

The History of the Family

ISSN: 1081-602X (Print) 1873-5398 (Online) Journal homepage: www.tandfonline.com/journals/rhof20

Settlement and the intergenerational dispersion of kin as a spatial process in the nineteenth century US

Caglar Koylu, Alice B. Kasakoff & Maryam Torkashvand

To cite this article: Caglar Koylu, Alice B. Kasakoff & Maryam Torkashvand (2025) Settlement and the intergenerational dispersion of kin as a spatial process in the nineteenth century US, The History of the Family, 30:4, 706-729, DOI: 10.1080/1081602X.2025.2583972

To link to this article: https://doi.org/10.1080/1081602X.2025.2583972

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

View supplementary material

Published online: 14 Nov 2025.

Submit your article to this journal

Article views: 760

View related articles

View Crossmark data

Citing articles: 1 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=rhof20

THE HISTORY OF THE FAMILY
2025, VOL. 30, NO. 4, 706–729
https://doi.org/10.1080/1081602X.2025.2583972

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Settlement and the intergenerational dispersion of kin as a spatial process in the nineteenth century US

Caglar Koylu[a], Alice B. Kasakoff[b] and Maryam Torkashvand[a]

[a]School of Earth, Environment, and Sustainability, University of Iowa, Iowa City, IA, USA; [b]Department of Geography, University of South Carolina, Columbia, SC, USA

ABSTRACT

This study examines how families dispersed or remained clustered across three generations during the nineteenth-century American settlement process. While kin networks are known to influence migration decisions, economic outcomes, and reproductive behavior, most historical studies rely on snapshots of co-residence or proximity within a single generation. We extend this research by introducing a new methodological framework, the Intergenerational Symmetry Index, which takes the form of two measures, Stayer Symmetry and Migrant Symmetry, that capture kin retention and attraction across multiple generations using population-scale genealogical data. These measures are built from family-weighted and gendered lineal paths that trace grandparent (G1), parent (G2), and grandchild (G3) birthplaces to assess how rooted or dispersed family lines become over time and space. We apply these measures to eight grandparent (G1) birth cohorts that range between 1785 and 1866 across the US. Our results reveal a westward-shifting generational pattern of kin cohesion and dispersion during the settlement process, gradually giving way to more neutral patterns of dispersion as areas matured. Despite data limitations, including underrepresentation of women and non-white individuals in genealogical sources, our method captures consistent intergenerational patterns and confirms that family-based migration was a key feature of nineteenth-century US settlement. Our findings support and extend classic theories of demographic change, including Turner's regional settlement stages, Easterlin's fertility-inheritance dynamics, and Zelinsky's mobility transition. They also underscore the importance of considering kinship, not just individuals, as a unit of spatial analysis. Our multigenerational approach reveals spatial processes invisible in census-based migration measures and, as genealogical datasets expand, offers a path for integrating family history with broader questions of migration, settlement, and demographic change.

ARTICLE HISTORY
Received 11 June 2025
Accepted 29 October 2025

KEYWORDS
Kin cohesion; kin dispersion; kinship networks; migration; family trees; nineteenth-century US settlement

1. Introduction

Geographic dispersion of families is key to understanding a diversity of processes, such as migration and demographic change (Hacker et al., 2021; Hey, 2010), the impact of health

CONTACT Caglar Koylu caglar-koylu@uiowa.edu
Supplemental data for this article can be accessed online at https://doi.org/10.1080/1081602X.2025.2583972

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

and genetics on communities (Black et al., 2023; Daelemans et al., 2013), the role of kin networks in economic outcomes and migration decisions (Castillo & Helgertz, 2025; Otterstrom et al., 2022), and the diffusion of information and innovation (Carlson & Dillman, 1983; Casterline, 2001; Godley, 2001). Kin networks influence economic outcomes, migration decisions, and provide critical social support, affecting elder care and child survival rates and fertility (Beise & Voland, 2002; Healy & Dunifon, 2025).

Despite growing recognition of the importance of kin cohesion and dispersion, i.e. how widely families are spread across space; measuring kin proximity, or the physical closeness of relatives, remains a major challenge due to limitations in traditional household surveys and censuses, which rarely capture relationships beyond the household or the neighborhood. Past efforts have relied on indirect proxies (Hacker, 1999), selective survey data (Sear & Coall, 2011), or genealogical records with limited geographic detail (Jennings et al., 2012), which limits our understanding of kin proximity and the relevant demographic processes. Hacker and Roberts (2017) addressed this gap by leveraging complete-count census data to estimate the proximity of potential mothers-in-law and showed a modest but positive association with fertility, reinforcing the role of nearby kin in shaping demographic behavior. However, that study and many other existing studies focus on static snapshots of co-residence or local proximity and treat kin presence as an explanatory variable, rather than asking how kin networks disperse or persist across geographic space and time. As a result, we still lack a multigenerational perspective on how kin networks cluster or scatter geographically, which is essential for understanding long-term migration trends, demographic transitions, and the relationship between family structure and spatial settlement.

The processes of migration, kin clustering and dispersion, and demographic change were not only social and demographic but also deeply spatial. In the nineteenth-century United States, the ways families remained rooted or spread out shaped the kin geography of the entire country, from older eastern states to newly developing regions in the West and South. Whether families stayed in place, dispersed into emerging areas, or transformed older regions through renewed attraction influenced how kin networks developed and how the geography of settlement evolved across the nation. By tracing these multigenerational patterns, this study provides a new framework for understanding settlement as a spatial process, one that places kinship, not just individuals, at the center of nineteenth-century American migration history. Building directly on this aim, we draw on population-scale kinship networks derived from crowdsourced genealogical data generated by Koylu et al. (2021), which link successive generations across time and space. With a connected component of 51 million individuals, this dataset provides the scale and generational continuity necessary to trace kin clustering and dispersion across the nineteenth-century settlement era.

In this article, we introduce the Intergenerational Symmetry Index (ISI), a three-generation framework that follows grandparents (G1), parents (G2), and grandchildren (G3) within a population-scale kinship network. We apply the ISI to eight eleven-year grandparent (G1) birth cohorts spanning 1785 to 1866 across the US to examine how kin geography evolved during the settlement era. Anchored on the parental (G2) birthplace, the ISI has two implementations. Stayer Symmetry (ISI-S) compares the share of families whose grandchildren (G3) remain in the parental birthplace with the share whose grandparents (G1) had also remained there. It shows whether a place is becoming more or less

rooted across generations. Migrant Symmetry (ISI-M) compares the share of families whose grandparents (G1) were born outside the parental birthplace with the share whose grandchildren (G3) were born outside it. It shows whether a place functions as a net attractor of kin from elsewhere or a net exporter of kin to other places. To capture differences in descent, the ISI is further divided into four gender-specific lineal paths that trace paternal and maternal links separately through paternal-paternal (grandfather-father-child), maternal-paternal (grandmother-father-child), paternal-maternal (grandfather-mother-child), and maternal-maternal (grandmother- mother-child) links.

The ISI is particularly useful because it relies only on birthplace information, which are far more complete than mortality records in family tree data, and because it weights families rather than individuals, avoiding biases from sibship size. More importantly, it enables consistent measurement of kin clustering and dispersion across multiple generations and provides a new lens for connecting family geography to the broader process of nineteenth-century US settlement. As genealogical datasets continue to grow, this approach offers a promising path for integrating family history with broader questions of migration, settlement, and demographic change.

2. Literature review

2.1. Kin dispersion

Kin ties influence a variety of demographic and social processes from child survival to elder care, fertility behaviour, migration decisions, and the transmission of information and resources. For instance, Engelhardt et al. (2019) demonstrated that daughters whose mothers are alive raise more surviving children, but this 'grandmother effect' falls sharply the farther the grandmother lives from her reproducing daughters. However, the distance effect may be confounded by age, as daughters with living mothers are typically younger than those with deceased mothers. By calculating non-random isonymy within an enumeration district using complete-count US census data, Nelson (2020) showed that such kin propinquity has declined precipitously since the early nineteenth century likely as a result of migration and falling fertility. Hacker et al. (2021) used the measure Nelson developed to reveal that proximity to maternal versus paternal grandmothers shapes reproductive success in distinct ways.

Historical demography explains why kin availability can shift so dramatically. Pre-industrial America's late marriage and child bearing late in life, and high fertility limited the number of three-generation households; elderly parents typically resided with a single child while other siblings founded separate nuclear households elsewhere (Ruggles, 1994). As commerce and wage labor expanded, the economic logic of multigenerational coresidence weakened, accelerating geographic dispersal even without a cultural preference for nuclear families (Ruggles, 2003). In the Netherlands, inheritance systems and parental migration histories shaped this process further: children of mobile mothers were more likely to disperse, while partible inheritance often delayed departure but eventually led to smaller, more scattered households (Bras & Tilburg, 2007; Kok & Bras, 2008).

These studies illustrate why having kin nearby matters and how economic, demographic, and institutional forces affect kin proximity. Yet they also expose a critical gap: most rely on single generation coresidence, surname clustering, or individual life

events. While there are studies of siblings (Bratti et al., 2020; Kok & Bras, 2008); few have traced families across multiple generations to understand the larger process of family dispersion or quantifying whether kin groups stayed rooted or fanned out during the nineteenth century or transplanted themselves to the growing cities. The period of settlement meant a move away from kin in the sending area and, due to high fertility in the US, the rapid development of families in the newly settled areas. Hacker and Roberts (2017) used complete-count US 1880 census data to construct a measure of kin proximity. Specifically, they examined the relationship between the presence of a potential mother-in-law living within ten nearby households and net marital fertility. Their findings showed that proximity to kin facilitated reproductive behavior. However, their analysis was limited to a single point in time and a narrow window of local proximity, without linking kin across generations or broader geographic scales. Nelson (2020) used decennial census data to measure the probability that people bearing the same surname lived geographically close, and identified a life cycle pattern in kin propinquity with the highest rates corresponding to periods of greatest generational overlap. Because of the large temporal frame of our study, spanning three generations, it is difficult to confirm the detailed spatial patterns of closely located multi-generational households observed by Nelson (2020). Our generational approach smooths these short-term dynamics of kin proximity and instead highlights longer-term settlement trends. Nevertheless, our approach allows us to confirm the general patterns characterizing the settlement process, occurring over several decades. Specifically, we address Turner's concept of frontier settlement stages (Turner & Wilson, 1920), Zelinsky's idea that migration patterns are linked to other demographic transitions (Zelinsky, 1971), and Easterlin's perspective that generational fertility behaviors are shaped by inheritances (Easterlin et al., 1978).

2.2. Fertility and migration dynamics in early settlement

Geographic variations in kinship networks are closely tied to fertility and migration patterns. Zelinsky's (1971) hypothesis of the mobility transition parallels the demographic transition, which suggests that as societies modernize, migration increases and later stabilizes in parallel with the stages of population growth. Populations shift from low mobility with high fertility and mortality (Phase I), to mass rural out-migration and frontier settlement as mortality declines and populations surge (Phase II), to slowing frontier flows and more urban-to-urban moves as fertility begins to fall (Phase III), and finally to high but steady circulation in advanced societies with low fertility and mortality (Phase IV) (Zelinsky, 1971). Zelinsky's framework highlights how fertility decline and migration are intertwined elements of modernization, with rapid population growth, land scarcity, and shifting opportunities driving both.

Migration westward over the course of the nineteenth-century US also involved substantial restriction of fertility. Yasuba (1962) and Easterlin (1976a) demonstrated that child-woman ratios fell earlier in the East than in the West, shaped by land availability and inheritance practices. Migrants in the West had more land and lower mortality, supporting larger families, while constrained land in the East incentivized smaller families (Easterlin, 1976b). Migration timing and socioeconomic status also mattered: migrants during reproductive years had higher fertility (Mineau et al., 1989), whereas moving

before childbearing aligned migrants with native-born fertility norms (Maloney et al., 2014). Families in urban, non-manual occupations increasingly practiced fertility control (Alter & Hacker, 2024).

Frontier mobility aligned with Turner's 'safety-valve' hypothesis (Turner & Wilson, 1920), drawing land-poor, lower-status families west (Ferrie, 1997). Early settlers often accumulated above-average wealth (Galenson & Pope, 1989), and blue-collar or immigrant groups pursued wage opportunities in emerging frontier cities (Glover & Towner, 2009; Stewart, 2006, 2009, 2012). Economic opportunity and kin networks were key pull factors in both trans-Atlantic (Williamson, 1974) and internal westward migration in the US (Hudson, 1976; Malin, 1935).

Early migration studies in the US relied on aggregated census data, especially from 1850 onward, when recorded birthplaces enabled studies of lifetime migration by comparing state of birth and state of residence (Lee & Lee, 1960). Price (1953) introduced the 'survival method' to estimate net interstate migration in the United States from 1870 to 1940 by contrasting expected populations with actual census counts at each decade's end. Price's census-based survival method, however, was limited by its assumption of uniform mortality and accuracy across states, and it required individuals to be at least ten years old. Local studies using cohort survival (Hinde, 2004) and census record linkage (Thernstrom & Knights, 1970) revealed high population turnover. However, simple persistence measures overstated migration by conflating it with death or enumeration error (Guest, 1987; Parkerson, 1982). While these historical methods confirmed high east-to-west and rural-to-urban migration rates (Molloy et al., 2011), they often overlooked intergenerational family dynamics and movements occurring between censuses.

2.3. Population-scale family trees

Population-scale family trees provide unique insight into multigenerational migration and kin dispersion. Leveraging this potential, Spencer et al. (2024) employed a migration symmetry metric: $MS = (births - deaths)/(births + deaths)$, where births refer to the number of people born within a location but who died elsewhere, while deaths refer to those born outside the location but who died within it. When a small number of locations lose population to several others ($MS < 0$), it indicates dispersion, typical of frontier expansion. In contrast, when many locations send people to fewer, more centralized blocks ($MS > 0$), it reflects population concentration, characteristic of more established or settled regions and/or urbanization. A key limitation of the migration symmetry measure based on individuals' birth and death locations is that death information is available for only about 20% of individuals on average. Therefore, the use of birth-death information produces incomplete and potentially biased measures, especially across time periods with uneven coverage. Finally, this measure reflects the lifetime movement of a single generation, but does not capture intergenerational flows, which are a distinctive strength of genealogical data.

Existing measures based on family trees such as birth-death (Spencer et al., 2024), life-course (Adams et al., 2002), and child-ladder (Koylu & Kasakoff, 2022) methods are often used to capture kin dispersion from genealogical data. On the other hand, census-based approaches either link household records to track location changes across enumerations (Koylu, Kasakoff, & Torkashvand, 2025) or use surname proximity to infer how closely

households with the same surname reside (Nelson, 2020)). While useful for showing who is present at a given point in time, these methods do little to reveal how entire family lines accumulate or dissipate in a place across successive generations. To fill that gap, we organize grandparents into birth cohorts and compute simple, family-weighted ratios that (i) compare the share of newcomers among grandparents to the share of leavers among grandchildren and (ii) contrast descendant retention with ancestral presence. Because these ratios rely only on the far more complete birthplace information in family tree data, they sidestep the missing-death problem, avoid overweighting large sibships, and pinpoint which generation moves. The result is a multidecadal, lineage-specific portrait of kin geography that complements short-interval migration studies and allows direct tests of the long-run settlement theories advanced by Turner, Easterlin, and Zelinsky.

3. Data and methods

We use the population-scale family tree data generated by Koylu et al. (2021), who cleaned, connected and deduplicated crowdsourced family tree records from Rootsweb.com. The largest connected component in this dataset consists of 51 million individuals who could be linked to their common ancestor in a giant connected family tree. Working at the state level, we address questions such as: To what extent do states act as net attractors or dispersers of kin over time? How persistent are kin ties within states across successive generations? How do these dynamics shift as European-descended settlement moved westward? In this section, we outline the steps employed to preprocess, analyze, and derive intergenerational cohorts from this population-scale family tree data.

3.1. Generational cohorts

We examine kin dispersion by tracing families across three generations: Generation 1 (G1: grandparents), Generation 2 (G2: parents), and Generation 3 (G3: grandchildren). We base our analysis on G1 couples whose birth years fall within designated eleven-year birth cohorts, beginning with 1785–1796 and continuing through 1855–1866. Because genealogical birth years often cluster at years ending in 0 or 5, strict 10-year bins create boundary artifacts when these clusters fall on a cutoff. To avoid splitting such clusters and to minimize overlap between adjacent cohorts, we use half-open 11-year windows [y, y +11). We restrict G1 to couples where at least one spouse is born in the United States and falls within the cohort birth year range. After applying the additional constraints described below, however, our final samples include only G1 couples in which both partners were born in the US. Three generations provide a natural unit at least in Western cultures and include the relatives one usually can know during one's lifetime. G2 serves as the focal generation for our analysis: we examine the kin ties that connect G2 both upward to G1 and downward to G3.

Because many trees are incomplete and sometimes contain inaccurate data, we filter the cases to reduce inconsistencies. We first check for improbable birth spacing. While variation is expected, extreme outliers can create unlikely generational overlaps and may be the result of inaccurate information. At the sibling level, we exclude any family where a sibling pair has a birth gap greater than 10 years. At the family level, we also remove

cases where the span between the first and last child exceeds 30 years. Historically, siblings are typically born one to five years apart; longer intervals often reflect remarriage or non-biological ties. We deliberately set very large cut-offs well above typical fertility limits to include mis-dated births commonly found in crowdsourced family tree data. Otherwise, some trees with uncertain birth years would be unnecessarily discarded.

To ensure a robust, three-generation sample, we first require that every G1 couple have at least two G2 children with complete birth year and birthplace. We further restrict to cases where at least two of those G2 children each have at least two G3 offspring with complete birth year and birthplace. These dual thresholds guarantee that each family in the sample includes at least two well-documented G2 members, each with at least two well-documented G3 descendants.

These criteria also capture G1 parents at roughly comparable ages, which is on average around 30 at the birth of their children for women and about 35 for men. More importantly, because our analysis begins with grandparents grouped into narrow eleven-year birth cohorts, this already constrains G1 to a common age structure and sets the basis for consistent comparisons. This helps reduce bias from the fact that the cumulative chance of being an interstate migrant increases with age, a point emphasized by Hall and Ruggles (2004), who measured migration at ages 50–59 to standardize comparisons across cohorts. Our goal, however, is not to estimate individual migration rates or trace life-course exposure. Instead, our analysis focuses on intergenerational migration, where the relevant comparison is not age but historical era. Following Hall and Ruggles (2004), we recognize that migration dynamics changed significantly across the nineteenth and twentieth centuries, and our three-generation design captures these broader temporal changes across successive cohorts.

3.2. Birthplace sequences and lineal paths

To capture both gender and lineal effects, we begin by examining the directed parent – child transitions within each family: from each G1 ancestor to each G2 descendant (grandfather-father, grandfather-mother, grandmother-father, grandmother-mother) and from each G2 parent to each G3 child (father-child, mother-child). For example, the 'grandfather-father' transition asks how often a G2 father is born in the same state as his G1 father, while the 'mother-child' transition asks how often a G3 child is born in the same state as their G2 mother. Analyzing all transitions separately allows us to detect whether paternal versus maternal lines, and male versus female descendants, exhibit different patterns of geographic continuity or change in birthplaces.

To trace three-generation birthplace sequences, we define four distinct lineal paths, each linking a G1 grandparent to a G3 grandchild through a specific G2 parent. The labels always follow the order G1 > G2 > G3:

- Paternal-Paternal (g1fg2f): from G1 grandfather to G2 father to G3 child.
- Maternal-Paternal (g1mg2f): from G1 grandmother to G2 father to G3 child.
- Paternal-Maternal (g1fg2m): from G1 grandfather to G2 mother to G3 child.
- Maternal-Maternal (g1mg2m): from G1 grandmother to G2 mother to G3 child.

Within a family, each grandchild contributes to exactly two of these paths (depending on whether the G2 parent is male or female). However, multiple siblings may share the same birthplace sequence (G1_state, G2_state, G3_state). To avoid inflating the contribution of large sibships, we index each family f and path $p \varepsilon \{g1fg2f, g1mg2f, g1fg2m, g1mg2m\}$. For family f and path p, let

$$T_{f,p} = \{ (g1bplace_i, g2bplace_i, g3bplace_i) : i \varepsilon \text{ grandchildren of } f \text{ along } p \}$$

be the set of unique three-generation birthplace triplets observed. We then assign each triplet $s \in T_{f,p}$ a family-path weight

$$w_{f,p}(s) = \frac{1}{|T_{f,p}|}$$

so that

$$\sum_{s \in T_{f,p}} w_{f,p}(s) = 1 \text{ for every } (f, p).$$

In other words, regardless of how many siblings share the same triplet, or how many distinct triplets exist, each family contributes exactly one unit of weight to each path, split evenly across its distinct sequences.

Finally, we aggregate across all families to produce weighted birthplace-sequence counts for each path-state combination:

$$M_p(s) = \sum_f w_{f,p}(s) \text{ for each triplet } s = (g1b, g2b, g3b)$$

These $M_p(s)$ values form the entries in our $(g1b, g2b, g3b, p)$ table and serve as the normalized flow volumes for our symmetry metrics. By construction, this equal-sequence weighting removes any bias from sibship size and multiple G3 siblings sharing the same sequence do not inflate the count. In cases where both parents descend from G1 couples in the cohort, the same grandchild will appear once through each parental line. This does not constitute double counting, because our weighting scheme ensures that each family-path contributes exactly one unit regardless of sibship size.

3.3. Intergenerational kin dispersion

We introduce the Intergenerational Symmetry Index (ISI) to capture kin dispersion across three generations. We implement ISI in two complementary forms:

- ISI-S (Stayer Symmetry) focuses on persistence.
- ISI-M (Migrant Symmetry) focuses on net attraction and dispersion.

Both versions are anchored on the birth state of the middle (G2) generation and compare two opposing proportions of family-level flows. To avoid confounding by fertility, each family contributes exactly one unit of weight to each of four lineal paths (paternal – paternal, paternal – maternal, maternal – paternal, maternal – maternal), split equally among its distinct grandchild sequences.

Let X be a given G2 birth state and F_x the set of weighted family-path records anchored on X. For any implementation of ISI we compute two shares. While $P_A(X)$ is the ancestral

share defined by an 'ancestral' condition, $P_D(X)$ is the descendant share defined by a 'descendant' condition.

$$P_A(X) = \frac{\sum_{i \in F_X:\ ancestral\ cond(i)} w_i}{\sum_{i \in F_X} w_i}, P_D(X) = \frac{\sum_{i \in F_X:\ descendant\ cond(i)} w_i}{\sum_{i \in F_X} w_i}$$

We then set the Intergenerational Symmetry Index as:

$$ISI(X) = \frac{P_A(X) - P_D(X)}{P_A(X) + P_D(X)}, ISI(X) \in [-1, 1].$$

By construction, $S(X) > 0$ indicates *ancestral dominance* (net attraction or persistence), while $S(X) < 0$ indicates negative values indicates descendant dominance (net dispersion). These two directional tendencies are captured by distinct implementations: Stayer Symmetry focuses on evaluating immobility, which is especially important given that stayers are far more prevalent in real-world and population-scale genealogical data; in contrast, Migrant Symmetry is designed to uncover asymmetries in directional movement that Stayer Symmetry alone cannot detect.

3.3.1. Stayer Symmetry

Stayer Symmetry assesses generational persistence by comparing how often grandparents and grandchildren remain in the same state as G2. For each G2 birth state X, we define:

- Descendant Retention Share (DRS) as the proportion of families where grandchildren (G3) remain in the G2 birth state X:

$$DRS(X) = \frac{\sum_{i:G2_i = X,\ G3_i = X} w_i}{\sum_{i:G2_i = X} w_i}$$

- Ancestral Retention Share (ARS) as the proportion of families where the grandparents (G1) remained in their own birth state, which is also the birth state of the parents (G2), denoted as state X:

$$ARS(X) = \frac{\sum_{i:G2_i = X, G1_i = X} w_i}{\sum_{i:G2_i = X} w_i}$$

The stayer symmetry index is:

$$ISI - S(X) = \frac{DRS(X) - ARS(X)}{DRS(X) + ARS(X)}$$

Positive values indicate that descendant retention exceeds ancestral presence, which suggests growing rootedness over generations. Negative values imply declining persistence, where ancestors were more likely to stay than their descendants.

3.3.2. Migrant Symmetry

Migrant Symmetry contrasts the origins of grandparents and the destinations of grandchildren, anchored on the G2 generation. For each G2 birth state X, we define

- *Ancestral Inflow Share*(AIS) as the proportion of families where grandparents (G1) were born outside the G2 birth state (i.e. inflow into X):

$$AIS(X) = \frac{\sum_{i:G2_i=X,\, G1_i \neq X} w_i}{\sum_{i:G2_i=X} w_i}$$

- Descendant Outflow Share (DOS) as the proportion of families where grandchildren (G3) were born outside the G2 birth state (i.e. outflow from X):

$$DOS(X) = \frac{\sum_{i:G2_i=X,\, G3_i \neq X} w_i}{\sum_{i:G2_i=X} w_i}$$

The resulting index

$$ISI - M(X) = \frac{AIS(X) - DOS(X)}{AIS(X) + DOS(X)}$$

Values near +1 indicate strong net attraction in which there are more grandparental inflow than grandchild outflow while values near −1 reflect net dispersion, where descendants are more likely to leave than ancestors were to arrive.

We apply the two indexes of Stayer Symmetry (ISI-S) and Migrant Symmetry (ISI-M) to eight overlapping eleven-year G1 cohorts spanning 1785 – 1796 through 1845 – 1856 to trace how settlement dynamics evolved over time.

4. Results

4.1. Cohort statistics

We used interquartile range (IQR) plots to show how the birth-year distributions of parents (Figure 1 for G2) and grandchildren (Figure 2 for G3) vary across successive eleven-year grandparent (G1) cohorts. In each plot, a horizontal line spans the middle 50% of birth years (the 25th–75th percentiles), and a small tick marks the median. We observe three key patterns: median birth years advance by roughly a decade with each successive G1 cohort, G2 interquartile ranges remain steady around 12–14 years, and G3 IQRs decline from 16 to 12 years over time.

Figure 1 and Figure 2 both confirm that birth-year distributions shift predictably from one generation to the next while revealing a subtle generational compression in the timing of grandchildren's births that probably reflects the demographic changes that occurred at the time, i.e. fertility decline, increased clustering of the ages of the parents when their children were born. For example, the IQR range of the birth date of the G2 generation of earliest grandparent cohort is 13 years while the G3 for the same cohort is 16 years. One would expect more variation in G3 than G2 birth dates. On the other hand, there were definitely changes over time: the medians for each grandparental birth cohort should be 10 years apart but they start to narrow: in G2, the last cohort is only 8 years different from the previous one. In G3, the narrowing starts 2 cohorts earlier. These medians help to anchor the data in the experience of a person born at a certain date. We look back at his/her grandparents and forward to his/her children.

Some of the consistency across cohorts may also reflect our sample requirements: each G1 couple must have at least two G2 children with complete birth information, and at

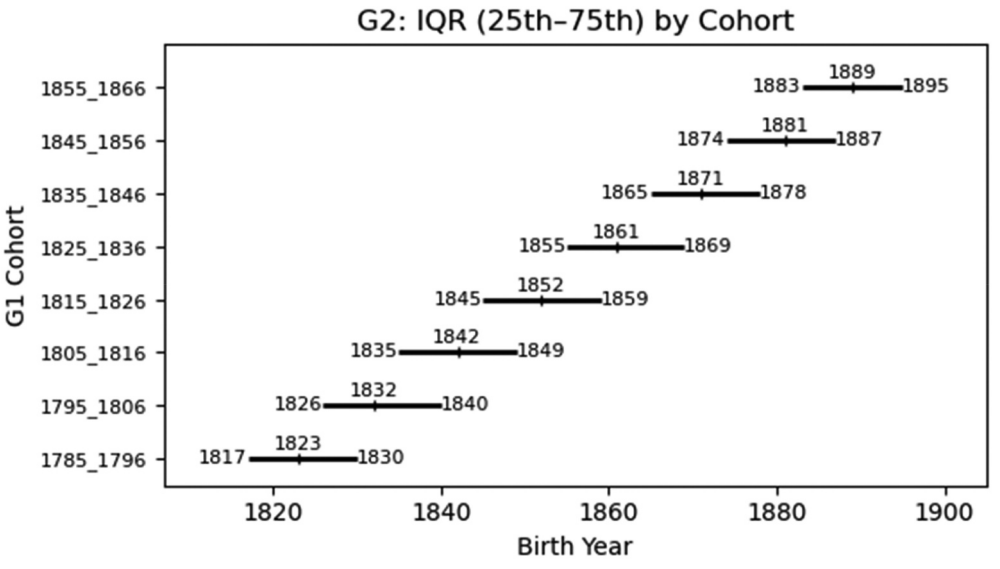


Figure 1. Interquartile range (IQR) plots of G2 birth year distribution by G1 cohort.

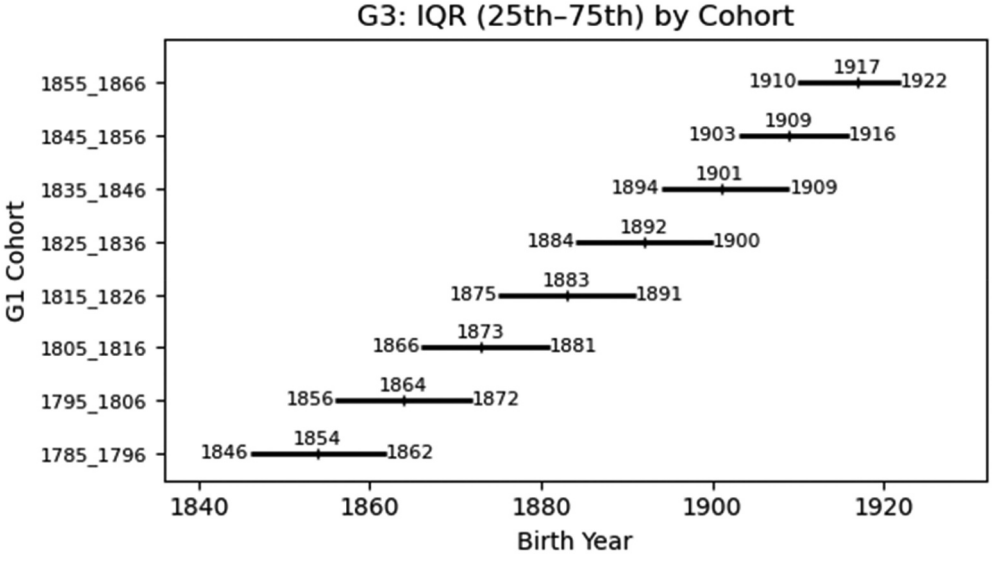


Figure 2. Interquartile range (IQR) plots of G3 birth year distribution by G1 cohort.

least two of those G2 children must each have two or more G3 offspring with complete records. These criteria ensure consistent lineage depth but favor larger and better-documented families, which likely suppresses variation in birth timing. Even so, the consistent shifts between cohorts and narrowing G3 ranges suggest that the overall pattern primarily reflects demographic changes rather than artifacts of data completeness.

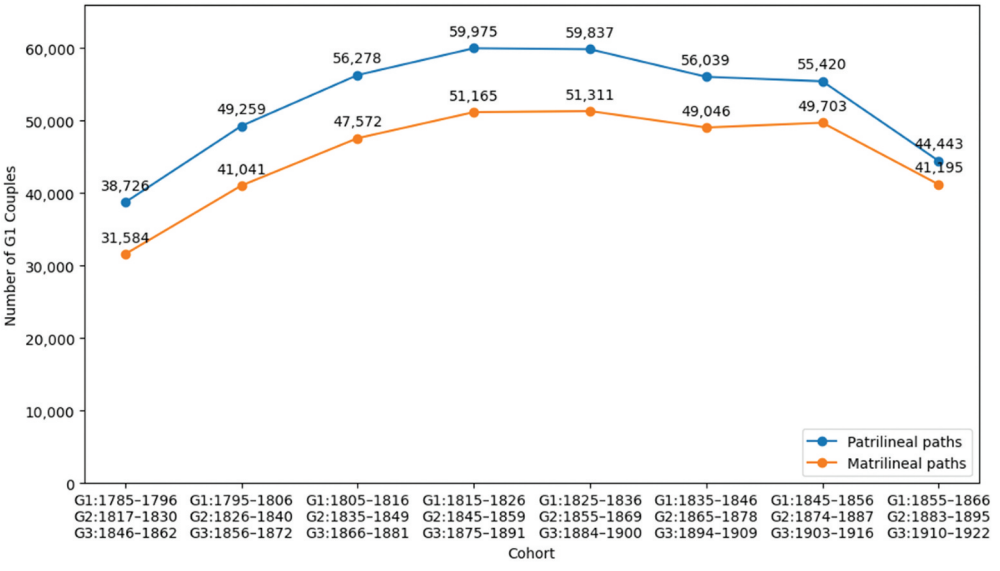


Figure 3. Frequency of patrilineal vs. matrilineal paths by cohort.

Figure 3 illustrates the frequency of patrilineal and matrilineal G1 couples across successive cohorts. On average across the eight cohorts, the patrilineal (G2-son) line sits about 15–16% above the matrilineal (G2-daughter) line. Early cohorts show the largest gap, but by 1855–66 it's fallen to under 10%. If we had truly complete, unbiased family data, the two lines would nearly overlap in a way that sons and daughters would be equally traceable. The steady narrowing of the gap likely reflects improving record-keeping and the Civil War's heavy toll on males, as well as the availability of full count census data to genealogists, which together made daughters almost as visible as sons in the genealogical record.

4.2. Changes in family migration over time

Before turning to the measures of symmetry, we begin with simple indicators of intergenerational persistence. Figure 4 illustrates the concentric rings showing intergenerational migration and persistence for all cohorts between 1785 and 1866 for the paternal-paternal path. In each panel, the inner ring depicts the proportion of families that stayed in the same state (stayer) versus moved to a new state (migrant) between G1 and G2. The outer ring similarly shows 'stayer' versus 'migrant' proportions for the transition from G2 to G3. Cohort labels correspond to the birth intervals used for G1, and the interquartile ranges of G2 and G3 birth years for that cohort. Overall, these concentric rings illustrate how interstate migration declined over successive cohorts and how staying increased from G1 to G2 through G2 to G3. Throughout all cohorts, G2 to G3 migration was at a lower rate than G1 to G2. Both transitions, however, converge toward high stayer rates by the grandparent cohorts born mid-century: among grandparents born 1845–1856 roughly 61.4% of G1 to G2 and 73.9% of G2 to G3 families stayed in-state; by the final cohort, grandparents born 1855–1866, those figures had risen to 65.6% and 74.5%,

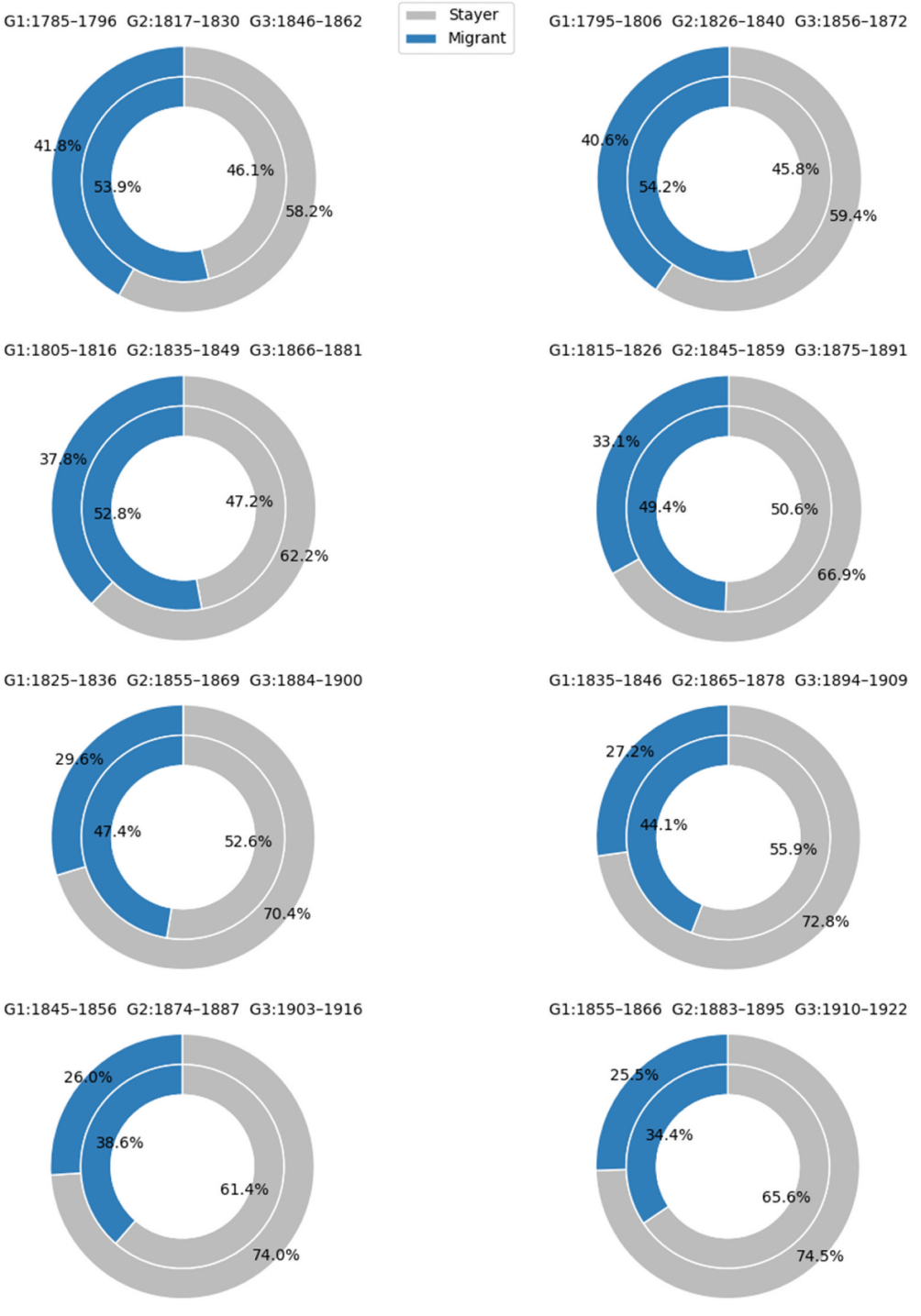


Figure 4. Concentric rings for each cohort of the g1fg2f (paternal-paternal) lineal path. The inner ring shows the share of families that stayed versus moved between G1 and G2; the outer ring shows the same for G2 to G3.

Table 1. Intergenerational migrant percentages for the first vs. the last cohorts, by path.

Lineal Path	G1: 1785–1796 (G1 to G2 / G2 to G3)	G1: 1855–1866 (G1 to G2 / G2 to G3)
Paternal-Paternal (g1fg2f)	53.884% / 41.840%	34.427% / 25.489%
Maternal-Paternal (g1mg2f)	47.689% / 41.835%	29.421% / 25.487%
Paternal-Maternal (g1fg2m)	54.101% / 37.423%	35.463% / 23.622%
Maternal-Maternal (g1mg2m)	48.050% / 37.454%	30.278% / 23.635%

respectively. In other words, interstate movement along the purely paternal line declined across three generations and ended in cohorts where a large majority of families remained within their birth state.

Table 1 reports the percentages of migrants from G1 to G2 and from G2 to G3 for the earliest cohort (1785–1796) and the latest cohort (1855–1866) along each lineal path. Comparing across all four lineal paths (Table 1), migration declines over generations and cohorts. Lineages traced through grandfathers show higher migration levels in the earlier cohorts, whereas those traced through grandmothers remain lower. Also, the four lineal paths produce nearly identical results, which shows that the sex of G2 has little to no effect on migration. Appendix Figure A1.1-A1.3 present these patterns in detail.

4.3. Spatiotemporal patterns of settlement

We present maps for the two Symmetry measures for each G1 cohort by state and discuss how they changed over time (Figures 5 and 6). The maps display the Symmetry indices calculated for each state based upon where G2 was born. The inner and outer ring proportions in Figure 4 form the components of the symmetry metrics: the stayer shares correspond to the retention proportions used in Stayer Symmetry (Ancestral Retention Share for G1 to G2 and Descendant Retention Share for G2 to G3), while the migrant shares correspond to the inflow and outflow proportions used in Migrant Symmetry (Ancestral Inflow Share for G1 to G2 and Descendant Outflow Share for G2 to G3).

To help readers interpret the Intergenerational Symmetry measures in Figures 5 and 6, Table 2 summarizes what each measure compares and how to read its values. Positive ISI-S scores (blue) signal growing rootedness: grandchildren (G3) are staying in their parents (G2)' state more often than grandparents (G1) did. Conversely, negative ISI-S scores (red) indicate declining persistence, where ancestral roots outnumber descendant stayers. A positive ISI-M value (blue) denotes net attraction, where grandparental inflow exceeds grandchild outflow, while a negative value (red) reveals net dispersion, indicating that descendants depart more frequently than ancestors. Finally, white indicates near-zero balance where persistence or migration is about the same across generations.

We structure our discussion of the Stayer- and Migrant-Symmetry maps into two complementary subsections. The first tracks how newly settled areas move from intense attraction and retention to equilibrium as successive waves push westward. The second follows the early settled areas or older core to emphasize how long-term kin dispersion eventually gives way to selective urban or resource-driven in-migration. Grouping the discussion this way lets us align each geographic regime with the specific mechanisms proposed by Easterlin, Turner, and Zelinsky while making the interpretation of Figures 5–6

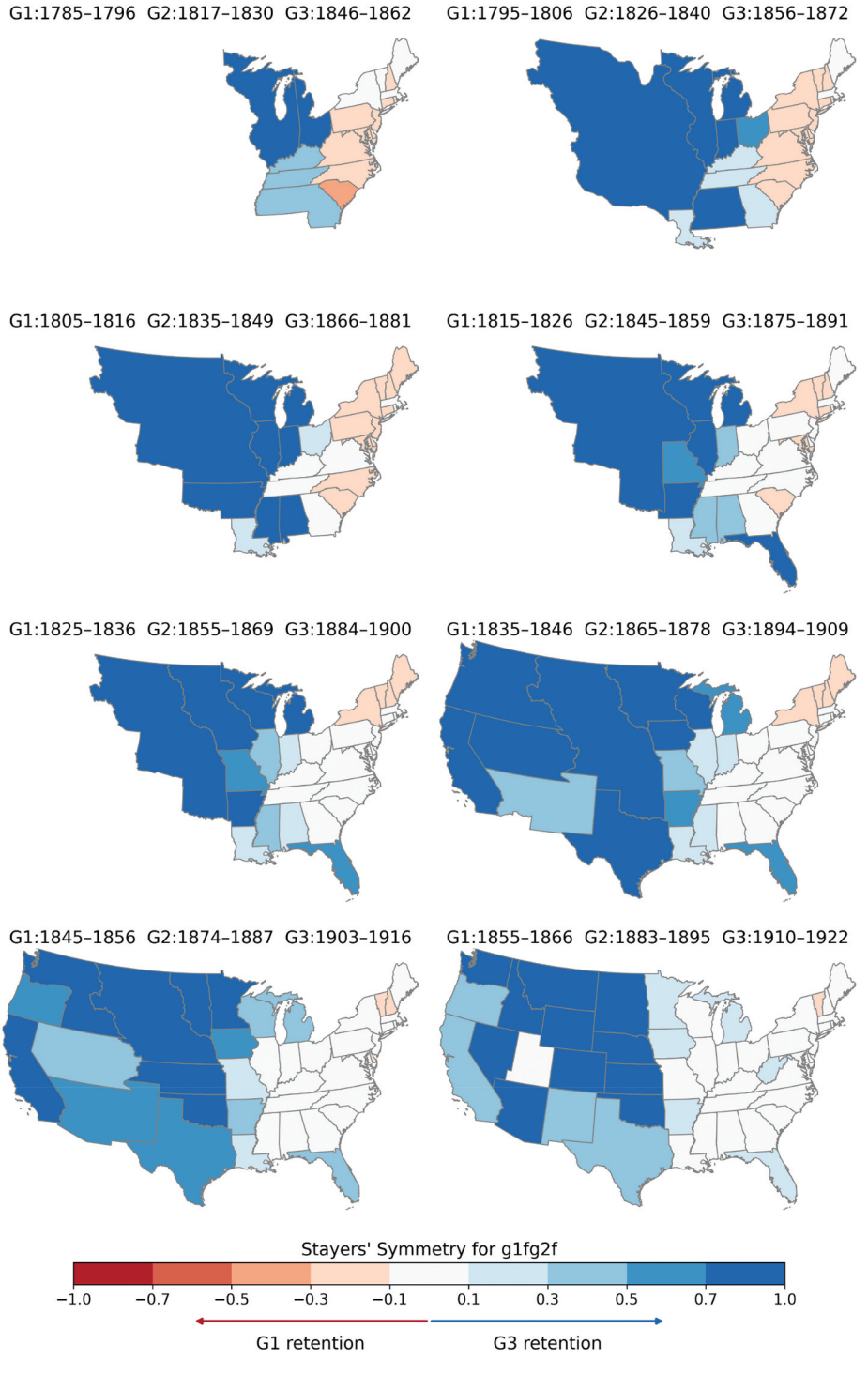


Figure 5. Paternal-paternal Stayer Symmetry for G1 (grandfather) cohorts 1785–1855.

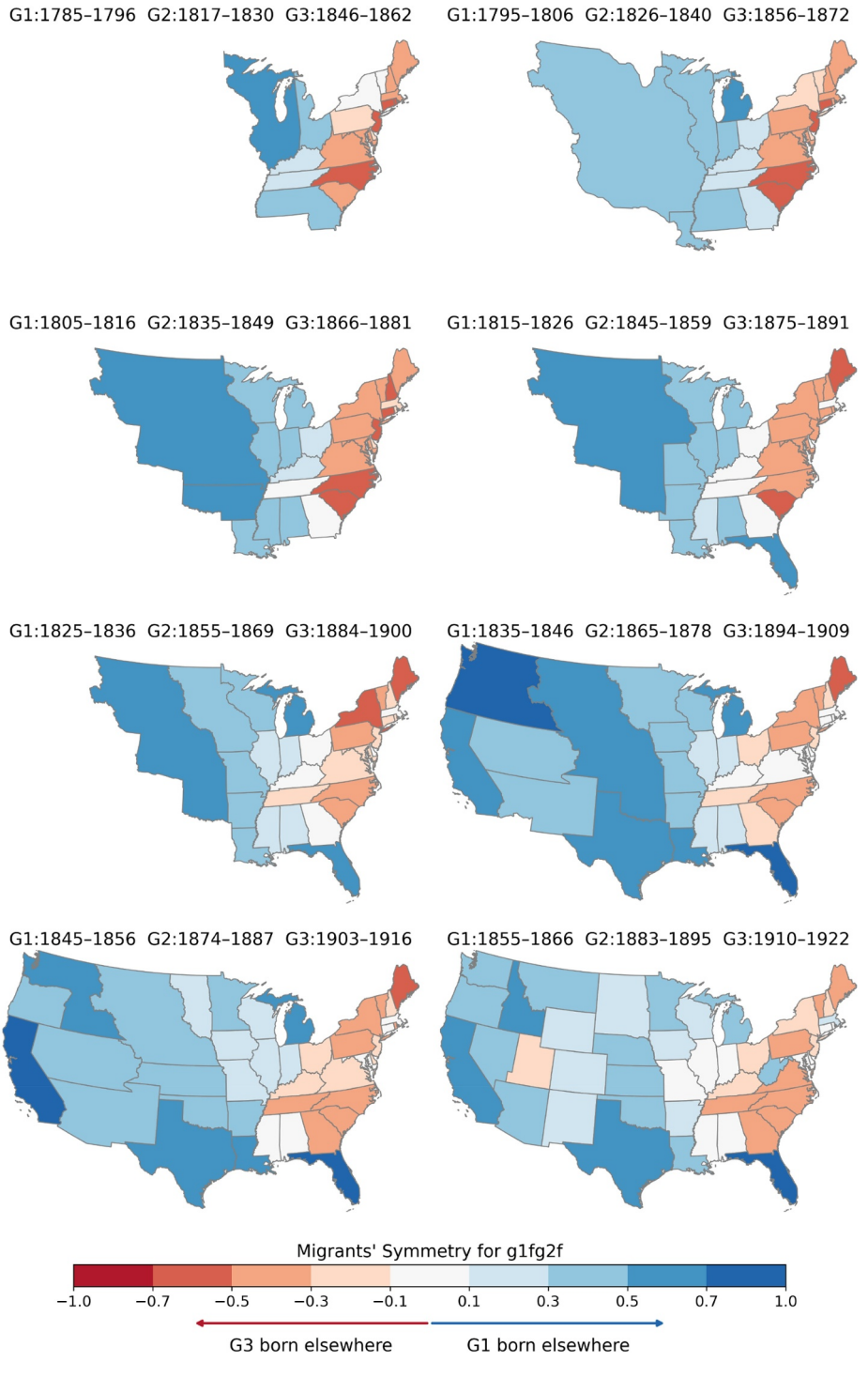


Figure 6. Paternal-paternal Migrant Symmetry for G1 (grandfather) cohorts 1785–1855.

Table 2. Stayer and Migrant Symmetry interpretation.

Index	What it compares	Positive (> 0)	Negative (< 0)	~ 0
Stayer Symmetry (ISI-S)	Share of G3 who stay vs. share of G1 who stayed, in the same G2 state	**Growing rootedness**: more G3 remain than G1 did	**Fading roots**: G3 disperse more than G1	No net change
Migrant Symmetry (ISI-M)	Share of G1 who moved in vs. share of G3 who moved out	**Net magnet:** more G1 arrived than G3 departed	**Net dispersal:** more G3 depart than G1 arrived	In- and out-flows balanced

easier to follow. The patterns we describe should be read as informed speculations that illustrate how the maps might align with classic hypotheses, rather than as definitive, confirmatory tests of those theories.

4.3.1. Early settlement, consolidation, and successive waves

The maps in Figures 5 and 6 show a consistent east-west gradient of settlement across successive cohorts. On Stayer Symmetry, frontier zones appear in blue, reflecting that grandchildren (G3) were more likely than grandparents (G1) to remain in the parental state (G2), while older zones fade to white or red as persistence weakens. On Migrant Symmetry, the same gradient appears but is weaker, since inflows of grandparents were generally less intense than the rootedness of grandchildren. The two measures capture an origin-destination dynamic: regions with different darkness of red/orange hues act as sources of out-migration, blue regions as destinations where kin concentrate, and white zones as transitional places where inflow and outflow balance. The origin-destination relationship is not a one-to-one mapping because families originate from and spread across multiple origins and destinations rather than moving in a single stream. Still, the overall gradation of darker to lighter reds as sources and darker to lighter blues as destinations remains clear. By the final two cohorts, the frontier effectively ends around the parent generation (G2), yet positive Stayer Symmetry persists, which suggests that migration within already settled regions continued, much like the flux seen in contemporary migration systems. The fading of Migrant Symmetry into lighter blues at the interior may provide tentative evidence for Turner's and Easterlin's claim that each frontier drew most strongly from the tier just behind it rather than directly from the far East.

What stands out across the long run is how little the darkest blues in Stayer Symmetry change after the second cohort: much of the West remains deeply rooted until the very end, only lightening in the seventh and eighth cohorts. This long persistence suggests that the West functioned as a kin frontier far longer than earlier regions, perhaps helping explain the durability of the western frontier. Whether sustained by continued inflow or by movement within the West, the resilience of kin rootedness contrasts with Migrant Symmetry, where attraction weakens earlier. Indeed, the successive wave story is clearest in the Migrant Symmetry maps: the progressive lightening of western regions tracks Turner's idea of a frontier closing around 1890, when G2 of the last cohort was born. By contrast, Stayer Symmetry shows only a partial break-up of the dark blues, with places like Texas and California shifting lighter, which indicates that G3 was moving out as opposed to those states' earlier role as frontier destinations.

4.3.2. Evolution of the settled areas and the emergence of urban magnets

The eastern seaboard tells a story that complements the narrative of the early western settlement. From the very first cohort the Atlantic core is already dispersing kin: in the 1817–1830 parental cohort (G2) almost every state from Maine to Georgia records negative Migrant Symmetry (medium to dark red) and negative (light red) or neutral (white) Stayer Symmetry. Easterlin's land-scarcity logic fits these colors: high prices and limited acreage made it impossible to endow large sibships with viable farms. In the next two cohorts, dispersion intensified in the next cohorts, especially across the Carolinas, Connecticut, and New Jersey, where grandchildren leave at over twice the rate of grandparent inflow. Meanwhile, urban New England stays relatively stable (near white), likely due to urban employment, factories and ports limiting kin outflow.

A clearer turning point emerges in the fourth cohort (parents born 1845–1859; children born 1875–1891). On Stayer Symmetry the first interior tier begins to neutralize: Pennsylvania, New Jersey, North Carolina and Maine shift from light red to white. Ohio follows suit after decades of positive symmetry with decreasing lightness of blue for both Stayer and Migrant Symmetry. The same states, however, stay red with decreased lightness on Migrant Symmetry, indicating that out-migration is slowing, yet in-migrants have not arrived in sufficient numbers to offset past losses.

Farther south, Alabama transitioned from high retention to neutrality, while Florida emerged as a new magnet, attracting large inflows as railroads opened the interior. By the late nineteenth century, industrial and resource frontiers such as Washington, Florida, California, and the Great Plains, posted the strongest positive Migrant Symmetry values, which reflects new economic opportunities in citrus, timber, mining, and rail industries. Overall, the eastern states evolved from kin exporters to demographically balanced regions, while new industrial and agricultural zones in the South and West became the primary attractors of family migration. The divergence between Stayer and Migrant Symmetry along this gradient captures the broader transition from a land-based to an urban-industrial demographic regime.

5. Discussion

The changing relationship between Stayer Symmetry (ISI-S) and Migrant Symmetry (ISI-M) reflects successive stages in the demographic transformation of the long-settled eastern states. Red on both sets of maps (Figures 5 and 6) signals kin loss; when Stayer Symmetry (ISI-S) turns white but Migrant Symmetry (ISI-M) turns blue, inflows begin to offset dispersing lineages; when ISI-S remains blue but ISI-M fades to white, retention is still high but attraction is declining; and when ISI-M stays red while ISI-S shifts to white, out-migration slows while inflows remain low. These changes illustrate Easterlin's fertility-inheritance gradient (Easterlin, 1976a), Turner's moving frontier (Turner & Wilson, 1920), and Zelinsky's mobility transition as overlapping layers of the same continental process (Zelinsky, 1971). In Zelinsky's terms, persistent out-migration from older eastern states reflects Phase II (mass rural and frontier migration), lessening frontier migration and rapid settlement correspond to Phase III, and eventual stabilization in East and West marks Phase IV (high but steady urban and circular movement). By the final cohort the country presents a clear split: the trans-Mississippi West remains blue on both symmetry measures, showing that kin attraction still

outweighs dispersal, whereas the eastern seaboard is mostly white on ISI-S yet still shades light red to orange on ISI-M. Rootedness in the East has stabilized, but a modest net out-migration of parental lines persists. These patterns suggest that eastern states continue to exhibit net dispersal even after overall inflow and outflow have moved closer to equilibrium. Overall, these maps capture the end point of America's long transition from a land-rich agrarian republic to an urban-industrial nation of highly mobile kin networks.

Because our three-generation sample includes only US-born grandparents, parents, and grandchildren, the symmetry measures capture strictly domestic kin dispersion and omit the substantial inflows of foreign-born immigrants. In the initial data, a G1 couple could include one US-born and one foreign-born partner. However, once we applied the sample selection criteria that require that each G1 couple have at least two US-born G2 children with complete birth records and that at least two of those G2 children each have two US-born G3 offspring, the remaining sample consisted entirely of native-born G1 couples. In other words, the lineage requirements filtered out foreign-born grandparents, leaving only fully US-born family lines. As a result, the migration patterns we observe reflect strictly internal movements of native-born populations, and states that were major immigrant gateways (e.g. New York, Pennsylvania, Illinois) appear here only in terms of the mobility and retention of established American lineages, not the demographic effects of immigration.

Data incompleteness remains a challenge, with many individuals lacking exact birth or death dates and some kin ties, especially among mothers, under-documented. To ensure reliability, we include only families with well-defined paternal, maternal, and sibling relationships, focusing on core kin units with sufficient data to characterize dispersion. Despite efforts to trace maternal lines, Figure 3 and Table 2 reveal a persistent male-line bias in genealogical records. This likely reflects both name-change barriers and broader tracing challenges for women. Women adopt their husband's surname at marriage, which makes it harder to trace without marriage records that link maiden and married names. Unlike men, who keep the same surname, women's records in the US are further fragmented by state-level variation in marriage registration. As a result, women's birth and death information are more often missing from genealogical data, and women who migrated after marriage are systematically harder to follow. This absence can contribute to lower observed maternal dispersion: while Ravenstein's hypothesis that women migrate shorter distances (Ravenstein, 1885) may partly explain the pattern, missing data on maternal birthplaces likely also play a role. Linking genealogical records with nineteenth-century census data offers one way to disentangle whether maternal underrepresentation reflects actual migration behavior or data limitations.

Our data emphasize family-based migration, since many moves in the nineteenth century occurred at the household level (Adams & Kasakoff, 1992; Darroch, 1981). Our intergenerational approach is designed to capture the dispersion of kin across generations rather than the movements of single or non-familial migrants. Individuals who appear to have migrated alone may in fact have traveled with in-laws or extended kin not fully traceable in genealogical data. Because the relative balance of family versus individual migration cannot be determined from genealogical sources alone, linking these records to census data that enumerate all household members and coresidents provides an important avenue for future research.

Although we use genealogical data in this article, our approach could also be applied to census sources with birthplace information. Linked census data such as IPUMS Multigenerational Longitudinal Panel (MLP) connect individuals across two consecutive censuses, which makes it possible to reconstruct multi-generation birthplace paths. In census years when the birthplace of parents is recorded, one could compare the birthplaces of coresident parents (G2) and children (G3) with the reported birthplace of the grandparents (G1). A key limitation, however, is that MLP begins by linking men, so gendered lineal paths may be incomplete, and those that can be fully traced represent only a subset of the census population.

Easterlin observed that it took longer for eastern states to reach settlement saturation than the West: New York required 60 years to go from 10% to 90% settled, while Nebraska reached that threshold in just 30 years, likely due to better transportation and a larger available population. Although our data begin after early eastern settlement was underway, our generational time steps align with Easterlin's finding that out-migration began roughly one generation after initial settlement, and this interval shortened further west. Turner's claim that the frontier closed in 1890 is visible in our final G2 cohort (born 1883–1895), where dark blue on the Stayer Symmetry map fades in the West, and western states begin to resemble eastern ones in retention patterns. Migrant Symmetry similarly shows a stabilization across the Midwest, with inflows and outflows coming into closer balance.

We find little difference in intergenerational kin dispersion between the North and South, which may reflect data limitations: genealogies disproportionately represent White landholding families in both regions. Although Easterlin focused on the North due to greater data availability and its more homogeneous agricultural structure, southern agriculture also included substantial family migration, especially westward with cotton cultivation. Differences in inheritance, such as the rarity of subdividing large plantations, did not prevent southern children, including those from wealthy households, from migrating west.

6. Conclusion

This study repositions kinship at the center of studying kin dispersion and migration by introducing two intergenerational dispersion measures: Stayer Symmetry and Migrant Symmetry. By tracing three linked generations within a population-scale kinship network, the Intergenerational Symmetry Index (ISI) reveals how family lineages, rather than isolated individuals, structured the geography of nineteenth-century settlement.

Because genealogical records contain far more birthplaces than deathplaces, our approach spans a broad temporal frame and captures kin dispersion across three generations. Centering on parental birth cohorts, we trace lineages backward to grandparent birthplaces and forward to those of grandchildren to study demographic change, spatial mobility, and the evolving geography of family persistence and movement.

Applying the intergenerational dispersion indices to eight eleven-year grandparent birth cohorts (G1 = 1785–1866) across the contiguous United States reveals a westward-moving ridge of high kin cohesion, followed by neutralization, consistent with Turner's frontier stages, Easterlin's fertility dynamics and Zelinsky's demographic transitions of the frontier. Although Easterlin argued that fertility-migration dynamics applied mainly in the

North (Easterlin et al., 1978), our results show similar patterns across the entire US, including the plantation South, at least for White families. As we have shown elsewhere, the genealogical trees used here substantially underrepresent Black individuals (Koylu & Kasakoff, 2025).

States trailing the frontier's leading edge tend to neutralize kin retention within roughly one generation (~30 years) after peak inflow. Meanwhile, resource and industrial frontiers such as California, Florida, and the Pacific Northwest experienced renewed attraction despite lower stayer rates. Future work at the county level could reveal finer-scale dynamics obscured by state averages, such as urban – rural contrasts, intra-state regional differences, or corridor-based patterns tied to railroads or early industrialization. Combining our spatial metrics with historical covariates (e.g. land prices, infrastructure development, local fertility rates) offers a promising direction for testing the causes of kin dispersion in history.

More broadly, our results underscore the value of treating kinship networks, not just individuals, as a unit of spatial demographic analysis. Standard census-based measures capture only snapshots of migration between enumerations, often missing kin relationships and movement patterns across generations. By contrast, genealogical data allow us to reconstruct lineage-based birthplace sequences that reveal kin processes of co-location, and dispersal. As genealogical datasets continue to grow, the kin-centered framework developed here provides a bridge between family history and broader questions of migration, settlement, and demographic transformation.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by The National Science Foundation (NSF) under Award No. [2215568].

Data availability statement

The dataset for this study is publicly available through OPENICPSR at https://doi.org/10.3886/E237845V2. It includes three-generation birthplace sequences for the United States from 1785–1922, derived from population-scale family trees, as well as state-level intergenerational metrics: Stayer Symmetry, and Migrant Symmetry.

References

Adams, J. W., & Kasakoff, A. B. (1992). The farm family economy in the American North, 1775-1875. An exploration of sibling differences. *Continuity and Change*, *7*(3), 357–375. https://doi.org/10.1017/S0268416000001715

Adams, J. W., Kasakoff, A., & Kok, J. (2002). Migration over the life course in XIXth century Netherlands and 632 the American north: A comparative analysis based on genealogies and population registers. *Annales de Démographie Historique 2002*, *104*(2), 5–27. https://doi.org/10.3917/adh.104.0005

Alter, G., & Hacker, J. D. (2024). The impact of multiple births on fertility: Stopping and spacing in the United States during the demographic transition. *Demography*, *61*(5), 1509–1533. https://doi.org/10.1215/00703370-11577526

Beise, J., & Voland, E. (2002). A multilevel event history analysis of the effects of grandmothers on child mortality in a historical German population (Krummhörn, Ostfriesland, 1720-1874). *Demographic Research*, *7*, 469–498. https://doi.org/10.4054/DemRes.2002.7.13

Black, S. E., Duzett, N., Lleras-Muney, A., Pope, N. G., & Price, J. (2023). *Intergenerational correlations in longevity*.

Bras, H., & Tilburg, T. V. (2007). Kinship and social networks: A regional analysis of sibling relations in twentieth-century Netherlands. *Journal of Family History*, *32*(3), 296–322. https://doi.org/10.1177/0363199007300244

Bratti, M., Fiore, S., & Mendola, M. (2020). The impact of family size and sibling structure on the great Mexico-USA migration. *Journal of Population Economics*, *33*(2), 483–529. https://doi.org/10.1007/s00148-019-00754-5

Carlson, J. E., & Dillman, D. A. (1983). Influence of kinship arrangements on farmer innovativeness. *Rural Sociology*, *48*(2), 183.

Casterline, J. B. (2001). *Diffusion processes and fertility transition: Selected perspectives*.

Castillo, M., & Helgertz, J. (2025). The role of migrant networks on economic outcomes during the age of mass migration: Swedish immigrants in the US 1900-1920. *The History of the Family*, *30*(2), 193–216. https://doi.org/10.1080/1081602X.2024.2441813

Daelemans, S., Vandevoorde, J., Vansintejan, J., Borgermans, L., & Devroey, D. (2013). The use of family history in primary health care: A qualitative study. *Advances in Preventive Medicine*, *2013*, 1–8. https://doi.org/10.1155/2013/695763

Darroch, G. (1981). Migrants in the nineteenth century: Fugitives or families in motion? *Journal of Family History*, *6*(3), 257–277. https://doi.org/10.1177/036319908100600302

Easterlin, R. A. (1976a). Factors in the decline of farm family fertility in the United States: Some preliminary research results. *Journal of American History*, *63*(3), 600–614. https://doi.org/10.2307/1887347

Easterlin, R. A. (1976b). Population change and farm settlement in the northern United States. *The Journal of Economic History*, *36*(1), 45–75. https://doi.org/10.1017/S002205070009450X

Easterlin, R. A., Alter, G., & Condran, G. A. (1978). Farms and farm families in old and new areas: The northern states in 1860. In T. K. Hareven & M. A. Vinovskis (Eds.), *Family and population in 19th century America* (pp. 22–84). Princeton University Press. https://doi.org/10.1515/9781400869398-004

Engelhardt, S. C., Bergeron, P., Gagnon, A., Dillon, L., & Pelletier, F. (2019). Using geographic distance as a potential proxy for help in the assessment of the grandmother hypothesis. *Current Biology*, *29*(4), 651–656.e3. https://doi.org/10.1016/j.cub.2019.01.027

Ferrie, J. P. (1997). *Migration to the frontier in mid-nineteenth century America: A re-examination of Turner's 'safety valve'*.

Galenson, D. W., & Pope, C. L. (1989). Economic and geographic mobility on the farming frontier: Evidence from Appanoose County, Iowa, 1850–1870. *The Journal of Economic History*, *49*(3), 635–655. https://doi.org/10.1017/S0022050700008792

Glover, S., & Towner, M. (2009). Long-distance dispersal to the mining frontier in late 19th century Colorado. *Behaviour*, *146*(4–5), 677–700. https://doi.org/10.1163/156853908X395558

Godley, J. (2001). Kinship networks and contraceptive choice in Nang Rong, Thailand. *International Family Planning Perspectives*, *27*(1), 4–41. https://doi.org/10.2307/2673799

Guest, A. M. (1987). Notes from the national panel study: Linkage and migration in the late nineteenth century. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, *20*(2), 63–77. https://doi.org/10.1080/01615440.1987.9955260

Hacker, J. D. (1999). Child naming, religion, and the decline of marital fertility in nineteenth-century America. *The History of the Family*, *4*(3), 339–365. https://doi.org/10.1016/S1081-602X(99)00019-6

Hacker, J. D., Helgertz, J., Nelson, M. A., & Roberts, E. (2021). The influence of kin proximity on the reproductive success of American couples, 1900-1910. *Demography*, *58*(6), 2337–2364. https://doi.org/10.1215/00703370-9518532

Hacker, J. D., & Roberts, E. (2017). The impact of kin availability, parental religiosity, and nativity on fertility differentials in the late 19th-century United States. *Demographic Research*, *37*(34), 1049. https://doi.org/10.4054/DemRes.2017.37.34

Hall, P. K., & Ruggles, S. (2004). "Restless in the midst of their prosperity": New Evidence on the internal migration of americans, 1850-2000. *Journal of American History*, *91*(3), 829–846. https://doi.org/10.2307/3662857

Healy, O., & Dunifon, R. (2025). Grandchildren's spatial proximity to grandparents and intergenerational support in the United States. *Demographic Research*, *52*, 1097–1110. https://doi.org/10.4054/DemRes.2025.52.34

Hey, D. (2010). *The Oxford companion to family and local history*. OUP Oxford.

Hinde, A. (2004). The use of nineteenth-century census data to investigate local migration. *Local Population Studies*, *73*(Autumn), 8–28.

Hudson, J. C. (1976). Migration to an American frontier*. *Annals of the Association of American Geographers*, *66*(2), 242–264. https://doi.org/10.1111/j.1467-8306.1976.tb01088.x

Jennings, J. A., Sullivan, A. R., & Hacker, J. D. (2012). Intergenerational transmission of reproductive behavior during the demographic transition. *The Journal of Interdisciplinary History*, *42*(4), 543–569. https://doi.org/10.1162/JINH_a_00304

Kok, J., & Bras, H. (2008). Clustering and dispersal of siblings in the North-Holland countryside, 1850-1940. *Historical Social Research/Historische Sozialforschung*, *33*(3(125)), 278–300. http://www.jstor.org/stable/20762310

Koylu, C., Guo, D., Huang, Y., Kasakoff, A., & Grieve, J. (2021). Connecting family trees to construct a population-scale and longitudinal geo-social network for the U.S. *International Journal of Geographical Information Science*, *35*(12), 2380–2423. https://doi.org/10.1080/13658816.2020.1821885

Koylu, C., & Kasakoff, A. (2022). Measuring and mapping long-term changes in migration flows using population-scale family tree data. *Cartography and Geographic Information Science*, *49*(2), 154–170. https://doi.org/10.1080/15230406.2021.2011419

Koylu, C., & Kasakoff, A. B. (2025). Ethical challenges in analyzing and mapping historical demographic changes and migration using population-scale family trees. *Cartographic Perspectives*, (105), 55–62–55–62. https://doi.org/10.14714/CP105.1945

Koylu, C., Kasakoff, A. B., & Torkashvand, M. (2025). Us migration from 1850 to 1920: A comparison of family trees with linked census data. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, *58*(3), 160–174. https://doi.org/10.1080/01615440.2025.2516022

Lee, E. S., & Lee, A. S. (1960). Internal migration statistics for the United States. *Journal of the American Statistical Association*, *55*(292), 664–697. https://doi.org/10.1080/01621459.1960.10483367

Malin, J. C. (1935). *The turnover of farm population in Kansas*.

Maloney, T. N., Hanson, H., & Smith, K. R. (2014). Occupation and fertility on the frontier: Evidence from the state of Utah. *Demographic Research*, *30*, 853. https://doi.org/10.4054/DemRes.2014.30.29

Mineau, G. P., Bean, L. L., & Anderton, D. L. (1989). Migration and fertility: Behavioral change on the American frontier. *Journal of Family History*, *14*(1), 43–61. https://doi.org/10.1177/036319908901400103

Molloy, R., Smith, C. L., & Wozniak, A. (2011). Internal migration in the United States. *Journal of Economic Perspectives*, *25*(3), 173–196. https://doi.org/10.1257/jep.25.3.173

Nelson, M. A. (2020). The decline of patrilineal kin propinquity in the United States, 1790-1940. *Demographic Research*, *43*, 501–532. https://doi.org/10.4054/DemRes.2020.43.18

Otterstrom, S. M., Price, J. P., & Van Leeuwen, J. (2022). Using linked census records to study shrinking cities in the United States from 1900 to 1940. *Professional Geographer*, *74*(1), 88–101. https://doi.org/10.1080/00330124.2021.1952882

Parkerson, D. H. (1982). How mobile were nineteenth-century Americans? *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, *15*(3), 99–109. https://doi.org/10.1080/01615440.1982.10594085

Price, D. O. (1953). Estimates of net migration in the United States, 1870-1940. *American Sociological Review*, *18*(1), 35–39. https://doi.org/10.2307/2087846

Ruggles, S. (1994). The transformation of American family structure. *American Historical Review*, *99* (1), 103–128. https://doi.org/10.1086/ahr/99.1.103

Ruggles, S. (2003). Multigenerational families in nineteenth-century America. *Continuity and Change*, *18*(1), 139–165. https://doi.org/10.1017/S0268416003004466

Sear, R., & Coall, D. (2011). How much does family matter? Cooperative breeding and the demographic transition. *Population & Development Review*, *37*(s1), 81–112. https://doi.org/10.1111/j.1728-4457.2011.00379.x

Spencer, R. W., & Otterstrom, S.M., (2024). Massive genealogies distinguish frontier from steady-state internal migration. *arXiv*. https://doi.org/10.48550/arXiv.2410.18235

Stewart, J. I. (2006). Migration to the agricultural frontier and wealth accumulation, 1860-1870. *Explorations in Economic History*, *43*(4), 547–577. https://doi.org/10.1016/j.eeh.2005.08.002

Stewart, J. I. (2009). Economic opportunity or hardship? The causes of geographic mobility on the agricultural frontier, 1860-1880. *The Journal of Economic History*, *69*(1), 238–268. https://doi.org/10.1017/S0022050709000394

Stewart, J. I. (2012). Migration to US frontier cities and job opportunity, 1860-1880. *Explorations in Economic History*, *49*(4), 528–542. https://doi.org/10.1016/j.eeh.2012.06.003

Thernstrom, S., & Knights, P. R. (1970). Men in motion: Some data and speculations about urban population mobility in nineteenth-century America. *The Journal of Interdisciplinary History*, *1*(1), 7–35. https://doi.org/10.2307/202408

Turner, F. J., & Wilson, W. (1920). *The frontier in American history*. H. Hold and Company.

Williamson, J. G. (1974). Migration to the new world: Long term influences and impact. *Explorations in Economic History*, *11*(4), 357–389. https://doi.org/10.1016/0014-4983(74)90025-4

Yasuba, Y. (1962). *Birth rates of the white population in the United States, 1800-1860 : An economic study*. Johns Hopkins Press.

Zelinsky, W. (1971). The hypothesis of the mobility transition. *Geographical Review*, *61*(2), 219–249. https://doi.org/10.2307/213996