Internal Migration in the United States: Rates, Selection, and Destination Choice, 1850–1940

ARIELL ZIMRAN

I study native-born white men's internal migration in the United States over all possible 10- and 20-year periods between 1850 and 1940. Inter-county migration rates—after implementing a new method to correct for errors in linkage—were stable over time. Migrant selection on the basis of occupational status was neutral or slightly negative and also largely stable. But the orientation of internal migration changed over time, declining in distance and increasingly driving urbanization. In the 1930s, migration became less common and less urban oriented. These results provide a clearer understanding of historic U.S. internal migration than previously possible.

Internal migration is fundamental to the American narrative. It has been seen for centuries as a tool for individuals to improve their economic situation (de Tocqueville 1835 [2000]; Turner 1921; Ward 2022), with "great opportunities [lying] just over the horizon" (Brooks 2003, p. WK15). It was also a fundamental driver of the United States's transformation into the world's industrial and economic powerhouse, providing an avenue for labor to reallocate to the most productive sectors (Caselli and Coleman 2001; Kuznets 1966).[1] Because of its importance on both the micro and macro scales, internal migration has been the subject of large literatures in economics, economic history, and history, which have

[1] See also Lewis (1954). Conversely, Cheremukhin et al. (2017) and Hayashi and Prescott (2008) argue that restrictions on internal migration were responsible for delaying growth in Russia and Japan, respectively.

provided insights into the nature, causes, and effects of internal migration in specific contexts in U.S. history.[2] But constraints on the data previously available to study internal migration have severely limited scholars' ability to study its long-run patterns, leaving significant and fundamental blind spots in economic historians' understanding of this formative phenomenon in U.S. economic and social history.

This paper documents and describes, for the first time, the rates of, selection into, and destination choice patterns of native-born white men's inter-county migration in the United States over the period 1850–1940. My analysis is enabled by recent advances in the availability of complete-count data from the U.S. censuses of this period (Ruggles et al. 2021) and in the technology by which to make links between them (Abramitzky et al. 2021a; Bailey et al. 2020). Building on these advances, I construct 13 datasets linking native-born white men, aged 18–40 when first observed, over all possible 10- and 20-year spans in the period 1850–1940.[3] These datasets enable me to overcome limitations faced in prior studies of the long-run trends in U.S. internal migration: only with linked data is it possible to observe inter-county migration, to separate the flow of migration from its stock, and to measure the selectivity of migration.

I find that the rates of inter-county migration of the native-born adult white male population were remarkably stable between the 1850s and the 1920s, at about 33 percent for 10-year spans and about 40 percent for 20-year spans. Selection into migration on the basis of occupational rank was also largely constant over time, with migrants either neutrally or slightly negatively selected.

This constancy contrasts with substantial changes over time in the orientation of internal migration, coming from changes in internal migrants' origins and destination choice patterns. Both the deterrent effect of distance in destination choices and the relative attractiveness of the west increased over the study period. At the same time, the average distance of a move declined, and intra-state moves grew to comprise a

---

[2] By this, I mean that papers tend to focus on internal migration surrounding a particular event, in a relatively short span of time, or in a specific region within the United States. Studies of specific instances of internal migration in U.S. history include Boone and Wilse-Samson (2023), Boustan (2009, 2010), Boustan, Fishback, and Kantor (2010), Boustan, Kahn, and Rhode (2012), Caselli and Coleman (2001), Collins and Wanamaker (2014, 2015), Derenoncourt (2022), Eli, Salisbury, and Shertzer (2018), Eriksson and Niemesh (2016), Ferrie (2005), Fishback, Horrace, and Kantor (2006), Fouka, Mazumder, and Tabellini (2022), Hatton and Williamson (1992), Hornbeck (2012, 2023), Hornbeck and Naidu (2014), Long and Siu (2018), Rosenbloom (2002), Salisbury (2014), Sichko (2024), Steckel (1989), Stewart (2006, 2009, 2012), Stuart and Taylor (2021), Vigdor (2002), and Ward (2022).

[3] Spans beginning or ending in 1890 must be excluded because the vast majority of the 1890 census records were destroyed by fire.

greater share of inter-county moves, implying that a focus on inter-state moves alone would miss an increasingly large share of migration.[4] Most strikingly, the relationship between internal migration and urbanization changed over my study period. Urbanites were initially more likely than observationally similar ruralists to migrate, but by the twentieth century were either less likely or approximately as likely, depending on the definition of an urban place. The attractiveness of urban areas as destinations for internal migrants also increased over time. The combination of these patterns resulted in a steady increase from the beginning of my study period through the 1920s in the degree to which internal migrants' increase in urbanization over a linkage span exceeded that of stayers. That is, internal migration increasingly became a force driving the urbanization of the economy.

The 1930s marked a change in these patterns in all respects. Selection on the basis of both occupational rank and initial urban status was moderated relative to the earlier twentieth century. More dramatically, a substantial decline in the rates of inter-county migration occurred for the first time, with 10-year migration rates declining by nearly 8 percentage points, or about 25 percent. This decline was coupled with a substantial retrenchment in the degree to which the urbanization growth or labor demand growth experienced by internal migrants exceeded that of stayers as urban areas became less attractive as destinations. That is, whereas internal migration was a force driving urbanization in the earlier parts of the twentieth century, this was not true in the 1930s.

The contribution of this paper is predicated on the advantages arising from the newfound ability to make links between all complete-count U.S. censuses from 1850 to 1940. But recent scholarship (e.g., Abramitzky et al. 2021a; Bailey et al. 2020) has brought attention to potential bias due to false links arising from automated linking methods. This challenge is particularly apposite in studying internal migration because any incorrect match will, in all likelihood, appear as an observation of inter-county migration. This challenge will spuriously increase observed migration rates, conflate selection into migration with selection into false matching, and confound true destination choice patterns with spurious ones generated by false matches. To address this concern, I repeat the main results with alternative matching methods of various strictness and draw only conclusions that are robust to the choice of method. For my estimates of the rates of inter-county migration, which are most sensitive to this danger, I also propose a method to estimate the rate of false matches

---

[4] Consistent with this suspicion, I find that inter-state migration rates declined over time.

directly and to correct my estimates for them. The principle of this method is that the difference between two estimates of the same quantity, one of which depends on linkage and the other of which does not, is informative of the rate of false linkage. Specifically, I use information on the ages and birthplaces of children in the household to generate an alternate measure of inter-state migration that does not require linkage (Collins and Zimran 2019; Rosenbloom and Sundstrom 2004), though it can be applied only to a select sample of individuals.[5] Comparing this measure to that arising from linkage enables me to estimate the rate of false matches for each linkage method and to correct my estimated inter-county migration rates. The resulting estimates are largely invariant to the strictness of the linkage method.

   The main contribution of this paper is to update, deepen, and expand existing descriptions of U.S. internal migration in the period 1850–1940 (Ferrie 1997a, 2006a, 2006b; Hall and Ruggles 2004; Rosenbloom and Sundstrom 2004).[6] Prior work in this vein has relied on unlinked census data and on information on individuals' state of birth. It has therefore only been able to quantify the stock of inter-state migrants or the rate of inter-state migration of individuals with young children over spans of 10 years or shorter, and has been extremely limited in its ability to describe the selection and destination choice patterns of migrants. By using linked census data, I am able to provide the first comprehensive description of the rates, selection, and sorting of U.S. inter-county migration. Indeed, given the constraints on prior studies, this is the first description of the rates, selection, and sorting of internal migration flows at any geographic level that is not limited to families with children. On the whole, my findings provide an entirely new view of U.S. internal migration, documenting, for the first time, facts that are interesting and important for their own sake and for the better understanding of U.S. history that they provide. Importantly, the story of internal migration arising from my analysis differs from the one arising from earlier studies of U.S. internal migration over this period—constant rather than declining in frequency before the 1930s and neutrally or slightly negatively rather than positively selected—providing a different interpretation of U.S. internal migration and its interaction with the development of the U.S. economy.

   The main limitation of this contribution is that I constrain attention to native-born white men. This is in part necessitated by the shortcomings

---

[5] I show that selection into this sample is unlikely to result in meaningful bias in estimated false match rates.

[6] See also Adams and Kasakoff (1985), Eldridge and Thomas (1964), Oberly (1986), and Villaflor and Sokoloff (1982).

of the data. Name changes at marriage make linkage impractical for women,[7] and the vast majority of the black population was enslaved in 1850 and 1860 and therefore was not included in the census and cannot be linked. Beyond data constraints, I exclude immigrants from the analysis because Zimran (2022c) has already studied immigrants' internal migration in the United States in general and in comparison to that of natives.[8] I exclude black men from the analysis, although in principle they could be included from 1870 onward, for three main reasons. First, doing so enables me to restrict attention to the same population throughout the study period. Second, internal migration patterns among blacks have already received substantial scholarly attention, including with linked census data.[9] Finally, as Zimran (2022c) shows for immigrants, the unique questions surrounding the internal migration of blacks and the comparison of these patterns to those of whites merit attention beyond the scope of the present paper.

In establishing fundamental facts regarding long-run patterns in the internal migration of native-born white males, this paper also adds to several literatures beyond that seeking to describe long-run trends in internal migration. First, it contributes to the large literature studying specific instances of internal migration in the United States, such as frontier migration (Ferrie 1997b; Stewart 2006), the Great Migration of African Americans (Collins and Wanamaker (2014, 2015), and Dust Bowl and Great Depression migration of the 1930s (Boone and Wilse-Samson 2023; Fishback, Horrace, and Kantor 2006; Gutmann et al. 2016; Hornbeck 2023; Long and Siu 2018; Sichko 2024). Although it has been possible to study these specific instances of migration in detail, limited understanding of the broad patterns of U.S. internal migration implies that the context into which the findings of this literature fit has not been clear. This paper brings this backdrop into sharper focus. This paper also dovetails with papers describing the rates of, selection into, and sorting patterns of modern internal migration (e.g., Greenwood 1975; Jia et al. 2023; Molloy, Smith, and Wozniak 2011; Sprung-Keyser, Hendren, and

---

[7] It is not impossible, however, to link women. Price et al. (2019) show that this is possible using family tree information (see also Abramitzky et al. 2023), but, at the time of writing, these methods have not been applied to the nineteenth century. The ability to include women in the analysis is also one of the advantages of studies using unlinked census data to study internal migration (e.g., Hall and Ruggles 2004).

[8] See also Aaronson, Davis, and Schulze (2020) and Abramitzky, Boustan, and Connor (2020).

[9] See, for instance, Boustan (2009), Collins (1997), Collins and Wanamaker (2014, 2015), Derenoncourt (2022), Eldridge and Thomas (1964), Fouka, Mazumder, and Tabellini (2022), Hornbeck and Naidu (2014), Stuart and Taylor (2021), and Vigdor (2002). This is not to say that a systematic linked data-based analysis of black men's internal migration is not called for, but it is beyond the scope of the current paper.

Porter 2022). In combination with them, it enables a description of internal migration patterns in the United States over nearly 175 years. Moreover, given the weight assigned to internal migration in the literatures on U.S. economic growth (Kuznets 1966) and on intergenerational mobility (e.g., Ward 2022), the clearer understanding that this paper provides of internal migration adds to these literatures.

Finally, in exploiting linked data, seriously considering the biases that might characterize them, and proposing a method to estimate and adjust for bias due to linking error, this paper also relates to recent work taking advantage of the increased ease in making links across complete-count censuses to better describe patterns in mobility in U.S. history, broadly defined to include intergenerational mobility (e.g., Pérez 2019; Ward 2023) and immigrant assimilation (e.g., Abramitzky et al. 2021b; Collins and Zimran 2019, 2023; Zimran 2022c) in addition to internal migration.

Ultimately, this paper addresses basic questions with relatively simple answers. But these questions and the new answers that I provide are fundamental and essential to a complete economic history of internal migration in the United States specifically and to a complete economic history of the United States more generally.

BACKGROUND

Most of what is known about U.S. internal migration before 1935 is based on the census's question on individuals' state of birth, which, through the census of 1930, was the only systematically available information on internal migration. The simplest application of these data uses information in census publications to determine the share of the population in each census year that lived outside of the state of birth—that is, the stock of inter-state migrants. Ferrie (1997a, 2006a, 2006b) and Hall and Ruggles (2004) report the results of such an analysis.[10] They find that the stock of white male inter-state migrants was effectively constant (as a share of population) from 1850 to 1940. Some improvement over these tabulations is possible with census microdata. These enable a focus on particular age cohorts, reducing the impact of concerns such as changing age composition.[11] Hall and Ruggles (2004) and Rosenbloom and Sundstrom (2004) do this, finding a strong decline in the inter-state migration rates of native-born white men throughout the nineteenth

---

[10] Eldridge and Thomas (1964) also use these data to compute estimated net population movements.

[11] For instance, an increase in life expectancy can cause the stock of internal migrants to increase since the elderly would have had a longer time period in which to potentially migrate.

century, with a slight increase from 1900 to 1920 for the young and a continuing decline in this period for the old.

The foremost shortcoming of analyses of this type is that they describe only the stock of inter-state migrants, not the flow.[12] Relatedly, if an individual moved several times between birth and observation, only one move would be observed. As a result, conclusions drawn from these data regarding migration are largely incomparable to those in, for instance, the literature on the Age of Mass Migration (Abramitzky and Boustan 2017; Hatton and Ward 2019; Hatton and Williamson 1998) or the literature on modern internal migration (e.g., Jia et al. 2023; Molloy, Smith, and Wozniak 2011; Sprung-Keyser, Hendren, and Porter 2022). Rosenbloom and Sundstrom (2004) attempt to overcome this constraint. They use information on the birthplaces and ages of children in order to determine whether their parents had moved over the prior decade, enabling them to observe migration rates rather than stocks. The main limitation of this approach is that it can be applied only to families with young children and can be used only for relatively short spans due to the tendency of children to leave their parents' household around age 18.[13] This analysis, like those focusing on individuals' birth state-residence state comparisons, yields evidence of a sharp decline in the migration rates of native-born white men through the nineteenth century, followed by a slight increase through the twentieth. Estimates of internal migration based on birth state-residence state comparisons also do not permit the observation of intra-state moves. Such moves are likely to be particularly important in studying rural-to-urban migration (e.g., Department of Commerce 1933, p. 135; Ferrie 2005) and thus to shedding light on the role of such flows in U.S. development.

The available census data also limit what can be learned even about the moves that can be observed because individuals are observed only after any migration has taken place. It is, therefore, not possible to study selection into migration on any but the most basic characteristics, or without assuming (implicitly or explicitly) that post-migration education and occupational information reflect an individual's pre-migration characteristics. Hall and Ruggles (2004) and Rosenbloom and Sundstrom (2004) interpret the available information to indicate positive selection into inter-state migration throughout U.S. history. The ability to study migrants' destination choices is also limited. While individuals' destinations can,

---

[12] A "synthetic cohort" method can observe the net migration flow of a particular group, but not gross migration. Gross migration, however, is the typical interest of the economics of migration.

[13] Collins and Zimran (2019) use a similar approach to divide Irish immigrants into those arriving before or during the Great Irish Famine, but by focusing on the outcomes of children rather than adults, they do not face the same representativeness issues.

of course, be observed, the absence of detailed data on the prior place of residence limits the extent to which the distance of a move can be determined, meaning that it is difficult to determine the drivers of destination choice while accounting for the cost of migration. Moreover, the inability to determine the timing of migration means that conditions in the destination and potential alternative destinations at the time that the move occurred are not known.

All of these limitations can be overcome by using linked census data, which provide an alternative way to measure individuals' internal migration simply by comparing their places of residence in the initial and final census. Such data make it possible to bound the timing of a move, meaning that the flow rather than the stock can be observed, and the finer residence data enable the observation of intra-state inter-county moves. The pre-migration information available in the initial census enables the direct measurement of migrant selection and the determination of the distance of the move.

Several studies have exploited such data to study internal migration.[14] Steckel (1988, 1989), for instance, uses data on census records linked between 1850 and 1860 to analyze westward migration. A number of other scholars (e.g., Collins and Wanamaker 2014, 2015; Ferrie 1999; Long and Siu 2018; Stewart 2006, 2009) also study the rates, selection, and sorting of specific instances of internal migration with linked data. But until recently, it was not possible to construct linked datasets with sufficient coverage to reveal the broad, long-term patterns of U.S. internal migration. As a result, there is no systematic study of inter-county migration in the United States for the period before 1935, and even what is known about internal migration is extremely limited as a result of the constraints described earlier.

Fully digitized data on the characteristics of every individual in every census have only recently become available, enabling for the first time linkage of the white male population as a whole from 1850 to 1940 and equipping me to improve on the limitations of existing research.

Beyond Native-Born White Men

To maintain comparability to my empirical analysis, the preceding discussion has focused on the internal migration of native-born white

---

[14] The earliest iterations of this approach are the so-called "community studies," which infer migration through the failure to locate individuals observed in a particular location in one census in that same location in subsequent censuses (e.g., Coleman 1962; Curti 1959; Galenson and Pope 1989; Malin 1935; Throne 1959).

men. The existing literature shows, however, that internal migration patterns for other groups may have been different in important ways.

The group most overlooked by internal migration studies based on birth state-residence state comparisons is immigrants, for whom birthplace information is not informative of prior place of residence in the United States. Zimran (2022c) uses linked census data covering 1850–1930 to study immigrant men's internal migration over the Age of Mass Migration and to compare it to that of natives. He finds that immigrant men were at least as likely as non-southern native-born white men to make an inter-county move, were more likely than natives to remain in and move to urban areas, and, like natives, became increasingly attached to urban areas over time.

Differences have also been documented between the internal migration of black and white men through the course of U.S. history, and in particular during the Great Migration (1910–1965). Unlinked census data reveal a substantially stronger pattern of rising migration rates for blacks after 1910 than for whites (Hall and Ruggles 2004), and it has been argued that the end of mass European immigration was responsible for this surge (Collins 1997). This movement was, as I will find for whites, slowed in the 1930s due to the Great Depression (e.g., Derenoncourt 2022). Collins and Wanamaker (2014, 2015) have studied the peak of the Great Migration in the pre-1940 period using data linked between the censuses of 1910 and 1930. They find that selection into migration for both black and southern white men was neutral, and that, in their destination choices, blacks were "more deterred by distance, attracted to manufacturing, and responsive to labor demand" (Collins and Wanamaker 2015, p. 947) than were whites.

Perhaps the main advantage of unlinked census data in the study of internal migration is the ability to include women in the analysis, though their experience is generally not emphasized. Hall and Ruggles (2004, figure 3, p. 836) show that inter-state migration stocks of white and black women evolved similarly to those of men of the same race, though they were always slightly lower, indicating somewhat less frequent inter-state migration for women.[15]

On the whole, the limited application of linked census data to the study of U.S. internal migration prior to 1940 leaves a number of fundamental questions unanswered. This paper takes a first step toward answering them.

---

[15] Rosenbloom and Sundstrom's (2004) household classification method can also be applied to women, but in their analysis, Rosenbloom and Sundstrom (2004) restrict attention to families, meaning that the findings for men and women are the same by construction.

Modern Internal Migration

In contrast to the limited picture that exists of internal migration before 1935, internal migration since then is extremely well documented due both to the prior-place-of-residence question that has been included in the census since 1940 and to the availability of more detailed data, such as the CPS, the ACS, and tax records, some of which directly link individuals' residences over time. As in the historical literature, there is a substantial body of work studying specific instances of internal migration in the United States or using internal migration as a setting in which to study other questions.[16]

There is also a literature describing the basic characteristics of U.S. internal migration, though the more comprehensive data in this regard imply that measurement is more straightforward than in historical settings. The broad findings in this regard in terms of rates, selection, and sorting are summarized by Jia et al. (2023) and Molloy, Smith, and Wozniak (2011). The modern literature largely begins around 1980 and shows evidence of a decline in internal migration rates since then. Combined with results from Ferrie (2006a, 2006b), Hall and Ruggles (2004), and Rosenbloom and Sundstrom (2004), which describe the intervening decades, the post-WWII picture of internal migration is of an increase until about 1980, followed by a decrease since then. There is also evidence of positive selection into internal migration (Wozniak 2010). By creating the first comprehensive series of migration rates (rather than stocks) with known timing for the period before 1940, this paper enables, for the first time, the dovetailing of modern and historical internal migration rate series and, therefore, the construction, for the first time, of a series of internal migration rates spanning nearly 175 years of U.S. history.

DATA

My analysis is based on 13 datasets making all possible 10- and 20-year links between the U.S. censuses of 1850–1940.[17] I begin the analysis in 1850 because that year's census was the first to enumerate the entire free population. The analysis ends in 1940 because, at the time of writing, this was the most recent complete-count census that had been fully digitized. The 1870–1890, 1880–1890, 1890–1900, and 1890–1910 spans

---

[16] See, for instance, Amior (2023), Conway and Rork (2016), Huang, South, and Spring (2017), Huttunen, Møen, and Salvanes (2018), Jauer et al. (2019), Kennan and Walker (2011), Molloy, Smith, and Wozniak (2017), Oswald (2019), and Treyz et al. (1993).

[17] Online Appendix B.1 presents rates of successful linkage.

are omitted because the vast majority of the 1890 census records were destroyed by fire and are unavailable for linkage or analysis.

I created these datasets by merging complete-count census records provided by Ruggles et al. (2021) with the "basic" linkage crosswalks provided by Zimran (2022a).[18] Zimran's (2022a) method, which implements suggestions made in the literature (Abramitzky et al. 2021a; Bailey et al. 2020) to increase the quality of matches by requiring the uniqueness of individuals in a particular age band and using an orthographic distance measure to compare the names of potential matches, provides my preferred links, referred to as the *Main* links in the analysis later. The datasets created by merging the census records with the linkage crosswalks provide information on an individual's county of residence in each of the two censuses, which enables me to determine whether an individual made an inter-county move between them.[19]

Throughout my analysis, I restrict attention to native-born white men aged 18–40 in the initial census. Beyond the restriction to native-born white men discussed previously, the restriction to those aged 18–40 in the initial census is intended to ensure that men are observed in the labor force in the initial census of the linkage span (to enable construction of the occupational rank measure) while ensuring that they are also not so old that mortality is an important concern. It also limits attention to ages in which there are likely to be children in the household at the time of the final census, enabling the alternate household structure-based method for estimating migration and thus for estimating false match rates.[20]

There are two main concerns that arise in the use of linked census data to study internal migration. The first is the danger of false matches—that is, the concern that the linked datasets may not actually describe the same person in the two census years. This concern has been highlighted recently by Abramitzky et al. (2021a) and Bailey et al. (2020) and touches on all aspects of the analysis of internal migration.[21] Since nearly all false

---

[18] Zimran (2023) provides data and code to replicate the results presented in this paper.

[19] To address changing county boundaries, I use county boundary shapefiles from Manson et al. (2019) to define a move to have occurred if an individual is observed living in a final-year county that does not overlap geographically with his initial-year county. An example of how this definition is applied is given in Online Appendix Figure A.1.

[20] If, for instance, the range were extended to age 50, it would be necessary to use individuals as old as 64 in the latter census to make estimates of household structure—an age by which many would likely no longer have children in the household.

[21] The concern is also different than in studies of specific cases of internal migration. For instance, in these studies, the goal is often not measurement but the establishment of causal relationships. In these cases, the presence of linkage errors attenuates estimates, strengthening conclusions that certain relationships existed. In this paper, the goal is measurement, meaning that the challenge coming from false matches is more fundamental.

matches are to individuals living in a different county, the observed rate of migration conflates true migration with false matches. Selection into migration is also, therefore, conflated with selection into false matching. Finally, under the assumption that a false match links an individual in one census to a random individual in a subsequent one, true destination choice patterns are conflated with a tendency for false matches to show spurious migration toward more populous areas.

I address the danger posed by false matches in two ways. First, I repeat all of my analysis using data sets constructed by four alternate linkage methods, drawing only conclusions that are robust to the choice of linkage method. Two of the alternative linkage methods (referred to as *ABEE* and *ABEN*) are simply different from the main method in their linkage parameters, requiring uniqueness of links in a somewhat smaller age band than my preferred linkage method and requiring either an exact match of names or a match of the NYSIIS standardization of names rather than making an orthographic distance comparison. The other two linkage methods (referred to as *Int* and *Int+*) are stricter, reducing the danger of false matches at the potential risk of a less representative sample by limiting matches to the intersection of the set of matches by the main method and the ABEE and ABEN methods, and additionally by requiring agreement across censuses of information not used in linkage. Online Appendix B.1 presents further details of the samples constructed by these alternative linkage methods. In studying the rates of internal migration, I also propose a method, described in detail in the next section, to estimate the rate of false matches for each linkage method and to correct my estimated internal migration rates.

The second main concern in the use of linked data is that they may not be representative of all individuals at risk for linkage. Indeed, this is one reason why, even though it is tempting to simply use the strictest linkage method to minimize the danger of false matches, I do not do so—this would increase the danger of constructing an unrepresentative sample. To address this concern, I reweight each linked dataset so that its observable characteristics match (as closely as possible) the distribution of observable characteristics of those at risk for linkage in the initial census. Online Appendix B.1 provides details.

The initial census of each span also provides data on a variety of individuals' pre-migration characteristics, including occupation, literacy, and initial urban status, which I use to study migrant selection. Literacy, though flawed and potentially changing in definition over time, is the only consistently available measure of human capital. Occupation is the only measure of economic status that is available in a consistent way over the complete 1850–1940 period. As described in Online Appendix B.2, I use Ruggles

et al.'s (2021) occupational codes to construct a measure of an individual's occupational rank relative to the white male population aged 18–64 in each census. To determine whether an individual resided in an urban area in the initial and final census of each linkage span, I use the official Census Bureau definition of an urban place as one with at least 2,500 inhabitants, as well as an alternative definition using a population cutoff of 25,000.

Finally, I construct a number of measures of the characteristics of an individual's move and of his counties of initial and final residence. These include the distance of the move and whether it crossed state or regional (i.e., census divisions) boundaries; the share of each county's population residing in urban areas under various definitions; and a Bartik (1991)-type measure of labor demand growth. I address changing boundaries using Hornbeck's (2010) method.

INTER-COUNTY MIGRATION RATES, 1850–1930

I begin by answering the most basic, but perhaps the most fundamental, question about U.S. internal migration—what was the rate of inter-county migration? That is, how likely was it that an individual observed in a given census year would move to a different county over the next 10 or 20 years?

Uncorrected Estimates

Figure 1 presents the estimated uncorrected rates of inter-county migration over 10- and 20-year spans. The results vary depending on the linkage method—the stricter methods show uniformly lower rates and the patterns over time are also different, with the more permissive methods finding a slight decline over time before a sharper decline in the 1930s, and the stricter methods showing general stability before a final decline. However, the danger of bias from false matches challenges the validity of these results.[22]

Quantifying and Correcting for False Matches

To address concerns about spurious migration due to false matches, I propose a method to estimate the rate of false matches and to correct for the bias that they induce.[23] The principle of this method is that, where a quantity can be estimated for the same sample by two methods—one

---

[22] Note, however, that more than half of individuals are found living in the same county in both censuses, which could not occur without some validity in linkage: if all matches were random, the share of stayers would be below 1 percent.

[23] Zimran (2022c) also applies this method.

(a) 10-year inter-county migration
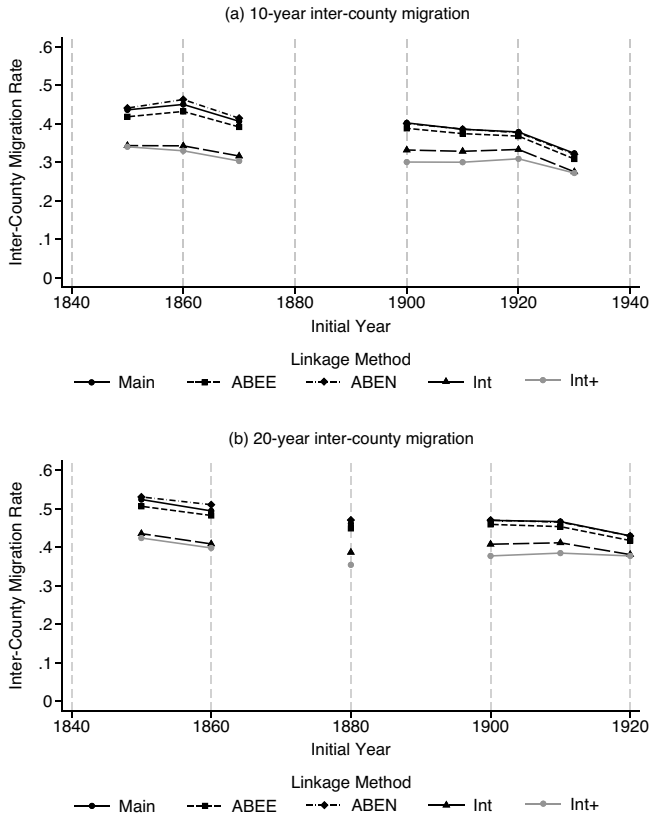
(b) 20-year inter-county migration

FIGURE 1
INTER-COUNTY MIGRATION RATES BY LINKAGE METHOD AND SPAN

Note: Each figure shows the probability that an individual in the linked sample beginning in the year on the x-axis was observed living in a different county 10 or 20 years later, according to each linkage method. All observations are weighted by inverse linkage probability. No correction is made for false matches.
Source: See the text.

based on linkage and the other not—the difference between the two estimates is informative of the rate of false matches.

In this case, for individuals in the linked sample whose household structure enables it, I use the birthplace and age composition of children in an individual's household to create a non-linkage-based measure of inter-state migration over the previous 10 years.[24] Movers are defined

---

[24] In principle, a similar analysis could be done with outcomes other than inter-state migration. For instance, it could be done by comparing the share of individuals whose father's birthplace changes over censuses. In principle, this should not happen except in cases of errors in linkage. Of course, one complication is that this information is not available in all censuses. Indeed, except for race, which may change over censuses for reasons other than linkage error (Abramitzky et al. 2023; Dahis, Nix, and Qian 2019), no corroborating variable is consistently available across censuses, further motivating my use of inter-state migration.

as those with a child aged less than 10 years old born in a different state than the current state of residence and no older child born in the state of residence. Stayers are defined as those with a child born in the state of residence at least 10 years old and no children younger than 10 born in a different state. This categorization is performed for the latter census of each 10-year span (i.e., in 1860 for the 1850–1860 span) for the linked sample; it is not applied to 20-year spans because children are unlikely to remain in their parents' household for over 20 years. This procedure results in a subsample of the linked dataset—which I refer to as the *corroboration sample*—composed of individuals whose inter-state migration status over a 10-year span can be measured in two ways—by comparing their state of residence in the initial and final censuses and according to their household composition.[25]

According to the law of total probability, the probability that an individual in the corroboration sample is observed to have made an inter-state move (whether he truly moved or not) according to his residence state in the initial and final census (i.e., according to linkage), which I denote as P(moved state), can be written as

$$P(\text{moved state}) = P(\text{moved state}|\text{true match})[1 - P(\text{false match})] \quad (1)$$
$$+ P(\text{moved state}|\text{false match})P(\text{false match}).$$

Rearranging Equation (1), I can express the probability of a false match as

$$P(\text{false match}) = \frac{P(\text{moved state}) - P(\text{moved state}|\text{true match})}{P(\text{moved state}|\text{false match}) - P(\text{moved state}|\text{true match})}. \quad (2)$$

In calculating Equation (2), I use the following quantities from the corroboration sample. As P(moved state), I use the linkage-based estimated rate of inter-state migration according to the comparison of the initial- and final-year residence state. As P(moved state|true match)—the true probability of an inter-state move—I use the estimated rate of inter-state migration according to the household composition method. Finally, as P(moved state|false match)—the probability of observing inter-state migration in the case of a false match—I use an individual's birthplace and age to determine the average probability that a person to whom he

---

[25] As with the definition of inter-county migration, I define an individual as having made an inter-state move only if his final-year and initial-year states of residence share no geographic overlap.

could be linked would live in a different state, which is straightforward to estimate.[26]

Figure 2 presents the estimated rates of false linkage for each method and initial census year. The estimates presented in Figure 2 fit expectations. The estimated false match rates are higher for the more permissive linkage methods and decline over time from about 15 percent to under 10 percent by the end of the study period.[27] For the more restrictive methods, the estimated false match rates are initially about 5 percent and fall to approximately zero in the twentieth century.[28]

Having estimated *P*(false match) based on inter-state migration in the corroboration sample, I use this estimate to correct my estimates of inter-county migration for the full linked sample. I rearrange Equation (1) and replace inter-state moves with inter-county moves to yield

$$P(\text{moved county}|\text{true match}) \qquad\qquad (3)$$

$$= \frac{P(\text{moved county}) - P(\text{moved county}|\text{false match})P(\text{false match})}{1 - P(\text{false match})}.$$

To compute this value, I use the following quantities. As *P*(false match), I use the estimate computed from Equation (2); for 10-year spans, I use the estimate from the analogous span; for 20-year spans, I use the estimate for the 10-year span beginning in the same year; there is no 1880–1890 span, and so I must omit the 1880–1900 span. As *P*(moved county), I use the probability of observing an inter-county move in the full linked dataset. For inter-county migration, *P*(moved county|false match) is sufficiently close to one that, with minimal loss, I can write Equation (3) as

$$P(\text{moved county}|\text{true match}) = \frac{P(\text{moved county}) - P(\text{false match})}{1 - P(\text{false match})}. \qquad (4)$$

The estimates coming from computing Equation (4) are my benchmark estimates of inter-county migration rates that are corrected for false

[26] For instance, a 20-year-old Alabama-born man in the 1850 census can be linked, according to my benchmark linkage method, to any Alabama-born man aged 26–34 in the 1860 census. To determine the probability that a false match would create an observation of inter-state migration for this man, I determine the fraction of Alabama-born 26-34-year-old men in 1860 who lived in a different state than the man originally lived in in 1850. Performing analogous calculations for all men in the sample and averaging the results yields my estimate of *P*(moved state|false match).

[27] These estimates are similar to estimates of false match rates for similar algorithms computed by Abramitzky et al. (2021a) and Bailey et al. (2020). Zimran (2022c) also finds that the rate of false matches estimated by this method for nineteenth-century immigrants is similar to that estimated by Anbinder et al. (2023).

[28] Imprecision in some cases results in estimates that are below zero by a very small margin. The largest in magnitude is –0.4 percent for spans beginning in 1920 by the *Int+* method.
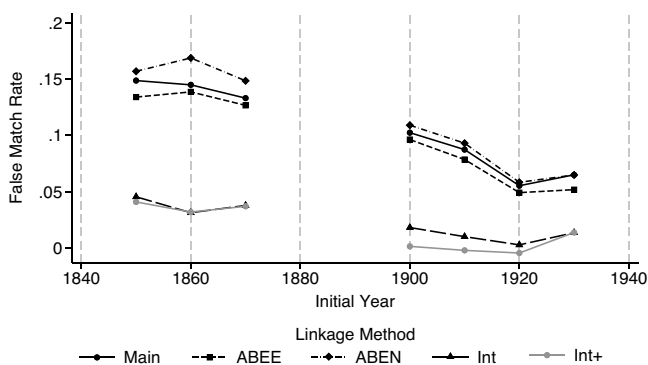
FIGURE 2
ESTIMATED RATES OF FALSE LINKAGE

Note: This figure shows the estimated rate of false matching for each linkage method and 10-year span according to the comparison of inter-state migration estimates by the linkage method and by the method based on the birth places and ages of children. The rate is computed according to the law of total probability, as explained in the text.
Source: See the text.

matches. Despite my efforts to correct these estimated migration rates, it is inevitable that there will remain some error. There are many possible causes. One in particular is that the corroboration sample is not representative of the broader linked sample. This is a somewhat less severe issue than when using the household composition method to directly measure migration (Rosenbloom and Sundstrom 2004). Although the internal migration of men with children is likely different from that of all men, an issue arises for this method only insofar as the probability of a false match differs between these groups. I investigate the degree to which selection into the corroboration sample is potentially problematic in Online Appendix B.3, where I show that such selection is unlikely to have led to biased estimates of false match rates.

This method also relies on the assumption that the true migration rates of the correctly linked and incorrectly linked are the same. Since it is well established that selection into linkage is non-random, it is not likely that this assumption holds in reality—characteristics determining success in linkage likely also determine the propensity to move. The method also assumes that there is no error (on average) in determining individuals' migration status using the household composition method. This method is also not informative as to *which* observations are incorrectly linked since it is used simply to deflate an aggregate quantity.[29] This implies that it cannot be used in more detailed analyses of selection and destination choice.

---

[29] For this reason, I do not study conditional changes in internal migration rates over time, as, for instance, Rosenbloom and Sundstrom (2004) do.
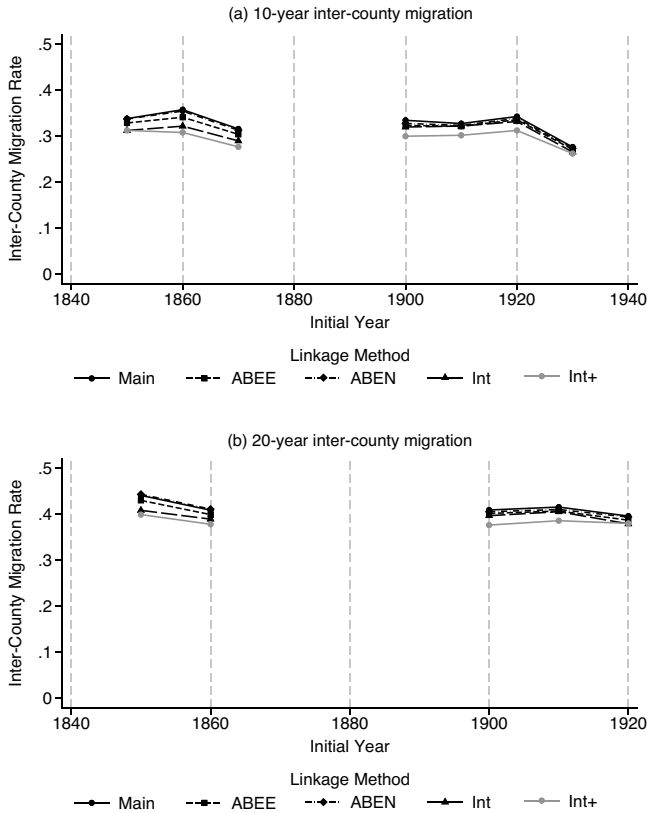
FIGURE 3
INTER-COUNTY MIGRATION RATES BY LINKAGE METHOD AND SPAN,
CORRECTED FOR FALSE MATCHES

Note: Each figure shows the probability that an individual in the linked sample beginning in the year on the x-axis was observed living in a different non-overlapping county 10 or 20 years later, according to each linkage method. All observations are weighted by inverse linkage probability. The estimates are then adjusted for false matches according to the method presented in the text.
Source: See the text.

Corrected Estimates

Figure 3 presents the estimated corrected inter-county migration rates for native-born white men for each linkage method and span. For both 10- and 20-year spans, the resulting predictions of inter-county migration rates and their changes over time are similar across the linkage methods.[30]

---

[30] It is notable that the estimates, even for the less restrictive methods, are similar to those arrived at with the stricter matching methods without correction, suggesting that the danger of unrepresentativeness due to strict matching is minimal. It is, of course, still non-zero. The remaining differences between methods after correction could be due to selection of less-mobile types into the stricter linked sample.

For 10-year spans, the estimated migration rates by my preferred linkage method are consistently between 31.5 and 35.7 percent from the 1850–1860 span to the 1920–1930 span, with no clear trend. I discuss the 1930s in more detail later. The estimated 20-year migration rates are higher, at 39.6 to 44.1 percent, and evolve largely without trend, save for a slight decline between the two spans of the nineteenth century. The main take-away is that the frequency of inter-county migration over the period 1850–1930 was largely constant.[31]

MIGRANT SELECTION, 1850–1930

Characterizing migrant selection is crucial to a basic description of any flow of migration. It provides insight into the potential effects of internal migration on individuals, on the economies of sending and receiving areas, and on the broader economy. By providing information on the pre-migration characteristics of prospective migrants, linked data enable me to delve further into this question than has previously been possible. However, it remains possible that errors in linkage may influence conclusions if the probability of a false match varies along a dimension on which I study selection. I present here the results using only my preferred linkage method, limiting my conclusions to those that are robust to the use of alternate and stricter linkage methods (results in Online Appendix C).

I focus first on two measures of migrant selection—occupational rank and literacy in the initial year of the span—that speak to whether internal migrants were positively or negatively selected. These correspond roughly to the typical focus in the economics of migration on earnings (e.g., Borjas 1987; Chiquiar and Hanson 2005) and education (e.g., Card 2005). To measure selection, I estimate, separately for each linked dataset, the equation

$$y_{it} = \alpha_t + \beta_t r_{it} + \gamma_t \ell_{it} + \delta_t u_{it} + \mathbf{x}'_{it} \varphi + \varepsilon_{it}, \tag{5}$$

where y_{it} is an indicator equal to one if individual i migrated across county lines in span t, r_{it} is individual i's initial-year occupational rank, ℓ_{it} is individual i's initial-year literacy, u_{it} is individual i's initial-year urban residence, and $\mathbf{x}_{it}$ is a vector of initial-year controls. The β_t, γ_t, and δ_t coefficients provide measures of migrant selection in span t, indicating whether

---

[31] Online Appendix Figure A.2 presents analogous results for inter-state migration. A decline in migration rates over time is evident. The nearest existing estimate to which these estimates can be compared come from Rosenbloom and Sundstrom (2004, figure 4.1, p. 309), which estimates 10-year migration rates at just under those estimated by my stricter linkage methods. I am not aware of comparable estimates for my 20-year inter-state rates.

greater occupational rank, literacy, or urban status were associated with a greater probability of migration.[32] To ensure that changes in the β_t, γ_t, and δ_t coefficients across censuses are the product of actual changes in selection patterns rather than of changing availability of controls in the census, and to avoid confounding the interpretation of the coefficients, I limit the vector $\mathbf{x}_{it}$ to the variables listed in Online Appendix Table A.1 that are available in all initial censuses.[33]

Figure 4 presents results on migrant selection on the basis of occupational rank and literacy. Each panel contains two sets of estimates—one with no controls and one with controls, including initial-county fixed effects and the other dimensions of selection.[34] Panels (a) and (b) focus on selection on the basis of occupational rank. Unconditional selection into inter-county migration was consistently negative, moving from very close to zero in the beginning of the study period to more strongly negative in the late nineteenth century and in the twentieth.[35] Such unconditional comparisons, however, ignore differences in age, sector, distance to potential destinations, and other factors likely to influence migration. When conditioning on all observables, including the initial county of residence, selection patterns were more stable over time. For spans beginning in 1850, conditional selection on the basis of occupational rank was zero or positive. A notable decline in spans beginning in 1860 then occurred, indicating more negative selection into migration in this span than earlier. From this point through the 1920s, the coefficients hover around –0.05. With the occupational rank measure ranging from zero to one with a standard deviation of roughly 0.25, these coefficients are small but not negligible: a one-standard deviation increase in occupational rank was associated with about a 1.25-percentage point decline in migration probability on a base ranging from 38 to 52 percent.[36] The

---

[32] Note that, despite the time subscripts in Equation (5), this is not a panel regression. Instead, the subscripts indicate that the equation is estimated separately for each linkage span, with separate coefficient estimates for each.

[33] I do not include every observable in Online Appendix Table A.1, as some of these might confound the interpretation of others. For instance, I do not wish to include both an indicator for farm residence and for urban residence, as these contain largely the same information. Similarly, I do not include multiple measures of occupational status. Ultimately, the controls are age, urban residence, literacy, occupational rank, household size, and whether the individual was a head of household.

[34] These results are based on estimating Equation (5) with u_{it} defined based on the 2,500-person cutoff.

[35] These results are largely insensitive, even in magnitude, to the strictness of the linkage method, meaning that linkage concerns are not a major issue.

[36] The appropriate reference point is the uncorrected migration rate by my preferred linkage method since the data used for the analysis are not corrected for false matches. The more negative coefficient over time is inconsistent with the results being driven by false matches since these became less common over time.
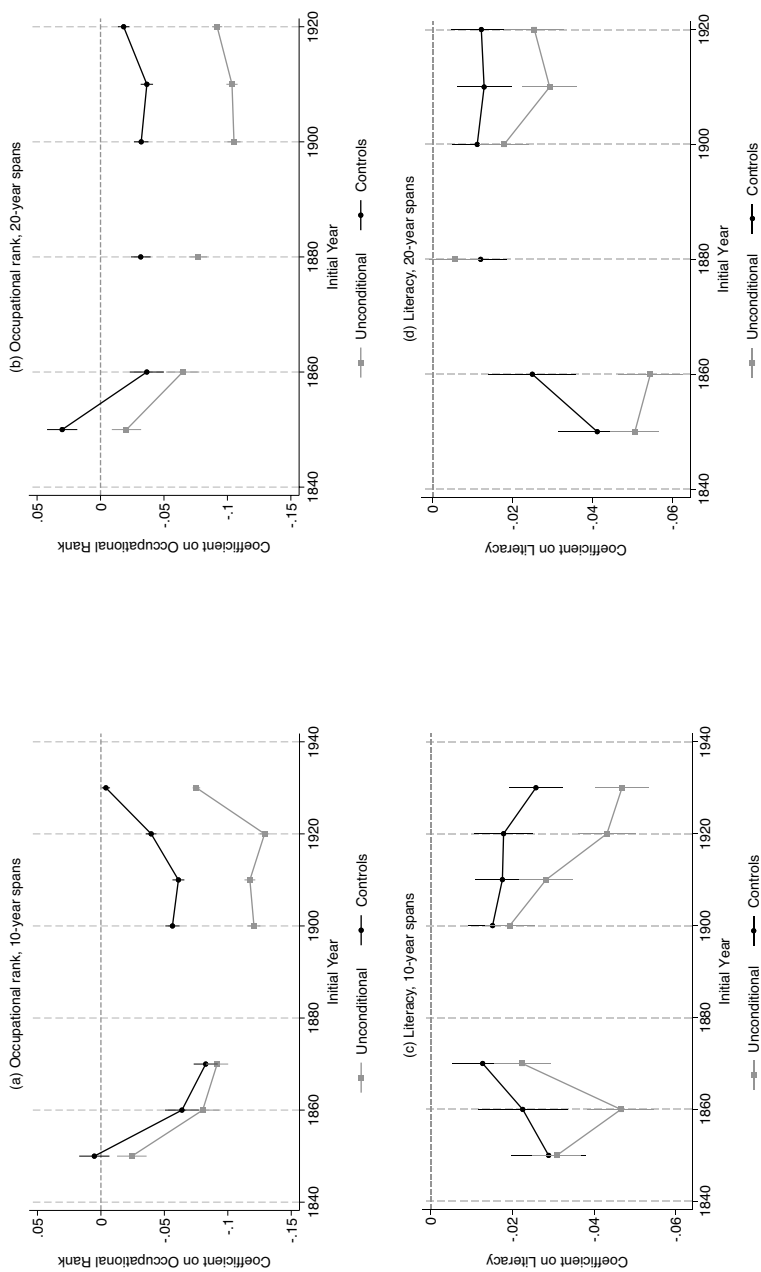
FIGURE 4

MIGRANT SELECTION BY MEASURE AND SPAN, INDIVIDUAL CHARACTERISTICS

Note: Each figure shows the coefficient on the variable in question in a selection regression for migration over the stated span. All observations are weighted by inverse linkage probability. Vertical lines indicate 95-percent confidence intervals. Controls are all controls available in all census years, as well as county fixed effects.
Source: See the text.

variation in these coefficients over time is even smaller.[37] In sum, beginning in the 1860s, individuals of greater occupational rank were somewhat less likely or as likely to make an inter-county move than otherwise similar individuals from the same initial county but of lower rank,[38] and this difference remained largely constant over time.

Panels (c) and (d) of Figure 4 focus on selection into internal migration on the basis of literacy. All estimates indicate that literate individuals were less or as likely as otherwise similar individuals to make an inter-county move (including those for alternate linkage methods). The patterns are similar across 10- and 20-year spans: a decline in the magnitude of selection through the nineteenth century and stabilization in the twentieth. The coefficients are again small but non-negligible, with literate individuals about 2 to 4 percentage points less likely to migrate than illiterates.

On the whole, these results paint a picture of migration that was neutrally or slightly negatively selected, with this selection, by most indications—including on the basis of the best available measure of socioeconomic status—largely constant from the 1860s through the 1920s.[39]

CHANGING ORIGINS AND ORIENTATION, 1850–1930

In this section, I study selection into internal migration on the basis of initial urban residence as well as the destination choices of internal migrants. I find that the results of general constancy in the rates and selection of internal migration conceal important changes in the nature of U.S. internal migration. Again, I present results here for my main linkage method, with results from alternate methods presented in Online Appendix C.

Figure 5 focuses on selection into inter-county migration on the basis of urban residence in the initial year of each span using two different measures of urbanization—the 2,500-inhabitant cutoff (Panels (a) and (b)) and an indicator for being in a city of at least 25,000 (Panels (c) and (d)).[40] For each measure and span, I include two sets of estimates—an unconditional estimate and one that conditions on all controls, but I use

---

[37] For 10-year spans, the variation is no more than 0.043 (about a 1-percentage point change in migration probability for a standard deviation change in rank), whereas for 20-year spans, the variation is no more than 0.011 (a 0.03-percentage point change in migration probability for a standard deviation change in rank).

[38] That is, individuals of greater occupational rank were less likely to move according to the main estimates, but stricter linkage methods show a coefficient closer to zero.

[39] Online Appendix Figure A.3 plots the unconditional and conditional age profiles of migration separately for each linkage span. The age profiles are also largely unchanged over time.

[40] Online Appendix Figure A.4 presents results for population density and urban residence, defined as the 20 largest cities in the country.
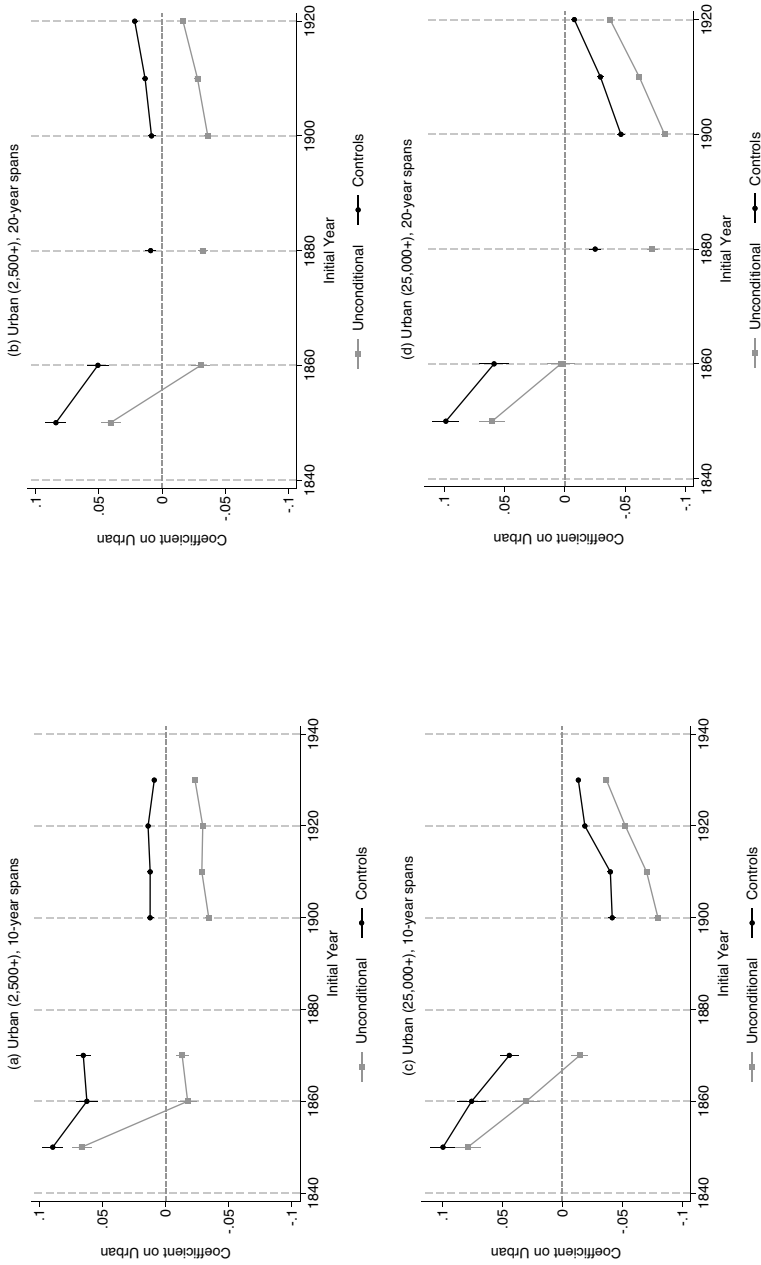
FIGURE 5

MIGRANT SELECTION BY MEASURE AND SPAN, URBAN RESIDENCE

Note: Each figure shows the coefficient on the variable in question in a selection regression for migration over the stated span. All observations are weighted by inverse linkage probability. Vertical lines indicate 95-percent confidence intervals. Controls are all controls available in all census years, as well as state fixed effects.
Source: See the text.

state fixed effects instead of county fixed effects.[41] In spans beginning in 1850, urbanites were 5 to 10 percentage points more likely to migrate than otherwise similar ruralists, depending on the linkage method. This conditional urban migration premium declined over the nineteenth century. By spans beginning in 1880 for 20-year spans or in 1900 for 10-year spans, this pattern reversed, with urbanites approximately as likely (for the 2,500-person definition) or up to 5 percentage points less likely (for the 25,000-person definition) than otherwise similar ruralists from the same state to migrate. Ruralists' conditional migration premium then declined over the twentieth century. Thus, in contrast to selection on the basis of occupational rank and literacy, selection on the basis of urban residence was relatively larger in magnitude, changed throughout the 1850–1930 period, and for larger cities changed in sign such that urbanites were initially more and then less likely than ruralists to move.[42]

Changes over time are also evident in internal migrants' destination choice patterns. Figure 6 focuses on the distance distribution of moves. Panels (a) and (b) divide moves into inter-county but intra-state, inter-state but intra-region, and inter-region. This division is particularly important as the fraction of moves that are intra-state measures the extent to which internal mobility is not observed when using state of birth to determine migration. Inter-county intra-state moves were an important component of U.S. internal migration, accounting for about 40 percent of moves at the beginning of the study period and rising over time to about 60 percent by the end of the study period. This increase came at the expense of inter-region moves, with the share of inter-state but intra-region moves remaining largely constant in frequency over time. Panels (c) and (d) present violin plots for distance of moves. The distribution of move distance has a consistent peak below 50 miles. In the earlier periods, a second peak around 500 miles is also evident and fades over time.

To what extent were these unconditional changes the product of changing individual characteristics over time? Using a pooled dataset of

---

[41] I do not include county fixed effects because many of these measures vary only across counties. Including county fixed effects would lead identification to come from only a small number of observations—that is, from counties that have both urban and rural residents, and not from completely urban or completely rural counties.

[42] There are also changes over time in migrants' origins. In Online Appendix Figures A.5 and A.6, I compute migration rates for individuals living in each county at the start of each 10- or 20-year span. These maps omit any county for which the sample contains fewer than 30 observations. The high-emigration area moves westward over time. The Northeast consistently shows low rates of emigration, whereas migration rates from the South and the West were consistently relatively high. The Midwest shows a change over time, however; its migration rates were relatively high in the nineteenth century but relatively low by the end of the study period. Online Appendix Figure A.7 formalizes these regional comparisons by estimating Equation (5) with region fixed effects.
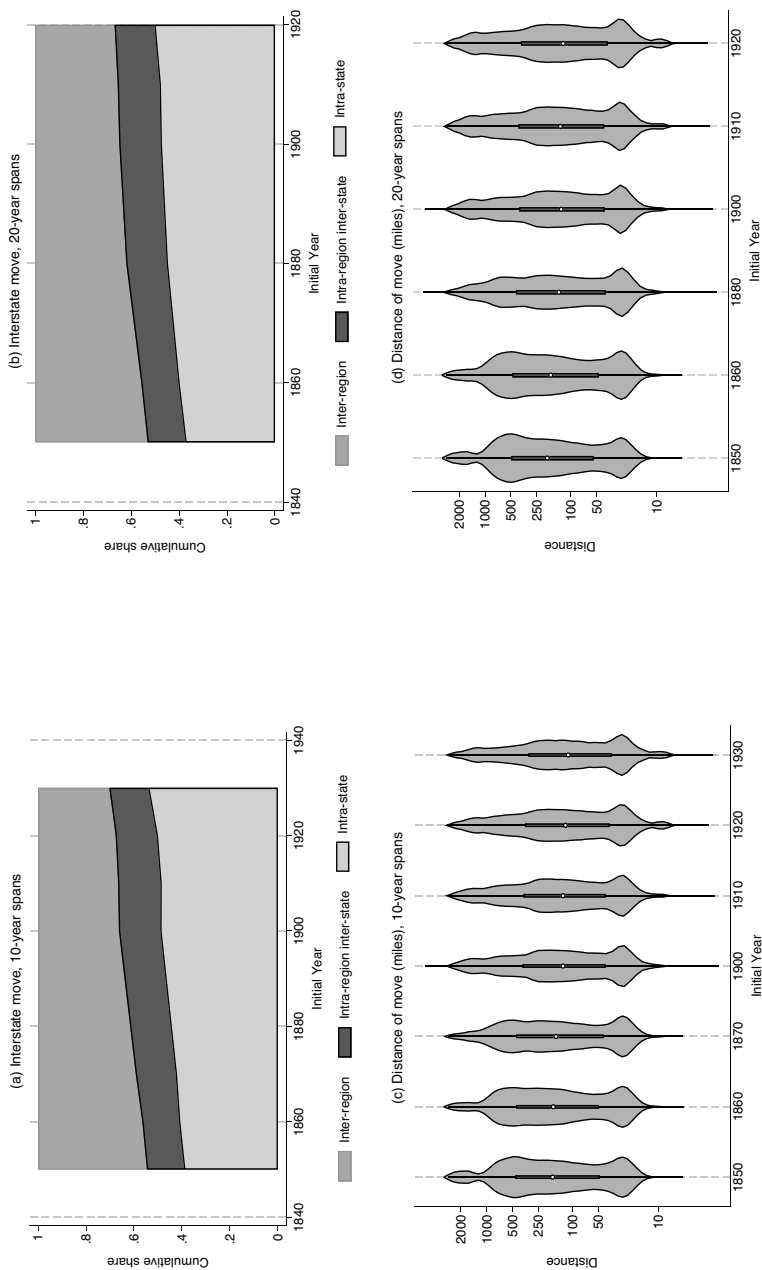
FIGURE 6
DISTANCE OF MOVES

Note: Panels (a) and (b) show the cumulative distribution of intra-state, inter-state but intra-region, and inter-region moves. Panels (c) and (d) present the distribution of move distance for each linkage span with box plots presenting the mean and quartiles.
Source: See the text.

all migrants from all linkage spans of a particular length, I estimate a regression of the form

$$y_{it} = \beta_t + \mathbf{x}'_{it}\delta + \varepsilon_{it}, \tag{6}$$

where y_{it} is some measure of individual i's move distance in span t, β_t is a series of span fixed effects, and $\mathbf{x}_{it}$ is a set of initial-year controls, including all observables available in all census years and either initial-year state or initial-year county fixed effects. Figure 7 plots the β_t resulting from estimating Equation (6) with no controls, with controls and initial-state fixed effects, and with controls and initial-county fixed effects. Omitting the indicator for spans beginning in 1850, these estimates show how migration distance changed over time. Paralleling Figure 6, all measures show a large, unconditional decline in the distance of the move or in the probability of moving across states or regions over time. The decline, however, was moderated by changes in the demographics of migrants, meaning, for instance, that the move made by an individual in 1850–1860 was about 100 miles further than that made by an observationally similar person in 1920–1930.

To determine which characteristics of destinations attracted migrants, I restrict attention to movers and estimate a conditional logit model of the form

$$P_{iotj} = \frac{\exp(\beta_t d_{oj} + \mathbf{z}'_{jt}\delta_t)}{\sum_k \exp(\beta_t d_{ok} + \mathbf{z}'_{kt}\delta_t)}, \tag{7}$$

where P_{iotj} is the probability that migrant i initially from county o in linkage span t chose destination j, $\mathbf{z}_{jt}$ is a vector of characteristics of county j in the initial year of span t, d_{oj} is the distance between counties o and j, and k indexes all potential destinations. The vector $\mathbf{z}$ includes a county's initial-year urbanization (i.e., the fraction of the population living in urban areas) and census-division fixed effects (with New England as the excluded category).[43] The coefficients β_t quantify the deterrent effect of distance in span t, and the coefficients δ_t quantify the attractiveness of the various county characteristics in that period. The coefficients have the usual interpretation of coefficients in a logit model—the marginal effect of the variable in question on the log odds of selecting a particular destination, conditional on its distance; where P_{iotj} is near zero (as is the case for most county pairs), the coefficient can also be interpreted as the

---

[43] I limit the contents of the vector $\mathbf{z}$ to avoid confounding the interpretation of the variables that I do include.
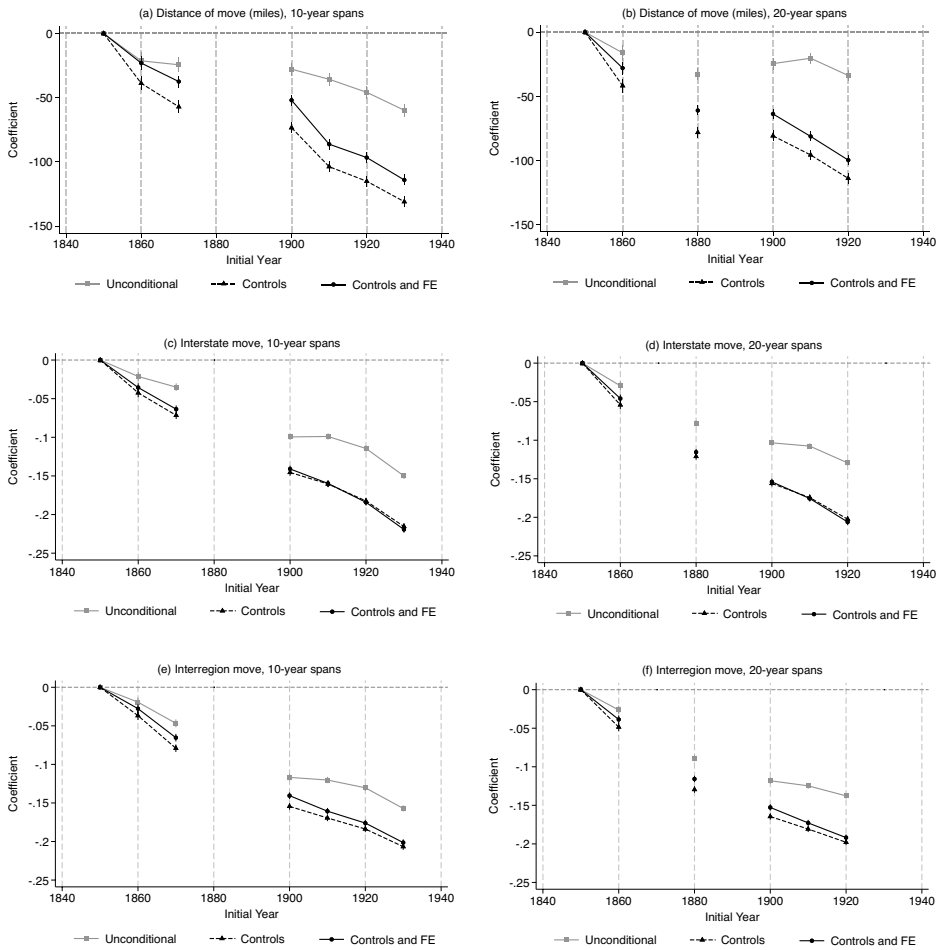
FIGURE 7
CHANGES IN DISTANCE OF MOVES

Note: These figures show the coefficients on initial-year fixed effects in regressions of destination characteristics with and without controls, with 1850 as the excluded year. "Controls" includes all controls available in all censuses, including state fixed effects. "Controls and FE" also includes county fixed effects. Vertical lines are 95-percent confidence intervals. Observations are weighted to correct for selection into linkage.
Source: See the text.

percent change in the probability of choosing a particular destination for a unit increase in the regressor. Online Appendix B.4 presents additional details on the estimation of this model.

Figure 8 presents the results of this estimation. Panels (a)–(d) focus on the deterrent effect of distance and on the attractiveness of urban areas.[44]

[44] Online Appendix Figure A.8 presents results for other measures of urban status, population density, and labor demand growth.
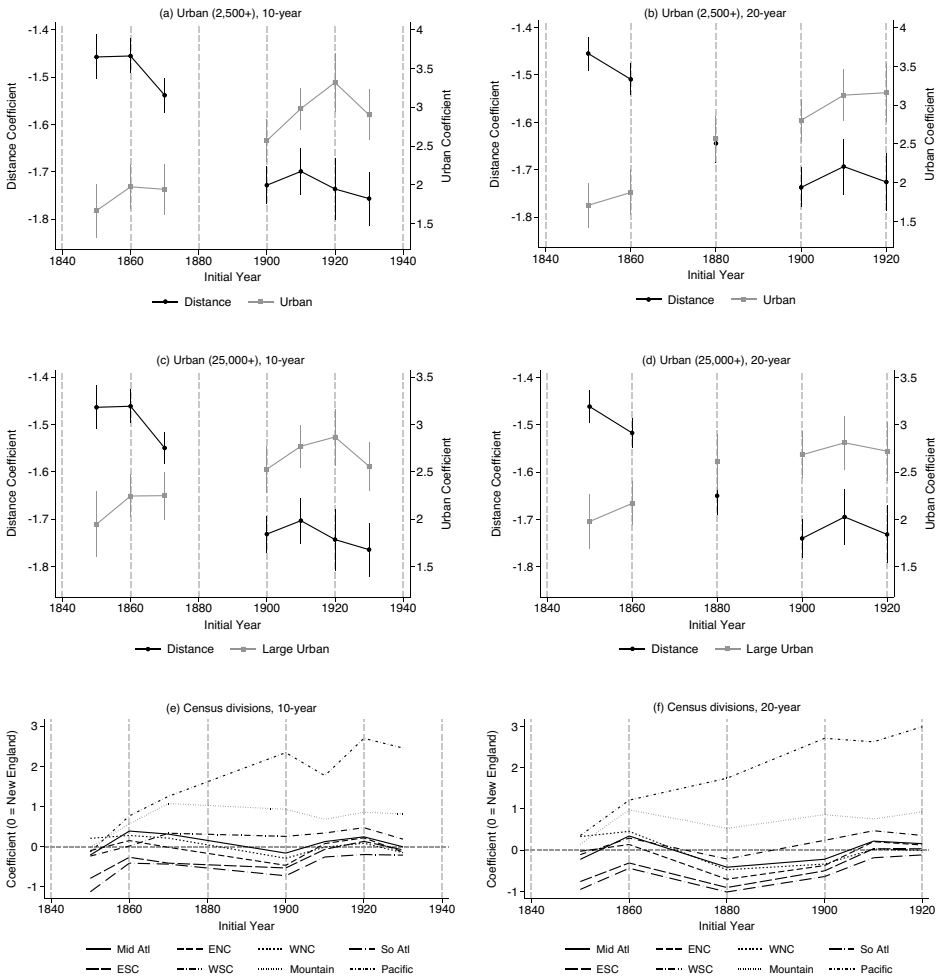
FIGURE 8
CONDITIONAL LOGIT RESULTS

Note: Panels (a)–(d) present coefficients on distance and the measure of urban residence from the conditional logit estimation. Panels (e) and (f) present the coefficients from the regressions of Panels (a) and (b), respectively, on the census division indicators, with New England excluded. Note that there is no data for spans that would include 1890; the lines span this period simply to enable a clearer reading of the trend over time.
Source: See the text.

An increase over time in the deterrent effect of distance is evident from the increasing magnitude of its negative coefficient over time: in the nineteenth century, the elasticity of migration probability with respect to distance was about –1.5; by the twentieth century, it was about –1.7. This result is consistent with the declining distance of moves described earlier. A large increase over time in the attractiveness of urban destinations is also clear through the 1920s. In the nineteenth century, a completely

urban county had a 1.5 greater log odds of being selected than an entirely rural destination.[45] By the 1920s, this figure had risen to about 3 or 3.5, depending on the definition of an urban place.[46]

Panels (e) and (f) of Figure 8 compare the attractiveness of each census division as a destination relative to New England. For the most part, there is little change over time. The main exceptions are the Mountain and Pacific divisions, however, which increased substantially in their relative attractiveness, indicating that there was some characteristic of these areas other than their urbanization that increasingly attracted internal migrants.[47]

Finally, I compare the growth over the linkage span of the urbanization or labor demand growth of movers' and stayers' residence counties by estimating an equation of the form

$$y_{it_2} - y_{it_1} = \alpha_t + \beta_t m_{it} + \mathbf{x}_{it}\gamma_t + \varepsilon_{it}$$

where y_{it_2} is the value for individual i's residence county in the final year of span t, y_{it_1} is the residence county in the initial year, m_{it} is an indicator equal to one if individual i moved during span t, and $\mathbf{x}_{it}$ is a vector of individual i's observables in the initial year of span t, including indicators for initial county of residence. For stayers, the y_{it} values describe the same county in both periods, meaning that stayers capture the change in the urbanization of their home county. A positive β_t implies that movers experienced a greater gain in urbanization than stayers from the same county—that is, that they moved to areas growing more quickly.

Figure 9 presents the results.[48] In essence, these results combine the selection and sorting patterns described previously, in which urbanites became relatively less likely to migrate as compared to ruralists and urban areas increased in attractiveness over time. In the nineteenth century, I find that movers and stayers experienced approximately the same increase in urban status and labor demand growth. By the twentieth century, the growth experienced by movers was greater than that experienced by stayers. Thus, whereas nineteenth-century migration was relatively neutral with respect to urbanization, twentieth-century migration was a force driving the contemporaneous urbanization of the economy by shifting population into more urban (or at least more rapidly urbanizing) areas.

---

[45] This corresponds to an increase in the odds by a factor of about 4.5.

[46] This corresponds to a 20- to 30-fold increase in the odds.

[47] Online Appendix Figure A.9 presents results controlling for labor demand growth instead of urbanization.

[48] Online Appendix Figure A.10 presents results for population density and the country's 20 largest cities.
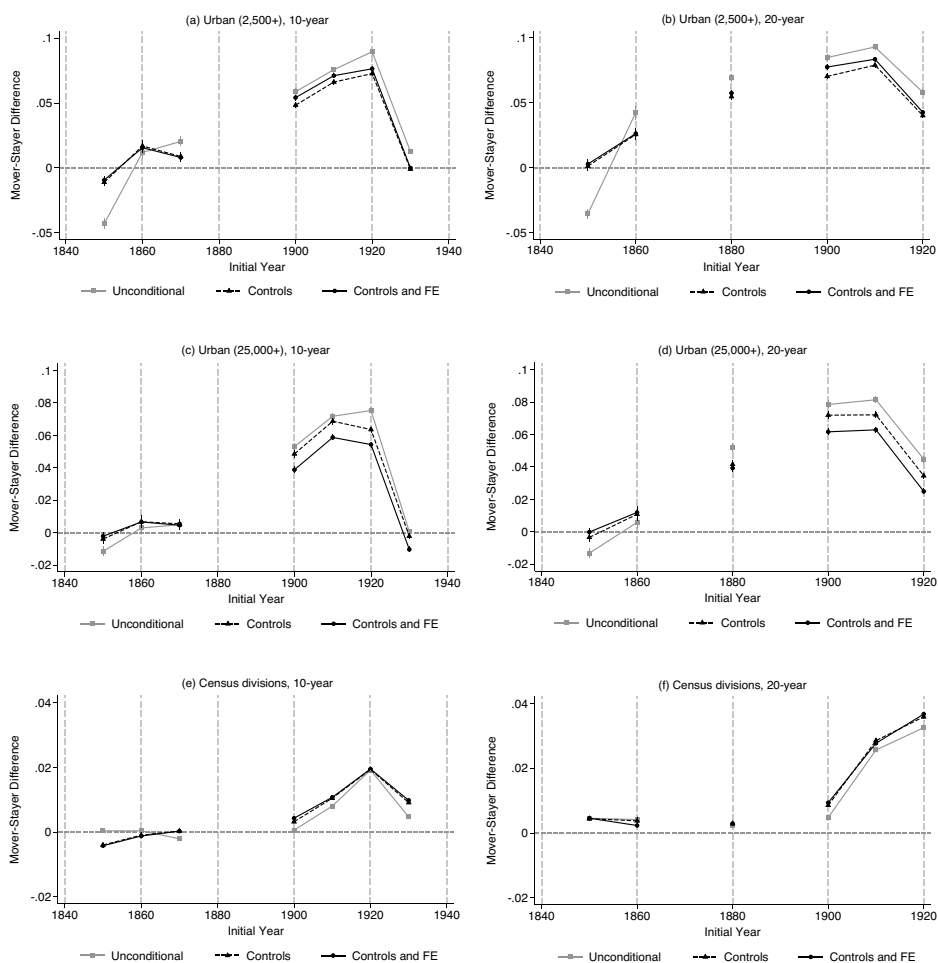
FIGURE 9
MOVERS' CHANGE IN URBANIZATION AND LABOR DEMAND GROWTH

Note: This figure compares the growth in urbanization or labor demand growth experienced by movers to that experienced by stayers, with either no controls, controlling for all observables, or controlling for all observables and county fixed effects.
Source: See the text.

THE 1930S

All of these patterns changed in the 1930s. For the first time, this decade was marked by a substantial decline in the rate of internal migration (Figure 3 Panel (a)), which fell from 34.2 percent for 1920–1930 to 27.6 percent for 1930–1940 according to my preferred linkage method. The 1930s also marked a moderation in selection into migration on the basis of occupational rank (Figure 4 Panel (a)). Relative to the 1920s, there was a considerable increase in the coefficient that was larger than its

change in any other period in the twentieth century. This change marked either a moderation of negative selection or a transition from somewhat negative to somewhat positive selection, depending on the linkage method; that is, internal migrants were less negatively selected in the 1930s than in the 1920s. A similar, though somewhat weaker, moderation is present in terms of selection into internal migration on the basis of urban residence in that urbanites and ruralists were more similar in terms of migration propensity than ever before in the twentieth century, though this was the product of a trend that evolved throughout the twentieth century (Figure 5 Panels (a) and (c)). Similarly, the distance of moves was lowest in the 1930s, again continuing a declining trend (Figure 7 Panels (a), (c), and (e)).

A particularly striking change concerns the attractiveness of urban areas as destinations for internal migrants. From the 1850s to the 1920s, the attractiveness of urban areas as destinations for internal migrants increased. But the 1930s marked a reversal of this trend, with a decline in this measure to levels approximately equal to those of the 1910s (Figure 8 Panels (a) and (c)). Even more dramatically, the 1930s marked a sharp reversal in the degree to which the urbanization growth or labor demand growth of movers exceeded that of stayers (Figure 9 Panels (a), (c), and (e)). This figure had climbed from the 1850s to the 1920s, but reversed sharply in the 1930s. For instance, in the 1920s, movers' growth in the likelihood of living in a city of 25,000 or more inhabitants was about 8 percentage points greater than that of stayers. By the 1930s, this figure had fallen to zero—a level not seen since the late nineteenth century—implying that the urbanization growth of movers and stayers was nearly identical and that internal migration was no longer a force driving urbanization.

SUMMARY OF ROBUSTNESS CHECKS

In addition to verifying, in Online Appendix C, that the conclusions that I draw are robust to the choice of linkage method, I also verify the robustness of my results to a number of other permutations of the sample or definitions of variables. In Online Appendix D, I redefine migration such that an individual must both cross county lines and move at least 150 miles to be considered an internal migrant. Naturally, the estimated migration rates are lower, in the vicinity of 15 percent over 10-year spans and 20 percent over 20-year spans. There is also a slight decline (about 2 to 3 percentage points) in these rates in the nineteenth century, which on the whole results in a slight downward trend, even when omitting the 1930s. Selection into migration was also different, with urbanites more

likely than ruralists to make longer-distance moves. Other results are largely unaffected.

In Online Appendix E, I omit individuals with foreign-born fathers from the sample. The intention of this exercise is to ensure that the continued assimilation of second-generation immigrants does not affect the results. In Online Appendix F, I repeat the main results using imputed occupational codes from Zimran (2022b) in cases where the occupations given by Ruggles et al. (2021) are listed as "Not Yet Classified."

DISCUSSION AND CONCLUSION

Internal migration is one of the fundamental forces that contributed to the development of the American economy and identity. But a lack of suitable data has made it difficult to establish even the most basic facts of internal migration in U.S. history. In this paper, I exploit recent advances in the availability of complete-count census data and in the technology to make links between censuses in order to describe the trends in the rates, selection, and sorting of the inter-county migration of native-born white men over the period 1850–1940. I find that the rates of and selection into inter-county migration were generally constant over time, with largely neutral or slightly negative selection into migration. But the distance of moves declined over time, migrants became increasingly attracted to the west, and the origins and orientation of migration shifted considerably over this period to become increasingly oriented toward driving a flow of population toward urban areas. The 1930s then marked a change in all regards, with migration declining in frequency, becoming somewhat more neutrally selected, and sharply declining in the degree to which it drove increases in urbanization.

These findings deepen our understanding of U.S. internal migration. Regarding the rates of internal migration, existing research (Ferrie 1997a; Hall and Ruggles 2004; Rosenbloom and Sundstrom 2004) has identified a decline in inter-state migration over the period 1850–1940, though the studies differ on the precise timing of this decline. My findings for inter-county migration paint a different picture—one of stable migration rates until the 1930s. In part, this difference in results can be attributed to the greater ability of linked data to bound the timing of migration. But it can also be attributed to my finding that inter-county but intra-state migration became increasingly important as average migration distances declined, meaning that a focus solely on inter-state migration would overlook an increasingly large share of internal population movements over time.

My results also give an understanding of the selection and sorting of internal migration that goes into far greater depth than has previously been possible. There are few existing estimates of migrant selection over a broad span to which my results can be compared. But the estimates that do exist (Hall and Ruggles 2004; Rosenbloom and Sundstrom 2004) point to positive selection—something for which I find no evidence in my analysis. The existing descriptions of migrants' destination choices in the long run (Hall and Ruggles 2004; Rosenbloom and Sundstrom 2004) are even more limited, generally focusing on the region of destination or the urban or rural status of the destination. My analysis provides a richer description than has previously been possible.

The results of this paper also help to better understand the development of the U.S. economy over the period that I study and, in particular, shed new light on Turner's (1921) influential interpretation of internal migration in U.S. history and later critiques of this interpretation. Turner (1921) famously argued that the United States's high internal migration rates in global perspective in the nineteenth century were the product of the availability of land on the frontier, which provided opportunities for surplus labor from eastern cities.[49] Turner (1921) also predicted that these high migration rates would be temporary, declining when the frontier ceased to exist. Others (e.g., Shannon 1945; Weber 1899) challenged claims of uniquely high migration rates in the nineteenth century, arguing that the growth of urban areas would draw migration, counteracting this decline.

Consistent with Turner (1921), I find evidence that urbanites were more likely to migrate in the nineteenth century. But I find no evidence of strongly negative selection into internal migration,[50] nor do my conditional logit results indicate that rural areas were particularly attractive in the nineteenth century (i.e., the coefficient on urban is always positive). I also find no evidence of a decline in migration rates after the closing of the frontier, contrary to findings for inter-state migration. Instead, I find that the closing of the frontier was followed by a twentieth century marked by a shift of internal migration into a force driving the urbanization of the economy from the perspective of both selection and destination choice. Beyond validating critiques of Turner's (1921) claims that the nineteenth century was exceptional in its high rates of internal migration, this finding speaks to the structural transformation of the economy as it shifted from agricultural to industrial (Caselli and Coleman 2001;

---

[49] Although these claims were published in book form in 1921, they date to 1893.

[50] Of course, this does not consider selection on unobservables (e.g., Ferrie 1997a; Stewart 2006).

Kuznets 1966), with internal migration helping to allocate labor to the nation's growing industrial sector in urban areas, as evidenced by my results regarding the labor demand growth experienced by movers.

My results also shed light on the unique nature of internal migration in the 1930s. The unique economic and climatic circumstances of the 1930s, combined with the U.S. census's first direct query on migration in the 1940 census, have attracted substantial attention to internal migration in this period.[51] Despite the common view of prevalent migration in this period in response to the unique shocks of the 1930s, my results show that migration rates were, in fact, lower during this period than in preceding decades (see also Long and Siu 2018). This result is consistent with Saks and Wozniak's (2011) finding that internal migration is generally pro-cyclical. My findings regarding the relationship between urbanization and migration also complement existing results. Fishback, Horrace, and Kantor (2006), for instance, use data on counties' aggregate population growth to show that more urban areas experienced relatively greater outmigration in the 1930s—a phenomenon due in part to "moves of despair" from depressed urban areas to farms (Boone and Wilse-Samson 2023; Boyd 2002).[52] My analysis complements this literature by putting these patterns into the broader context of U.S. history, comparing them to prior decades characterized by prior economic shocks.

Finally, the results of this paper also introduce or deepen a number of puzzles. The first concerns the relatively constant frequency of internal migration over the first 80 of the 90 years that I study. The structural transformation of the U.S. economy, combined with changes in transportation technology, land availability, international immigration, and labor market integration from 1850 to 1930, would be expected to lead to some change in the frequency or selectivity of internal migration. Indeed, there is evidence that these forces were associated with substantial changes in immigrant assimilation (Collins and Zimran 2023) and intergenerational mobility (Long and Ferrie 2013; c.f., Ward 2023). Yet such a change in the frequency of internal migration did not occur until the basket of shocks of the 1930s. A potential explanation is that this expected change came in the form of the increase in black migration after the onset of World War I, combined with changes in the frequency of international immigration (Collins 1997). But, in general, it is possible that linked data-based studies of the internal migration of groups other than native-born

---

[51] See, for instance, Boone and Wilse-Samson (2023), Boustan, Fishback, and Kantor (2010), Boyd (2002), Feigenbaum (2015), Fishback, Horrace, and Kantor (2006), Gutmann et al. (2016), Hornbeck (2012, 2023), Long and Siu (2018), and Sichko (2024).

[52] Feigenbaum (2015) documents similar patterns, which were initially noticed by Galbraith and Black (1938).

white men can illuminate this issue. Another puzzle concerns the increase over time in the deterrent effect of distance in the destination choices of internal migrants. This increase came despite improvements over the study period in transportation technology. This specific result can be rationalized by observing that the share of migrants who moved shorter distances increased,[53] but a larger question concerns why the increase in short-distance migration occurred. Explaining either of these puzzles is beyond the scope of this paper, but my documentation of them lays them out as the targets of future research.

REFERENCES

Aaronson, Daniel, Jonathan Davis, and Karl Schulze. "Internal Immigrant Mobility in the Early 20th Century: Evidence from Galveston, Texas." *Explorations in Economic History* 76 (2020): 101317.

Abramitzky, Ran, and Leah Platt Boustan. "Immigration in American Economic History." *Journal of Economic Literature* 55, no. 4 (2017): 1311–45.

Abramitzky, Ran, Leah Platt Boustan, and Dylan Connor. "Leaving the Enclave: Historical Evidence on Immigrant Mobility from the Industrial Removal Office." NBER Working Paper 27372, Cambridge, MA, June 2020.

Abramitzky, Ran, Leah Platt Boustan, Katherine Eriksson, James J. Feigenbaum, and Santiago Pérez. "Automated Linking of Historical Data." *Journal of Economic Literature* 59, no. 3 (2021a): 865–918.

Abramitzky, Ran, Leah Platt Boustan, Eisa Jácome, and Santiago Pérez. "Intergenerational Mobility of Immigrants in the US over Two Centuries." *American Economic Review* 111, no. 2 (2021b): 580–608.

Abramitzky, Ran, Jacob Conway, Roy Mill, and Luke Stein. "The Gendered Impacts of Perceived Skin Tone: Evidence from African-American Siblings in 1870–1940." NBER Working Paper 31016, Cambridge, MA, March 2023.

Adams, John W., and Alice Bee Kasakoff. "Wealth and Migration in Massachusetts and Maine: 1771–1798." *Journal of Economic History* 45, no. 2 (1985): 363–68.

Amior, Michael. "The Contribution of Immigration to Local Labor Market Adjustment." *Journal of Labor Economics,* forthcoming (2023).

Anbinder, Tyler, Dylan Connor, Cormac Ó Gráda, and Simone Wegge. "The Problem of False Positives in Automated Census Linking: Evidence from New York's Irish Immigrants as a Case Study." *Historical Methods* 56, no. 4 (2023): 240–59.

Bailey, Martha, Connor Cole, Morgan Henderson, and Catherine Massey. "How Well Do Automated Linking Methods Perform? Lessons from US Historical Data." *Journal of Economic Literature* 58, no. 4 (2020): 997–1044.

Bartik, Timothy J. *Who Benefits from State and Local Economic Development Policies?* Kalamazoo: W. E. Upjohn Institute for Employment Research, 1991.

Boone, Christopher D. A., and Laurence Wilse-Samson. "Structural Change and Internal Labor Migration: Evidence from the Great Depression." *Review of Economics and Statistics* 105, no. 4 (2023): 962–81.

---

[53] In this sense, it is not strictly correct to call this a deterrent *effect* of distance.

Borjas, George J. "Self-Selection and the Earnings of Immigrants." *American Economic Review* 77, no. 4 (1987): 531–53.

Boustan, Leah Platt. "Competition in the Promised Land: Black Migration and Racial Wage Convergence in the North, 1940–1970." *Journal of Economic History* 69, no. 3 (2009): 756–83.

———. "Was Postwar Suburbanization 'White Flight'? Evidence from the Black Migration." *Quarterly Journal of Economics* 215, no. 1 (2010): 417–43.

Boustan, Leah Platt, Price V. Fishback, and Shawn Kantor. "The Effect of Internal Migration on Local Labor Markets: American Cities during the Great Depression." *Journal of Labor Economics* 28, no. 4 (2010): 719–46.

Boustan, Leah Platt, Matthew E. Kahn, and Paul W. Rhode. "Moving to Higher Ground: Migration Response to Natural Disasters in the Early Twentieth Century." *American Economic Review, Papers and Proceedings* 102, no. 3 (2012): 238–44.

Boyd, Robert L. "A 'Migration of Despair': Unemployment, the Search for Work, and Migration to Farms During the Great Depression." *Social Science Quarterly* 83, no. 2 (2002): 554–67.

Brooks, David. "The Triumph of Hope over Self-Interest." *New York Times* CLII (January 12, 2003): WK15.

Card, David. "Is the New Immigration Really So Bad?" *Economic Journal* 115 (2005): F300–F323.

Caselli, Francesco, and Wilbur John Coleman. "The US Structural Transformation and Regional Convergence: A Reinterpretation." *Journal of Political Economy* 109, no. 3 (2001): 584–616.

Cheremukhin, Anton, Mikhail Golosov, Sergei Guriev, and Aleh Tsyvinsky. "The Industrialization and Economic Development of Russia through the Lens of a Neoclassical Growth Model." *Review of Economic Studies* 84 (2017): 613–49.

Chiquiar, Daniel, and Gordon H. Hanson. "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States." *Journal of Political Economy* 113, no. 2 (2005): 239–81.

Coleman, Peter J. "Restless Grant County: Americans on the Move." *Wisconsin Magazine of History* 46 (1962): 16–20.

Collins, William J. "When the Tide Turned: Immigration and the Delay of the Great Black Migration." *Journal of Economic History* 57, no. 3 (1997): 607–32.

Collins, William J., and Marianne H. Wanamaker. "Selection and Economic Gains in the Great Migration of African Americans: New Evidence from Linked Census Data." *American Economic Journal: Applied Economics* 6, no. 1 (2014): 220–52.

———. "The Great Migration in Black and White: New Evidence on the Selection and Sorting of Southern Migrants." *Journal of Economic History* 75, no. 4 (2015): 947–92.

Collins, William J., and Ariell Zimran. "The Economic Assimilation of Irish Famine Migrants to the United States." *Explorations in Economic History* 74 (2019): 101302.

———. "Working Their Way Up? US Immigrants' Changing Labor Market Assimilation in the Age of Mass Migration." *American Economic Journal: Applied Economics* 15, no. 3 (2023): 238–69.

Conway, Karen Smith, and Jonathan C. Rork. "How Has Elderly Migration Changed in the Twenty-First Century? What the Data Can—and Cannot—Tell Us." *Demography* 53, no. 4 (2016): 1011–25.

Curti, Merle E. *The Making of an American Community: A Case Study of Democracy in a Frontier Community*. Stanford: Stanford University Press, 1959.

Dahis, Ricardo, Emily Nix, and Nancy Qian. "Choosing Racial Identity in the United States, 1880–1940." NBER Working Paper 26465, Cambridge, MA, November 2019.

de Tocqueville, Alexis. *Democracy in America*, edited by Harvey C. Mansfield and Delba Winthrop. Chicago: University of Chicago Press, 1835 [2000].

Department of Commerce. *Fifteenth Census of the United States: 1930, Population*. Vol II. Washington: GPO, 1933.

Derenoncourt, Ellora. "Can You Move to Opportunity? Evidence from the Great Migration." *American Economic Review* 112, no. 2 (2022): 369–408.

Eldridge, Hope T., and Dorothy Swaine Thomas. *Population Redistribution and Economic Growth, United States, 1870–1950, Vol. 3: Demographic Analyses and Interrelations*. Philadelphia: American Philosophical Society, 1964.

Eli, Shari, Laura Salisbury, and Allison Shertzer. "Ideology and Migration after the Civil War." *Journal of Economic History* 78, no. 3 (2018): 822–61.

Eriksson, Katherine, and Gregory Niemesh. "Death in the Promised Land: The Great Migration and Black Infant Mortality." Mimeo, Miami University, Oxford, OH, 2016.

Feigenbaum, James J. "Intergenerational Mobility during the Great Depression." Mimeo, Boston University, Boston, MA, 2015.

Ferrie, Joseph P. "Long-Term Perspectives on Geographic and Economic Mobility, 1800–1900." Mimeo, Northwestern University, Evanston, IL, 1997a.

———. "Migration to the Frontier in Mid-Nineteenth Century America: A Re-Examination of Turner's 'Safety Valve'." Mimeo, Northwestern University, Evanston, IL, 1997b.

———. *Yankeys Now: Immigrants in the Antebellum United States, 1840–1860*. New York: Oxford University Press, 1999.

———. "How Ya Gonna Keep 'Em Down on the Farm [When They've Seen Schenectady]? Rural-to-Urban Migration in 19th Century America, 1850–70." Mimeo, Northwestern University, Evanston, IL, 2005.

———. "Internal Migration." In *Historical Statistics of the United States*, edited by Susan B. Carter, Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright, 1.489–1.494. Cambridge: Cambridge University Press, 2006a.

———. "Table Ac1-42: Interstate Migration—Native-Born Population, by Race and Residence within or outside the State of Birth: 1850–1990." In *Historical Statistics of the United States*, edited by Susan B. Carter, Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright, 1.495–1.497. Cambridge: Cambridge University Press, 2006b.

Fishback, Price V., William C. Horrace, and Shawn Kantor. "The Impact of New Deal Expenditures on Mobility during the Great Depression." *Explorations in Economic History* 43 (2006): 179–222.

Fouka, Vasiliki, Soumyajit Mazumder, and Marco Tabellini. "From Immigrants to Americans: Race and Assimilation during the Great Migration." *Review of Economic Studies* 89, no. 2 (2022): 811–42.

Galbraith, John K., and John D. Black. "The Maintenance of Agricultural Production during Depression: The Explanations Reviewed." *Journal of Political Economy* 46, no. 3 (1938): 305–23.

Galenson, David W., and Clayne L. Pope. "Economic and Geographic Mobility on the Farming Frontier: Evidence from Appanoose County, Iowa, 1850–1870." *Journal of Economic History* 49, no. 3 (1989): 635–55.

Greenwood, Michael J. "Research on Internal Migration in the United States; A Survey." *Journal of Economic Literature* 13, no. 2 (1975): 397–433.

Gutmann, Myron P., Daniel Brown, Angela R. Cunningham, James Dykes, Susan Hautaniemi Leonard, Jani Little, Jeremy Mikecz, Paul W. Rhode, Seth Spielman, and Kenneth M. Sylvester. "Migration in the 1930s: Beyond the Dust Bowl." *Social Science History* 40 (2016): 707–40.

Hall, Patricia Kelly, and Steven Ruggles. "'Restless in the Midst of Their Prosperity': New Evidence on the Internal Migration of Americans, 1850–2000." *Journal of American History* 91, no. 3 (2004): 829–46.

Hatton, Timothy J., and Zachary Ward. "International Migration in the Atlantic Economy 1850–1940." In *Handbook of Cliometrics*, edited by Claude Diebolt and Michael Haupert, 301–29. London: Springer, 2019.

Hatton, Timothy J., and Jeffrey G. Williamson. "What Explains Wage Gaps Between Farm and City? Exploring the Todaro Model with American Evidence, 1890–1941." *Economic Development and Cultural Change* 40, no. 2 (1992): 267–94.

———. *The Age of Mass Migration: Causes and Economic Impact*. New York: Oxford University Press, 1998.

Hayashi, Fumio, and Edward C. Prescott. "The Depressing Effect of Agricultural Institutions of the Prewar Japanese Economy." *Journal of Political Economy* 116, no. 4 (2008): 573–632.

Hornbeck, Richard. "Barbed Wire: Property Rights and Agricultural Development." *Quarterly Journal of Economics* 125 (2010): 767–810.

———. "The Enduring Impact of the American Dust Bowl: Short- and Long-Run Adjustments to Environmental Catastrophe." *American Economic Review* 102, no. 4 (2012): 1477–507.

———. "Dust Bowl Migrants: Environmental Refugees and Economic Adaptation." *Journal of Economic History* 83, no. 3 (2023): 645–75.

Hornbeck, Richard, and Suresh Naidu. "When the Levee Breaks: Black Migration and Economic Development in the American South." *American Economic Review* 104, no. 3 (2014): 963–90.

Huang, Ying, Scott J. South, and Amy Spring. "Racial Differences in Neighborhood Attainment: The Contributions of Interneighborhood Migration and In Situ Change." *Demography* 54, no. 5 (2017): 1819–43.

Huttunen, Kristiina, Jarle Møen, and Kjell G. Salvanes. "Job Loss and Regional Mobility." *Journal of Labor Economics* 36, no. 2 (2018): 479–509.

Jauer, Julia, Thomas Liebig, John P. Martin, and Patrick A. Puhani. "Migration as an Adjustment Mechanism in the Crisis? A Comparison of Europe and the United States 2006–2016." *Journal of Population Economics* 32, no. 1 (2019): 1–22.

Jia, Ning, Raven Molloy, Christopher L. Smith, and Abigail Wozniak. "The Economics of Internal Migration: Advances and Policy Questions." *Journal of Economic Literature* 61, no. 1 (2023): 144–80.

Kennan, John, and James R. Walker. "The Effect of Expected Income on Individual Migration Decisions." *Econometrica* 79, no. 1 (2011): 211–51.

Kuznets, Simon. *Modern Economic Growth: Rate, Structure, and Spread*. New Haven: Yale University Press, 1966.

Lewis, Arthur W. "Economic Development with Unlimited Supplies of Labor." *The Manchester School* 22, no. 2 (1954): 139–91.

Long, Jason, and Joseph P. Ferrie. "Intergenerational Occupational Mobility in Great Britain and the United States since 1850." *American Economic Review* 103, no. 4 (2013): 1109–37.

Long, Jason, and Henry Siu. "Refugees from Dust and Shrinking Land: Tracking the Dust Bowl Migrants." *Journal of Economic History* 78, no. 4 (2018): 1001–33.

Malin, James C. "The Turnover of Farm Population in Kansas." *Kansas Historical Quarterly* 4 (1935): 339–72.

Manson, Steven, Jonathan Schroeder, David Van Riper, and Steven Ruggles. *IPUMS National Historical Geographic Information System: Version 14.0* [Database]. Minneapolis: University of Minnesota, 2019.

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

———. "Job Changing and the Decline in Long-Distance Migration in the United States." *Demography* 54, no. 2 (2017): 631–53.

Oberly, James W. "Westward Who? Estimates of Native White Interstate Migration after the War of 1812." *Journal of Economic History* 46, no. 2 (1986): 431–40.

Oswald, Florian. "The Effect of Homeownership on the Option Value of Regional Migration." *Quantitative Economics* 10, no. 4 (2019): 1453–93.

Pérez, Santiago. "Intergenerational Occupational Mobility across Three Continents." *Journal of Economic History* 79, no. 2 (2019): 383–416.

Price, Joseph, Kasey Buckles, Jacob Van Leeuwen, and Isaac Riley. "Combining Family History and Machine Learning to Link Historical Records." NBER Working Paper 26227, Cambridge, MA, September 2019.

Rosenbloom, Joshua L. *Looking for Work, Searching for Workers: American Labor Markets during Industrialization*. New York: Cambridge University Press, 2002.

Rosenbloom, Joshua L., and William A. Sundstrom. "The Decline and Rise of Interstate Migration in the United States: Evidence from the IPUMS, 1850–1990." *Research in Economic History* 22 (2004): 289–325.

Ruggles, Steven, Sarah Flood, Sophia Foster, Ronald Goeken, Jose Pacas, Megan Schouweiler, and Matthew Sobek. *Integrated Public Use Microdata Series: Version 11.0* [dataset]. Minneapolis, University of Minnesota, 2021.

Saks, Raven E., and Abigail Wozniak. "Labor Reallocation over the Business Cycle: New Evidence from Internal Migration." *Journal of Labor Economics* 29, no. 4 (2011): 697–739.

Salisbury, Laura. "Selective Migration, Wages, and Occupational Mobility in Nineteenth Century America." *Explorations in Economic History* 53 (2014): 40–63.

Shannon, Fred A. "A Post-Mortem on the Labor-Safety-Valve Theory." *Agricultural History* 19 (1945): 31–37.

Sichko, Christopher. "Migrant Selection and Sorting during the Great American Drought."*World Development* 181 (2024): 106632.

Sprung-Keyser, Ben, Nathanial Hendren, and Sonya Porter. "The Radius of Economic Opportunity: Evidence from Migration and Local Labor Markets." CES Discussion Paper 22-27, Washington, DC, July 2022.

Steckel, Richard H. "Census Matching and Migration: A Research Strategy." *Historical Methods* 21, no. 2 (1988): 52–60.

———. "Household Migration and Rural Settlement in the United States, 1850–1860." *Explorations in Economic History* 26 (1989): 190–218.

Stewart, James I. "Migration to the Agricultural Frontier and Wealth Accumulation, 1860–1870." *Explorations in Economic History* 43 (2006): 547–77.

———. "Economic Opportunity or Hardship? The Causes of Geographic Mobility on the Agricultural Frontier, 1860–1880." *Journal of Economic History* 69, no. 1 (2009): 238–68.

———. "Migration to US Frontier Cities and Job Opportunity, 1860–1880." *Explorations in Economic History* 49 (2012): 528–42.

Stuart, Bryan A., and Evan J. Taylor. "Migration Networks and Location Decisions: Evidence from US Mass Migration." *American Economic Journal: Applied Economics* 13, no. 3 (2021): 134–75.

Throne, Mildred. "A Population Study of an Iowa County in 1850." *Iowa Journal of History* 57 (1959): 305–30.

Treyz, George I., Dan S. Rickman, Gary L. Hunt, and Michael T. Greenwood. "The Dynamics of US Internal Migration." *Review of Economics and Statistics* 75, no. 2 (1993): 209–14.

Turner, Frederick Jackson. *The Frontier in American History*. New York: Henry Holt and Company, 1921.

Vigdor, Jacob L. "The Pursuit of Opportunity: Explaining Selective Black Migration." *Journal of Urban Economics* 51, no. 3 (2002): 391–417.

Villaflor, Georgia C., and Kenneth L. Sokoloff. "Migration in Colonial America: Evidence from Militia Muster Rolls." *Social Science History* 6, no. 4 (1982): 539–70.

Ward, Zachary. "Internal Migration, Education and Upward Rank Mobility: Evidence from American History." *Journal of Human Resources* 57, no. 6 (2022): 1981–2011.

———. "Intergenerational Mobility in American History: Accounting for Race and Measurement Error." *American Economic Review* 113, no. 12 (2023): 3213–48.

Weber, Adna Ferrin. *The Growth of Cities in the Nineteenth Century: A Study in Statistics*. New York: MacMillan, 1899.

Wozniak, Abigail. "Are College Graduates More Responsive to Distant Labor Market Opportunities?" *Journal of Human Resources* 45, no. 4 (2010): 944–70.

Zimran, Ariell. "Census Linkage Crosswalks for US Censuses 1850–1940." Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 7 July 2022a. http://doi.org/ 10.3886/E171701V3.

———. "Miscellaneous US Census Crosswalks, 1850–1930." Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 3 June 2022b. http://doi.org/10.3886/E171861V2.

———. "US Immigrants' Secondary Migration and Geographic Assimilation during the Age of Mass Migration." *Explorations in Economic History* 85 (2022c): 101457.

———. "Data and Code for 'Internal Migration in the United States: Rates, Selection, and Destination Choice, 1850–1940.'" Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 21 July 2023. http://doi.org/10.3886/E192868V1.