude and statistical significance with the exclusion of outliers. Finally, I document that the costs of constitutional change were relevant for both adoption and repeal: states with lower costs of constitutional change were over twenty times more likely to adopt alien voting, and eight times more likely to repeal alien voting, in a given decade. The salience of the costs of constitutional change has been overlooked in the alien voting literature, and, to the extent that amending procedures do not vary systematically across states, suggests that the timing of repeal of alien voting may be (partially) exogenous to underlying attitudes toward immigrants.

Hazard models cannot account for fixed factors within states. Therefore, my second empirical strategy uses newly assembled county-level election data from thirteen states in order to explore the determinants of *within-state* support for disenfranchisement.[2] I further leverage the fact that Missouri voted twice for alien voting – first, unsuccessfully, in 1912, and then successfully in 1924 – to control for unobserved, time-invariant, county-level factors. I find that, consistent with high social costs of immigration in this period, support for repeal was generally greater in counties with a higher fraction of the population that was foreign born, although this is not statistically significant and there is evidence of heterogeneity among states.

This paper contributes to four literatures. First, it adds to the large existing body of work on the causes of democratization, ranging from the enfranchisement of working class men in Europe (Acemoglu and Robinson 2000 and 2001; Conley and Temimi 2001; Lizzeri and Persico 2004) to men in colonial British America (Nikolova 2015), to women in Europe (Bertocchi 2011) to women in the United States (Braun and Kvasnicka 2013). This paper not only studies a original setting and population – immigrants in late 19th and early 20th century America – but also is among the first, to my knowledge, to systematically address causes of franchise contraction in *any* setting.

Second, this paper complements the growing literature on immigration in the late 19th and early 20th centuries in the United States., and, particularly, the nascent literature on endogenous policy responses to immigration. Scholars have leveraged the widespread digitization of historical data to understand more about immigrant socio-economic assimilation in this period, ranging from labor market assimilation, to social assimilation, to return migration behavior (Abramitzky et al 2012; Abramitzky et al 2014; Abramitkzy et al 2016; Bandiera et al 2013; Ward and Greenwood 2016). There has been less work on the interaction of immigration and

[2] Amending a constitution requires that the proposed amending pass a popular vote for final ratification. I collected these county-level vote totals for 13 of the 14 alien voting states that repealed via constitutional amendment. I was unable to locate county-level returns for Oregon.

political outcomes, however, and scholars are only just beginning to address the impact of immigration on *policies* of the era (Bandiera et al 2015; Abramitzky and Boustan 2016).

Third, this paper relates more broadly to institutional change in U.S. economic history. Broadly construed, institutions shape economic interaction, and scholars have long recognized their importance for economic development. Recent evidence highlights the responsiveness of institutions to economic and demographic forces in the 19th and early 20th centuries (Rosebloom and Sundstrom 2009; Wallis 2005; Benmelech and Moskowitz 2005), and the role of institutions shaping public finance and infrastructure investment (Lamoreaux and Wallis 2016). By focusing on the expansion and contraction of alien voting rights, I am able to analyze determinants for institutional change within a given set of institutions *over time*.

Finally, this contributes to the growing literature on the effects of constitutions. This has, until recently, been relatively unexamined in economics (Persson and Tabellini 2003), although it has long been the subject of political science research (Sartori 1994, Powell 1982, Lijphart 1994). Results from this paper point to the importance of constitutional amending procedures in determining the pace of constitutional change, which, in turn, suggests that research using identification based on constitutional change should take these procedures into account.

In section 2, I discuss the context for the expansion and contraction of alien voting, and in section 3, present narrative evidence on causes of alien voting adoption and repeal. Section 4 rationalizes the competing hypotheses in a median voter model, and Section 5 presents empirical evidence for the state-level determinants of adoption and repeal. Section 6 explores support for repeal within-state, and Section 7 concludes.

2.2 Historical Context

2.2.1. Adoption

Starting with Wisconsin in 1848, twenty-three other states and territories from 1848-1890 adopted provisions in their state or territorial constitutions that allowed non-citizen immigrants the right to vote in local, state, and federal elections (Hayduk 2006, Keyssar 2000). These provisions enfranchised those immigrants that had completed the first of a two-step naturalization procedure by declaring their intent to become citizens (also known as "filing first papers").[3] There was substantial variation in terms of the timing and geography of this franchise expansion: Midwestern states and territories enacted these provisions in the late 1840s and 1850s; Western territories during the Civil War, and Southern states following the Civil War. No states in the Northeast adopted alien voting. See figure 2.1.1 for a map of the adoption of alien voting.

Of the twenty-four alien voting states and territories, ten *existing* states adopted alien voting as part of a new state constitution; six new states continued the existing policy of alien voting enacted when the state was a territory; two states that, as territories, did not allow alien voting, adopted it upon admission to the Union; and six territories allowed declarant aliens to vote, but ended the practice upon their admission to the union. This is summarized in appendix table 2.1.1, along with detailed sources for establishing the date of adoption.[4] Two features are notable: first, of the ten existing states that adopted alien voting, seven were Southern states that

[3] As opposed to the wholesale enfranchisement of all immigrants, this "intermediate" enfranchisement of immigrants was "framed in the spirit of compromise" (Page 230, *The Convention of 1846*). Given the low costs of declaring an intent for citizenship, however – this could happen at any point after arrival, at any one of 5,000 "competent courts" around the county – this compromise may have been broader than was initially intended.

[4] In the territories, voting qualifications were initially specified by a Congressional Organic acts, with subsequent policy left to the jurisdiction of the territorial legislatures (Porter 1918).

did so in the immediate aftermath of the Civil War, when all former confederate states were required to draft new constitutions as a condition of re-joining the Union (Hill and Hill, 2008). Second, none of the existing states adopted alien voting via constitutional amendment. I discuss both points in greater detail below.

2.2.2 Repeal

All twenty-four alien voting states and territories eventually disenfranchised declarant aliens, returning to the standard of citizen-voting. The timing does not mirror that of alien enfranchisement: the correlation coefficient between the year of adoption and year of repeal is -.19, implying that later-adopting states tended to repeal earlier, and there is considerably more variance in the timing of repeal (standard deviation of 17.9) than the timing of adoption (standard deviation 10.5). See figure 2.1.2 for a map of the year of repeal, by state. The modes of repeal also differed from the modes of adoption (see appendix table 2.1.2 for details). Of the twenty-four states and territories that contracted their electorates, fourteen did so via constitutional amendment – that is, replacing only a portion of the relevant constitutional article, rather than creating an entirely new constitution - and four as part of an entirely new constitution. Four territories disenfranchised non-citizen aliens upon entry to the Union, and two did so via territorial legislation while still a territory.

2.3 Narrative Evidence for Determinants of Adoption and Repeal

2.3.1 Adoption

Legal scholars and historians agree that one of the primary reasons for alien enfranchisement was economic self-interest: immigrants were expected to respond to the amenity of voting by migrating in large numbers, thus facilitating economic and agricultural development (Raskin 1993, Rosberg 1977).[5] The contemporary evidence for this view abounds. For example, in the Michigan constitutional convention of 1850, Delegate W.V. Morrison of Calhoun County contended:

"I consider that by extending the right of suffrage to foreigners, we are advancing the best interests of the State. We have expended thousands to induce emigrants to reside with us, and what have we effected? Wisconsin [which enacted non-citizen voting in 1848] has opened her doors - she has extended to them the right of suffrage, and thousands have poured in, developing the resources and adding to the riches of the State. Thousands have gone 'round by the Lakes; and if you question the travelers upon the Central Railroad, they will tell you they are going to Wisconsin. If any of them remain, their friends in Wisconsin tell them that rights are given there that are withheld in Michigan; and at the first opportunity they leave us and take up their abode there. And we may expend money by our emigration agents, but if we persist in our narrow policy, it will have but little effect."

The enthusiasm with which immigrants were embraced as engines of economic development is mirrored in the convention support for alien voting adoption: it passed unanimously in Wisconsin in 1848 and Minnesota in 1857, 89-10 in Indiana in 1850, and lost by only nine votes in Illinois (Porter 1918). Similar sentiment can be found in the proceedings of other constitutional conventions - for example, Montana in 1889[6] – as well as academic work and judicial decisions from the era (McCulloch 1929; Porter 1918; Harper-Ho 2000).[7] This

---

[5] It is worth noting that scholars in the late 19th and early 20th centuries noted that this expected surge of migration due to voting did not materialize: Chaney (1894) writes that "immigrants flocked into Illinois, which required citizenship, and in doing so passed through Indiana which was satisfied with a declaration of intention." This is echoed in Williams (1912): "Alien suffrage was a drawing card used by western states to attract foreign immigration… the failure of alien suffrage to accomplish its original purpose appears from the fact that foreigners today do not hesitate to leave Nebraska [in which alien suffrage still held] for Colorado, Montana, Wyoming and other states in which they cannot vote, provided they can better their economic condition."

[6] Mr Middleton, of Custer: "It does not seem to me that that amendment offered by the gentleman from Silver Bow should be made. The qualification of requiring a man to be a full citizen of the United States [in order to vote]… does not seem to me to be conducive to that kind of immigration that we want to have. … I do not believe that it is desirable in the interests of a new state as sparsely settled as this is at the time, and of such vast territory, to make that sort of qualification." Page 352, *Proceeding and Debates of the Constitutional Convention Held in the City of Helena, Montana, July 4th 1889, August 17th 1889*. McCulloch (1929) writes: "when… much-needed settlers were pouring into the central part of the United Sates, states in that section began to offer extra inducements by allowing aliens to vote after taking out first papers." Porter (1918) concurs: "to offer [immigrants] elective franchise very soon after their arrival seemed to be an effective way of attracting them."

[7]In Spragins v Houghton (1840), the Illinois Supreme court noted that the intent of granting the right to vote to immigrants was to "induce a flood of emigration to the state, and cause its early and compact settlement."

motivation was not limited to Midwestern states: scholars have hypothesized that the proliferation of alien voting in the postbellum South was due to the need to attract cheap immigrant labor following the emancipation of slaves (Hayduk 2006).

The prospect of a flood of in-migration would have conferred benefits in addition to that driven by the marginal product of the additional labor force. Congressional apportionment in this era was based on the "number of [free] persons", suggesting that increased Congressional representation may have been another benefit of in-migration (Porter 1918). Furthermore, as indicated in the above-quoted passage, jurisdictional competition in this era may have heightened the perceived need for, and benefit from, immigration.

It is possible that the economic case for alien voting was bolstered by the state budgetary crises of the early 1840s. Eight states and the territory of Florida defaulted on their debt, and three nearly did; in the wake of this crisis, states amended their constitutions to restrict modes of financing "internal improvements" (Wallis 2005). States, therefore, may have viewed the possibility of migration induced by alien voting as a relatively riskless means of development. This was implied during Illinois' constitutional convention of 1847, where one delegate explicitly linked the recent budgetary crisis with the prospect of a large-scale migration response to enfranchisement: "Is it our policy, as a state burdened with debt and sparsely settled, to restrict the right of suffrage, and thus prevent immigration to our soil?" (Keyssar 2000, pg 38).

Finally, states may have expanded their franchise – or not – for political reasons. Keyssar (2000, Pg 40) writes that "the Democratic embrace of alien suffrage… was unmistakably motivated in part by the party's desire to enroll and win the support of immigrant voters." Neuman (1993) echoes this, citing the dynamics of alien suffrage in Wisconsin: "in the Wisconsin Territory of the 1840s, the Democratic Party sought the support of the large immigrant population." Raskin (1993) arrives at a similar conclusion regarding Michigan's expansion of alien suffrage in 1850: delegates "linked alien suffrage to both democratic and Democratic politics."[8]

2.3.2 Repeal

There is little narrative evidence to suggest that political motivations drove the repeal of alien voting. A sole opinion piece - in the *Detroit Free Press* in 1894 - suggests a role for politics in the repeal of alien voting, arguing that the Michigan amendment is "Republican in origin", and stems from the fact that "it has been unable to secure [the foreign voting bloc's] support except to a very limited extent." An exhaustive search of *America's Historical Newspapers* for articles relating to alien voting, however, reveals no other evidence that commenters imputed political motivations to alien disenfranchisement.[9] I was also unable to find any suggestion – scholarly or contemporary - that the timing of repeal was related to economic forces. However, this is not evidence of absence: the broad time pattern of repeal suggests that the economic crisis of the 1890s may have been a motivating factor – four states (FL, MI, MN, LA) repealed from 1894-1898 - and Goldin (1994) documents that support for immigration restriction was higher in areas with a greater negative impact of immigration on native wages.

[8] While the relatively immigrant-heavy Northeastern states did not adopt alien voting, there is little to support the argument of Braun and Kvasnicka (2013), who argue that women's suffrage originated in the American West because a "scarcity" of women in the region implied that allowing women to vote imposed little cost upon male voters by potentially changing the balance of power. Indeed, several early adopting Midwestern alien voting states had among the highest proportion of adult male immigrants in the country in 1850: Wisconsin was 52% immigrant, Minnesota 38%, Michigan 19% (Ruggles et al, 2015).

[9] Given that the majority of repeals were effected via constitutional amendment, rather than as part of new constitutions – with the attendant written proceedings of each constitutional convention - there is far less record of public debate concerning the repeal of alien voting.

Modern scholars generally concur that the repeal of declarant alien voting was due to the rising anti-immigrant sentiment of the late 19th and early 20th centuries, although they differ in their opinions as to the sources of this changing sentiment. Factors that have been mentioned in the literature range from the assassination of President McKinley by a second-generation immigrant in 1901 (Rosberg 1977), to the shifting ethnic origins of immigrants from Northwest Europe to Southeast Europe in the late 19th century (Neuman 1993), to "the hysteria attending World War One" (Raskin 1993).

In general, contemporary observers opposed declarant alien voting on the grounds that allowing non-citizen immigrants to vote presented a significant social cost to existing voters. Public opposition to alien voting tended to focus on a perceived lack of immigrant assimilation. As early as 1888, an article in the *Topeka Weekly Capital* suggests a modification of existing immigration law in order to foster immigrant assimilation, and exhorts lawmakers to "exclude every such alien from exercising the right to vote at any election until the full period of probation has expired and he has become a citizen." A piece in the *Morning Oregonian* in 1893 serves as part warning, part public service announcement, apprising the reader that a number of "Western states… have extended the franchise to foreigners in advance of their acquisition of citizenship" and warns that the "ignorant Italian or Pole" has undue political influence because of these alien voting provisions, concluding by reassuring the reader that repealing alien voting would mean that "the evils of unrestricted immigration would be greatly diminished."[10] The tone of surprised indignation is echoed in the Washington Post in 1902, which, again, informs readers that "there are many States where an immigrant is not obliged to become a full-fledged citizen in order to be

[10] Kansas would go on to repeal alien voting in 1918, Oregon in 1914.

the peer of the native on election day" and suggests the introduction of a language qualification for voting to ensure that voting immigrants are committed to American ideals.[11]

2.3.3 *Constitutional Change*

Scholars of alien suffrage have universally neglected to consider the role of the processes guarding constitutional change when attempting to explain the timing of adoption and repeal.[12] This is puzzling for two reasons: first, the legal basis for alien suffrage was *constitutional*, not statutory, and thus requiring constitutional overhaul or amendment to both enact and repeal. Second, states differed significantly in the manner by which they can change their constitutions, with the costs of constitutional change varying accordingly. In this sub-section, I briefly review the procedures by which constitutions can be altered, and discuss the potential relevance to the timing of alien voting adoption and repeal.

In this era, states changed their constitutions in one of two ways: amending a part of an existing constitution, or crafting an entirely new constitution in a constitutional convention.[13] The constitutions themselves contain provisions on how to amend or how to call a convention, and the structure of the amending procedures tends to be stable across states: a proposed amendment must have originated in the state legislature and have achieved a certain fraction of legislative support (1/2, 3/5, or 2/3), *and* have passed with that level of support either once or

[11] This is echoed in academic work of the era. Chaney (1894) writes that Michigan's expansion of the electorate in 1850 to include declarant aliens "now looks so queer… and makes people wonder why a mere declaration of intention should ever have been thought enough of a preliminary for the most important right that can be exercised of any qualified citizen."

[12] In their study of the timing of adoption of state-level women's suffrage, Braun and Kvasnicka (2013) acknowledge that this may have been a relevant factor.

[13] Several states also have initiative procedures in which a proposed amendment may start as a petition and, given enough popular support, may be subject to a popular vote for ratification; however, these are rare even today, and virtually nonexistent in the early 20th century (Friedman 1988)

twice. If the amendment so passed, it was then referred to the public for a popular vote, in either a general or special election, and was required to achieve a majority of public support in order to become enshrined in the constitution. Crucially, states differed in the effective fraction of support in the popular ratification required for passage in the next general election.[14] Specifically, certain states require that the fraction of the popular vote needed for ratification constitutes a majority of all votes at that election, and other states require only a majority of votes from those voting *on that particular amendment.* Given the practice of "roll-off voting", in which individuals are less likely to vote for ballot questions of lower prominence on the physical ballot, or about which they are less informed, a larger number of votes will be cast for President, for example, than a proposed constitutional amendment (Bullock and Dunn 1996, Augenblick and Nicholson 2016). The stronger is this phenomenon, the higher the effective threshold required for popular passage in states that require a majority of all voters for ratification.[15]

The procedures for adopting an entirely new constitution are relatively more heterogeneous. Certain states, in early versions of their constitutions, allowed a convention to be called by the state legislature alone;[16] most – but not all – states required a certain standard of

[14] Certain states also placed a cap on the number of amendments that may be voted on in any one election, although this is relatively rare.

[15] Certain states implicitly acknowledge the importance of how a popular majority is calculated. Florida, in article 17 of its 1868 constitution, requires a "majority of the voters voting at such election" in order to call a constitutional convention, and mandates that "in determining what is a majority of the electors voting at such election, reference shall be had to the highest number of votes cast at such election for the candidates for any office or on any question." Delaware, in Article 9 of its 1831 constitution, requires that a constitutional convention may be formed only after achieving a "majority of all the Citizens in the State, having the right to vote for representatives", and then clarifies that "the majority of all citizens in the State having the right to vote for representatives shall be ascertained by reference to the highest number of votes cast in the State at any one of the three general elections next preceding the day of voting for a convention, except when they may be less than the whole number of voted voted both for and against a convention, in which case the said majority shall be ascertained by reference to the number of votes given on the day of voting for or against a convention…"

[16] Alabama, 1861; Florida, 1838, 1861, 1865; Georgia 1864 and 1877; North Carolina, 1776 and 1868; South Carolina, 1790, 1861, 1865; Teas, 1861 and 1866

legislative support for a proposal to call a convention, either to be passed once or twice in successive legislative sessions, and then required that the proposal be submitted to votes for approval or rejection. If approved, the constitutional convention convenes in order to draft an entirely new constitution, and the resulting document is either approved or rejected by a popular vote. Nine states did not specify a process for constitutional convention in their constitutions from 1830-1930, and at least six states during this era had a further stipulation that required popular vote for a proposed constitutional convention at fixed intervals.[17]

Despite the heterogeneity, the costs of the modal procedure for constitutional replacement exceed those for constitutional amendment. Procedurally, voters were required to agree to an "extra step": the act of calling a convention. Moreover, this was no guarantee that the new constitution – the product of the convention – would be approved, and, in fact, conventions often failed: from 1790-1992, states held constitutional conventions 233 times, but only adopted 145 new constitutions (Dinan 2010, Berkowitz and Clay 2004). Constitutional amendments, however, are a much more frequent phenomenon: as of 2000, state constitutions had been amended almost 12,000 times (Wallis 2000). States did not, generally, amend their constitutions prior to 1860: Michigan, for example, attempted two amendments in the 1850s, twelve in the 1860s, thirteen in the 1870s, eighteen in the 1880s, and thirteen from 1890-1897 (*Michigan Legislative Manual* 1897).

Research indicates that the details of the amendment process matter for the frequency of amendments: states that require a higher fraction of legislative support, or multiple successive legislative approvals (or both) have lower rates of constitutional amendment than states that

[17] Virginia, every 20 years; Oklahoma, every 20 years; New York, every 20 years; Michigan, every 16 years, Maryland, every 20 years; Iowa, every 10 years.

require only a single majority vote (Lutz 1994). There is also narrative evidence to suggest that the implied standard for majority popular vote was salient for the repeal of alien voting: Arkansas tried to abolish alien suffrage in a proposed constitution in 1918, which subsequently failed; then, a constitutional amendment proposed in 1920 passed (87,237-49,751), but not with the support needed to constitute a majority of all votes cast in that election (190,113). The issue went to the Supreme Court of Arkansas, which, ultimately, passed the amendment, thus disenfranchising aliens, in 1926 (*Biennial Report of the Secretary of State of Arkansas*, 1919-1920, 1925-1926; Aylsworth 1931). Similarly, Nebraska in 1910 attempted to disenfranchise non-citizen immigrants and the referendum passed, 100,450-74,878, but failed to achieve a majority of the 243,390 electors voting in the election (Nebraska Blue Book and Historical Register, 1918). Therefore, all else equal, states in which the costs of constitutional change are lower – perhaps because of requiring only a majority of voters voting on the particular proposed amendment, rather than a majority of all voters at a given election - should be more likely to either adopt or repeal alien voting.

2.4 What Drives Changes in the Franchise?

In this section, I sketch a simple model based on the insight that, in this setting, expansion and contraction of the electorate is a *voluntary* act undertaken – or, at least, approved - by popular vote.[18] Therefore, the decision can be modeled through the costs and benefits to a median existing voter. This occurs within a given jurisdiction, with a given method of constitutional amendment. Consider an economy with a population of P individuals, C of which

[18] This was inspired by Bertocchi and Strozzi (2010) who use a version of the median voter model to model the decision of countries to offer either jus solis or jus sanguinis citizenship.

are existing voters (either because they are native-born, or naturalized immigrants) and N of which are not permitted to vote: P = C + N, and C>N.

Both voters and non-voters obtain utility from consumption of a private good, c^i, and bear the social costs of either exclusion or inclusion into the franchise in the following manner:

$$u^i = c^i - k * 1(Exclude) - j * 1(Include),$$

Society bears costs of voter exclusion, k, and inclusion, j. Let k be the social cost of excluding non-voters (that is, non-citizens) from the franchise. This cost increases with the share of non-citizens in the population, reflecting the possibility that a large pool of non-eligible voters may lead to social unrest or reduced immigrant assimilation.[19] I define the exclusion cost, k, as

$$k = h * \frac{N}{P}$$

where h > 0. The social cost of *including* new voters in the electorate, j, depends on a function of the socio-cultural distance between existing voters, C, and non-voters N. This distance is positive and increasing in "dissimilarity" of the prospective voters and existing voters, along, for example, ethnic or religious lines. Additionally, the cost of inclusion also depends positively on the fraction of non-citizens in the population since the larger the population of new voters, the greater the threat to the existing power structure if those new voters vote as a bloc:

$$j = d(N, C) + f * \frac{N}{P}$$

where f > 0. Individual *i* faces a budget constraint of:

[19] In the mid-19th century United States, the possibility of social unrest of the disenfranchised was not purely academic: in Rhode Island in 1841, opponents of the property qualification for suffrage elected their own governor, and attempted to seize control of the state arsenal (Keyssar 2000, pages 71-74).

$$c^i \leq y^i.$$

Prior to expanding the franchise, the expression for the indirect utility of the median citizen voter, with income y^{c_*}, is given by:

$$v^{c_*} = y^{c_*} - k.$$

Following expansion, the indirect utility for that *same* voter – who, depending upon the dimension under consideration, may no longer be the median voter – with income y^* is given by:

$$v^* = y^* - j.$$

The median citizen voter, therefore, faces a simple set of costs and benefits when considering the decision to expand voting rights to non-citizens. If non-citizens cannot vote, the citizen voter incurs the social exclusion cost *k*, and earns income y^{c*}. If non-citizens vote, then the median citizen voter avoid the social cost of exclusion but incurs the social cost of *inclusion*, and earns income y^*. Thus, the median citizen voter will decide to grant voting rights to noncitizens if and only if:

$$y^{c*} - k \leq y^* - j \tag{1}$$

where the left-hand side of the inequality is the median voter's indirect utility when non-citizens cannot vote, and the right-hand side is his utility when they can vote. Conversely, the median voter will decide to *contract* the electorate is the opposite of equation (1) holds.[20]

The basic model can be extended to consider potentially relevant economic and political factors. An economic motivation for franchise expansion can be rationalized if it is *expected* that

[20] Given that C > N, disenfranchisement could occur even if no noncitizen voted for their own disenfranchisement: the median voter would not be disenfranchised.

allowing non-citizens to vote would result in mass in-migration. If labor is scarce, and so the marginal product of additional labor is high, then this new labor pool may serve to increase income for existing voters (that is, y^* rises). Similarly, a political motivation for franchise expansion can be modeled as a reduction in j: if the median voter believes that the additional voters added to the electorate would vote in a similar fashion to himself – by, for example, joining the same political party – then the costs of including new voters falls, and franchise expansion is more likely. Conversely, if this distance is large – manifested by the presence of anti-immigrant or nativist sentiment – then franchise expansion is less likely.

The median voter is more likely to vote for franchise contraction – that is, excluding non-citizens from the electorate – when exogenous factors increase the benefits (or decrease the costs) of exclusion. Shifts in the ethnic origins of new immigrants may increase the socio-cultural distance between non-citizen and citizen voters, thereby increasing the costs of inclusion; a higher fraction of the population that is non-citizen would both increase the cost of exclusion and may increase the costs of inclusion. Finally, if the non-citizen labor ceases to complement citizen labor, and instead competes with it, then the median voter's income would be lower under non-citizen voting, and the franchise is more likely to be contracted.

2.5 Empirical Evidence on Determinants of Adoption and Repeal of Alien Voting

2.5.1 Data

I test these hypotheses using a variety of data on economic, demographic, and political characteristics of the era. I assembled dates of adoption and repeal from a variety of sources, including (Keyssar 2000), Hayduk (2006), state constitutions, blues books, legislative manuals, territorial Organic acts, and the territorial legislative record. These dates, with primary sources, are presented in appendix tables 2.1.1 and 2.1.2 Data on gubernatorial elections are from Burham et al (1991); data on the partisan composition of state legislatures are from Burnham (1989); data on farm values are from Haines et al (2016); data on improved acreage and manufacturing output are from Haines (2010); demographic data are generated from aggregated Census microdata (Ruggles et al, 2015). I create an original dataset of amending procedures of all state constitutions based on from the original text of all state constitutions that apply at each decade from 1830-1930. See the data appendix 2.1 for more detail on data construction.

2.5.2. Method

In general, data availability poses a challenge to analysis, since many state-level covariates are typically not available until the state is admitted to the union. Even for established states, certain covariates are unavailable for earlier years: for example, the population of foreign born males is not available until 1850 (whereas, Wisconsin adopted alien voting in 1848). Given that the fourteen of states and territories adopted alien voting as either territories (ID, KS, MN, MT, NE, NV, ND, OR, OK, SD, WA, WY) or upon admission to the union (CO, WI), pre-adoption data is not available for the majority of states or territories that ever expanded the franchise to include declarant aliens.

I account for this by estimating the hazard of alien voting adoption *as of* a particular decade, rather than estimating the hazard of adoption in the *next* decade (for example, Wisconsin, which enacted alien voting in 1848, is coded as having adopted alien voting by 1850). This allows for the inclusion into the sample of the fourteen states that adopted alien voting prior to, or upon, admission to the Union, but may, to some degree, confound fast-moving responses to adoption with determinants of adoption. I address this concern in greater detail below. Territories that had enacted alien voting prior to admission to the union, and continued that upon admission, are coded as having adopted alien voting upon admission.[21]

I create a pooled cross-section of states from 1850-1890 and link on the economic, demographic, and political data mentioned above. In order to estimate determinants of adoption of alien voting, I estimate a hazard model: that is, a model of the probability of a state adopting alien voting laws, given that it has not adopted alien voting laws up until that point. I use a proportional hazard model, as it has the advantage of remaining agnostic as to the actual shape of the underlying baseline hazard of adoption (Fishback and Kantor 1998; Braun and Kvasnicka 2013). I implement this by estimating a complementary log-logistic specification using maximum likelihood, as this is the discrete time analog of the underlying continuous proportional hazards model (Prentice and Gloeckler 1978). Specifically, I estimate:

$$h_{it} = 1 - \exp\big(-exp(a_t + \beta X_{it})\big) \qquad (2)$$

I do not impose a functional form on the baseline hazard, and instead allow it to vary non-parametrically by including decade indicators (a_t). Given that a state is no longer at risk of *adopting* alien voting following adoption, states drop out of the sample following adoption. I include a variety of covariates (explained more below), and β is the vector of parameters of interest. I standardize continuous covariates (calculated from pre-adoption census year), so estimates should be interpreted as the effect of a one standard deviation increase of a given covariate. Estimates are scaling factors of an underlying baseline hazard of adoption: coefficients greater than one imply that, all else equal, a one standard deviation leads to a greater (or earlier) risk of adoption, and coefficients smaller than one imply a reduced (or later) hazard of adoption.

[21] Given that admission to the union required territories create a state constitution (to be approved by Congress), if a state includes existing alien voting in the constitution, it implicitly re-approved the policy.

Hypothesis tests, therefore, use a null hypothesis that a particular covariate is equal to one. Standard errors are clustered at the state level.[22]

2.5.3 Results on Determinants of Adoption

Results are presented in table 2.1. First, I test for the impact of various demographic characteristics on the hazard of adoption. If labor scarcity was a motivating factor in the adoption of alien voting, then a lower population, a lower percentage of the population that are adult males, and slower population growth should all be associated with an increased hazard of adoption. A larger foreign-born population may either increase or decrease the hazard of adoption, depending on whether it is perceived that the disenfranchised immigrants may agitate for inclusion, or, once included, may threaten the existing power structure. Finally, a larger percentage of immigrants from "new" sending countries should serve to increase the socio-cultural distance of these news voters if included in the electorate, and should, therefore, reduce the hazard of adoption.

These coefficients are reported in Column (1). Consistent with an economic motivation for adoption, there is evidence that states with a lower proportion of the population that are adult males were more likely to adopt alien voting: a one standard deviation increase in the percentage of adult males in the population reduces the hazard of adoption by 76.2%.[23] A larger population was also associated with a lower probability of adoption, but this effect is not statistically significant. Interestingly, states with faster population growth were significantly *more* likely to

---

[22] If jurisdictional competition drove the timing the adoption, then the assumption of independence of adoption between state clusters may be violated. However, clustering standard errors at the region level does not meaningfully change the statistical significance of any estimates hazard for both adoption and repeal.

[23] This is based on the coefficient .238. Relative to a null hypothesis of 1, this is a reduction in the hazard of (1-.238)/1, which equals 76.2%.

adopt alien voting: a one standard deviation increase in population growth is associated with a 215% increase in the hazard of adoption. This suggests that states may have used alien voting to accelerate already existing growth.

There is no evidence that the fraction of the population that was foreign-born affected the timing of adoption, but, contrary to the model, a greater fraction of “new” immigrants is *positively* associated with the hazard of adoption. This appears to be due to the influence of South Dakota which, in 1890, had the highest fraction of its immigrant population from new sending countries (15.31%). Excluding this observation reduces the magnitude of the estimate so that it is statistically indistinguishable from zero (estimate = .305, p value = .357).

Column (2) further suggests that there was an economic motivation for franchise expansion: states that had more agriculturally improved land area were significantly *less* likely to adopt alien voting. Specifically, an additional standard deviation of improved acreage is associated with a roughly 80% reduction in the hazard of adoption. Moreover, states with a *less* developed manufacturing sector were also more likely to adopt alien voting: a one standard deviation increase in the natural log of the real value of per capita manufacturing output is associated with a 49% reduction in the hazard of adoption. Together with the results from column (1), this is consistent with the hypothesis that alien voting was used as an economic development strategy meant to accelerate economic growth.

Contrary to the qualitative evidence, though, there is no evidence that alien voting was politically motivated. Column (3) presents coefficients from various measures of the political strength of the Democratic Party. I use state-level measures of political presence – statehouse composition and gubernatorial election results – rather presidential vote share in order to capture state-level political dynamics. The Democratic presence in the upper and lower chambers of the state legislature is intended to estimate the effect of *elected* Democrats on the hazard of adoption (and, later, repeal), while the percentage of the gubernatorial vote going to the Democratic candidate is intended to reflect the general support for the Democratic Party among voters. I find no evidence that either measure of strength of the Democratic Party affected the hazard of adoption of alien voting.

Finally, I test for the role of procedural costs in altering the constitution. I test for the effects of three covariates which measure different aspects of the amending process: the fraction of legislative support needed to introduce a constitutional amendment, the number of successive approvals required by the legislature for amendment, and an indicator for whether popular ratification requires only a simple majority of voters voting on that particular ballot question (coded as a 1), or a majority of all voters voting at the election (coded as a 0). Finally, I include an indicator variable for whether the state re-wrote its constitution for exogenous reasons – either to re-enter the Union following the Civil War, or in order to transition from territorial status to statehood. To the extent that states did not strategically enter the Union in order to adopt alien voting, this captures an exogenous reduction in the fixed costs of constitutional change: if the costs of altering the constitution deterred adoption of alien voting, then this should have a positive effect on the hazard of adoption.

Results are presented in column (4). I find that the details of the constitutional amending process had no impact on the hazard of adoption of alien voting. This is not surprising, given that states did not tend to alter their constitutions via amendment until the late 19th century. There is strong evidence that the fixed costs of constitutional change deterred adoption of alien voting: states that had were either writing constitutions for the first time, or in order to re-enter the Union following the Civil War, were over twenty times more likely to adopt alien voting than established states that did not re-write their constitutions.

Using this estimation strategy, it is possible that fast-moving consequences of alien voting may be conflated with causes of alien voting. For example, Indiana, which adopted alien voting in 1851, is coded as having adopted *as of* 1860; if Indiana adopted alien voting in order to attract settlers because of a dearth of adult males, and if, upon adoption of alien voting in 1851, adult males flooded into the state, then by 1860, it would appear as if the fraction of the population that is adult male *positively* predicts adoption. The empirical distribution of dates of adoption, however, suggests that this concern may be somewhat mitigated: only three out of twenty-two states in my estimation sample adopted alien voting in the first half the decade, and the median base 10 modulus for year of adoption in the estimation sample is eight (e.g., 1868, 1878, and so on). As a further check for this, I re-estimate specification (1) *without* the alien voting states that adopted in the first half of a given decade, to minimize the scope for immediate response affecting the estimated results. Results are presented in appendix table 2.2, and do not differ substantively from those in table 2.1.

2.5.4 Results on Determinants of Repeal

I re-estimate specification (2) on a cross-section of states from 1890-1920 in order to test for the role of economic, demographic, and constitutional factors on the hazard of repeal. Given that repeal occurred, on average, 38 years following adoption, data availability is not as binding a constraint – therefore, I estimate hazard rates for repealing alien voting within the *next* decade.[24]

[24] Two states – Montana and Wisconsin – included grandfathering provisions in the repeal of alien voting rights: Montana repealed alien voting right upon entry to the Union in 1889, but continued to allow existing voters to vote until 1894; Wisconsin voted to repeal alien voting in 1908, but allowed existing voters to vote until 1908. Given that this paper models the *decision* to repeal, I use the earlier dates in the analysis. However, results do not substantively change using the grandfathered dates. Results available upon request.

However, the sample size is significantly smaller than in the previous section, as, necessarily, this sample includes only those states that had ever adopted alien voting: these are the only states that were at risk of repeal.

Results are presented in table 2.2. There is suggestive evidence for the role of social costs in determining the timing of repeal. States with higher proportions of the population that were foreign-born were relatively more likely to repeal alien voting earlier: one additional standard deviation of foreign-born population is associated with a 47% increase in the hazard of repeal. There is some scope for outliers influencing these results: excluding Georgia – which repealed early, in 1877, and had only 1% of the population that was foreign-born – increases both the size and statistical significance of this estimate. Specifically, in this trimmed sample, a one standard deviation increase in the fraction of the population that is foreign born leads to a 146% increase in the hazard of adoption ($p = .07$). In the full sample, states with a higher proportion of foreign-born individuals from Central or Eastern Europe were more likely to repeal, but this effect is not precisely estimated.

Column (2) presents estimated hazard rates for economic covariates. States with a greater fraction of improved acreage were significantly less likely to repeal alien voting, implying that less agriculturally developed states were *more* likely to repeal alien voting within a given decade. This is a striking result, given that states with less agriculturally developed land were significantly *more* likely to adopt alien voting, and may reflect a broad shift in attitudes against immigrants that was strongest in states in which immigrants did not play a significant role in the agricultural sector. That this may be driven by attitudes, rather than economic reality, is borne out by the lack of significant results for either agricultural values or manufacturing output.

Estimates of hazard rates for political factors are presented in Column (3). If the repeal of alien voting was motivated – even in part – by the Republican Party seeking to reduce support for the Democratic Party, then, ceteris paribus, a higher fraction of Democratic Party support should make states less likely to repeal alien voting. However, I find no statistical evidence for *any* effect of Democratic Party support on the hazard of repeal.

Finally, in Column (4), I estimate the effects of constitutional amending procedures on the hazard of repeal. I find strong evidence that procedural costs mattered: states that required a simple popular majority for ratification of the amendment – as opposed to a majority of all those voting at that particular election – were eight times more likely to repeal alien voting in a given decade. This accords with the narrative evidence for the salience of the amending procedures, and, to the extent that states do not strategically adopt amendment procedures, lends credibility to research designs that exploit state-level constitutional change (as in Henderson 2017).

2.6 County-level Support for Repeal of Alien Voting

The state-level analysis supports the hypotheses that alien voting was adopted for primarily economic purposes, with the goal of inducing much needed in-migration, and repealed due to the social costs of immigration. Moreover, the evidence suggests a significant role for procedural factors in the timing of both adoption and repeal. However, it does not (and, mechanically, can not) control for within-state forces leading to change in the franchise. Therefore, in this section, I supplement the previous analysis by exploiting the fact that constitutional amendments had to pass a popular referendum for ratification. Using newly collected county-level vote totals on the repeal of alien voting, I estimate the roles of the demographic, economic, and political factors for popular support for the repeal of alien voting *within* alien voting states. This analysis is descriptive, not causal; however, this is still a useful exercise insofar as I am able to test for correlates of support for repeal controlling for unobserved within-state factors.

I collected vote totals for alien voting repeal from 1,122 counties in thirteen states. The states and totals are listed in table 2.3; the counties are mapped by the support for alien voting repeal, in figure 2.2. There is considerable variation in support for repeal, ranging from .06 (Duval County, Florida, 1894) to 1.0 (Bennett County, South Dakota, 1918). The average level of support across counties is 63.3% - which is consistent with the fact that all of these measures passed at the state level – and a majority of the voting population supported repeal in 926 of 1,122 counties. I test for effect of a parsimonious set of county-level economic, political, and demographic covariates on the fraction of the voters supporting alien voting repeal using the following specification:

$$Fraction\ Yes_{cs} = \alpha_s + \boldsymbol{\beta} X_{cs} + \varepsilon_{st} \qquad (3)$$

Fraction Yes_{cs} is the measure of support for the proposed amendment in county c, state s, and X_{cs} is a vector of county-level covariates: the fraction of the population that is foreign-born, the fraction of the immigrant population that are from "new" sending counties, the fraction of votes for the Democratic in the most recent gubernatorial election, and the log of the real value manufacturing output per capita. With the exception of the gubernatorial election data, which in general is from the election coinciding with, or immediately prior to, the year of repeal, all covariates are taken from the decadal census prior to the repeal election. I include state fixed effects in order to capture unobservable time-invariant differences across states in attitudes toward immigrants. Standard errors are calculated using wild bootstrap in order to account for the small number of state clusters (Cameron et al 2008).

Results are presented in table 2.4. Consistent with social costs of immigration driving the repeal of alien voting, I find that counties with a higher fraction of the population that are foreign born tend to support the repeal of alien voting more strongly; however, the effect sizes are small and imprecisely measured. Contrary to the narrative evidence that suggests that the Democratic party had an interest in keeping non-citizen immigrants in the electorate, support for the Democratic gubernatorial candidate is *positively* related with support for repeal of alien voting, but, again, the effect sizes are imprecisely estimated. Finally, there is no evidence that real manufacturing values are related to the intensity of support for alien voting.

These results mask heterogeneity among states. I re-estimate specification (3) using all four covariates on each state individually, and present each coefficient in the corresponding quadrant in figure 2.3 (with underlying results in appendix tables 2.3.1, 2.3.2, and 2.3.3). The fraction of the population that is foreign born is positively associated with support for repeal in Missouri and Texas, and negatively so in Kansas and Nebraska; moreover, the fraction of the immigrant population that is from "new" sending countries is significantly negatively associated with support for alien voting in Wisconsin and South Dakota, and nearly so in North Dakota and Michigan, but positively related in Arkansas. There is considerable variance among states in the relationship between support for the Democratic Party and support for repeal of alien voting: the two are generally positively related in Southern states, negatively related for states bordering the Great Lakes, and positive for more western states. Finally, the relationship of manufacturing values to support for repeal varies significantly as well: it is (statistically significantly) positive in Arkansas, Missouri, Wisconsin, and Colorado, but smaller in magnitude and less precisely estimated in all other states. The heterogeneity of effects across states suggests that the costs and benefits of restricting the franchise are local, and may explain why scholars have in general failed to document significant state-level predictors for the timing of women's suffrage: the factors that predict support in one state may not do so in another (Miller 2008).

While these regressions are informative as to the differing correlates of support for repeal of alien voting within state, this can not account for unobserved county-specific factors which may influence attitudes toward alien voting (and which may also be correlated with the covariates). In order to control for county fixed effects, I leverage the fact that Missouri attempted to repeal alien voting twice: first, unsuccessfully, in 1912, and then again, successfully, in 1924. I re-estimate specification (3) on these years for only Missouri, and include both year fixed effects and county fixed effects. I present these results, together with the cross-sectional results from 1924, in panels A and B of Table 2.5.

The results suggest the importance of unobserved, time-invariant factors in determining support for alien voting repeal. From Panel A, it is clear that, in the 1924 cross-section, counties in Missouri with a higher fraction of foreign born individuals were significantly more likely to support the repeal of alien voting. However, increasing the fraction of the population that is foreign born *within* a county does *not* increase the support for alien voting repeal and, if anything, weakly reduces it: the coefficient on the fraction foreign born ranges from -.445 to -.535 in Panel B. Similarly, counties in Missouri in 1924 with higher per capital real manufacturing output value were more likely to support alien voting repeal, but, again, there is no statistical evidence for the within-county effect of an increase in manufacturing values on support for alien voting repeal.

2.7 Conclusion

This paper studies the determinants of a little-known expansion and contraction of the electorate that took place from 1848 to 1926 and that affected 24 states and territories: the political inclusion, and exclusion, of non-citizen immigrants. I find evidence consistent with the hypothesis that less agriculturally developed states enacted alien voting in order to induce immigration: a lower fraction of adult males in the population, less improved agricultural acreage, and a smaller manufacturing sector all predicted earlier adoption of alien voting. Notably, despite speculation in the academic literature, there is no evidence that states enfranchised aliens in order to bolster support for the Democratic Party.

There is also no evidence that economic hardship drove the exclusion of immigrants from the electorate, nor is there evidence that the Republican Party strategically excluded immigrants in order to reduce the political power of Democrats. Instead, the results are consistent with social costs driving disenfranchisement: states in which immigrants comprised a greater fraction of the population, and, of the immigrant pool, where Central and Eastern European immigrants were more populous, were more likely to repeal alien voting earlier. I corroborate this using newly collected data on county-level support for the repeal of alien voting: support for repeal of alien voting is positively related to the fraction of the population that is foreign-born, although there is substantial heterogeneity in the correlates of support across states.

The strongest state-level results indicate that the details of the constitutional amending process significantly affected the timing of adoption and repeal. States that re-wrote their constitutions for exogenous reasons were more than 20 times more likely to adopt alien voting than established states, suggesting that the fixed costs of constitutional overhaul presented a significant barrier to adoption. Moreover, states in which ratification of a proposed constitutional amendment required only a majority of voters voting on that particular amendment, instead of a majority of *all* voters voting at the ratification election, were eight times more likely to repeal alien voting in a given decade. This finding implies that identification strategies based on the timing of constitutional change may be promising in terms of exogeneity, although more work is needed to understand the distribution of amending procedures across states.

While alien voting may have been intended to induce in-migration and economic development, this paper does not address whether or not this strategy succeeded. The fact that population growth was a predictor of adoption suggests that economic development may have occurred even in the absence of alien voting. The demographic, economic, and political *consequences* of alien enfranchisement are important questions in economic history, and I leave it to future research for further exploration.

2.8 Works Cited

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

Abramitzky, Ran and Leah Platt Boustan. 2016. “Immigration in American History.” *Journal of Economic Literature.*

Abramitzky, Ran, Leah Platt Boustan and Katherine Eriksson. 2016. “Cultural Assimilation during the Age of Mass Migration.” NBER Working Paper 22381.

Abramitzky, Ran and Leah Platt Boustan. 2016. “Immigration in American History.” *Journal of Economic Literature.*

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

Acemoglu, Dron and James A. Robinson. 2000. “Why did the West extend the Franchise? Democracy, inequality, and Growth in a Historical Perspective.” *Quarterly Journal of Economic*s, 115(4): 497-551.

Acemoglu, Daron and James A. Robinson. 2001. “A Theory of Political Transitions.” *American Economic Review,* 91(4): 938-963.

Aidt, Toke S. and Bianca Dallal. 2008. “Female voting power: The contribution of women’s suffrage to the growth of social spending in Western Europe (1869- 1960)”, *Public Choice,* 134(3-4): 391-417.

Augenblick, Ned and Scott Nicholson. 2016. “Ballot Position, Choice Fatigue, and Voter Behavior,” *Review of Economic Studies*, 83(2): 460-480.

Aylsworth, Leon E. 1931. “The Passing of Alien Suffrage,” *The American Political Science Review*, 25(1):114-116.

Bandiera, Oriana, Myra Mohnen, Imran Rasul, and Martina Viarengo. 2016. “Nation-Building Through Compulsory Schooling During the Age of Mass Migration.” Manuscript.

Bandiera, Oriena, Imran Rasul, and Martina Viarengo. 2013. “The Making of Modern America: Migratory Flows in the Age of Mass Migration.” Journal of Development Economics, 102: 23-47

Benlemech, Efraim, and Toby Moskowitz. 2010. “The Political Economy of Financial Regulation: Evidence from U.S. State Usury Laws in the 19th Century.” *Journal of Finance,* 65(3): 1029-1073.

Berkowitz, Daniel, and Karen Clay. 2005. “American Civil Law Origins: Implications for State Constitutions.” *American Law and Economics Review,* 7(1): 62-84

Bertocchi, Graziella. “The Enfranchisement of Women and the Welfare State.” *European Economic Review,* 55(4): 535-553.

Bertocchi, Graziella and Chiara Strozzi. 2010. “The Evolution of Citizenship: Economic and Institutional Determinants.” *The Journal of Law & Economics,* 53(1): 95-136.

Braun, Sebastian and Michael Kvasnicka. 2013. “Men, Women, and the Ballot: Gender Imbalances and Suffrage Extensions in the United States.” *Explorations in Economic History,* 50: 405-426.

Bullock, Charles S. III and Richard E. Dunn. 1996. “Election Roll-Off: A Test of Three Explanations.” *Urban Affairs Review,* 32(1): 71-86.

Burnham, W. Dean. PARTISAN DIVISION OF AMERICAN STATE GOVERNMENTS, 1834-1985. Conducted by Massachusetts Institute of Technology. ICPSR ed. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor], 198?. http://doi.org/10.3886/ICPSR00016.v1

Burnham, W. Dean, Jerome M. Clubb, and William Flanigan. STATE-LEVEL CONGRESSIONAL, GUBERNATORIAL AND SENATORIAL ELECTION DATA FOR THE UNITED STATES, 1824-1972. ICPSR ed. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor], 1991. http://doi.org/10.3886/ICPSR00075.v1

Cameron, Colin A, Jonah B. Gelbach, and Douglas L. Miller. 2008. “Bootstrap-based Improvements for Inference with Clustered Standard Errors.” *Review of Economics and Statistics,* 90(3): 414-427

Cascio, Elizabeth and Ebonya Washington. 2014. “Valuing the Vote: The Redistribution of Voting Rights and State Funds Following the Voting Rights Act of 1965”, *Quarterly Journal of Economics*: 379-433

Conley, John P. and Akram Temimi. 2001. “Endogenous Enfranchisement when Groups’ Preferences Conflict.” *Journal of Political Economy*, 109(1): 79-102.

Dinan, John J. 2010. “The Political Dynamics of Mandatory State Constitutional Convention Referendums: Lessons from the 2000s Regarding Obstacles and Pathways to their Passage.” *Montana Law Review*, 71(2): 395-432

Fishback, Price and Shawn Kantor. 1998. “The Adoption of Workers’ Compensation in the United States.” The Journal of Law and Economics, 41(2): 305-342.

Goldin, Claudia. 1994. “The Political Economy of Immigration Restriction in the United States.” *The Regulated Economy: A Historical Approach to Political Economy*. University of Chicago Press.

Haines, Michael, Price Fishback, and Paul Rhode. United States Agriculture Data, 1840 - 2012. ICPSR35206-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2016-06-29. http://doi.org/10.3886/ICPSR35206.v3

Haines, Michael R., and Inter-university Consortium for Political and Social Research. Historical, Demographic, Economic, and Social Data: The United States, 1790-2002. ICPSR02896-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2010-05-21. http://doi.org/10.3886/ICPSR02896.v3

Harper-Ho, Virginia. 2000. "Noncitizen Voting Rights: The History, the Law and Current Prospect for Change*," Immigration and Nationality Law Review*, 21:477

Hayduk, Ron. *Democracy for all: Restoring Immigrant Voting Rights in the United States.* Taylor and Francis Group, 2006

Henderson, Morgan. 2017. "What Drives Expansions and Contractions of the Electorate? Evidence from Alien Voting in the United States." Manuscript.

Hill, Melvin B. Jr. and Laverne Williamson Hill. 2008. "Georgia: Tectonic Plates Shifting." In *The Constitutionalism of the American States*, eds. George E. Connor and Christopher W. Hammons. University of Missouri Press.

Inter-university Consortium for Political and Social Research. United States Historical Election Returns, 1824-1968 [Computer File]. ICPSR00001-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1999-04-26. http://doi.org/10.3886/ICPSR00001.v3

Keyssar, Alexander. 2000. *The Right to Vote: The Contested History of democracy in the United States*, Basic Books.

Lamoreaux, Naomi R. and John J. Wallis. 2016. "States, not Nations: The Sources of Poltical and Economic Development in the Early United States." Manuscript.

Lemke, Jayme S. 2016. "Why Statehood? A Note on Interpreting Jurisdictional Competition in U.S. History." Manuscript.

Lizzeri, Alessandro and Nicola Persico. 2004. "Why did the Elites Extend the Suffrage? Democracy and the Scope of Government, with an Application to Britain's "Age of Reform"." *Quarterly Journal of Economics*, 119(2): 707-765.

Lott, John and Lawrence Kenney. 1999. "Did Women's Suffrage Change the Size and Scope of government?" *Journal of Political Economy*, 107(6)

Lutz, Donald S. 1994. "Toward a Theory of Constitutional Amendment". *The American Political Science Review*, Vol. 88, No. 2 pp. 355-370

McCulloch, Albert J. 1929. *Suffrage and Its Problems*. Warwick and York, Inc: Baltimore.

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

Meltzer, A. H., and S. F. Richard. 1981. "A rational theory of the size of government." *Journal of Political Economy*, 89: 914-927

Meltzer, A. H., and S. F. Richard. 1983. "Tests of a Rational Theory of the Size of Government." *Public Choice*, 41(3):403-418

Neuman, Gerald. 1992. "We are the People: Alien Suffrage in German and American Perspective", *Michigan journal of International Law*

Nikolova, Elena. 2015. "Destined for Democracy? Labour Markets and Political Change in Colonial British America." *British Journal of Political Science,* 47: 19-45.

Persson, Torsten and Guido Tabellini. 2003. *The Economic Effects of Constitutions.* MIT Press: Cambridge, MA.

Porter, Kirk Harold. 1918. *A History of Suffrage of the United States*. University of Chicago Press: Chicago.

Prentice, R. L. and L. A. Gloeckler. 1978. "Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data." *Biometrics,* 34(1): 57-67.

Raskin, James. 1993. "Legal Aliens, Local Citizens: the Historical, Constitutional, and Theoretical Meanings of Alien Suffrage." *University of Pennsylvania Law Review*

Rosberg, Gerald M. 1977. "Aliens and Equal Protections: Why Not the Right to Vote?" Michigan Law Review.

Rosenbloom, Joshua L. and William A. Sundstrom. 2011. "Labor-Market Regimes in U.S. Economic History." In Paul W. Rhode, Joshua L. Rosenbloom and David F. Weiman, eds. *Economic Evolution and Revolutions in Historical Time,* Stanford, CA: Stanford University Press.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. *Integrated Public Use Microdata Series: Version 6.0* [dataset]. Minneapolis: University of Minnesota, 2015. http://doi.org/10.18128/D010.V6.0.

Shertzer, Allison. 2013. "Immigrant Group Size and Political Mobilization: Evidence from European Migration to the Unites States." *Journal of Public Economics,* 139: 1-12.

Wallis, John J. 2005. "Constitutions, Corporations, and Corruption: American States and Constitutional Change." *Journal of Economic History,* 65(1): 211-256.

Wallis, John Joseph, "NBER/University of Maryland State Constitution Project", www.stateconstitutions.umd.edu

Ward, Zach and Michael J. Greenwood. 2016. "Immigration Quotas, World War I, and Emigrant Flows from the United States in the Early 20th Century." *Explorations in Economic History,* 55(1): 76-96

Figure 2.1.1: Map of Adoption of Alien Voting

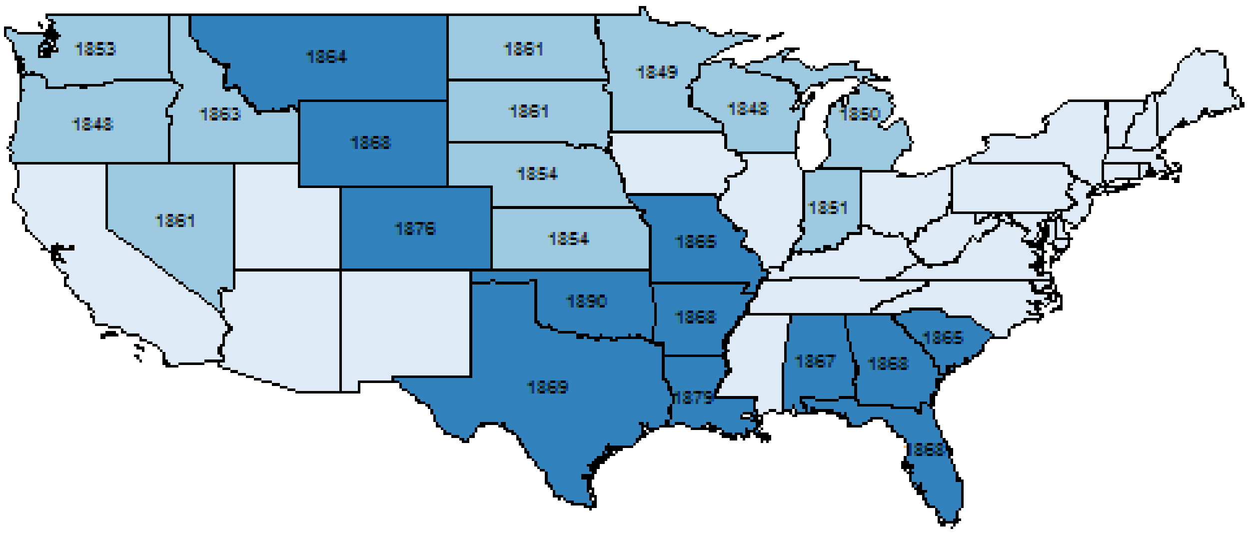
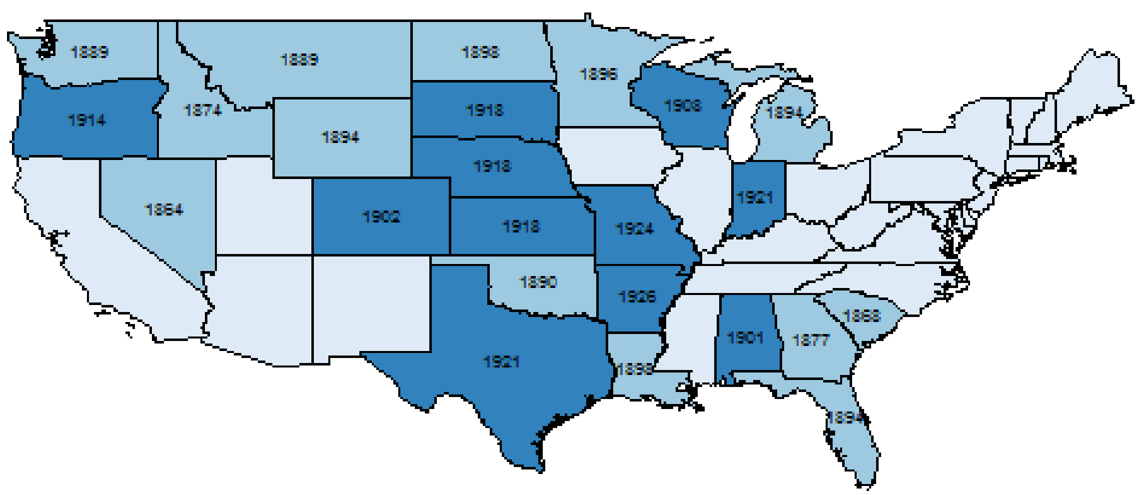
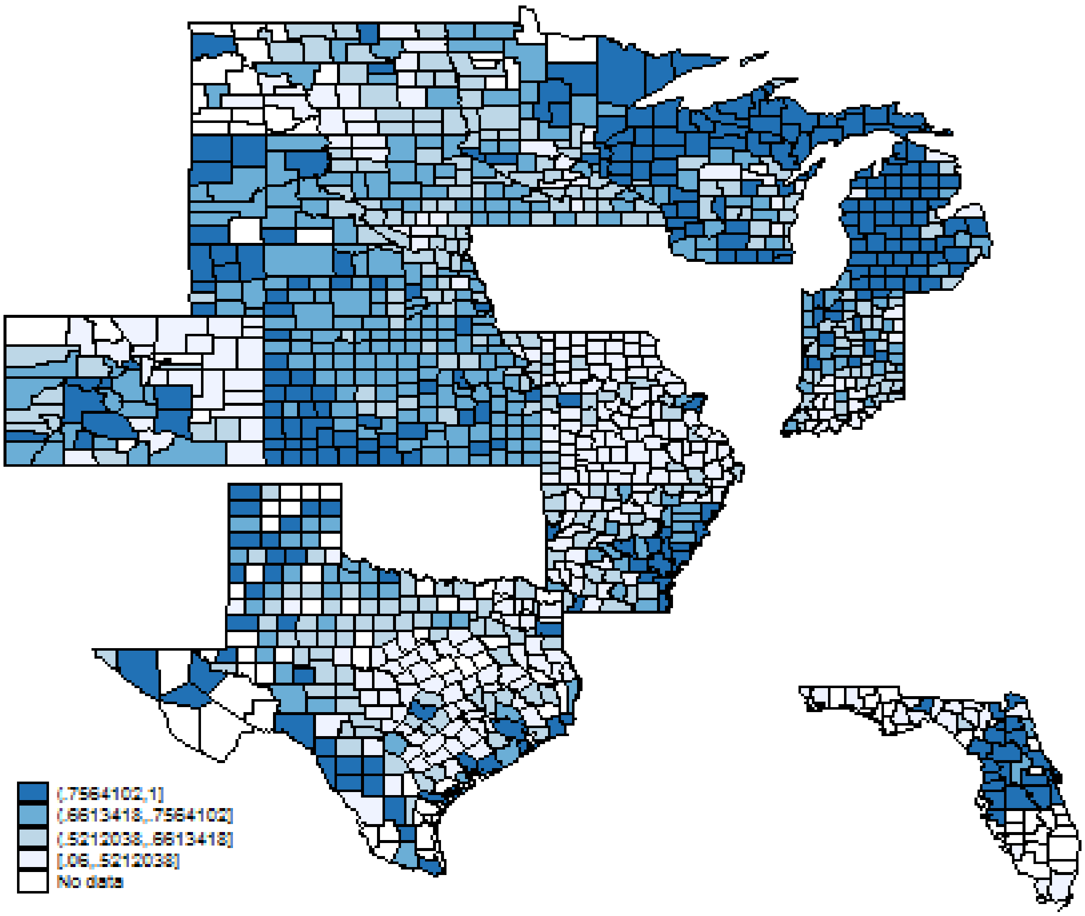


Figure 2.1.2: Map of Repeal of Alien Voting



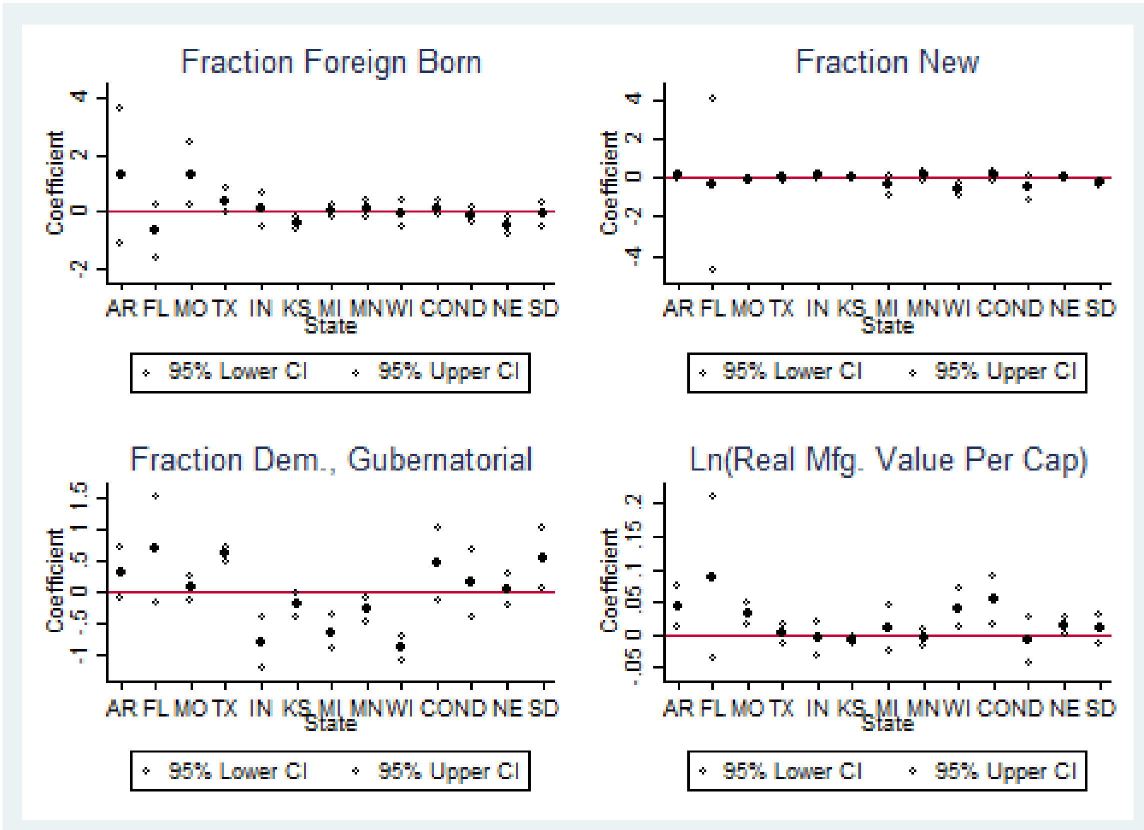
Notes: dates of adoption and repeal are from Keyssar (2000), Hayduk (2006), state constitutions, state blue books state legislative manuals, and territorial constitutions. See the appendix for details.

Figure 2.2: Support for Repeal of Alien Voting, County Level



Notes: election returns on constitutional amendments are from various state legislative manuals, blue books, and canvassing board reports. See appendix table 1B for the exact source for each state.

Figure 2.3: Correlates of County-level support for Repeal of Alien Voting, by State



Notes: each point in each graph is the coefficient for a state-specific regression of the fraction of voters supporting alien repeal on the fraction of the population that is foreign born, the fraction of immigrants from "new" sending countries, the fraction of support for the democratic gubernatorial candidate at the most recent election, and the natural log of the real value of manufacturing output per capita in the county. Underlying regression coefficients are presented in appendix tables 2A – 2C.

Data are from county-level elections for constitutional amendments repealing declarant alien suffrage in 13 states (AR, CO, FL, IN, KS, MI, MN, MO, ND, NE, SD, TX, WI). In general, gubernatorial election data are from the election corresponding, or just prior to, the date of repeal. I used gubernatorial election data from 1920 for Arkansas, since it voted to repeal alien voting in 1920. Moreover, gubernatorial election data was not available for Florida in 1892, so I used values from 1888 (the next most recent election). In North Dakota, the Democratic Party was not represented in the 1898 gubernatorial election, so values for the "Fusion" candidate were used in place of support for the Democratic candidate. Similarly, in Minnesota in the 1896 gubernatorial election, values for the "People's and Democrat" party were used. Standard errors are estimated using wild bootstrap, clustered at the state level.

Table 2.1: Determinants of Adoption of Alien Voting

	Adopt (1)	Adopt (2)	Adopt (3)	Adopt (4)
Ln(Population)	0.714			
	(0.77)			
Fraction Adult Male	0.238			
	(3.18)***			
Fraction Foreign Born	0.941			
	(0.19)			
Fraction 'New' Immigrants	1.800			
	(2.02)**			
Population Growth	3.149			
	(3.75)***			
Fraction improved acres		0.197		
		(2.97)***		
Ln(Real Value of Farms Per Acre)		1.273		
		(0.53)		
Ln(Real Mfg. Output p.c.)		0.508		
		(2.11)**		
Fraction of Gub. Votes Dem			0.885	
			(0.51)	
Fraction of Upper House Dem.			1.786	
			(0.55)	
Fraction of Lower House Dem.			0.501	
			(0.67)	
Exogenous Constitutional Change				23.865
				(4.27)***
% Leg. Approval				0.963
				(0.70)
Successive Approvals				0.327
				(1.08)
Simple Majority				0.908
				(0.80)
N	143	144	137	123

Z-statistics reported in parentheses. Estimates are from a non-parametric discrete-time hazard model, and so tests of significance are based upon a null hypothesis that the coefficient is 1. All continuous covariates are measured in effect sizes, so estimates should be interpreted as the effect of a one standard deviation increase. Sample sizes differ due to data availability – certain states did not have constitutional amendment procedures in place until the late 19th century. ID and OK are excluded from the sample, as they both adopted, and repealed, alien voting as territories. Decade fixed effects are included in all specifications, and standard errors are clustered at the state level. * p<0.1; ** p<0.05; *** p<0.01

Table 2.2: Determinants of Repeal of Alien Voting

	Repeal	Repeal	Repeal	Repeal
	(1)	(2)	(3)	(4)
Ln(Population)	0.581			
	(0.72)			
Fraction Adult Male	1.218			
	(0.29)			
Fraction Foreign Born	1.466			
	(0.92)			
Fraction 'New' Immigrants	1.091			
	(0.22)			
Population Growth	0.498			
	(1.46)			
Fraction improved acres		0.482		
		(2.26)**		
Ln(Real Value of Farms Per Acre)		1.07		
		(0.04)		
Ln(Real Mfg. Output p.c.)		1.572		
		(1.32)		
Fraction of Gub. Votes Dem			1.163	
			(0.35)	
Fraction of Upper House Dem.			0.357	
			(1.59)	
Fraction of Lower House Dem.			1.826	
			(0.92)	
Exogenous Constitutional Change				.
				.
% Leg. Approval				1.048
				(1.06)
Successive Approvals				1.694
				(0.67)
Simple Majority				9.000
				(2.19)**
N	72	71	44	36

Z-statistics are reported in parentheses. Estimates are from a non-parametric discrete-time hazard model, and so tests of significance are based upon a null hypothesis that the coefficient is 1. All continuous covariates are measured in effect sizes, so estimates should be interpreted as the effect of a one standard deviation increase. Montana repealed alien voting upon entry to the union in 1889, and so political and constitutional data are unavailable prior to 1890. I exclude Nevada and South Carolina from the estimation sample, as they both adopted and repealed alien voting within the 1860s: that is, as of 1860, they had not yet adopted alien voting, and so were not yet at risk of repeal. All specifications include decade fixed effects. Standard errors are clustered at the state level. * $p<0.1$; ** $p<0.05$; *** $p<0.01$

Table 2.3: Popular Support for Repeal of Alien Voting

State	Year of Repeal	Vote Result
AR	1926[1]	87,273-49,757
CO	1902	44.769-27,077
FL	1894	11,691-5,664
IN	1921	130,242-80,574
KS	1918	238,453-91,617
MI	1894	117,088-31,537
MN	1898	97,980-52,454
MO	1924	175,580-152,713
NE	1918	123,292-51,600
ND	1898	21,177-16,329
OR[2]	1914	164,879-39,847
SD	1918	49,318-28,934
TX	1921	57,622-53,910
WI	1908	86,576-36,733

For sources for county-level votes for repeal, see sources section in Appendix Table 1B.
1: The vote actually took place in 1920, but failed to obtain a majority of the voters voting at that election. The Supreme Court eventually abolished alien voting in 1926.
2. I was not able to locate county-level breakdowns for this election.

Table 2.4: County-Level Support for Repeal of Alien Voting

	Fraction Voting for Repeal	Fraction Voting for Repeal	Fraction Voting for Repeal	Fraction Voting for Repeal
Fraction Foreign Born	0.007	0.012	0.033	0.029
	(0.048)	(0.061)	(0.096)	(0.101)
Fraction 'New' Immigrants		-0.011	-0.012	0.007
		(0.066)	(0.061)	(0.065)
Fraction Dem., Gubernatorial			0.196	0.149
			(0.257)	(0.273)
Ln(Real Value Mfg. Output per Capita)				0.010
				(0.008)
R^2	0.31	0.31	0.33	0.37
N	1,108	1,108	1,104	978

Data are from county-level elections for constitutional amendments repealing declarant alien suffrage in 13 states (AR, CO, FL, IN, KS, MI, MN, MO, ND, NE, SD, TX, WI). In general, gubernatorial election data are from the election corresponding, or just prior to, the date of repeal. I used gubernatorial election data from 1920 for Arkansas, since it voted to repeal alien voting in 1920. Moreover, gubernatorial election data was not available for Florida in 1892, so I used values from 1888 (the next most recent election). In North Dakota, the Democratic Party was not represented in the 1898 gubernatorial election, so values for the "Fusion" candidate were used in place of support for the Democratic candidate. Similarly, in Minnesota in the 1896 gubernatorial election, values for the "People's and Democrat" party were used. Standard errors are estimated using wild bootstrap, clustered at the state level. * $p<0.1$; ** $p<0.05$; *** $p<0.01$

Table 2.5: County-Level Support for Repeal of Alien Voting in Missouri

	Fraction Voting for Repeal	Fraction Voting for Repeal	Fraction Voting for Repeal	Fraction Voting for Repeal
Panel A: 1924 Cross-Section				
Fraction Foreign Born	1.696	1.854	1.954	1.346
	(0.596)***	(0.620)***	(0.632)***	(0.564)**
Fraction 'New' Immigrants		-0.042	-0.044	-0.051
		(0.053)	(0.053)	(0.049)
Fraction Dem., Gubernatorial			0.107	0.081
			(0.108)	(0.101)
Ln(Real Value Mfg. Output per Capita)				0.032
				(0.008)***
R^2	0.07	0.07	0.08	0.18
N	114	114	114	114
Panel B: 1912 and 1924 County-Year Panel				
Fraction Foreign Born	-0.515	-0.535	-0.502	-0.445
	(0.694)	(0.710)	(0.703)	(0.730)
Fraction 'New' Immigrants		0.007	0.012	0.011
		(0.031)	(0.034)	(0.033)
Fraction Dem., Gubernatorial			-0.408	-0.368
			(0.398)	(0.401)
Ln(Real Value Mfg. Output per Capita)				0.040
				(0.053)
R^2	0.89	0.89	0.89	0.89
N	228	228	228	228

Data are from county-level elections for constitutional amendments repealing declarant alien suffrage in Missouri in 1912 and 1914. Standard errors are corrected for heteroskedasticity. * p<0.1; ** p<0.05; *** p<0.01

2.9 Appendix

Data Appendix 2.1: Data Sources and Construction

Population estimates from 1850 onward are made by aggregating Census microdata (Ruggles et al 2015). The value per acre of farmland and buildings is from Haines, Fishback, and Rhode (2016). This is scaled by the historical CPI from the Minneapolis Fed. Data on improved acreage and manufacturing output are from Haines (2010). Note that manufacturing output is missing for 1910, so is linearly interpolated from 1900-1920.

Data on statehouse composition are from Burnham (1989) (ICPSR 16). To account for the differential timing of statehouse elections, the fraction of the upper and lower houses that are democratic in each decade is estimated by averaging number of seats that are Democrat in the 10 years for the previous decade (for example, from 1840-1849 for the 1850 value), and dividing by the total number of seats available in that chamber averaged in the same way. Gubernatorial election data are from Burnham, Clubb, and Flanigan (1991) (ICPSR 75). To account for the differential timing of gubernatorial elections across states, the fraction of votes for the democratic candidate in each decade is estimated by averaging the fraction of votes going to the Democratic candidate in the 10 years in the previous decade (for example, from 1840-1849 for the 1850 value).

County level gubernatorial data are from ICPSR 1. The fraction of votes going to the Democratic candidate is generated using candidate-level vote totals. In instances where there was no Democratic candidate, I use the Democratic equivalent (for example, the Fusion party).

Features of state constitutions were coded from the original texts of every constitutional version, in every state, from 1830-1930. I relied heavily, but not exclusively, on Wallis (2000), which is a database of the text of constitutions from 1776 onward for (almost) every state.

The start and end dates of alien voting were culled from a variety of sources. To start, dates were obtained from Keyssar (2000) and Hayduk (2006). These did not entirely concur, however, and multiple sources had noted the difficulty in obtaining correct dates. Therefore, I cross-checked these with the suffrage provisions in every version of every state's constitution from 1830-1930, and also consulted state blue books, legislative manuals, and reports of secretary of state. For territories, I checked suffrage provisions in the compiled territorial statutes. I verified starting dates by documenting that prior to year t, the *previous* constitution allowed only citizen suffrage. This is noted in appendix table 2.1.1. Ending dates of alien voting were more straightforward, since 14 of these were subject to popular ratification votes, with much more by way of documentation. The precise sources used for both start and end dates each are noted in Appendix tables 2.1.1 and 2.1.2.

Appendix Table 2.1.1: Declarant Alien Voting Start Dates, with Sources

State	Year Start	Mode	Sources
AL	1867	New Constitution	Constitutions of 1865, Article 8 (citizen suffrage) and 1867, Article 7 (alien suffrage)
AR	1868	New Constitution	Constitutions of 1864, article 4 (citizen suffrage) and 1868, Article 8 (alien suffrage)
CO	1876	*New Constitution*	An Act to Provide a Temporary Government for the Territory of Colorado, February 28, 1861, Section 5 (Citizen suffrage); Constitution of 1876, Article 7 (alien suffrage)
FL	1868	New Constitution	Constitutions of 1865, Article 6 (citizen suffrage) and 1868, Article 14 (alien suffrage)
GA	1868	New Constitution	Constitution 1865, Article 5 (citizen suffrage) and 1868, Article 2 (alien suffrage)
ID	1863	Territorial Act	An Act to Provide a Temporary Government for Territory of Idaho, March 3, 1863, Section 5[1]
IN	1851	New Constitution	Constitutions of 1816, Article 6 (citizen suffrage) and 1851, Article 2 (alien suffrage)
KS	1854	Territorial Act	An Act to Organize the Territories of Nebraska and Kansas, May 30, 1854, Section 5 (alien suffrage); Kansas Constitution of 1859, article 5 (alien suffrage)
LA	1879	New Constitution	Constitution of 1868, Article 6 (citizen suffrage) and 1879 (Alien Suffrage)
MI	1850	New Constitution	Constitutions of 1835, Article 2 (Citizen suffrage) and 1850, Article 7 (alien suffrage)
MN	1849	Territorial Act	An Act to Establish the Territorial Government of Minnesota, March 3, 1849, Section 5 (alien suffrage); 1857 Constitution, Article 7 (alien suffrage)
MO	1865	New Constitution	Constitutions of 1820, Article 3 (citizen suffrage) and 1865, Article 2 (alien suffrage)
MT	1864	Territorial Act	An Act to Provide a Temporary Government for the Territory of Montana, May 26, 1864, Section 5 (alien suffrage)
NE	1854	Territorial Act	An Act to Organize the Territories of Nebraska and Kansas, May 30, 1854, Section 5 (alien suffrage); Constitution of 1859 Article 5 (alien suffrage)
ND	1861	Territorial Act	An Act to provide a Temporary Government for the Territory of Dakota, and to create the Office of the Surveyor General therein, March 2, 1861, Section 5 (alien suffrage); 1889 Constitution, Article 5 (alien suffrage)
NV	1861	Territorial Act	An Act to Organize the Territory of Nevada, March 2, 1861, Section 5 (alien suffrage)
OR	1848	Territorial Act	An Act to Establish the Territorial Government of Oregon, August 14, 1848, Section 5 (alien suffrage); Constitution of 1857, Article 2 (alien suffrage)

OK	1890	Territorial Act	An Act to Provide a Temporary Government for the Territory of Oklahoma, to Enlarge the Jurisdiction of the United States Court in the Indian Territory, and for other purposes, May 2, 1890, Section 5 (alien suffrage)
SC	1865	New Constitution	Constitutions of 1861, Article 1 (Citizen suffrage), and 1865, Article 4 (Alien suffrage)
SD	1861	<u>Territorial Act</u>	An Act to provide a Temporary Government for the Territory of Dakota, and to create the Office of the Surveyor General therein, March 2, 1861, Section 5 (Alien suffrage); Constitution of 1889, Article 7 (Alien suffrage)
TX	1869	New Constitution	Constitution of 1866, Article 3 (Citizen suffrage); Constitution of 1869, Article 3 (Alien suffrage)
WA	1853	Territorial Act	An Act to Establish the Territorial Government of Washington, March 2, 1853, Section 5 (Alien suffrage)
WI	1848	*New Constitution*	An Act Establishing the territorial government of Wisconsin, 1836, Section 5 (Citizen suffrage); Constitution of 1848, Article 3 (alien suffrage)
WY	1868	Territorial Act	An Act to Provide a Temporary Government for the Territory of Wyoming, July 25, 1868, Section 5 (alien suffrage)

Notes: for every state, it was verified that, prior to the adoption of alien voting, the standard of citizen voting was in place. This is noted in parentheses in the table. Italics denotes a new constitution upon entry into the Union. Underline denotes that the state continued alien voting upon entry to the union.
1: The territorial Act did not specify declarant alien voting; rather, it permitted all free white male inhabitants the right to vote, and, it specifies that state legislature will decide on voting qualifications in all subsequent elections.

Appendix Table 2.1.2: Declarant Alien Voting End Dates, with Sources

State	Year End	Mode	Sources
AL	1901	New Constitution	Constitution of 1901, Article 8
AR	1926	Amendment	Biennial Report of the Secretary of State, 1925-1926, 1919-1920
CO	1902	Amendment	Legislative Manual, 1903
FL	1894	Amendment	Report of the Secretary of State of Florida, 1893/1894
GA	1877	New Constitution	Constitution of 1877, Article 2
ID	1874	Territorial Act	Keyssar (2000) pg 138; General Laws of the Territory of Idaho (1874-1875)
IN	1921	Amendment	Yearbook of the State of Indiana 1921
KS	1918	Amendment	Biennial Report of the Secretary of State, 1917/1918
LA	1898	New Constitution	Constitution of 1898, Section 197
MI	1894	Amendment	Michigan Legislative Manual and Official Directory, 1895-96, 1897
MN	1896	Amendment	Legislative Manual of the State of Minnesota, 1897
MO	1924	Amendment	Official Manual of the State of Missouri, 1925-1926
MT	1889[1]	*New Constitution*	Compiled Statutes of Montana 1889, Section 1007; Constitution of 1889, Article 9
NE	1918	Amendment	Official Report of the Nebraska State Canvassing Board, 1918
ND	1898	Amendment	State of North Dakota Legislative Manual, 1899
NV	1864	*New Constitution*	Constitution of 1864, Article 2
OR	1914	Amendment	State of Oregon Blue Book and Official Directory, 1921-1922
OK	1890	Territorial Act	The Statutes of Oklahoma 1890, Chapter 33, Section 1s
SC	1868	New Constitution	Constitution of 1868, Article 8
SD	1918	Amendment	South Dakota Legislative Manual 1921
TX	1921	Amendment	Supplemental Biennial Report of the Secretary of State, 1922
WA	1889	*New Constitution*	Constitution of 1889, Article 6
WI	1908[1]	Amendment	The Blue Book of the State of Wisconsin, 1909
WY	1894	*New Constitution*	Constitution of 1890, Article 6

Italics denotes a new constitution upon entry into the Union.

1: Montana included provisions in repeal to allow existing non-citizen immigrants the right to vote until 1894; Wisconsin, until 1912.

Appendix Table 2.2: Determinants of Adoption of Alien Voting for States Adopting Late in a Given Decade

	Adopt	Adopt	Adopt	Adopt
	(1)	(2)	(3)	(4)
Ln(Population)	0.511			
	(1.55)			
Fraction Adult Male	0.344			
	(2.04)**			
Fraction Foreign Born	0.975			
	(0.08)			
Fraction 'New' Immigrants	2.141			
	(2.37)**			
Population Growth	3.176			
	(3.48)***			
Fraction improved acres		0.100		
		(3.88)***		
Ln(Real Value of Farms Per Acre)		1.000		
		(0.00)		
Ln(Real Mfg. Output p.c.)		0.621		
		(1.19)		
Fraction of Gub. Votes Dem			0.971	
			(0.11)	
Fraction of Upper House Dem.			1.729	
			(0.46)	
Fraction of Lower House Dem.			0.503	
			(0.57)	
Exogenous Constitutional Change				29.862
				(3.92)***
% Leg. Approval				0.955
				(0.85)
Successive Approvals				0.336
				(0.99)
Simple Majority				0.944
				(0.46)
N	137	138	131	118

Z-statistics reported in parentheses. Estimates are from a non-parametric discrete-time hazard model, and so tests of significance are based upon a null hypothesis that the coefficient is 1. All continuous covariates are measured in effect sizes, so estimates should be interpreted as the effect of a one standard deviation increase. Sample sizes differ due to data availability – certain states did not have constitutional amendment procedures in place until the late 19th century. ID and OK are excluded from the sample, as they both adopted, and repealed, alien voting as territories. Also excluded are IN (1851), KS (1861), and NV (1864), since they adopted alien voting in the first half of a given decade. Decade fixed effects are included in all specifications, and standard errors are clustered at the state level. * p<0.1; ** p<0.05; *** p<0.01

Appendix Table 2.3.1: Individual County-Level Regressions for Support for Repeal, by state - SOUTH

	Fraction Voting for Repeal Arkansas	Fraction Voting for Repeal Florida	Fraction Voting for Repeal Missouri	Fraction Voting for Repeal Texas
Fraction Foreign Born	1.265	-0.666	1.346	0.407
	(1.212)	(0.466)	(0.564)**	(0.210)*
Fraction 'New' Immigrants	0.138	-0.288	-0.051	-0.001
	(0.051)***	(2.242)	(0.049)	(0.084)
Fraction Dem., Gubernatorial	0.318	0.691	0.081	0.612
	(0.209)	(0.429)	(0.101)	(0.061)***
Ln(Real Value Mfg. Output Per Capita)	0.044	0.088	0.032	0.002
	(0.015)***	(0.062)	(0.008)***	(0.008)
R^2	0.30	0.13	0.18	0.38
N	74	34	114	132

* p<0.1; ** p<0.05; *** p<0.01

Appendix Table 2.3.2: Individual County-Level Regressions for Support for Repeal, by State – MIDWEST

	Fraction Voting for Repeal Indiana	Fraction Voting for Repeal Kansas	Fraction Voting for Repeal Michigan	Fraction Voting for Repeal Minnesota	Fraction Voting for Repeal Wisconsin
Fraction Foreign Born	0.089	-0.368	0.021	0.130	-0.071
	(0.313)	(0.123)***	(0.100)	(0.149)	(0.241)
Fraction 'New' Immigrants	0.117	-0.009	-0.346	0.120	-0.607
	(0.075)	(0.028)	(0.262)	(0.113)	(0.158)***
Fraction Dem., Gubernatorial	-0.804	-0.202	-0.639	-0.265	-0.878
	(0.205)***	(0.099)**	(0.138)***	(0.099)***	(0.098)***
Ln(Real Value Mfg. Output Per Capita)	-0.005	-0.008	0.011	-0.003	0.041
	(0.013)	(0.003)**	(0.018)	(0.007)	(0.015)***
R^2	0.22	0.28	0.35	0.07	0.61
N	92	96	80	75	70

* p<0.1; ** p<0.05; *** p<0.01

Appendix Table 2.3.3: Individual County-Level Regressions for Support for Repeal, by State – MIDWEST and WEST

	Fraction Voting for Repeal Colorado	Fraction Voting for Repeal North Dakota	Fraction Voting for Repeal Nebraska	Fraction Voting for Repeal South Dakota
Fraction Foreign Born	0.149	-0.104	-0.473	-0.085
	(0.139)	(0.122)	(0.151)***	(0.232)
Fraction 'New' Immigrants	0.139	-0.491	0.011	-0.255
	(0.118)	(0.322)	(0.035)	(0.072)***
Fraction Dem., Gubernatorial	0.463	0.144	0.056	0.547
	(0.290)	(0.277)	(0.133)	(0.241)**
Ln(Real Value Mfg. Output Per Capita)	0.053	-0.008	0.014	0.010
	(0.018)***	(0.017)	(0.007)**	(0.012)
R^2	0.32	0.11	0.23	0.38
N	54	25	77	55

* p<0.1; ** p<0.05; *** p<0.01

Chapter 3

How Well Do Automated Linking Methods Perform? Evidence from the LIFE-M Project[1]

(joint with Martha Bailey, Catherine Massey, and Connor Cole)

New large-scale longitudinal and intergenerational data are revolutionizing the study of U.S. history, the long-term effects of early life influences, and population health and aging. Innovations in automated linking methods have facilitated these developments. This paper uses new ground truth data to evaluate the performance of four commonly used automated record-linking algorithms and two phonetic name-cleaning methods. Our results show high match rates for each algorithm, but we document important shortcomings of each. First, no method (including the ground truth) appears representative. Second, automated linking results in a very high incidence of Type I errors (erroneous matches), ranging from 17 percent to over 60 percent. Third, the use of phonetic name cleaning such as NYSIIS or Soundex increases Type I errors by at least 60 percent. Finally, Type I errors are positively correlated with Type II errors (missed true links) suggesting that current methods used to increase match rate *decrease* the share of true matches. These findings hold in multiple ground truth data, including internally simulated ground truth, LIFE-M data, and the Early Indicator's genealogically linked Oldest Old sample of Union Army veterans. Measurement error introduced by different automated linking methods could have substantial (and difficult to sign) effects on inferences. As an illustration, we show that estimates of intergenerational income elasticities for father-son pairs in the 1940 Census vary by linking method, with some linking methods reducing them as much as 67 percent. We conclude with constructive suggestions for improving automated methods without using clerical review or genealogical methods.

3.1 Introduction

An explosion of large-scale record linkage is transforming the study of U.S. history, the long-term effects of early life influences, and population health and aging. Between 1980 and 2010, only 15 papers using linked data appeared in economics journals, but this number has been growing exponentially since 2010 (Figure 3.1).

[1] This project was generously supported by the National Science Foundation (SMA 1539228), the University of Michigan Population Studies Center Small Grants (R24 HD041028), the Michigan Center for the Demography of Aging (MiCDA, P30 AG012846-21), and the Michigan Institute on Research and Teaching in Economics (MITRE). We gratefully acknowledge the use of the services and facilities of the Population Studies Center at the University of Michigan (R24 HD041028). We are grateful to Dora Costa, Shari Eli, Adriana Lleras-Muney, Joseph Price, and the board members of the LIFE-M project, including Eytan Adar, George Alter, Hoyt Bleakley, Matias Cattaneo, William Collins, Katie Genadek, Maggie Levenstein, Bhash Mazumder, and Evan Roberts, for their helpful suggestions. We are also grateful to Sarah Anderson, Garrett Anstreicher, and Hanna Zlotnick for their excellent contributions to the LIFE-M project and assistance with this analysis.

Data availability and developments in automated linking methods have fueled this growth. The recent release of the 1940 full-count Census (the first census to contain educational attainment, income from wages, and many other outcomes of interest to economists), digitization and linkage of historical censuses to the 1880 full-count Census (Ruggles 2006, Ruggles et al. 2015), and new availability of indexed Vital Statistics records allow researchers to break new ground on old questions and are opening entirely new areas of inquiry.

These data facilitate several large-scale linking initiatives. The Census Bureau has linked the complete count 1940 Census to current administrative data (Census Longitudinal Infrastructure Project, CLIP). The Minnesota Population Center is planning to link the 1940 Census to historical censuses. The Longitudinal, Intergenerational Family Electronic Micro-Database (LIFE-M) is linking millions of vital records to the 1940 Census (Bailey et al. 2016). Entrepreneurial researchers have also combined large datasets to illuminate other questions of interest.

Despite these recent advances, less is known about the quality of these linked data. Although most papers that use linked data report match rates and evaluate the representativeness of their samples, the incidence and characteristics of Type I (false matches) and Type II errors (missed matches) remain unknown. Moreover, the role of linking-induced errors could have important effects on inferences based on these linked samples.

This paper uses new ground truth to evaluate the performance of the most commonly used automated linking algorithms.[2] We quantify performance in terms of (1) match rates, (2) representativeness, (3) Type I (false matches) errors, and (4) Type II errors (unmatched true links) by comparing the links made by Ferrie (1996), the Iterative Method (Abramitzky et al. 2012b), random disambiguation of multiple matches (Nix and Qian 2015), and one regression-based,

[2] "Ground truth" is used in record linkage to mean correctly linked data, typically obtained by direct observation.

supervised machine learning approach (Feigenbaum 2016a) to those in three ground truth samples: the Longitudinal Intergenerational Family Electronic Micro-database's (LIFE-M) sample of Ohio boys linked to the 1940 Census, a synthetic ground truth we create for testing purposes, and a genealogically linked sample of Union Army veterans from the Early Indicators Project (Costa et al. 2017). We also examine how two commonly used phonetic name-cleaning methods, Soundex and NYSIIS, affect performance.

The results of our analysis reveal high match rates for all automated methods, but less encouraging results for other performance measures. First, none of the linked samples (including LIFE-M) appears representative of the population. Second, automated methods yield *distressingly* high rates of incorrect links (Type I errors), ranging from 17 to 61 percent in the LIFE-M ground truth. Careful review suggests a large share of links that are matched by automated methods are *likely links* to the wrong person. Recent implementations of machine learning such as those by Feigenbaum (2016a) have not fixed the problem. Erroneous links also appear to be selected on different baseline characteristics of the sample, potentially suggesting a role for more systematic measurement error. Third, the incidence of Type I errors increases by 60 to *100* percent with the use of phonetic name cleaning, because these techniques clean out *meaningful* variations in spelling used by humans.[3] Randomly choosing among multiple matches Nix and Qian (2015) dramatically worsens the false positive rate. Fourth, Type I errors are positively correlated with Type II errors (missed true links) suggesting that current methods used to increase match rate tend to *decrease* the share of true matches.

[3] An automated matching method would, for example, link Rupert Hermann 65 to Robert Hermann 65 (an exact match using NYSIIS) over Rupert Hermann 66. Human trainers tend to select the latter match, treating the difference between Robert and Rupert as a meaningful variation in spelling but tolerating a one year age difference.

A final exercise demonstrates how the combined effects of Type I and Type II errors ultimately affects inference. Using a common sample of father-son pairs linked to the 1940 Census, we show that estimates of intergenerational income elasticities range from 0.10 to 0.30. The estimates do not vary with Type I error rates as one would expect with classical measurement error, suggesting a role for non-classical error or selection under linking methods. We conclude with constructive suggestions for low-cost ways of improving automated linking methods.

3.2 Advances in Historical Record Linking Methodology

Record linkage has long been a mainstay of historical demography and economic history. But recent developments in computational speed, data availability, and fuzzy linking algorithms have increased the scale and accuracy of record linkage and dramatically expanded what is possible. This section provides a brief overview of record linkage in historical contexts with a focus on the methods evaluated in this paper.

3.2.1 Innovations in Blocking and Population Indexes

The first set of innovations in historical linkage relates to "blocking." In computer science, blocking refers to the partition of a dataset into "blocks" (or clusters of records) using a record attribute (Michelson 2006). Blocking methods limit the number of potential matches according to the attribute, thereby improving computational efficiency while maintaining accuracy. For instance, blocking on place of birth and sex means that a linking algorithm would *only* search for Franklin Jones born in Kentucky within the set of candidate matches of men born in that state.

Historical methods have effectively used "blocking" techniques since their inception. The earliest methods involved identifying a group of individuals in a particular location (e.g., township, county, or state) in one census and manually searching for the same person within the same region (the block) in the subsequent census (Malin 1935, Curti 1959, Bogue 1963, Thernstrom 1964, Guest 1987). This blocking strategy, however, missed individuals who relocated or changed names between census years, generating unrepresentative samples (Ruggles 2006).

The creation of state population indexes[4] facilitated the next innovation in blocking. Improving on previous methods, researchers drew a random sample of households with children at least 10 years old in a historical census. They then searched for the same household in the previous census using the *birth state* of the child to narrow the search (Schaefer 1985, Steckel 1988). This technique, therefore, could find individuals who moved between the census years, but it was similarly limited by geographic mobility between birth and first enumeration in the census. It also restricted the sample of linked households to those with children surviving to age ten.

3.2.2 Innovations in Automated Link Generation

The second major innovation was the use of automated matching, effectively replacing the time-intensive human search with computer queries. Leveraging newly created national population indexes and Public Use Microdata Samples (PUMS), automated matching generated a set of candidate links based on record attributes. Atack et al. (1992) created probabilistic matching software, called PC Matchmaker, to link to census samples in a DOS environment. PC Matchmaker coded names using Soundex or NYSIIS phonetic codes and allowed for user-specified blocking and weighting schemes.[5] Before recent developments in probabilistic matching had become standard, the purpose of phonetic string cleaning programs was to increase match rates by correcting for orthographic differences: minor spelling differences, name Anglicization, and

[4] A state "index" is a list of individuals living in a state at a point in time.
[5] Soundex was developed in the early 20th century to examine census links. It groups similar sounding names like "Smith," "Smyth" and "Smythe" to the same code (in this example, S530). NYSIIS, the acronym for the New York State Identification and Intelligence System, was developed in 1970 as an improvement to the Soundex algorithm. NYSIIS transforms names like "Wilhem" and "William" to WALAN.

transcription errors. This software was then used to create a linked sample between the agricultural and population censuses between 1850 and 1880 (Atack 2004).

Building on this approach, Ferrie (1996) linked men 10 years and older between the 1850 and 1860 U.S. Census samples. To achieve large-scale record linkage in the age before recent advances in computational technology and machine learning, Ferrie (1996) began by selecting a sample of uncommon names from the 1850 Census.[6] The purpose of limiting the sample in this way was to save computation time for the records that would likely *not* be disambiguated later in the process (for example, "John Smith"). Like Atack et al. (1992), Ferrie's method transformed the last name strings using NYSIIS codes, and also truncated the untransformed first name after the fourth letter. Both of these procedures aimed to correct for minor spelling variations, diminutives, etc. Ferrie then linked forward to the 1860 Census, eliminating as potential links all individuals not born in the same state and not living with the same family. To account for misreporting and age heaping, Ferrie did not require an individual's year of birth to be exactly the same between a record and its candidate match but allowed for a ±5 difference in age. If more than two links remained, Ferrie chose the link with the smallest age difference.[7] This method produced a linked sample of 4,938 men (9 percent of the total population of men, 19 percent of the population of men with uncommon names) from 1850 to 1860 and, since, has become the foundation for much of the historical linking literature.

More recently, Abramitzky et al.'s (2012b) iterative method refined the Ferrie method to use a more inclusive set of names. As summarized in their web appendix (Abramitzky et al. 2012a),

---

[6]Ferrie (1996) notes that he only searched for 25,586 of the 55,852 men in the 1860 Census whose surname and first name appeared ten or fewer times in 1850.

[7] Ferrie (1996) does not specify his process for multiple match disambiguation – in his linking from 1850-1860, there were no ties after minimizing the difference in age.

they select a sample of boys ages 3 to 15 with unique names in the 1865 Norwegian Census, standardize first and last names using the NYSIIS, and look for exact, unique matches. If there are multiple exact matches, they do not link the observation. If they do not identify a link in the first step, they search for an exact match within a ±1 year band in the second step and, if they do not find a link, repeat this process up to a ±2 band around the reported birth year. This method's major innovation is to select the potential match for a census record that has the smallest difference in year of birth as opposed to dropping all records with multiple potential matches (regardless of age). This iterative method drops records only if the record has multiple exact candidate matches with identical years of birth or a year of birth falling outside a specified interval. They ultimately link a sample of 2,613 migrants and 17,833 non-migrants of a baseline sample of 71,644 individuals with unique names for a match rate of 29 percent.

The Ferrie and Iterative methods form the basis for record linkage in a number of high profile papers. Abramitzky et al. (2012b) show that the returns to migration were relatively low, and Abramitzky et al. (2013) show that immigrants from Norway to the U.S. were negatively selected based upon parents' wealth. To study migrant assimilation for 16 sending countries, Abramitzky et al. (2014) make use of the iterative method to link men between the ages of 18 and 35 across the 1900, 1910, and 1920 U.S. Censuses, producing a sample of 20,225 immigrant and 1,650 native-born men.[8] In another application of the iterative method, Boustan et al. (2012) produce a sample of men ages 30 to 40 years old linked from the 1920 to the 1930 U.S. Censuses to study migration responses to natural disasters.[9] Hornbeck and Naidu (2014) use the same linked sample restricted to 1920 and 1930 links to study black migration out of the Mississippi Delta after

[8] Abramitzky et al. (2014) achieve match rates across the 1900, 1910, and 1920 Censuses of 12 percent for immigrants and 16 percent for natives.
[9] Boustan et al. (2012) achieves a match rate of 24 percent across the 1920 and 1930 Censuses.

the 1927 Flood. Aizer et al. (2016) use an updated Ferrie method to link male children of applicants to the mother's pension programs between 1911 and 1935 to the Social Security Death Master File to study the effect of cash transfers on later-life outcomes. With a match rate of 52 percent to mortality records, they find that mother's pension programs increased longevity of recipient children by one year relative to children of mothers who were denied benefits.

Nix and Qian (2015) propose an extension of the iterative method to achieve much higher match rates. Similar to Ferrie (1996) and the Iterative Method, they standardize names using a phonetic string cleaning software[10] and use name, birth place and age to generate candidate matches. They also allow a ±3 difference in age. An iterative procedure then generates a sample of potential matches for each individual, which is scored according to the "perfect", Ferrie (but using Soundex), or Jaro-Winkler string distance.[11] If, after applying these different scoring procedures there are ties (multiple matches), their key innovation is to *randomly* select among multiple candidate matches. This approach clearly increases match rates above what is typical in the historical linking literature, which they argue increases their samples sufficiently to measure of racial fluidity.

3.2.3 Innovations in Probabilistic Matching

The Minnesota Population Center (MPC) has pioneered incorporating the latest advances in probabilistic matching into historical record linkage as described in Goeken et al. (2011). MPC standardized surnames minimally but standardized first name strings that occurred at least 100 times (Vick and Huynh 2011). Within race and birth place blocks, MPC uses the Freely Extensible Biomedical Record Linkage (FEBRL) software to score age and name similarity. Next, MPC

[10] They use Phoenex which combines of Soundex and Metaphone methods (Snae 2007).
[11] This is a commonly used statistic capturing the "distance" between two strings.

constructed training data to use with a Support Vector Machine (SVM) classifier. Goeken et al. (2011) writes,

> *…we selected a random sample of potential links, and had a group of MPC data entry operators code each potential link as a "yes" or "no" based on a visual examination of names and ages of potential links (with yes indicating that it was in their opinion a true link). If a majority had the potential link as a "yes," then it was coded as a "yes" in the training data (with the remainder coded as "no")* (p. 3).

Using these training data, Support Vector Machine (SVM) classified all potential links as true or false. A final step eliminates cases with numerous potential links.

This approach creates the Integrated Public Use Microdata Series Linked Representative Samples of the 1880 Census, which contains matches to the 1850 to 1930 Census one-percent samples (Ruggles et al. 2010). Researchers have used these samples in a variety of research questions including the economic effects of racial fluidity (Saperstein and Gullickson 2013), long-term differences in black and white women's labor-force participation (Boustan and Collins 2014) and intergenerational co-residency (Ruggles 2011).

Other researchers have also incorporated probabilistic and machine learning into their record linkage. Mill and Stein (2016) and Mill (2013) describe and use a method that also employs string comparators and scoring of matches, using an Expectation Maximization (EM) algorithm and maximum likelihood estimation to determine the probability of a true match. Similar to the MPC, Antonie et al. (2014) and Wisselgren et al. (2014) use a text-string comparator and estimate probability scores using truth data and Vector Machine Learning.[12]

---

[12] Antonie et al. (2014) describe the linkage system they developed to link across historical Canadian census data and their application of this system to linkage of men from the 1871 Canadian Census to the 1881 Canadian Census. Their linkage rates range from 17.5 percent (Quebec) to 25.5 percent (New Brunswick). Wisselgren et al. (2014) use an approach similar to Antonie et al. (2014) and the MPC to link the 1890 and 1900 Swedish Censuses. Depending on their treatment of names (standardized names versus constructed name variables) they achieve match rates ranging from 18 percent to 70 percent for men and 24 percent to 66 percent for women (Wisselgren et al. 2014, pp. 148). They use Parish records as "truth" data to assess Type I error in their linkages, of which they find very little (less than 3 percent). Preliminary work by (Eriksson 2016) shows the error rate of the Swedish linked data increases by as much as 24 percent if linked using county of birth – a significantly more aggregated geography – than parish of birth, which contributes to the significantly lower error rates found by Wisselgren et al.

Another use of probabilistic matching makes use of Ancestry.com's search algorithm to link records. Bailey et al. (2011) link records of lynching to the 1900 to 1930 U.S. Censuses to determine which community characteristics were associated with lynching. Collins and Wanamaker (2015) enter information on name, age, and place of birth for black males ages 0 to 40 resident in Southern states in the 1910 Census into Ancestry to search for individuals who uniquely match these criteria. They match 19.4 percent (5,465/28,215) of individuals to a unique person in 1930.

Bleakley and Ferrie (2016) link individuals eligible for Georgia's Cherokee Land Lottery of 1832 (and their sons) to examine the long-run effects of income on the human capital and economic outcomes of their children. Using a variant of Ferrie's (1996) method, they link sons from 1850 using name, year of birth, birth place, and parents' birth places to the 1880 full-count Census. In addition, they used Ancestry's search algorithm to search for individuals who were not linked. For part of their analysis, they weighted observations by 1/n where n is the number of candidate matches to make use of multiple matches. Their combination of methods produces a match rate of 46 percent from the baseline sample of 32,738 men in 1850.

Recent unpublished papers continue to extend this frontier. Eli et al. (2016) use probabilistic matching to link military records from the War and Treasury Departments to study the effect of the Civil War on migration decisions for those living on the border of Union and Confederate states. Eli et al. limit the age window to ±3 years and also discard candidate matches with a Jaro-Winkler scores less than 0.9.

Most recently, Feigenbaum (2016b) extends the MPC methods by introducing a supervised, regression machine learning approach. Using single clerically reviewed data, he uses non-linear regression models to predict the probability of a match using observed features of the linked and unlinked data. These features include name similarity scores, differences in age, indicators for Soundex matches of first and last name, and whether truncated parts of the first and last name match. After calibrating his model using tuning parameters, he uses it to predict matches for a large number of men in the 1915 Iowa Census linked to the 1940 Census.[13]

In summary, existing methods involve complex procedures that may improve match rates and accuracy. Comparisons of methods are further complicated by the fact studies link different data (censuses, vital records, and administrative records) in different countries or geographic areas for different periods and samples. Many variations on linking methods across studies suggest the need for a more systematic analysis of these methods. Which method or combination of methods should researchers use in different historical contexts? What are the trade-offs when using different automated linking methods and phonetic name cleaning algorithms? The next sections address these questions using a common sample.

3.3 Metrics of Automated Method Performance

This paper uses new ground truth samples to evaluate the performance of commonly used linking methods. The reviewed methods are Ferrie (1996); the Iterative Method (Abramitzky et al. 2012a); random tie breaking (Nix and Qian 2015); and supervised, regression-based machine learning (Feigenbaum 2016b). Detailed web appendices, published articles, and posted code make replicating these methods straightforward. Ferrie (1996) describes the construction of his method in a step by step basis, which we implemented exactly as described; we also examine its robustness

[13] Computer scientists have also taken a recent interest in historical record linkage. Using each decennial Lancashire census from 1841 to 1901, Fu et al. (2014a) develop a model for an automated pairwise linking process that uses both a supervised SVM method and a similarity threshold model. To compare first and last name strings they employ Q-grams and Jaccard measures of string distance. Their match rates range from 44 to 48 percent depending on the two censuses used in the linkage (Fu et al. 2014, pp. 219). Building upon this work, Fu et al. (2014b) use graph-based methods to link households across U.K. census data. After data cleaning, the first step of their linkage process involves the scoring of pairwise links and the use of logistic regression to score record pair similarity, which is used in the selection of record pairs. They then construct a graph for each household and perform graph matching. Vertex matching and graph similarity calculations are used to select the final matched household pairs.

to inclusion of more common names.[14] Abramitzky et al. (2012a) posted their Stata code at the *American Economic Review*, which we use directly.[15] We implement Nix and Qian's (2015) method of random match disambiguation by randomly selecting one link in the 1940 Census from the set of (multiple) perfect ties. This method is implemented for the sample of Ferrie's (1996) matches using the sample of more common names.[16]

We evaluate the performance of each algorithm for *the same base sample* of Ohio born boys to the 1940 full-count U.S. Census. This eliminates the possibility that differences across datasets and periods affect this paper's comparisons across methods: only the differences in the design of the linking algorithm should influence the results. The links of each automated linking algorithm and LIFE-M ground truth are evaluated in the following four dimensions:

(1) Match rate: The match rate is calculated as the share of records that were successfully matched from the original sample. This dimension of match quality is reported in almost all linking papers.

(2) Representativeness: We compare the characteristics of the linked samples to the characteristics of the original sample (including non-links). In terms of characteristics on the birth records, these comparisons include the Kolmogorov-Smirnov tests for the equality of exact day of birth distributions and t-tests of the differences in the share of fathers and mothers who are foreign born, the number of siblings in the reconstituted family, the number of characters in the infant,

[14] Four other modest deviations reflect the fact that we do not limit links based on family continuity or limit our baseline sample on the basis of age (because this does not make sense to do with birth certificates). In addition, we treat records with multiple matches after the last step as having no link, although Ferrie reports having none of these instances and, therefore, does not indicate how he would have dealt with them. We limit our analysis of the 1940 Census to men born in Ohio, obviating the need for a restriction on birthplace.

[15] The code was retrieved in February 1, 2016, from this link: https://www.aeaweb.org/aer/data/aug2012/20100051_data.zip. We have made one major change this code. If implemented as written, the code deletes the best link if there are multiples and matches to the second best link. This results in a substantially higher Type I error rates (documented in earlier drafts of this paper). The results we present here alter the posted code to match from the birth certificate data *forward* to the 1940 census, which eliminates this problem. Our revised code is available upon request.

[16] As noted in Ferrie (1996), the uncommon name sample has almost no multiple matches to disambiguate.

mother and father's names, and the share of records with at least one misspelling or transcription error.

Although these first two measures are the most common statistics reported for linked samples, they are not sufficient statistics for link quality. This is easily illustrated in an example. Consider a matching algorithm that *randomly* links individuals between two datasets. This algorithm would perform very well in terms of match rates (criteria 1, 100 percent!) and representativeness (criteria 2, perfect performance!), because the entire matched sample would be matched and identical to the baseline sample in observed characteristics. Few researchers, however, would want to work with these data, because the incidence of false positives would approach 100 percent in large samples. If false links are random with respect the relationship of interest, they would tend to attenuate it. A systematic relationship between the false links and the independent variable of interest could have an unknown effect on an estimated relationship.

Two more criteria are, therefore, particularly important for assessing the quality of statistical inferences (Abowd and Vilhuber 2005, Kim and Chambers 2012):

(3) <u>False positives (Type I errors):</u> We compare links obtained from each automated method to the LIFE-M ground truth. If the method's link and ground truth link agree, we treat the link as correct. In the case of disagreements between the linking algorithm and the ground truth, we *re*-review these links by staging the equivalent of a "police line-up." In the line-up, reviewers see the LIFE-M link, the automated link, and a number of similar potential links arising from our probabilistic match procedure. This *re*-review gives each method a fair shot at having its link chosen; it also allows trainers the opportunity to identify potential errors in the ground truth. We report the share of links for each method that human trainers ultimately code as erroneous matches. The false positive rate for each method is defined as the fraction of links

deemed erroneous by trainers in the police line-up. As an additional check on the quality of the reviewers, we also examine how well automated linking methods perform against an objective ground truth sample of synthetic data.

(4) False negatives (Type II errors): this metric is intended to capture the fraction of “true” links that are not found by each method. We construct this as (1 – Match Rate*(1 – Type 1 Error Rate)). The Type I errors are those computed after the police line-up described in (3) above.

3.4 A Comparison of Links by Automated Methods and LIFE-M

The primary challenge to evaluating historical linking methods is that there is very little “ground truth” data against which to evaluate the quality of different linked samples. Because a central focus of a growing historical literature is linking to the newly available 1940 census, we begin our analysis by examining the performance of automated methods using LIFE-M sample of 19,090 Ohio boys linked to the 1940 census.

3.4.1 The LIFE-M Sample of Ohio Boys Linked to the 1940 Census

LIFE-M’s linked sample of Ohio boys is created by matching birth certificates to the 1940 full-count U.S. Census. The process is semi-automated, making use of both computer programming and human input. After cleaning and standardizing the data, we use a bi-gram matching procedure to generate a list of candidate matches based upon name similarity and age for men born in Ohio.

From this list of candidate links, LIFE-M creates links using an independent, blinded human review process. “Data trainers” first participate in a rigorous orientation process. During this period, they receive detailed feedback on their accuracy relative to an answer key. They continue this process (this takes 10 to 20 hours of data training) until their matches agree with the truth dataset 95 percent of the time.

After completing this orientation, trainers become part of a team that conducts independent, blinded clerical review. Each potential match is reviewed by two trainers who choose from a set of candidate matches generated using a probabilistic, bigram match on name, date of birth (or age), and birth state (Wasi 2014). In the cases where the two initial reviewers disagree, the records are *re*-reviewed by an additional three individuals to resolve these discrepancies. Our automated system randomly assigns batches among the 15 to 30 trainers who are employed at any time, so it is difficult for trainers to coordinate with peers. Any discrepancies between the two trainers results in additional reviews by three other trainers who also make independent determinations about whether the candidate records are true links. Random "audit batches" provide feedback to trainers about the accuracy of their decisions, and weekly meetings encourage discussions of difficult cases to help trainers achieve consistent and accurate matches. The result of this process is a highly vetted, hand-matched ground truth dataset for a random sample of 19,090 Ohio birth certificates—boys born between 1881 and 1940—linked to the 1940 Census (Bailey et al. 2016).

To validate the quality of the LIFE-M ground truth, the Family History and Technology Lab at Brigham Young University (BYU) employed research assistants to use genealogical methods to hand link a sample of 543 boys, 225 of which had been linked by LIFE-M. They used multiple sources of genealogical data (only a subset of which were used in the LIFE-M linking) to create correct record linkages and complete family trees. Although genealogical linking is cost (and time) prohibitive for larger projects, the advantage of the genealogical approach is that it produces a very low rate of false links. The BYU team had no knowledge of LIFE-M's links while doing this exercise, so their work can be viewed as independent and blinded.

BYU's genealogical method ultimately linked 392 of 543 boys for a match rate of 72 percent. This was 151 more links than the LIFE-M team found and can be attributed to the use of multiple data sources and more intensive searching to distinguish between seemingly similar matches. However, for the 225 links found by LIFE-M, the BYU team agreed with these matches 96 percent of the time. Only 16 of LIFE-M's links differed from those found by the genealogical method. Taking the genealogical method as the gold standard, this implies that LIFE-M's false positive rate would be 4 percent, which is substantively identical to rates calculated in section IV.D. of this paper.

3.4.2 Match Rates

Figure .1 describes the match rates for the LIFE-M ground truth and for each automated algorithm. These match rates represent the share of the baseline sample of boys born in Ohio from 1881 to 1940 who were successfully matched to the 1940 complete count Census. LIFE-M's ground truth matched 51 percent of the baseline sample. Ferrie's (1996) method matches between 19 and 38 percent of baseline sample depending on the use of phonetic name cleaning. For the Ferrie method, match rates fall as phonetic name cleaning creates more potential matches than permitted using the uncommon name sample restriction.[17] As expected, relaxing the uncommon name sample restriction (labeled "Ferrie 1996 – all potentials") results in higher match rates, ranging from 42 to 50 percent. Due to great similarities with Ferrie's method, Abramitzky et al.'s (2012a) iterative approach yields comparable match rates, ranging from 43 to 47 percent, depending upon the phonetic name cleaning procedure. As intended, the Nix and Qian (2015) method of randomly choosing a match for records with multiple matches effectively increases the match rate to 60 to 76 percent. This is very close to their reported match rate of 61 to 67 percent across the 1880-1940 Censuses (Nix and Qian 2015). These match rates are higher than the Ferrie

[17] This restricts the set records to be linked to names that have fewer than 10 potential links. Because the match relies on truncated first name and either last name or a phonetic code for last name and variation is lost using phonetic codes, the loss of variation in last name results in fewer matches when using phonetic codes.

method (with more common names) and highlight the importance of record disambiguation in historical linkage: methods that help choose among multiple matches can significantly improve match rates! Finally, Feigenbaum's (2016a) regression-based machine learning method matches approximately 26 percent of the baseline sample.

These results highlight how the use of phonetic name cleaning software can increase or decrease match rates, depending upon how it interacts with record disambiguation. While phonetic cleaning can correct for orthographic errors and increase the match rate, it may also removes meaningful spelling variation from names, making match disambiguation more difficult and decreasing links. This interaction is important for the Ferrie (1996) method. This method's uncommon name restriction means that the sample that Ferrie tries to link (i.e., that has fewer than 10 potential matches in the 1940 Census regardless of age) falls from 50, to 42, to 27 percent with the use of NYSIIS and Soundex, respectively. However, because the Nix and Qian (2015) method does not require unique matches, the use of phonetic name cleaning unambiguously *increases* match rates.

Also noteworthy is that our match rates are higher than published elsewhere. For instance, the Ferrie (1996) method matches between 19 and 38 percent of our sample versus his published figure of 9 percent of all men between 1850 and 1860. Similarly, the Abramitzky et al. (2012b) algorithm links 43 to 47 percent of our sample, whereas it links only 29 percent of Norwegian men in their published work. Higher match rates may occur here for several reasons. First, the LIFE-M sample has better information on the complete and correct name. This is the case because formal birth name—often including middle names—was recorded on the birth certificates was the exact date of birth. The presence of a formal name and birth date removes nicknames and measurement error in age from the primary record, which improves match rates for all methods.[18] Second, the LIFE-M data link the Ohio boys as men in the 1940 Census at fairly young ages when mortality and outmigration may have been lower than in other contexts. In summary, our context is conducive to higher match rates and, presumably, higher quality matches due to the higher quality of name and age reporting as well as a greater likelihood for the right link to be in the linked dataset.

3.4.3 Representativeness of the Linked Samples

We next compare the representativeness of the linked records. Because birth certificates do not contain socio-demographic measures found in the census (race, age, or incomes of the parents), we make use of alternative features of these data. First, we test whether the distribution of exact day of birth (1-366, due to leap year) in the baseline sample and in the linked subsamples are drawn from the same probability distribution using the Kolmogorov-Smirnov test. Exact day of birth is ideal because it is as close to a continuous measure as we can get in historical records, and season of birth is strongly correlated with socio-economic in modern data (Buckles and Hungerman 2013).

Figure 3.3 shows the p-values for each linking method for the null hypothesis that the distribution of day of birth is identical in a given linked sample relative to the baseline sample of males born in Ohio. Two linking methods appear to fail this distribution test at conventional levels of significance: Nix and Qian (2015)'s method using full names, and Nix and Qian (2015)'s method using Soundex. For the other linking methods, the Kolmogorov-Smirnov tests fail to reject equality at the 10 percent level.

[18] Massey (2017) shows that decreasing the noise in age results in higher match rates and lower Type I error rates

A second test examines differences in a handful of individual characteristics from the baseline records as well as features of the records themselves for the linked and unlinked portion of the baseline sample across linking methods. For the LIFE-M data, these features include the day of birth, share of mothers and fathers who were foreign born (parents' place of birth is reported on the birth certificate), and the number of siblings in the family. In addition, we examine the number of characters in the infants' (boys'), mothers', and fathers' names—a characteristic which is strongly positively correlated with years of schooling and income from wages in the 1940 Census. We also examine the share of family records in the family with a misspelled mother or father's name, which we expect to be negatively correlated with years of schooling and income (Aizer et al. 2016).[19]

The results for the LIFE-M sample are presented in Table 3.1. Each point estimate in the table is from a separate regression of the variable indicated in the column header on a binary variable for whether the linking method matched the observation. Standard errors are reported in parentheses below each estimate.

We find that *none* of the linking methods creates perfectly representative samples along the observable characteristics. The LIFE-M ground truth sample is less likely to link boys with foreign-born parents is more successful in linking records with more characters in the infant's, mother's, and father's names, and is somewhat less likely to link records in which the father's name is misspelled. All methods are more likely to link children with longer names, indicating that they tend to represent more affluent families, and in general, methods are more likely to code links for individuals with native-born parents (except for the Ferrie method, which is more likely to link

[19] We are able to calculate this by leveraging the unique structure of LIFE-M records: we have multiple observation of each parent's name (one for each sibling in the family).

children of foreign-born parents). Interestingly, Feigenbaum (2016a) is both more likely to link foreign-born parents and more likely to link longer names (positively correlated with education and income) and less likely to link boys in families where a father's name is misspelled on a record (negatively correlated with education and income).

In summary, these results provide little evidence that LIFE-M's clerical review or the automated linking methods provide representative samples of the population. This is consistent with findings in multiple papers (Abramitzky et al. 2012b, 2014, Collins and Wanamaker 2015). This could imply limited external validity of results using these samples, especially because the linked samples tend to include native born and more educated individuals.

3.4.4 False Positives (Type I Errors)

A third indicator of performance considers the incidence of Type I errors using different linking algorithms. Our first analysis compares the links of each automated algorithm to the LIFE-M clerically reviewed sample. Having "the same link" means that a particular individual in the Ohio birth certificate data is linked to the *same* individual in the 1940 Census using the automated method and in the LIFE-M ground truth. These agreements are treated as correct. However, discordant links are reviewed by two additional trainers in a "police line-up" process, where the ground truth and algorithm links are presented alongside close candidate matches. Without knowing which is the LIFE-M link or the link chosen by the automated method, trainers select the best match from the group using name, age, and place of birth information. This gives each method as well as LIFE-M an equal chance at being chosen.

Figure 3.2.2 presents the results of this process. In 2 percent of original LIFE-M matches, the trainers *reverse* the original LIFE-M linking decisions. This could happen because LIFE-M's choices to block eliminated a better link and because two trainers occasionally make *the* same mistake. Consistent with the BYU validation, these reversals are very rare.

In contrast, the incidence of false positives for the automated methods are distressingly high, ranging from 17 percent to 61 percent. In his 1996 paper, Ferrie used NYSIIS, which is associated with a 22 percent false positive rate in our sample.[20] The share of false positives using Abramitzky et al.'s (2012b) iterative method and Nix and Qian's (2015) random choice among multiples do not fare better. At least 19 percent of the links generated by Abramitzky et al.'s (2012b) method are false positives, increasing to 29 percent with the use of NYSIIS (as they used in their papers) and 40 percent with the use of Soundex.[21] In the case of Nix and Qian (2015), Type I error range from 35 to 61 percent. This high rate of type I error is not surprising because—by construction—this method randomly selects a link from sets of potential matches with *identical* names and ages, only one of which is correct. Although this method ensures a high match rate, this random disambiguation is unlikely to identify the *correct* match in the majority of cases. Feigenbaum's (2016b) supervised, regression-based machine learning model produces one of the lowest Type I error rates of 24 percent. After 20 years of research on historical linking, Ferrie's (1996) method still achieves the lowest Type I error rate.

Another important finding is that the incidence of Type I error universally increases with the use of phonetic name cleaning in our sample. Across methods, using Soundex *increases Type I error rates by roughly 60 to 100 percent* above those observed using an uncleaned name. This is the result of the phonetic codes removing *meaningful* variation from names, as well as orthographic differences. This name cleaning interacts with requirements in most automated methods that names match *exactly*, while allowing for differences in other characteristics like age. For instance, automated methods requiring an exact match on name would link William Alvin

[20] This error rate consistent with Massey (2017) who uses contemporary administrative data linked by Social Security Number as the ground truth. She finds that Ferrie and Long's method (2013) (which use Soundex) , which is associated with a 30 percent false positive rate in this sample.
[21] Using the uncorrected AER code results in a 29 to 58 percent Type I error rate, as reported in previous drafts of this paper.

Gibbons 45 to William A. Gibbons 47 rather than William Alvin Gbbons 45, whereas humans identify and adjust for the likely transcription error. Similarly, human reviewers are adept at recognizing nicknames and diminutives, whereas automated algorithms requiring exact name matches tend to link Margaret Alva Billingsworth 31 to Margaret Billingsworth 33 over Peggie Alva Billingsworth 31 (human reviewers tend to choose the latter).

A second analysis of these erroneous links examines their relationship to baseline sample and record characteristics. Table 3.2 repeats the representativeness tests presented in Table 3.1, this time coding only a method's *false* links as ones in the analysis and the remaining links as zeros. This reveals whether the false links are random errors (in which case they should exhibit no relationship with given characteristics) or systematically differ in terms of their relationships with baseline record features. The false positives in LIFE-M appear to be very similar to the rest of the core sample in terms of their baseline characteristics, with only two characteristics statistically different for LIFE-M (instance of misspelled mother's and father's name), which were also small in magnitude. On the other hand, the false positives for other methods appear systematically correlated with a variety of baseline characteristics, suggesting that these linking methods may introduce more systematic measurement error into analyses using these datasets.

3.4.5 False Negatives (Type II Errors)

A fourth indicator of performance is the incidence of Type II errors—records that are not made or made incorrectly. Across all methods, the horizontal axis in Figure 3.2.3 shows that Type II error rates are high, ranging from a low of 50 for LIFE-M to 88 percent for Ferrie (1996)-Soundex (with the original uncommon name restriction). Mortality and under-enumeration likely accounts for around 15 percent of missed links:[22] a fraction of the individuals in the birth records

[22] Based on life tables from 1939 to 1941, we calculate that 8.27 percent of our sample should be un-linkable due to death prior to

will not have survived until the 1940 Census, meaning that there is no "true" link for this individual. However, differences in mortality and under-enumeration are not likely driving the differences between methods in these indicators—all automated methods and LIFE-M should be affected equally.

Figure 3.2.3 also plots the relationship between Type I and Type II errors. In this plot, the best linking method would locate at the origin with a zero Type I and Type II error rate. This would mean that every individual was linked and linked correctly. One typically thinks about a trade-off between Type I and Type II errors, with Type II errors being higher with lower Type I errors and vice versa. Surprisingly, the plot exhibits a positive relationship, with Type II error rates tending to be higher where Type I error rates are higher. This suggests that existing methods that increase match rates have tended to worsen Type I errors and *increase* Type II errors! The (much more expensive) human reviewed data, LIFE-M, is attains both a lower Type I and Type II error rate relative to automated methods. Ferrie (1996) unrestricted to the uncommon name sample and Abramitzky et al. (2012b) without phonetic name cleaning are the second best in terms of proximity to the origin.

In summary, this analysis suggests that a large share of links used for inference in historical settings are likely erroneous. In addition to these samples not being representative, these findings imply that measurement error may play a potentially large role in the findings of studies using linked samples.

1940 (National Office of Vital Statistics 1948). Moreover, census analyses estimate that around 5.4 percent of individuals were missed in 1940 (West and Robinson 1999).

3.5 How Well Automated Methods Perform in Alternative Ground Truth Samples

Given high variability in record quality over time, differences in enumerated versus self-reported surveys, and the care with which microfilmed data are transcribed, the results from the LIFE-M sample may not generalize. Moreover, an obvious critique of the "police-line up" exercise in section IV is that the LIFE-M project trained the human reviewers. And, as a consequence, those reviewers make decisions more likely to favor the LIFE-M ground truth. To address both concerns, we use two alternative ground truth samples. First, we *simulate* a ground truth dataset, so that we know objectively the true link. Therefore, our objective truth should not be influenced by human reviewers at all. This exercise presents an internal validity check of our analysis. Second, we use the Oldest Old Sample from the Early Indicators Project that was linked by genealogists and is known to be highly accurate. This section presents the results of each of these analyses using our four performance criteria.

3.5.1 The Synthetic Ground Truth

One of the issues in historical linking is that the sample to be linked is often much smaller than the dataset of potentials (we will call this the "using" dataset for short). We construct our using dataset in two steps. First, we take all of our Ohio born boys, randomly drop 10 percent to reflect mortality and emigration and drop another 5 percent to reflect under-enumeration.[23] There is also considerable scope for orthographic error. The original Census enumerators may have misspelled respondent names, the household respondent may have reported incorrectly, or the individual may have changed their name (perhaps using a middle name or diminutive in place of the given name). Compounding this, digitization may have mis-keyed handwriting from original

[23] See footnote 22.

Census forms.[24] Consequently, we also add noise to names and ages to reflect age heaping and transcription error.[25]

Second, to mimic the size and scale of the 1940 census, we append an additional sample of boys born in Michigan and Indiana (states that neighbor Ohio) from the 1940 Census. We choose Michigan and Indiana to allow for regional similarity in names but preclude the possibility that these data contain any true links (our boys were all born in Ohio and not Michigan or Indiana). We then limit the number of these additional individuals so that our using dataset matches the size of the sample of all Ohio born boys of the relevant ages in the 1940 Census (3,133,982 individuals). We can then examine the performance of automated methods relative to a set of objective links.

Figure 3.4.1 presents our match rates. These simulation results closely match the results presented above: Ferrie's (1996) method links 45 percent of the master sample without phonetic name cleaning, and links 36 percent and 23 percent using NYSIIS and Soundex, respectively (compared to 38, 30, and 19 percent of the LIFE-M ground truth, respectively). Again, Abramitzky et al.'s (2012b) method yields higher match rates, ranging from 49 to 54 percent. And, as before, the Nix and Qian (2015) method of random match disambiguation results in very high match rates: 66 to 81 percent. By design, the "true" match rate is 85 percent, because 15 percent of the original

---

[24] Our comparison of two independently transcribed versions of the 1940 Census suggest that *25 percent* of names are not identical. Moreover, in a recent paper, Goeken et al. (2016) document that in two enumerations of residents of St. Louis in the 1880 Census, nearly 46 percent of records linked between the two enumerations are not perfect matches on first name.

[25] Age-heaping reflects the fact that respondents tend to round their age to the nearest multiple of 5 and has been used as a metric of historical innumeracy (A'Hearn et al. 2009)). To mimic this, we round 25 percent of observed ages to the closest multiple of 5. We introduce orthographic problems as follows: In 10 percent of cases, the first and middle name are transposed (if a middle name exists). In 5 percent of cases, the first and last name are transposed. In 3 percent of cases, the first character of the first name is randomly changed. In 3 percent of cases, the first character of the last name is randomly changed; in 4 percent of cases, the second character of the first name is randomly changed; in 4 percent of cases, the second character of the last name is randomly changed; in 4 percent of observations, the third character of the first name is randomly changed; in 4 percent of observations, the third character of the last name is randomly changed. In 5 percent of cases, the first names have a letter randomly repeated (for example, James -> Jamees). In 5 percent of last names, a letter is randomly repeated in the last name. In 5 percent of cases, a random letter is dropped from the first name (for example, Matthew -> Mathew). In another 5 percent of cases, a random letter is dropped from the last name. In 5 percent of first names, a random pair of adjacent letters are transposed (for example, William -> Willaim). And, in 5 percent of last names, a random pair of adjacent letters are transposed.

links are absent in the perturbed data. In short, these findings corroborate the findings made by humans in section IV.

Figure 3.4.2 presents rates of Type I error for the automated linking algorithms on the simulated data. False links are defined as those where the “true” link in the using dataset differs from the link identified by the automated method. The incidence of false matches is again high, and ranges from 18 to 58 percent over the matching methods. Ferrie’s (1996) method again performs well, with Type I error ranging from 21 to 27 percent. As above, error rates are highest for the Nix and Qian (2015) method, ranging from 38 to 58 percent. Feigenbaum (2016)’s method performs also well, with an 18 percent Type I error rate. As with the LIFE-M ground truth, Type I error rates increase universally with phonetic name cleaning.

Figure 3.4.3 shows the relationship between Type I and Type II error from each method using the synthetic truth data. Again, Type I and Type II errors exhibit a strong, *positive* relationship, suggesting that current methods aimed at increasing match rates do not reduce Type II errors. Methods with higher match rates tend to have both higher Type I and Type II errors.

In short, these results support the internal validity of the LIFE-M results. For a very similar sample (including errors in name and age, mortality, emigration, under-enumeration, and transcription errors), we find very similar patterns in method performance. We are able to replicate the approximate levels and ordering of match rates, Type I, and Type II errors as well as problems with phonetic name cleaning. The next section seeks to understand how well these results generalize to other samples and eras.

3.5.2 Early Indicators Oldest Old Sample of Union Army Veterans

The Oldest Old sample of Union Army veterans from the Early Indicators project provides another ground truth sample. Using genealogical methods and a rich set of supplementary information, Costa et al. (2017) created this oversample of 2,096 individuals at least 95 years old linked to the 1900 complete count Census. These veterans tended to report very complete and accurate information to ensure they would receive their army pensions and benefits. Moreover, sources such as gravestone databases, obituaries, newspaper accounts, veterans associations and pension files allow multiple cross-validation exercises to create an extremely high match rate of 90 percent. This exercise treats matches coded as the highest quality (quality 1 and 2) as another ground truth sample.[26]

Figure 3.5.1 describes the match rates of the oldest, old Union Army veterans and the 1900 complete count census. Ferrie's (1996) method matches a larger share of the Union Army veterans sample, ranging from 37 and 43 percent versus 19 to 38 percent in the LIFE-M data. Abramitzky et al.'s (2012b) iterative method also yields slightly higher match rates in this sample. As with both the LIFE-M ground truth and synthetic ground truth samples, Nix and Qian's (2015) method of randomly choosing a match among multiple matches substantially increases the match rates for the unrestricted sample: here, they range from 53 to 64 percent versus 60 to 76 percent in LIFE-M. Finally, Feigenbaum's (2016a) supervised, regression-based machine learning method matches a considerably higher percent of the surviving veterans, 43 percent, versus 26 percent in the LIFE-M data. These higher match rates may reflect the fact that we are linking a sample that has already been linked which, by construction, excludes records that were unlinkable for the Early Indicators team. Another interesting feature of the data is that the use of phonetic cleaning appears to have less of an effect on match rates than in the LIFE-M sample, perhaps reflecting differences in the representation of names over time.[27]

---

[26] The Early Indicators project scores matches on a scale of 1 to 4 to indicate their confidence in a match. Quality indicators equal to 1 or 2 indicate their most confident matches.

[27] The different effects of NYSIIS and Soundex could reflect differences in the quality of transcribed names in the baseline samples, differences in the types of names and spelling variations over time, or other differences in the socio-demographic composition of the samples (i.e., differences in country of origin of parents, socio-economic characteristics of individuals represented, etc.).

Table 3.3 presents an analysis of the representativeness of the linked samples by testing for balance in observable factors between the linked and unlinked samples for each method. Data availability requires that we use a set of different covariates than in the LIFE-M sample: age, whether the individual speaks English, owns a farm or house, is married, and is foreign born. These results show that all of the linked individuals are more likely to speak English across all methods, and are more likely to be married across most methods. In addition, the Abramitzky et al. (2012b)-name and the Feigenbaum links appear to be more affluent and older. In short, we find little support for the hypothesis that the linked samples are representative of the underlying sample of oldest, old veterans.

Type 1 error rates for the different methods are captured in Figure 3.5.2, which are calculated as the share of observations linked by a given method that differ from the Early Indicators links. (Because of the extensive genealogical work done to create this sample, we assume that links from the Union Army veterans to the 1900 Census are correct). As with the results from the LIFE-M ground truth, the incidence of false positives for the automated methods is high, ranging from 14 percent to 40 percent. The performance of Ferrie's (1996) method is comparable to in the LIFE-M ground truth, with false positive rates ranging from 20 (Name) to 23 percent (Soundex). The false positives using Abramitzky et al.'s (2012b) iterative method are also slightly better, ranging from 21 to 29 percent versus 19 to 40 percent in LIFE-M. As before, Nix and Qian's (2015) method appears to generate very high Type I error for high match rates, ranging from26 to 40 percent—high, although slightly better than in the LIFE-M sample. Feigenbaum's (2016a) supervised, regression-based machine learning model produces the lowest Type I error rate for the sample of 14 percent versus 24 percent in the LIFE-M sample.

As a summary of these findings, figure 3.5.3 plots Type I error rates against Type II error rates by method. Type II errors – which can be read off the x-axis - range for all but the ground truth range from 62 to 73 percent—rates that are relatively close to those from the LIFE-M and synthetic ground truth samples. Because differential mortality is not an issue (all individuals in the Oldest Old Union Army sample survived to be at least 95 years old), these numbers imply that roughly two out of every three possible links is not captured in automated methods versus 1 in 10 using genealogical methods. Moreover, as in the LIFE-M and synthetic ground truth, we document a strong positive relationship between Type I and Type II errors. Overall, differences in the types of records and the period affect linking and error rates in the LIFE-M and Union Army ground truths, but both samples reveal similar patterns in match rates, representativeness, and Type I errors. This analysis also supports the conclusion that phonetic name cleaning universally increases Type I errors.

In summary, this section shows that different samples for different periods in history and in different contexts provide very similar results. Across the board, automated methods tend to produce higher than desired Type I error rates as well as very high Type II error rates. Phonetic name cleaning almost universally increases Type I error rates while also increasing Type II error rates.

3.6 How Automated Methods Affect Inferences

Following the literature on intergenerational mobility (Solon 1999, Black and Devereux 2011), this analysis considers the following benchmark specification,

$$\log(y_1) = \pi \log(y_0) + \varepsilon \qquad (1)$$

where this equation refers to deviations from population means. The dependent variable, $\log(y_1)$, refers to the log of son's income in adulthood in the 1940 Census. The key independent variable, log(y_0), refers to the parent's log income in the 1940 Census. Within this framework, π is interpreted as the intergenerational earnings elasticity. A higher π indicates that parents' incomes play a stronger role in determining their children's income. Intergenerational mobility is often measured as 1- π.

3.6.1 How Type I and Type II Errors May Affect Inferences

Within this framework, it is well known that mean zero measurement error in son's income (the dependent variable in the regression) will still allow us to estimate π consistently using OLS, though the estimates will be less precise. We, therefore, focus on three other types of measurement error in father's income due to linking that occurs for some portion of the sample, α^{m}, where the Type I error rate varies by linking method, *m*.

At first consideration, linking to the wrong son may seem like measurement error in the dependent variable (which we note should only reduce precision). However, linking a boy on a birth certificate to the wrong adult in the 1940 census means that the son's observed income is, in fact, linked to the wrong father—the cases for the following three variants of Type I errors.

First, classical measurement error would occur when a father's income is measured with error, or z_1= log(y_0)+u, where u is an additive error term drawn from N(0,1). This functional form assumes that z_0 is correlated with log(y_0), but that the father's observed income is a noisy estimate of a father's actual income.[28] With classical measurement error, the estimand of interest converges in probability to the following quantity,

$$\text{plim}\ \widehat{\pi^{m}} = \alpha^{m}\left(\frac{\sigma_{y0}^{2}}{\sigma_{y0}^{2}+\sigma_{u}^{2}}\right)\pi + (1-\alpha^{m})\pi. \tag{2}$$

[28] This correlation could occur if fathers are linked within the same birth state, race, and class of names.

Second, we consider measurement error where $z_2 = \varepsilon$, where $\varepsilon \sim N(0,1)$. In this case, $z_2 \perp \log(y_0)$ and captures the instance where the linked observation for father's income is independent of his actual income. That is, an erroneous link is a random draw from the father's income distribution. In this case, the estimand of interest converges in probability to the following quantity,

$$\text{plim}\ \widehat{\pi^m} = (1 - \alpha^m)\pi. \tag{3}$$

Third, we assume a special case of non-classical measurement error with $z_2 = \log(y_0) + \mu$, where $\sigma_{y0,\mu} = \text{cov}(\log(y_0), \mu) > 0$. In this case, the estimand of interest converges in probability to

$$\text{plim}\ \widehat{\pi^m} = \alpha^m \left(\frac{\sigma_{y0}^2 + \sigma_{y0,\mu}}{\sigma_{y0}^2 + \sigma_\mu^2 + 2\sigma_{y0,\mu}}\right)\pi + (1 - \alpha^m)\pi. \tag{4}$$

We use these three formulae as benchmarks against which to compare our actual regression estimates.

The role of Type II error enters through the selective representation of different groups in the linked sample. For instance, this paper finds that being linked is positively associated with the length of name and, by extension, income (Olivetti and Paserman 2015). This differential representation matters to the extent that different groups have heterogeneous intergenerational income elasticities. For instance, they may differ between whites and blacks (Duncan 1968, Margo 2016, Collins and Wanamaker 2016) or farmers and non-farmers (Hout and Guest 2013, Xie and Killewald 2013). If less mobile groups are over-represented in the linked data, this would tend to overstate estimates of π and understate the historical rate of intergenerational mobility.

To make this point concretely, assume there is a high mobility, *h*, and low mobility, *l*, group, with intergenerational income elasticities of π^h and π^l, respectively. The share of the population in each group is $\emptyset^h$ and $\emptyset^l$. In the absence of Type I (α^m=0) and Type II errors ($\beta^m = 0$), the population intergenerational income elasticity can be written as the weighted average of elasticities in the two groups converges in probability to the following quantity,

$$\text{plim}\hat{\pi} = \emptyset^{h}\pi^{h} + \emptyset^{l}\pi^{l} \,. \tag{5}$$

Now suppose that expected population shares of the high and low groups change differentially as Type II errors grow or, more specifically, that $\emptyset^{h} \neq E(\text{Link}|h = 1\,,\beta > 0) > E(\text{Link}|l = 1\,,\beta > 0) \neq \emptyset^{l}$. This implies that estimates of π should tend to be too small (high mobility implies lower π) relative to the population parameter. This could lead to an overstatement of mobility in linked historical samples. The reverse could be true if increases in Type II errors disproportionately reduced link rates for the high mobility group. Moreover, variation in the type and incidence of Type II errors by linking method could lead to an under- or overstatement of the parameter of interest.

3.6.2 Results: Intergenerational Elasticity Estimates from the 1940 Census

We quantify the role of Type I and Type II errors by comparing intergenerational income elasticities resulting from different linking methods. Our sample consists of 19,090 birth certificates for Ohio-born boys which we link to the 1940 census. We supplement these data with a single sample of their fathers linked to the 1940 census using the LIFE-M method as generation 0. Therefore, the only differences in outcomes across automated linking methods are driven by differences in the link for sons.[29]

We report estimates of the intergenerational elasticity of income in Figure 3.6.1 using sons matched from birth certificates to the 1940 census by each linking method. From the LIFE-M ground truth data, we estimate an income elasticity of 0.27 between fathers and sons. While similar to Chetty et al. (2014) estimate of 0.33 for the late 20th century, we expect these estimates misstate the true intergenerational income elasticity for several reasons.

---

[29] Fathers are on the Ohio boys birth certificates but there is no income information on the birth certificate. We therefore link these fathers to the 1940 Census using the LIFE-M clerical review process. It may seem odd to think of measurement error in father's income when the set of linked fathers is the same across linking methods. However, an incorrect link for the son means that his income is attached to the wrong father. This problem is, therefore, analogous to the Type I error discussion in section IV.A.

First, the age structure in the data likely results in considerable life-cycle bias (Mazumder 2005, Haider and Solon 2006, Black and Devereux 2011), which may attenuate the intergenerational elasticities regardless of matching method (Mazumder 2015). Second, wage income observed in the 1940 Census is an imperfect measure of permanent income, and we expect the single observation of income for both generations allows its transitory component to generate sizable attenuation (Solon 1992, Zimmerman 1992, Mazumder 2005). On the other hand, the absence of farm and self-employed income in 1940 may lead this analysis to overstate mobility by excluding father-son pairs of famers—an occupation that tends to be highly persistent across generations (Hout and Guest 2013, Xie and Killewald 2013). Although these data limitations will lead us to misstate the true (population) intergenerational income elasticity, they should nonetheless affect all methods equally. The purpose of this exercise is to make comparisons of estimates obtained using different linking methods for this period and sample rather than making comparisons of these estimates to other periods or samples.

Generally, methods with higher Type I error rates have smaller elasticities. For instance, the Nix and Qian (2015) method which results in Type I error rates from 0.35 to 0.61 yields intergenerational income elasticity estimates of 0.12 to 0.20, whereas the Ferrie (1996) name method, which results in Type I error rates from 0.17 to 0.30, delivers intergenerational income elasticity estimates of 0.19 to 0.24. Consistent with phonetic name cleaning introducing considerable Type I errors of the classical variety (equations 2 and 3, Figure 3.2.2), Figure 3.6.1 shows that intergenerational income elasticity estimates tend to fall with the use of NYSIIS and Soundex, with Soundex generally reducing the estimates relative to LIFE-M by 30 (Ferrie 1996) to 55 percent (Nix and Qian 2015). Most results are *very* similar to what would be expected with the second variety of Type I error—orthogonal draws of fathers income in equation 3. Ferrie

(1996) with Soundex implies a 30 percent Type I error rate and the intergenerational elasticity estimate obtained from these links is 0.19—a 30 percent (0.08/.27) attenuation of the LIFE-M estimate (from equation 2, 0.70 is the set of true links × 0.27=0.19). Similarly, Nix and Qian (2012) with Soundex implies a 61 percent Type I error rate. The intergenerational elasticity estimate obtained from these links is 0.12—roughly 0.39 × 0.27=0.11).

Owing to small sample sizes, many of the methods produce s elasticities statistically indistinguishable from the LIFE-M matched sample. Estimates from the Ferrie (1996) method are statistically indistinguishable from the LIFE-M estimate of 0.27, with p-values ranging from 0.14 to 0.35. Results from the Ferrie (1996) method without the uncommon name restriction, however, statistically differ from the LIFE-M estimate (p-value of 0.07 using name, 0.04 using NYSIIS, and 0.001 using Soundex). Similarly, the estimates derived from Abramitzky et al.'s (2012b) method increasingly differ from the LIFE-M estimate with the use of phonetic name cleaning (p-value of 0.12 using name, 0.07 using NYSIIS, and 0.001 using Soundex). All estimates from Nix and Qian's (2015) method differ from the LIFE-M estimate (all p-values under 0.001). Finally, the estimate of 0.23 from Feigenbaum's (2016) method is statistically indistinguishable from the LIFE-M estimate.

In terms of deviations from these predictions in section 3.6.1 there may be several explanations: (1) The incomes of falsely linked fathers are uncorrelated with the incomes of the true father (i.e., the Type I error structure is similar to classical measurement error). (2) The measurement error in falsely linked fathers' incomes is non-classical and positively correlated with fathers' incomes. (3) The Type II errors introduce upward bias in the estimates (i.e., individuals with lower mobility, or higher intergenerational income elasticities, are overrepresented in our linked sample).

It is difficult to assess the first two possibilities. However, a final analysis addresses the third. To what extent do variations across these methods reflect differences in representativeness due to Type II errors? To answer this question, we use inverse propensity-score weights (DiNardo et al. 1996, Heckman et al. 1998) to reweight the linked sample to look like the population as a whole.[30] Figure 3.6.2 presents the reweighted intergenerational income elasticities. Across all methods, the propensity-score reweighted intergenerational income elasticities are *larger* in magnitude than the unweighted estimates in Figure 3.6.1, which is consistent with linked samples underrepresenting low-mobility individuals (e.g., groups with higher $\boldsymbol{\pi}$). Of course, this reweighting does not address Type I errors, which are associated with substantial downward bias in these estimates, as they were in Figure 3.6.1. In Figure 3.6.2, we cannot reject the hypothesis at the 10-percent level that estimates from the Ferrie (1996) method (using name, NYSIIS, or Soundex) and the Feigenbaum (2016) method equal the reweighted LIFE-M estimate of 0.30. All other linking methods, however, yield estimates that are statistically different than the reweighted LIFE-M estimate at the 10-percent level. P-values range from 0.001 to 0.08 for the Ferrie (1996) using all names, 0.02-0.07 for the Abramitzky et al. (2012b) method, and 0.0001-.003 for the Nix and Qian (2015) method.

3.7 Conclusion

Large-scale record linkage is one of the only means to produce longitudinal and intergenerational microdata for the 20th century US, but current approaches to this process are imperfect at best, and at worst, the choice of method contributes heavily to inferences using these

[30] To construct these weights, we first run a probit model of matched status (per each method) on an indicator variable for presence of milled name, length of first, middle, and last name, polynomials in day of birth, polynomials in age, an index for first name commonness, an index for last name commonness, indicators for whether parents were born in the US or abroad, number of siblings, an indicator variable for presence of siblings, and length of father's and mother's names. Our probit regressions estimate the propensity of being matched or unmatched (p) for each method, which then use to reweight the matched cases by $(p/(1-p) * (1-q)/q)$, where q is the share of matched cases.

data. We document staggering problems with false links in the most popular historical linking algorithms and demonstrate how these errors may have important implications for inferences using these linked samples. Our findings recommend several changes to current standard practice in historical record linkage.

First, we recommend comparing linked samples across several automated methods (Stata do-files are available with this paper). This approach could be used to diagnose Type I errors and their causes, allowing researchers to improve their matching algorithms. It is also a low-cost method to discard problematic cases and reduce Type I error rates.

Second, we strongly recommend against using NYSIIS and Soundex to improve match rates as these name cleaning algorithms increase Type I errors dramatically without reducing Type II error rates. In many cases, they lead to substantial attenuation in estimates of intergenerational income elasticities. Moreover, the errors arising from these name-cleaning algorithms appear systematically related to a number of record characteristics, making it unclear how they should affect inferences.

Third, we recommend careful consideration of record features to assess and improve sample representativeness. Many papers do this, but this should become even more important if researchers attempt to purge false matches from their samples. Making greater use of common record features such as name length or other socio-demographic information also allows researchers to use of standard survey research methods to create weights. These features may also be used to create inverse-propensity weights for linked samples, which could help balance both observed and potentially unobserved characteristics (DiNardo et al. 1996, Heckman et al. 1998).

3.8 References

A'Hearn, B., J. Baten and D. Crayen (2009). "Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital." *The Journal of Economic History* 69(3): 783-808.

Abowd, J.M. and L. Vilhuber (2005). "The Sensitivity of Economic Statistics to Coding Errors in Personal Identifiers." *Journal of Business and Economic Statistics* 23(2): 133-165.

Abramitzky, R., L. Boustan and K. Eriksson (2012a). "Web Appendix: Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration." Retrieved October 18, 2016. Available at https://assets.aeaweb.org/assets/production/articles-attachments/aer/data/aug2012/20100051_app.pdf.

Abramitzky, R., L. Platt Boustan and K. Eriksson (2012b). "Europe's Tired, Poor, and Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration." *American Economic Review* 102(5): 1832-1856.

Abramitzky, R., L. Platt Boustan and K. Eriksson (2013). "Have the Poor Always been Less Likely to Migrate? Evidence from Inheritance Practices during the Age of Mass Migration." *Journal of Development Economics* 102: 2-14.

Abramitzky, R., L. Platt Boustan and K. Eriksson (2014). "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration." *Journal of Political Economy* 122(3): 467-506.

Aizer, A., et al. (2016). "The Long Term Impact of Cash Transfers to Poor Families." *American Economic Review* 106(4): 935-971.

Antonie, L., et al. (2014). "Tracking People Over Time in 19th Century Canada for Longitudinal Analysis." *Machine Learning* 95(1): 129-146.

Atack, J., F. Bateman and M.E. Gregson (1992). "Matchmaker, Matchmaker, Make Me a Match." *Historical Methods* 25(2): 53-65.

Atack, J. (2004). "A Nineteenth-Century Resource for Agricultural History Research in the Twenty-First Century." *Agricultural History* 78(4): 389-412.

Bailey, A.K., et al. (2011). "Targeting Lynch Victims: Social Marginality or Status Transgressions?" *American Sociological Review* 76(3): 412-436.

Bailey, M.J., S. Anderson and C.G. Massey (2016). "Creating LIFE-M: The Longitudinal, Intergenerational Family Electronic Micro-Database." Retrieved September 29, 2016. Available at University of Michigan Working Paper.

Black, S.E. and P.J. Devereux (2011). "Recent Developments in Intergenerational Mobility." In *Handbook of Labor Economics*, edited by C. David and A. Orley. (Amsterdam, Elsevier. Volume 4B**:** 1487-1541).

Bleakley, H. and J. Ferrie (2016). "Shocking Behavior: Random Wealth in Antebellum Georgia and Human Capital Across Generations." *Quarterly Journal of Economics* 131(3): 1455-1495.

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

Boustan, L.P., M.E. Kahn and P.W. Rhode (2012). "Moving to Higher Ground: Migration Response to Natural Disasters in the Early Twentieth Century." *American Economic Review: Papers and Proceedings* 102(3): 238-244.

Boustan, L.P. and W. Collins (2014). "The Origins and Persistence of Black-White Differences in Women's Labor Force Participation from the Civil War to the Present." In *In Human*

Capital and History: The American Record, edited by L. Boustan, C. Frydman and R.A. Margo. (Chicago, IL, University of Chicago Press).
Buckles, K.S. and D.M. Hungerman (2013). "Season of Birth and Later Outcomes: Old Questions, New Answers." *Review of Economics and Statistics* 95(3): 711-724.
Chetty, R., et al. (2014). "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility." *American Economic Review* 104(5): 141-147.
Collins, W.J. and M.H. Wanamaker (2015). "The Great Migration in Black and White: New Evidence on the Selection and Sorting of Southern Migrants." *Journal of Economic History* 75(4): 947-992.
Collins, W.J. and M.H. Wanamaker (2016). "Up from Slavery? African American Intergenerational Economic Mobility Since 1880." Retrieved November 8, 2016. Available at
Costa, D.L., et al. (2017). "Union Army Veterans, All Grown Up." *Historical Methods* 50(2): 79-95.
Curti, M. (1959). *The Making of an American Community: A Case Study of Democracy in a Frontier County*. (Stanford: Stanford University Press).
DiNardo, J., N.M. Fortin and T. Lemieux (1996). "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica* 64(5): 1001-1044.
Duncan, O.D. (1968). "Patterns of Occupational Mobility among Negro Men." *Demography* 5(1): 11-22.
Eli, S., L. Salisbury and A. Shertzer (2016). "Migration in Response to Civil Conflict: Evidence from the Border of the American Civil War." Retrieved October 21, 2016. Available at https://economics.stanford.edu/sites/default/files/salisbury_feb-17.pdf.
Eriksson, B. (2016). "The Missing Links: Data Quality and Bias to Estimates of Social Mobility."
Feigenbaum, J.J. (2016a). "A Machine Learning Approach to Census Record Linking." Retrieved March 28, 2016. Available at http://scholar.harvard.edu/files/jfeigenbaum/files/feigenbaum-censuslink.pdf?m=1423080976.
Feigenbaum, J.J. (2016b). "A Machine Learning Approach to Census Record Linking." Retrieved March 28, 2016. Available at http://scholar.harvard.edu/files/jfeigenbaum/files/feigenbaum-censuslink.pdf?m=1423080976.
Ferrie, J.P. (1996). "A New Sample of Males Linked from the 1850 Public Use Micro Sample of the Federal Census of Population to the 1860 Federal Census Manuscript Schedules." *Historical Methods* 29(4): 141-156.
Ferrie, J.P. and J. Long (2013). "Intergenerational Occupational Mobility in Great Britain and the United States since 1850." *American Economic Review* 103(4): 1109-1137.
Fu, Z., et al. (2014a). "Automatic Record Linkage of Individuals and Households in Historical Census Data." *International Journal of Humanities and Arts Computing* 8.2: 204-225.
Fu, Z., P. Christen and J. Zhou (2014b). "A Graph Matching Method for Historical Census Household Linkage." In *Advances in Knowledge Discovery and Data Mining*, edited by V.S. Tseng, T.B. Ho, Z.-H. Zhou, A.L.P. Chen and H.-Y. Kao. (Tainan, Taiwan, 18th Pacific-Asia Conference).
Goeken, R., et al. (2011). "New Methods of Census Record Linking." *Historical Methods* 44(1): 7-14.

Goeken, R., et al. (2016). "Evaluating the Accuracy of Linked U. S. Census Data: A Household Approach."
Guest, A.M. (1987). "Notes from the National Panel Study: Linkage and Migration in the Late Nineteenth Century." *Historical Methods* 20(2): 63-77.
Haider, S.J. and G. Solon (2006). "Life-Cycle Variation in the Association between Current and Lifetime Earnings." *American Economic Review* 96(4): 1308-1320.
Heckman, J.J., et al. (1998). "Characterizing Selection Bias Using Experimental Data." *Econometrica* 66(5): 1017-1098.
Hornbeck, R. and S. Naidu (2014). "When the Levee Breaks: Black Migration and Economic Development in the American South." *American Economic Review* 104(3): 963-990.
Hout, M. and A.M. Guest (2013). "Intergenerational Occupational Mobility in Great Britain and the United States since 1850: Comment." *American Economic Review* 103(5): 2021-2040.
Kim, G. and R. Chambers (2012). "Regression Analysis under Probabilistic Multi-Linkage." *Statistica Neerlandica* 66(1): 64-79.
Malin, J. (1935). "The Turnover of Farm Pouplation in Kansas." *Kansas Historical* 20: 339-372.
Margo, R.A. (2016). "Obama, Katrina, and the Persistence of Racial Inequality." *Journal of Economic History* 76(2): 301-341.
Massey, C.G. (2017). "Playing with matches: An assessment of accuracy in linked historical data." *Historical Methods: A Journal of Quantitative and Interdisciplinary History*: 1-15.
Mazumder, B. (2005). "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data." *Review of Economics and Statistics* 87(2): 235-255.
Mazumder, B. (2015). "Estimating the Intergenerational Elasticity and Rank Association in the U.S.: Overcoming the Current Limitations of Tax Data." Retrieved November 10, 2015. Available at https://www.chicagofed.org/publications/working-papers/2015/wp2015-04.
Michelson, M.a.K., C. A (2006). "Learning Blocking Schemes for Record Linkage." *Proceedings of the 21st National Conference on Artificial Intelligence* AAAI-06.
Mill, R. (2013). "Record Linkage across Historical Datasets." Retrieved Available at Proquest.
Mill, R. and L.C. Stein (2016). "Race, Skin Color, and Economic Outcomes in Early Twentieth-Century America." Retrieved Available at http://www.public.asu.edu/~lstein2/research/mill-stein-skincolor.pdf.
National Office of Vital Statistics (1948). State and Regional Life Tables, 1939-1941. Edited by. Available at https://www.cdc.gov/nchs/data/lifetables/life39-41.pdf.
Nix, E. and N. Qian (2015). "The Fluidity of Race: 'Passing' in the United States, 1880-1940." *NBER Working Paper 20828*.
Olivetti, C. and M.D. Paserman (2015). "In the Name of the Son (and the Daughter): Intergenerational Mobility in the United States, 1850-1940." *American Economic Review* 105(8): 2695-2724.
Ruggles, S. (2006). "Linked Historical Censuses: A New Approach." *History and Computing* 14: 213-224.
Ruggles, S., et al. (2010). Integrated Public Use Microdata Series (Version 5.0) [Machine-readable database]. Minneapolis: University of Minnesota.
Ruggles, S. (2011). "Intergenerational Coredisence and Family Transitions in the United States, 1850-1880." *Journal of Marriage and the Family* 73(1): 138-148.
Ruggles, S., et al. (2015). Integrated Public Use Microdata Series: Version 6.0 [Machine-readable database]. Minneapolis: University of Minnesota.

Saperstein, A. and A. Gullickson (2013). "A Mulatto Escape Hatch? Examining Evidence of U.S. Racial and Social Mobility in the Jim Crow Era." *Demography* 50(5): 1921-1942.

Schaefer, D. (1985). "A Statistical Profile of Frontier and New South Migration: 1850-1860." *Agricultural History* 59: 563-567.

Snae, C. (2007). "A Comparison and Analysis of Name Matching Algorithms." *International Scholarly and Scientific Research and Innovation* 1(1): 107-112.

Solon, G. (1992). "Intergenerational Income Mobility in the United States." *The American Economic Review* 82(3): 393-408.

Solon, G. (1999). "Intergenerational Mobility in the Labor Market." In *Handbook of Labor Economics*, edited by O. Ashenfelter and D. Card. (Amsterdam, Elsevier**:** 1761-1800).

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

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

Vick, R. and L. Huynh (2011). "The Effects of Standardizing Names for Record Linkage: Evidence from the United States and Norway." *Historical Methods* 44(1): 15-24.

Wasi, N. (2014). RecLink3. Edited by M. Blasnik. Available at http://EconPapers.repec.org/RePEc:boc:bocode:s456876.

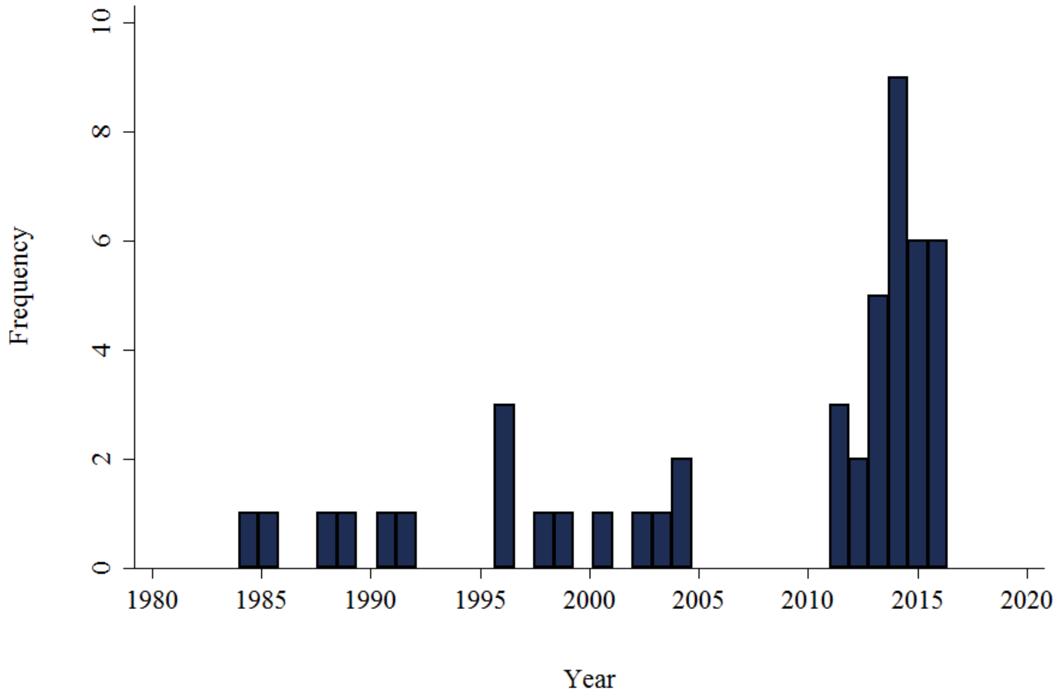
West, K.K. and J.G. Robinson (1999). "What do we know about the Undercount of Children?" *U.S. Census Bureau Population Division Working Paper* 39.

Wisselgren, M.J., et al. (2014). "Testing Methods of Record Linkage on Swedish Censuses." *Historical Methods* 47(3): 138-151.

Xie, Y. and A. Killewald (2013). "Intergenerational Occupational Mobility in Britain and the U.S. since 1850: Comment." *American Economic Review* 103(5): 2003-2020.

Zimmerman, D.J. (1992). "Regression Toward Mediocrity in Economic Stature." *American Economic Review* 82(3): 409-429.

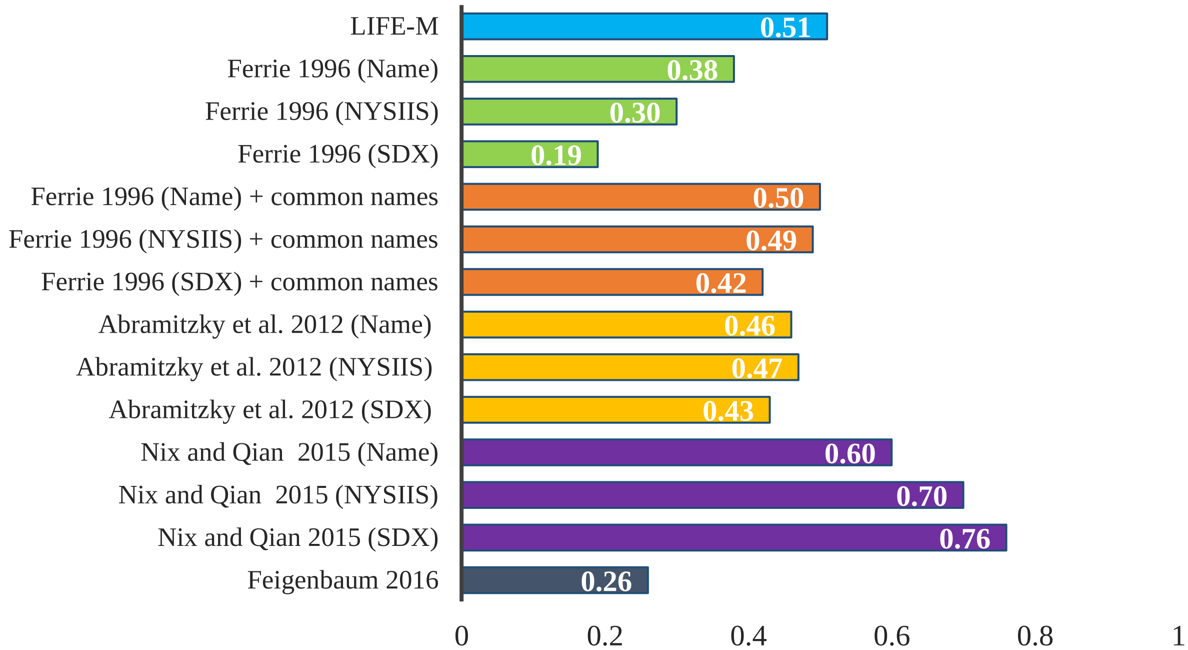
Figure 3.1: Papers in Economics Using Longitudinally Linked Samples, by Year



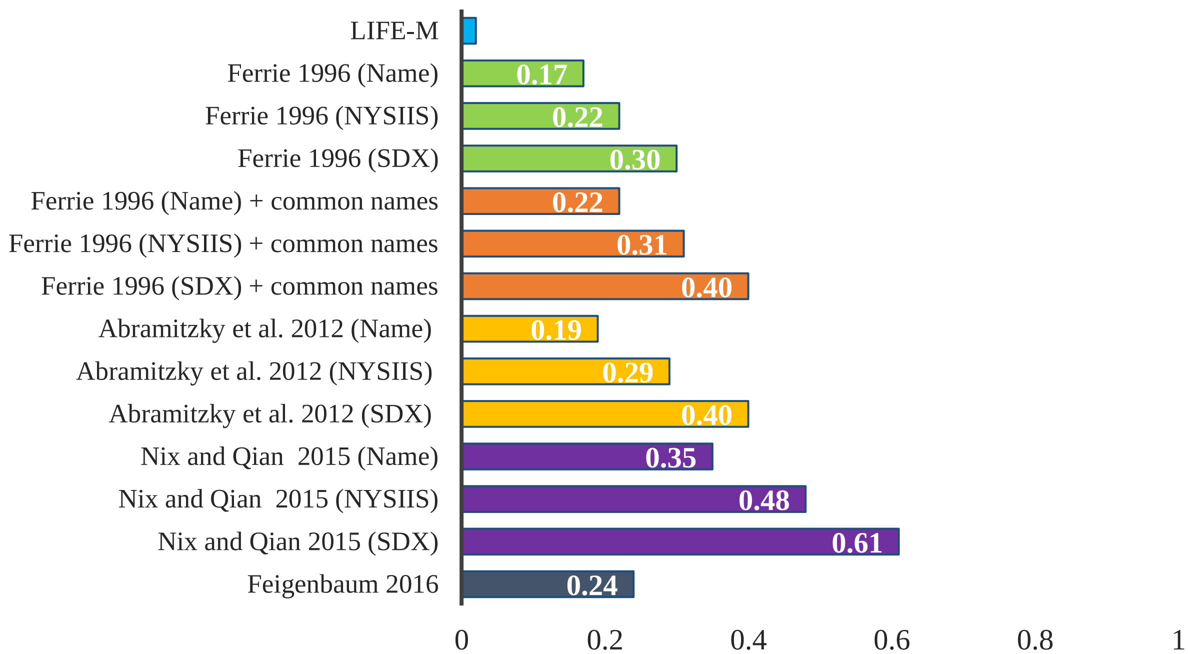
Source: Authors calculations of papers published in journals or as working papers by year.

Figure 3.2 Performance of Automated Linking Methods using the LIFE-M Ground-Truth

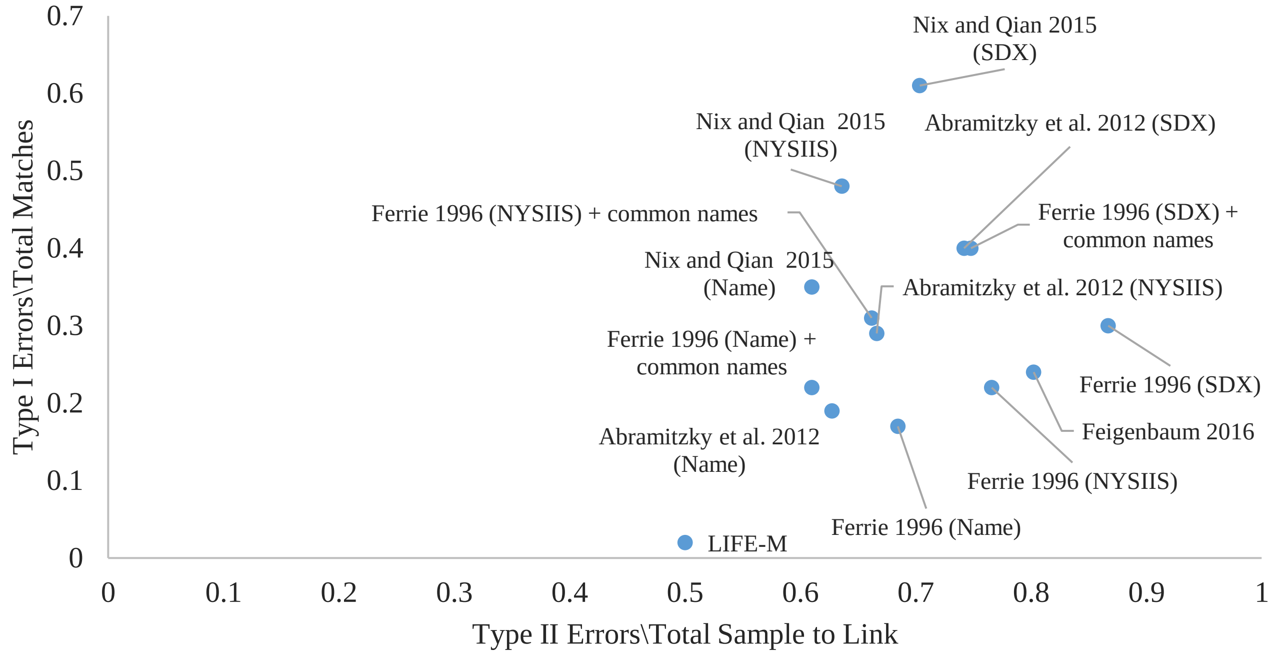
3.2.1 Match Rates



3.2.2 Type I Error Rates (Share of False Matches)

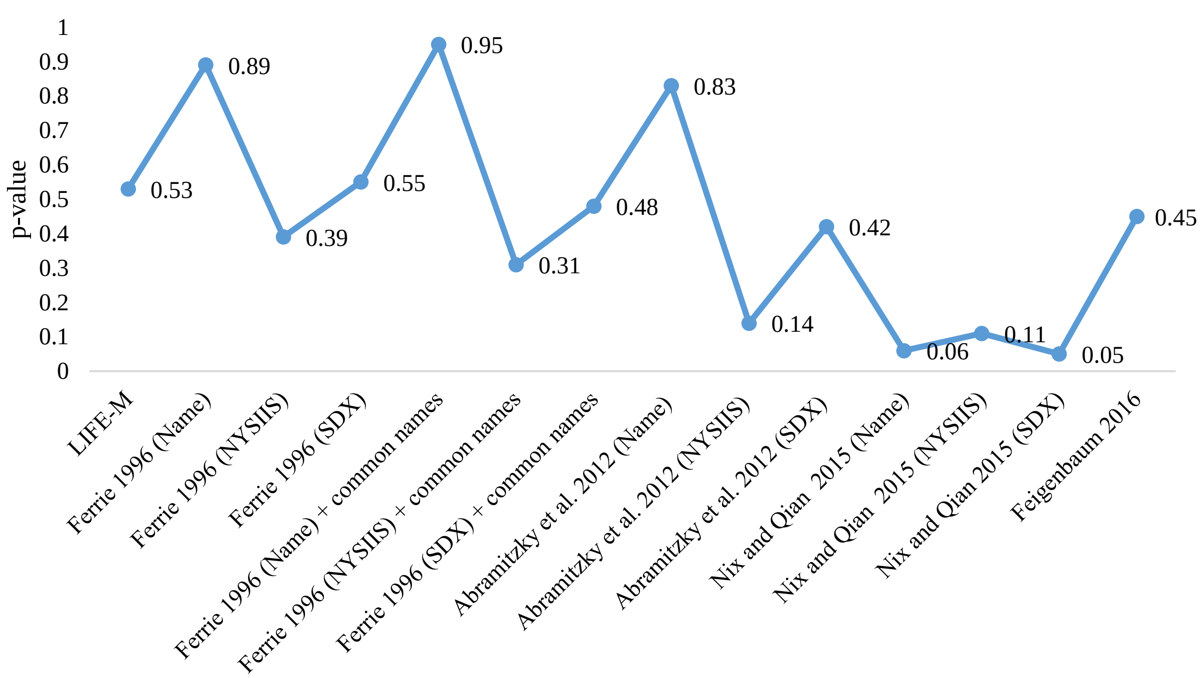


3.2.3 Share of Type I Errors (False Matches) versus Type II Errors (Missed Matches)



Notes: Panels A to C are based on the LIFE-M sample of Ohio boys linked to the 1940 Census using different linking methods as described in the text.

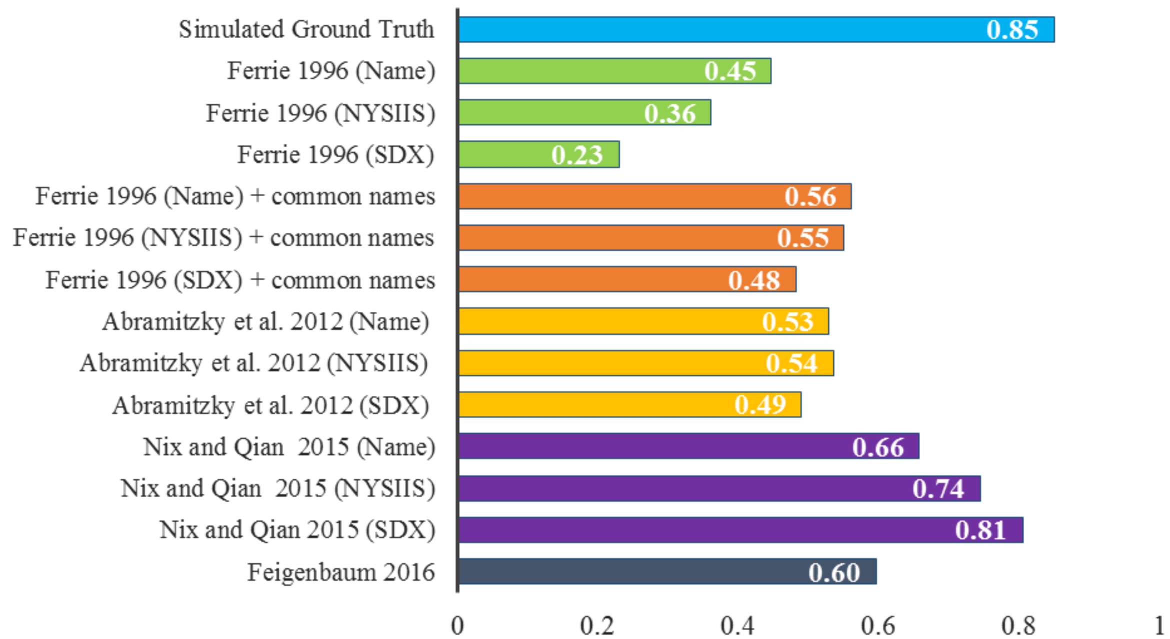
Figure 2.3: Tests of Equality of Linked and Baseline Day-of-Birth Distributions, by Linking Method



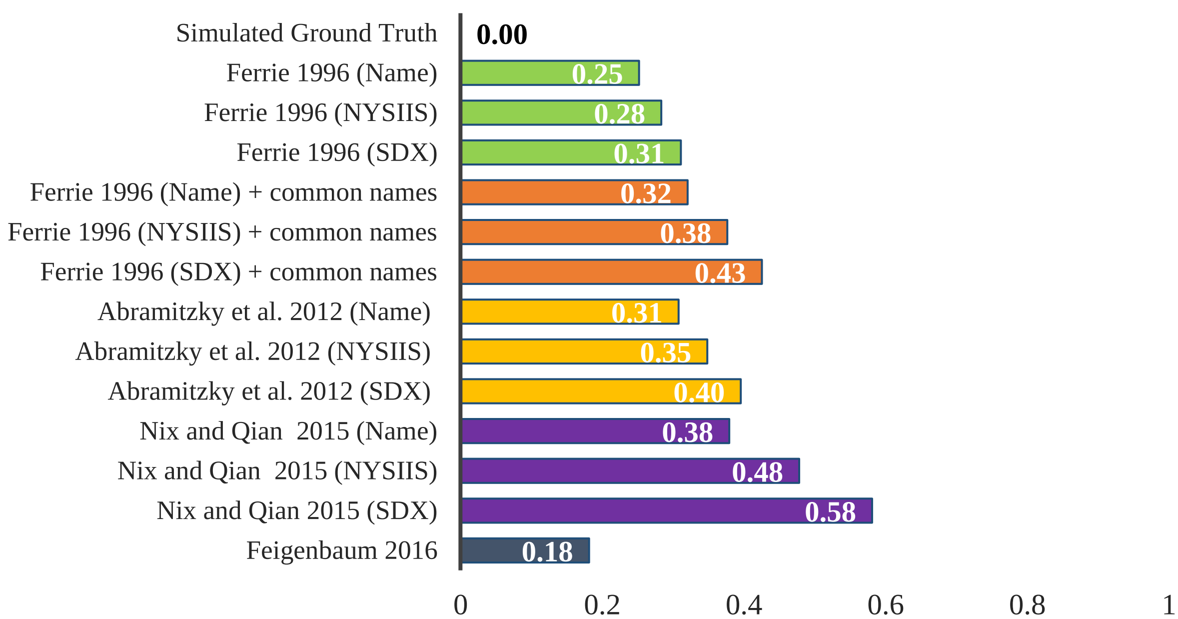
Notes: Figure plots p-values for a two-sample Kolmogorov-Smirnov test for equality of day-of-birth (1-366, including February 29 in leap year) distributions for the linked observations versus the unlinked observations.

Figure 3.4: Performance of Automated Linking Methods using Synethic Ground-Truth

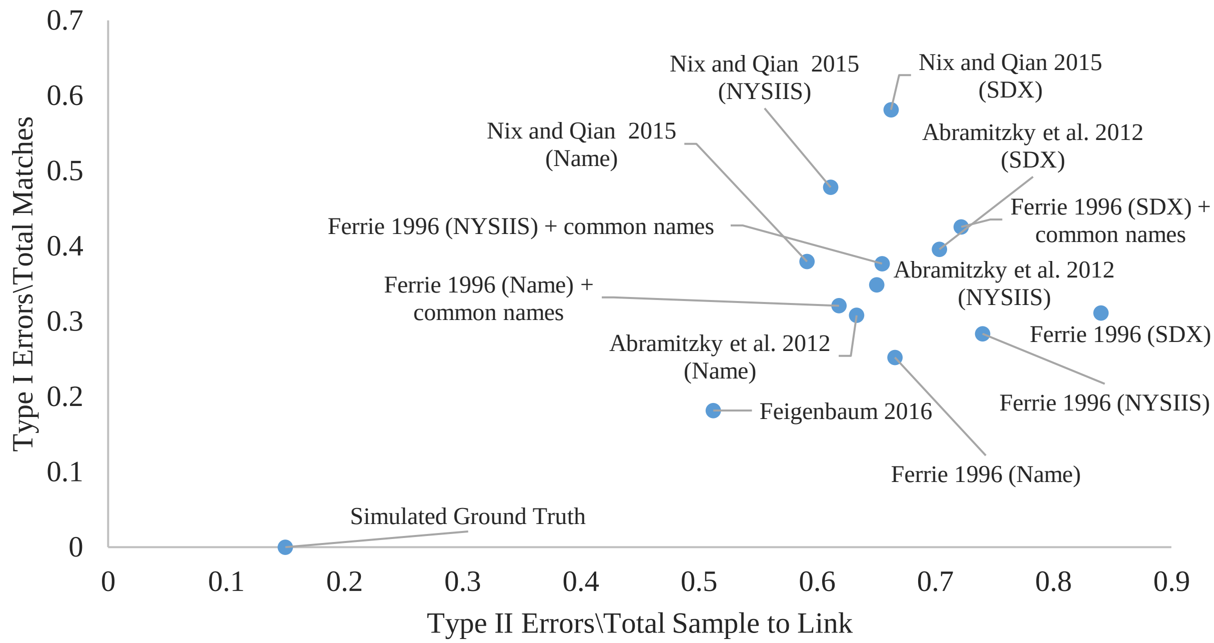
3.4.1 Match Rates



3.4.2 Type I Error Rates (Share of False Matches)



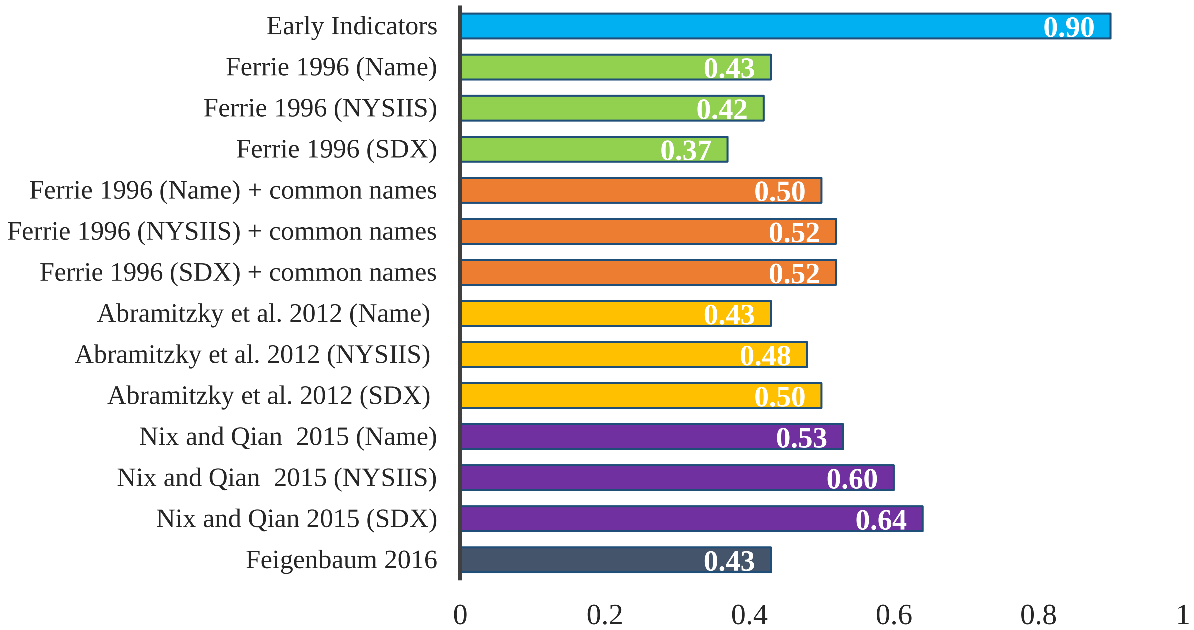
3.4.3 Share of Type I Errors (False Matches) versus Type II Errors (Missed Matches)



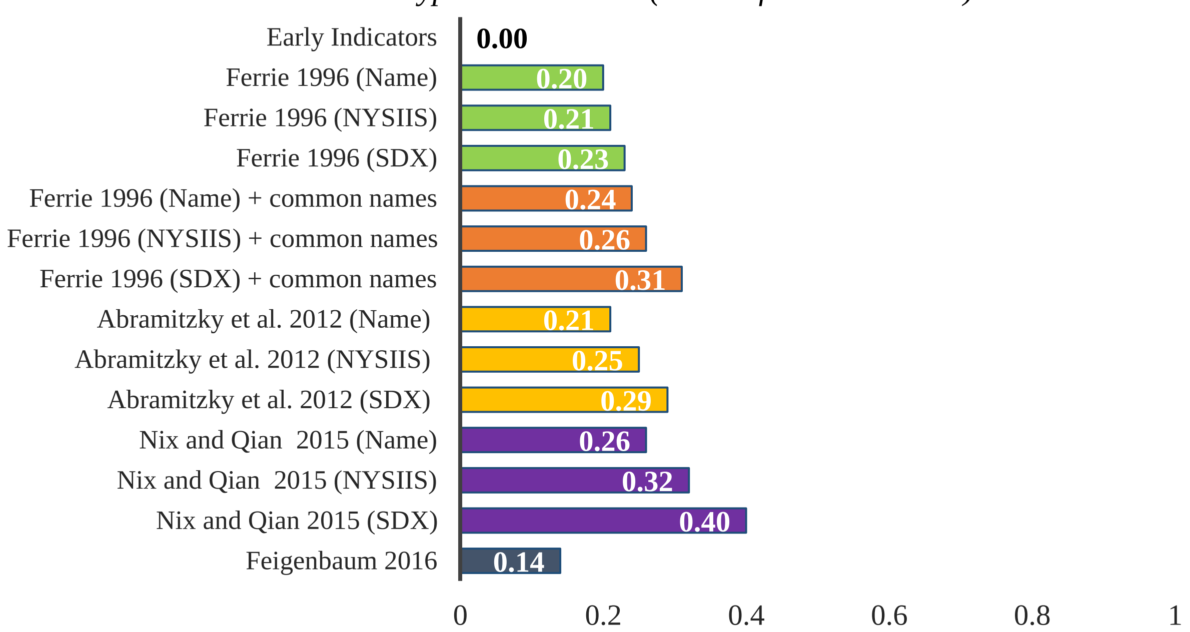
Notes: Panels A to C are based on the simulated ground truth sample. We assume that only 85 percent of the sample can be linked due to under-enumeration by the Census Bureau, mortality, and emigration.

Figure 3.5: Performance of Automated Linking Methods using the Early Indicators Ground-Truth

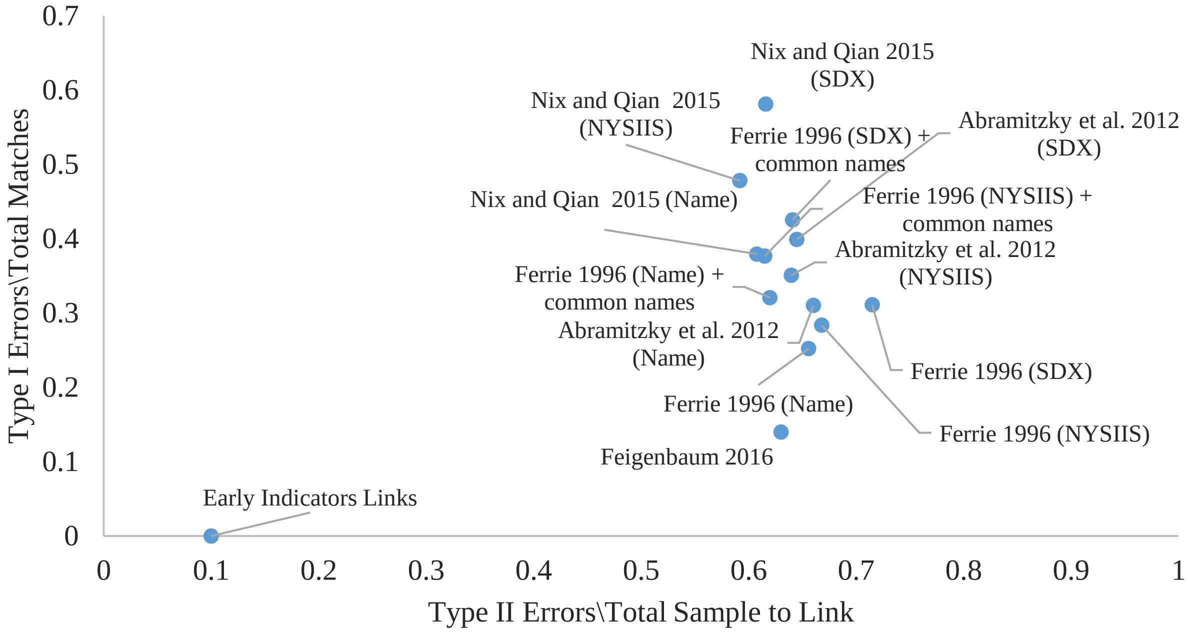
3.5.1 Match Rates



3.5.2 Type I Error Rates (Share of False Matches)



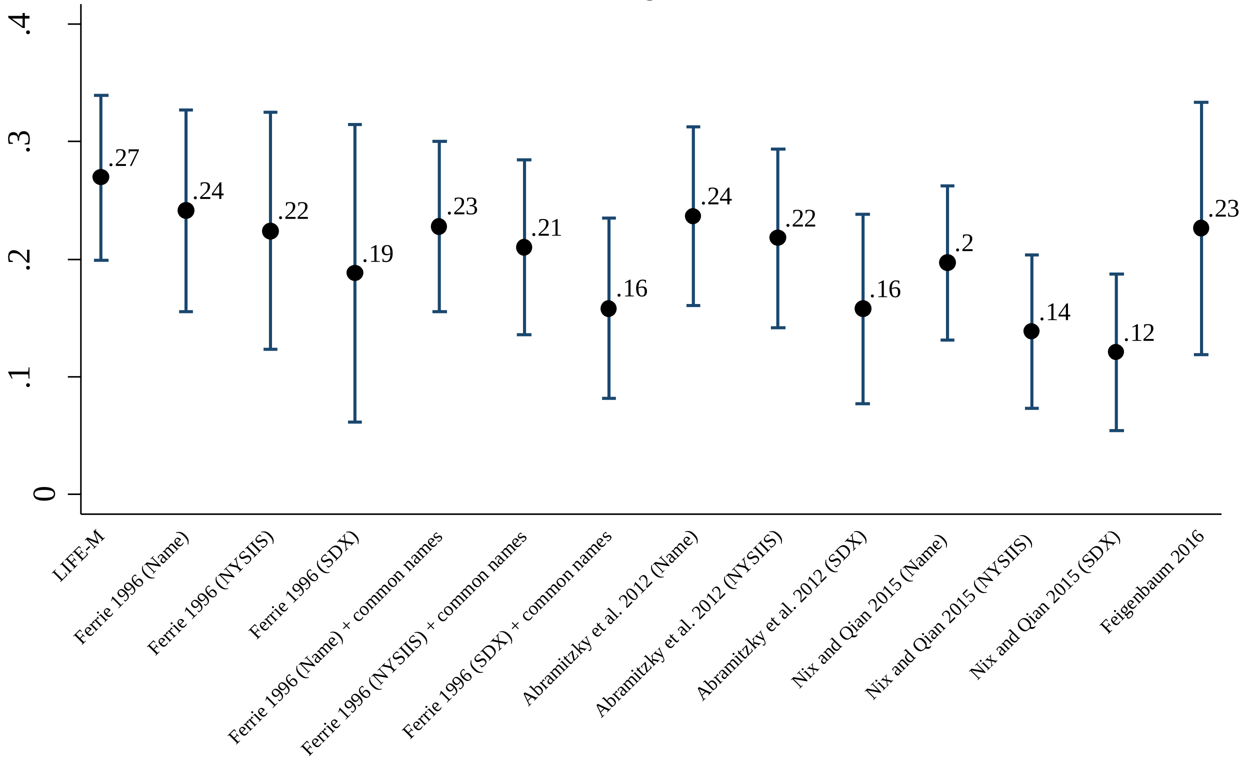
3.5.3 Share of Type I Errors (False Matches) versus Type II Errors (Missed Matches)



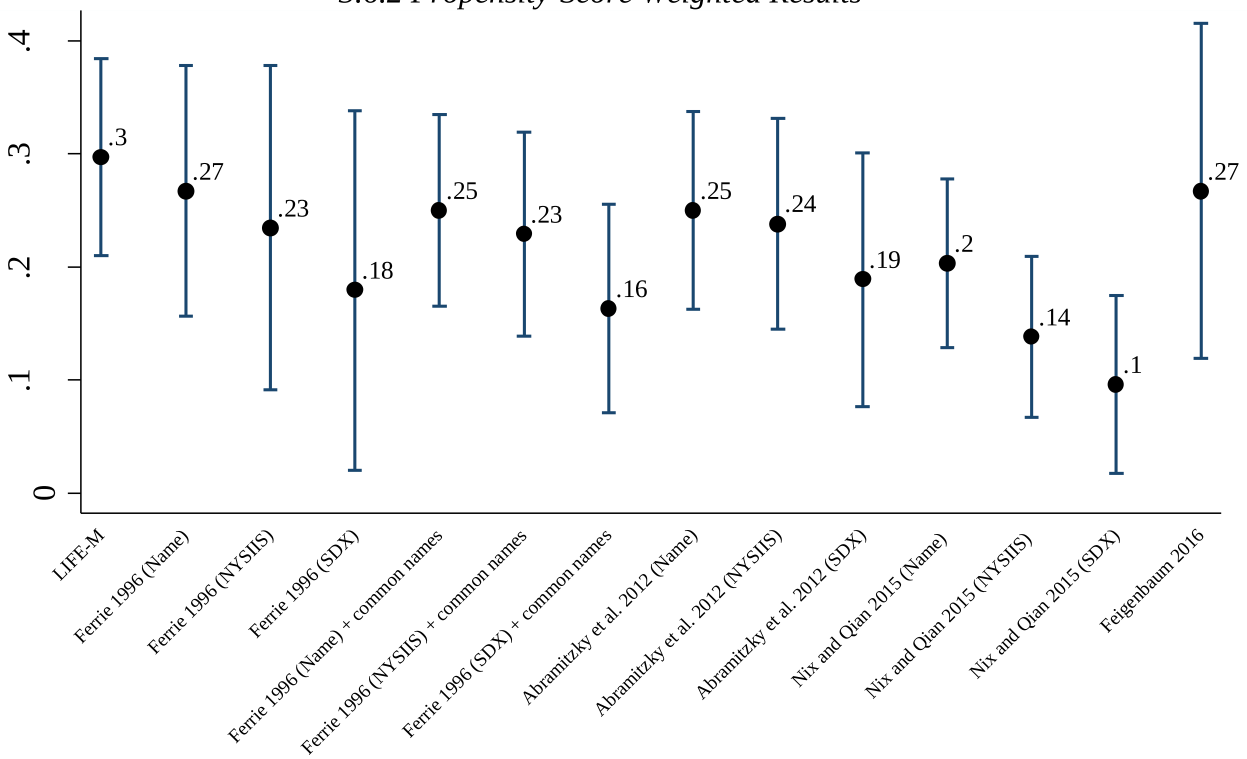
Notes: Panel A to C present results based on the genealogically linked Early Indicators Oldest-Old sample of Union Army veterans linked to the 1900 Census using different automated linking methods.

Figure 3.6: Intergenerational Income Elasticity Estimates

3.6.1 Unweighted Results



3.6.2 Propensity-Score Weighted Results



Notes: Wage incomes are from fathers and sons from the Ohio boys sample linked to the 1940 Census. Differences across estimates represent differences in the characteristics and incidence of Type I and Type II errors.

Table 3.1: Differences in Characteristics of Linked Samples from Baseline Sample of Ohio Boys, by Linking Method

	Day of Birth	Foreign Born		Number of Siblings	Length of Name of			Percent of Siblings with Mispelled	
		Mother	Father		Boy	Mother	Father	Mother's Name	Father's Name
LIFE-M	0.68	-0.018***	-0.012*	0.049	0.046***	0.267***	0.217***	-0.002	-0.005*
	(1.525)	(0.006)	(0.006)	(0.042)	(0.018)	(0.047)	(0.047)	(0.003)	(0.003)
Ferrie 1996 (Name)	-0.15	-0.014**	-0.009	0.145***	0.081***	0.130***	0.087*	-0.001	-0.001
	(1.580)	(0.006)	(0.007)	(0.044)	(0.018)	(0.048)	(0.049)	(0.003)	(0.003)
Ferrie 1996 (NYSIIS)	2.14	0.032***	0.045***	0.107**	0.088***	0.600***	0.558***	-0.006	-0.009***
	(1.659)	(0.007)	(0.007)	(0.046)	(0.019)	(0.051)	(0.051)	(0.004)	(0.003)
Ferrie 1996 (SDX)	2.09	0.039***	0.047***	0.144***	0.129***	0.463***	0.474***	-0.008**	-0.011***
	(1.931)	(0.008)	(0.008)	(0.053)	(0.023)	(0.059)	(0.060)	(0.004)	(0.004)
Ferrie 1996 (Name) + common names	-1.44	-0.063***	-0.067***	0.095**	0.068***	0.029	0.001	0.004	0.004
	(1.526)	(0.006)	(0.006)	(0.042)	(0.018)	(0.047)	(0.047)	(0.003)	(0.003)
Ferrie 1996 (NYSIIS) + common names	1.39	-0.010*	-0.008	0.060	0.070***	0.240***	0.222***	-0.002	-0.002
	(1.526)	(0.006)	(0.006)	(0.042)	(0.018)	(0.047)	(0.047)	(0.003)	(0.003)
Ferrie 1996 (SDX) + common names	-0.44	0.010	0.015**	0.082*	0.112***	0.333***	0.347***	-0.006*	-0.009***
	(1.543)	(0.006)	(0.006)	(0.043)	(0.018)	(0.047)	(0.048)	(0.003)	(0.003)
Abramitzky et al. 2012 (Name)	-0.97	-0.061***	-0.064***	0.090**	0.056***	0.036	0.017	0.007**	0.003
	(1.530)	(0.006)	(0.006)	(0.042)	(0.018)	(0.047)	(0.047)	(0.003)	(0.003)
Abramitzky et al. 2012 (NYSIIS)	1.98	-0.007	-0.005	0.039	0.097***	0.217***	0.230***	0.001	-0.006**
	(1.529)	(0.006)	(0.006)	(0.042)	(0.018)	(0.047)	(0.047)	(0.003)	(0.003)
Abramitzky et al. 2012 (SDX)	-0.08	0.011*	0.013**	0.066	0.095***	0.316***	0.362***	-0.003	-0.012***
	(1.542)	(0.006)	(0.006)	(0.043)	(0.018)	(0.047)	(0.048)	(0.003)	(0.003)
Nix and Qian 2015 (Name)	-0.73	-0.136***	-0.158***	0.182***	0.152***	-0.253***	-0.232***	0.012***	0.013***
	(1.616)	(0.006)	(0.007)	(0.045)	(0.019)	(0.049)	(0.050)	(0.003)	(0.003)
Nix and Qian 2015 (NYSIIS)	-1.68	-0.127***	-0.150***	0.108**	0.193***	-0.598***	-0.512***	0.011***	0.022***
	(1.855)	(0.007)	(0.008)	(0.051)	(0.022)	(0.057)	(0.057)	(0.004)	(0.004)
Nix and Qian 2015 (SDX)	-2.05	-0.081***	-0.099***	-0.067	0.294***	-0.281***	-0.271***	0.010**	0.023***
	(2.297)	(0.009)	(0.009)	(0.063)	(0.027)	(0.070)	(0.071)	(0.005)	(0.004)
Feigenbaum 2016	-0.39	0.042***	0.062***	0.135***	0.100***	0.358***	0.319***	-0.008**	-0.011***
	(1.739)	(0.007)	(0.007)	(0.048)	(0.020)	(0.053)	(0.054)	(0.004)	(0.003)

Notes: Each estimate is a from a separate regression of the variable indicated in the column header on a binary variable for whether the linking method matched the observation. The sample size for every regression is n = 19,090. Standard errors are reported in parentheses below. * indicates statistically significant at the 10-percent level, ** at the 5-percent level, and * at the 1-percent level.

Table 3.2: Differences in Characteristics of Erroneous Links for Ohio Boys by Linking Method

	Day of Birth	Foreign Born: Mother	Foreign Born: Father	Number of Siblings	Length of Name of: Boy	Length of Name of: Mother	Length of Name of: Father	Percent of Siblings with Mispelled Mother's Name	Percent of Siblings with Mispelled Father's Name
LIFE-M	-2.66	-0.029	0.005	0.142	-0.028	-0.061	-0.109	-0.028*	-0.028*
	(7.494)	(0.029)	(0.031)	(0.204)	(0.085)	(0.232)	(0.233)	(0.016)	(0.015)
Ferrie 1996 (Name)	-3.33	-0.043***	-0.035***	-0.035	-0.010	-0.150	-0.231**	-0.011	0.003
	(3.303)	(0.013)	(0.014)	(0.090)	(0.037)	(0.102)	(0.102)	(0.007)	(0.006)
Ferrie 1996 (NYSIIS)	-3.50	0.019	0.029**	-0.228**	-0.094**	-0.631***	-0.519***	-0.000	0.011
	(3.320)	(0.014)	(0.014)	(0.090)	(0.037)	(0.102)	(0.103)	(0.007)	(0.006)
Ferrie 1996 (SDX)	0.53	0.080***	0.100***	-0.445***	-0.042	-0.396***	-0.332***	-0.009	0.012
	(3.791)	(0.016)	(0.016)	(0.105)	(0.043)	(0.115)	(0.117)	(0.008)	(0.007)
Ferrie 1996 (Name) + common names	0.66	-0.034***	-0.037***	0.110**	0.102***	0.094**	0.062	0.005	-0.001
	(1.567)	(0.006)	(0.006)	(0.043)	(0.018)	(0.048)	(0.048)	(0.003)	(0.003)
Ferrie 1996 (NYSIIS) + common names	2.34	-0.011*	-0.006	0.141***	0.100***	0.518***	0.453***	-0.000	-0.008***
	(1.614)	(0.006)	(0.007)	(0.045)	(0.019)	(0.049)	(0.050)	(0.003)	(0.003)
Ferrie 1996 (SDX) + common names	0.05	-0.020***	-0.019***	0.170***	0.117***	0.460***	0.403***	-0.002	-0.008**
	(1.753)	(0.007)	(0.007)	(0.048)	(0.020)	(0.053)	(0.054)	(0.004)	(0.003)
Abramitzky et al. 2012 (Name)	-4.09	-0.042***	-0.045***	-0.132*	-0.060*	-0.181**	-0.238***	0.001	0.008
	(2.853)	(0.011)	(0.011)	(0.078)	(0.031)	(0.088)	(0.088)	(0.006)	(0.005)
Abramitzky et al. 2012 (NYSIIS)	-2.91	0.015	0.007	-0.218***	-0.136***	-0.809***	-0.656***	0.001	0.011**
	(2.459)	(0.010)	(0.010)	(0.067)	(0.027)	(0.075)	(0.076)	(0.005)	(0.005)
Abramitzky et al. 2012 (SDX)	-2.02	0.052***	0.049***	-0.190***	-0.138***	-0.446***	-0.305***	-0.003	0.001
	(2.395)	(0.010)	(0.010)	(0.065)	(0.028)	(0.073)	(0.073)	(0.005)	(0.005)
Nix and Qian 2015 (Name)	0.94	-0.014**	-0.015**	0.112**	0.103***	0.123**	0.129***	0.001	-0.006*
	(1.619)	(0.006)	(0.007)	(0.045)	(0.019)	(0.050)	(0.050)	(0.003)	(0.003)
Nix and Qian 2015 (NYSIIS)	3.01*	0.007	0.012*	0.161***	0.116***	0.563***	0.551***	0.001	-0.010***
	(1.719)	(0.007)	(0.007)	(0.047)	(0.020)	(0.052)	(0.053)	(0.004)	(0.003)
Nix and Qian 2015 (SDX)	1.23	-0.005	0.003	0.225***	0.092***	0.402***	0.411***	-0.005	-0.011***
	(1.979)	(0.008)	(0.008)	(0.055)	(0.023)	(0.060)	(0.061)	(0.004)	(0.004)
Feigenbaum 2016	-4.61	-0.001	-0.002	-0.003	-0.128***	-0.328***	-0.217**	-0.005	0.008
	(3.492)	(0.015)	(0.015)	(0.098)	(0.041)	(0.108)	(0.109)	(0.008)	(0.007)

Notes: Each estimate is a from a separate regression of the variable indicated in the column header on a binary variable for whether the linking method matched a false positive. In this table, we count false positives as “no link” (or zero), so this table captures the relationship of the “correct” links each method makes to observed characteristics. Standard errors are reported in parentheses below. * indicates statistically significant at the 10-percent level, ** at the 5-percent level, and * at the 1-percent level.

Table 3.3: Differences in Characteristics of Linked Samples from Baseline Sample of the Oldest Old, by Linking Method

	Age	Speaks English	Owns Farm	Owns House	Married	Foreign-Born
Early Indicators	0.092	0.058***	0.024	0.034**	0.059***	0.007
	(0.214)	(0.006)	(0.015)	(0.016)	(0.011)	(0.009)
Ferrie 1996 (Name)	0.011	0.038***	0.027	0.022	0.056***	-0.007
	(0.266)	(0.009)	(0.019)	(0.020)	(0.014)	(0.011)
Ferrie 1996 (NYSIIS)	-0.302	0.037***	0.032	0.014	0.040***	-0.011
	(0.270)	(0.009)	(0.019)	(0.020)	(0.015)	(0.012)
Ferrie 1996 (SDX)	-0.027	0.033***	0.029	0.013	0.038**	-0.016
	(0.282)	(0.010)	(0.020)	(0.021)	(0.015)	(0.012)
Ferrie 1996 (Name) + common names	0.166	0.030***	0.015	0.015	0.040***	-0.004
	(0.255)	(0.009)	(0.018)	(0.019)	(0.014)	(0.011)
Ferrie 1996 (NYSIIS) + common names	0.219	0.025***	0.018	0.006	0.025*	-0.003
	(0.257)	(0.009)	(0.018)	(0.019)	(0.014)	(0.011)
Ferrie 1996 (SDX) + common names	0.230	0.025***	0.009	0.010	0.021	-0.007
	(0.255)	(0.009)	(0.018)	(0.019)	(0.014)	(0.011)
Abramitzky et al. 2012 (Name)	0.128	0.036***	0.033*	-0.007	0.048***	-0.004
	(0.267)	(0.009)	(0.019)	(0.020)	(0.015)	(0.012)
Abramitzky et al. 2012 (NYSIIS)	0.196	0.032***	0.027	-0.002	0.033**	-0.004
	(0.262)	(0.009)	(0.019)	(0.019)	(0.014)	(0.011)
Abramitzky et al. 2012 (SDX)	0.208	0.029***	0.014	0.007	0.027*	-0.008
	(0.257)	(0.009)	(0.018)	(0.019)	(0.014)	(0.011)
Nix and Qian 2015 (Name)	0.064	0.036***	0.025	0.014	0.042***	-0.008
	(0.248)	(0.008)	(0.018)	(0.019)	(0.014)	(0.011)
Nix and Qian 2015 (NYSIIS)	0.137	0.034***	0.023	0.011	0.035***	-0.000
	(0.241)	(0.008)	(0.017)	(0.018)	(0.013)	(0.010)
Nix and Qian 2015 (SDX)	0.118	0.029***	0.017	0.008	0.030**	0.001
	(0.236)	(0.008)	(0.017)	(0.017)	(0.013)	(0.010)
Feigenbaum 2016	0.812***	0.034***	0.032*	-0.001	0.034**	-0.022*
	(0.277)	(0.009)	(0.019)	(0.020)	(0.015)	(0.011)

Notes: Each estimate is from a separate regression of the variable indicated in the column header on a binary variable for whether the linking method matched the observation. Standard errors are reported in parentheses below. * indicates statistically significant at the 10-percent level, ** at the 5-percent level, and * at the 1-percent level.

Table 3.4: Differences in Characteristics of Erroneous Links from Baseline Sample of the Oldest Old, by Linking Method

	Age	Speaks English	Owns Farm	Owns House	Married	Foreign-Born
Early Indicators	0.092	0.058***	0.024	0.034**	0.059***	0.007
Early	(0.214)	(0.006)	(0.015)	(0.016)	(0.011)	(0.009)
Ferrie 1996 (Name)	0.011	0.038***	0.027	0.022	0.056***	-0.007
	(0.266)	(0.009)	(0.019)	(0.020)	(0.014)	(0.011)
Ferrie 1996 (NYSIIS)	-0.302	0.037***	0.032	0.014	0.040***	-0.011
	(0.270)	(0.009)	(0.019)	(0.020)	(0.015)	(0.012)
Ferrie 1996 (SDX)	-0.027	0.033***	0.029	0.013	0.038**	-0.016
	(0.282)	(0.010)	(0.020)	(0.021)	(0.015)	(0.012)
Ferrie 1996 (Name) + common names	0.166	0.030***	0.015	0.015	0.040***	-0.004
	(0.255)	(0.009)	(0.018)	(0.019)	(0.014)	(0.011)
Ferrie 1996 (NYSIIS) + common names	0.219	0.025***	0.018	0.006	0.025*	-0.003
	(0.257)	(0.009)	(0.018)	(0.019)	(0.014)	(0.011)
Ferrie 1996 (SDX) + common names	0.230	0.025***	0.009	0.010	0.021	-0.007
	(0.255)	(0.009)	(0.018)	(0.019)	(0.014)	(0.011)
Abramitzky et al. 2012 (Name)	0.128	0.036***	0.033*	-0.007	0.048***	-0.004
	(0.267)	(0.009)	(0.019)	(0.020)	(0.015)	(0.012)
Abramitzky et al. 2012 (NYSIIS)	0.196	0.032***	0.027	-0.002	0.033**	-0.004
	(0.262)	(0.009)	(0.019)	(0.019)	(0.014)	(0.011)
Abramitzky et al. 2012 (SDX)	0.208	0.029***	0.014	0.007	0.027*	-0.008
	(0.257)	(0.009)	(0.018)	(0.019)	(0.014)	(0.011)
Nix and Qian 2015 (Name)	0.064	0.036***	0.025	0.014	0.042***	-0.008
	(0.248)	(0.008)	(0.018)	(0.019)	(0.014)	(0.011)
Nix and Qian 2015 (NYSIIS)	0.137	0.034***	0.023	0.011	0.035***	-0.000
	(0.241)	(0.008)	(0.017)	(0.018)	(0.013)	(0.010)
Nix and Qian 2015 (SDX)	0.118	0.029***	0.017	0.008	0.030**	0.001
	(0.236)	(0.008)	(0.017)	(0.017)	(0.013)	(0.010)
Feigenbaum 2016	0.812***	0.034***	0.032*	-0.001	0.034**	-0.022*
	(0.277)	(0.009)	(0.019)	(0.020)	(0.015)	(0.011)

Notes: Each estimate is from a separate regression of the variable indicated in the column header on a binary variable for whether the linking method matched the observation. Standard errors are reported in parentheses below. * indicates statistically significant at the 10-percent level, ** at the 5-percent level, and * at the 1-percent level.