

Inheriting Geography*

Marriage and the Intergenerational Spatial Transmission of Opportunity

Luca Looser[†]

October 2026

Updated regularly, please click [here](#) for the latest version.

Abstract

Children inherit not only resources from their parents but also a geography: where they grow up and the places to which their family is connected. I study how marriage shapes this inheritance across U.S. families from 1880 to 1950 and what it implies for children's access to opportunity. Marriage expands the set of family-connected places by 59%, while the geographic concentration of marriage explains most of the overlap between spouses' premarital family geographies. Within families, people are more likely to settle near a spouse's relatives and their children are more likely to settle in places introduced through the marriage, including places where neither parent ever lived. I develop a two-generation location model to quantify these effects. Family ties strongly anchor parents to inherited locations: removing family proximity makes couples about 20 percentage points less likely to settle in a spouse's premarital commuting zone and weakens the relationship between parents' socioeconomic status and the destination wage their children reach by about one third. The constraint is strongest for children of poorer parents. Most of the intergenerational effect operates through where parents settle rather than through the destinations children inherit directly.

*I am grateful to Elisa Giannone and Dávid Krisztián Nagy for continuous support. All errors are my own.

[†]Department of Economics, Universitat Pompeu Fabra, Carrer de Ramon Trias Fargas, 25–27, Sant Martí, 08005 Barcelona. Email: luca.looser@upf.edu.

1 Introduction

Where children grow up shapes economic opportunity and intergenerational mobility (Chetty & Hendren, 2018a, 2018b; Chetty, Hendren, Kline, & Saez, 2014). Yet the geography in which children are raised is itself partly inherited: parents choose where the household settles and family ties connect children to places beyond their childhood location. Existing work shows that barriers to mobility limit moves to higher-opportunity places and that social connections shape where migrants go (Bergman et al., 2024; Blumenstock, Chi, & Tan, 2025; Stuart & Taylor, 2021). Less is known about how these spatial conditions are formed within families in the first place and what they imply for the next generation.

This paper studies one central source of this inherited geography: marriage. Marriage combines not only the resources and family backgrounds of two spouses, but also their connections to places. Each spouse enters marriage with relatives living in different locations. Bringing these family geographies together can change where the couple settles and hence where their children grow up. It also connects children to destinations through both sides of the family. This raises a central question: how does marriage shape the geography children inherit and the opportunities they can reach?

I study this question using linked U.S. families from 1880 to 1950. I reconstruct the locations of each spouse’s blood relatives before marriage, observe where the couple subsequently settles and follow their children to their adult locations. Across marriages, combining the two spouses’ family geographies expands the set of connected commuting zones (CZs) by 59%. Most of the overlap between the two sides is geographic: reassigning spouses within the same premarital CZ increases their combined geography by only 1.4%, while broader reassignments generate much less overlap.

These new connections affect where both generations live. Comparing siblings from the same family, destinations connected only through the spouse are about 0.3 percentage points more likely to be chosen after marriage. Settlement also falls when the spouse’s only relative at a destination dies before the couple chooses where to live, but not when the death occurs afterward. The effect persists into the next generation. Children are 0.183 percentage points more likely to settle in destinations introduced through one side of their parents’ marriage, including places where neither parent ever lived.

I develop a two-generation location model to quantify how these family connections shape spatial persistence. Marriage matters through two channels: it gives children connections to new destinations and it changes where their parents establish the household. The model implies that family ties are a substantial source of persistence in place. Removing family proximity from both generations makes parents about 20 percentage points less likely to settle in a spouse’s premarital CZ and increases children’s access to higher-opportunity destinations. More importantly, the relationship between parental socioeconomic rank and the quality of

children's destinations falls by about one third and the gap in access to top-quartile destinations between children of high- and low-status parents essentially disappears.

These constraints are strongest for families from lower socio-economic backgrounds. Removing family ties raises access to above-median destinations by roughly 4–5 percentage points for children in the bottom three parental quintiles, but by less than 2 points near the top. The same gradient appears when family connections are broadened. When both spouses bring family connections from across the country instead of only from their own CZs, children reach better destinations and the intergenerational persistence of place quality falls by about 30%. The gains are particularly large for families starting in low-opportunity places.

Most of the intergenerational effect operates through where parents settle. When I broaden the family connections brought into the household through one spouse, changes in parental settlement account for nearly three quarters of the increase in children's access to top-quartile destinations. Directly inherited connections account for the remainder. Parents respond mainly because they value proximity to their own relatives rather than because they anticipate their children's future location choices. Family geography therefore becomes intergenerational largely because parents follow their own family connections and the place they choose becomes the child's origin.

Related literature These findings contribute to three literatures. First, the paper adds to work on spatial adjustment and the role of social connections in migration. Existing research shows that local ties can discourage moving away and that connections at destinations make migration easier (Bound & Holzer, 2000; Ganong & Shoag, 2017; Ilyés, Boza, Lórinicz, & Eriksson, 2023; Jia, Molloy, Smith, & Wozniak, 2023; Stuart & Taylor, 2021; Zabek, 2024). This work establishes that the location of social connections matters for mobility, but largely takes that geography as given. I study how marriage reshapes it by combining the relatives of two spouses across places. These new connections affect where couples settle and remain relevant for their children, showing how the geography of family ties is itself formed and transmitted across generations.

Second, the paper contributes to work on household migration and family proximity. Since Mincer (1978), migration has been studied as a joint household decision, with subsequent work examining joint search, spouses' gains from moving and proximity to parents and in-laws (Compton & Pollak, 2014; Costa & Kahn, 2000; Guler, Guvenen, & Violante, 2012; Kaplan, 2012; Konrad, Künemund, Lommerud, & Robledo, 2002). This work shows that both spouses and their families shape household location. I study how marriage changes the geography of those family connections. By observing both spouses' relatives before marriage and comparing siblings who marry spouses connected to different places, I identify destinations brought into the household through marriage. A spouse therefore affects migration not only through their

own location preferences, but also through family connections to new destinations.

Third, the paper contributes to research on marriage and intergenerational transmission. Existing work studies how marriage combines education, earnings, wealth and family background and how these characteristics are passed across generations (Chiappori, Costa Dias, Meghir, & Zhang, 2025; Eika, Mogstad, & Zafar, 2019; Fernández & Rogerson, 2001; Greenwood, Guner, Kocharkov, & Santos, 2014; Kremer, 1997). I study an additional dimension: geography. Marriage combines two sets of relatives living in different places, which changes where parents settle and the destinations to which their children remain connected. This extends the study of intergenerational transmission from the resources and characteristics parents bring into the household to the places the next generation starts from and can more easily reach.

The remainder of the paper proceeds as follows. Section 2 describes the linked family lineages and marriage sample. Section 3 establishes how marriage changes family geography and how couples and children respond to the resulting connections. Section 4 develops the two-generation location model and Section 5 describes its quantification. Section 6 studies the implications for spatial persistence and access to opportunity.

2 Family geography and marriage

The analysis requires observing three objects jointly: the family geography of each spouse before marriage, the location subsequently chosen by the couple and the adult locations of their children. I construct these objects from complete-count U.S. census records linked into extended blood-family lineages. Following Looser (2026), I track individuals across census waves and use household relationships to extend these lineages across households, generations and locations. By 1940, roughly 32 million individuals belong to at least one reconstructed lineage. The marriage analysis uses census observations from 1880 through 1950. Appendix A1 describes the family-tree construction, linkage files and sample coverage in detail.

Premarital family geography. For each future spouse, I identify eligible blood relatives observed in the premarital census and map their counties into harmonized commuting zones following the crosswalk files by Eckert, Gvrtz, Liang, and Peters (2020). Let $\mathcal{K}_i^s$ for $s \in \{H, W\}$, denote the set of commuting zones occupied by the premarital family of husband H or wife W . The baseline family includes parents, grandparents, siblings, children, aunts and uncles, nieces and nephews and first cousins.

The two maps are observed separately before marriage. I can therefore distinguish destinations already connected to both spouses from destinations supplied by only one side of the match. Once the couple has children, the two branches become part of the child’s extended blood family.

Marriage sample. A couple enters the sample when both future spouses are observed single before marriage and married to one another in the next usable census. I exclude couples who belong to the same reconstructed lineage, share a premarital household or have ambiguous spouse links. Future spouses must be aged 10–40 in the premarital census and 16–50 in the post-marriage census. Because the 1890 census is unavailable, the first marriage window spans 1880–1900; subsequent cohorts use adjacent decennial censuses. These restrictions yield 778,678 marriages for the descriptive analysis. The sibling and child comparisons require additional links, described with each comparison and documented in Appendix A1.

Table 1: Marriage sample

	Husband	Wife
Age at premarital census	19.05 (5.00)	16.08 (4.35)
Occupational score, premarital household head	20.34 (10.88)	20.74 (11.02)
Foreign-born	0.002	0.001
Premarital family map, CZs	2.21 (1.77)	2.07 (1.57)
Zones on both maps	0.77	(0.55)
Union of the two maps	3.52	(2.42)
Spouses from the same CZ	0.681	
Marriages	778,678	

Notes: Standard deviations are in parentheses. Characteristics are measured in the premarital census. The occupational score is the IPUMS occ-score of the premarital household head and is available for 692,954 husbands and 707,718 wives. Premarital family maps contain commuting zones occupied by parents, grandparents, siblings, children, aunts and uncles, nieces and nephews and first cousins. Means are unweighted.

Destination quality. I use two measures of destination quality. For the descriptive analysis, I rank commuting zones within census year by the mean male occupational-income score among individuals aged 20–60, using population weights to construct the national distribution. This measure is available consistently across the historical period and describes the quality of the geography brought into marriage.

The quantitative analysis uses a common destination-year wage index constructed from agricultural and manufacturing earnings,

$$Q_{dt} = \bar{s}_t \log w_{dt}^A + (1 - \bar{s}_t) \log w_{dt}^M, \quad (1)$$

where $\bar{s}_t$ is the mean agricultural share across the two sectors in year t . Destination ranks and

upper-quartile cutoffs are constructed within census year and held fixed across the quantitative exercises. Appendix A1 describes the sector shares, wage interpolation and treatment of missing destination-year observations.

Selection and scope. Reconstructing extended family lineages over several census waves generates selection. The linked sample is disproportionately white, U.S.-born, rural and farm-resident relative to the complete census. I therefore interpret the estimates as applying to the linked families represented in each comparison. I do not use individual inverse-propensity weights in the main analysis because the relevant sampling units are couples and extended families, while individual linkage probabilities do not identify their joint probability of observation. Selection into the linked sample is limited with respect to interstate mobility. Among native-born adults aged 15–40, the ratio of tree-membership rates for individuals living outside rather than inside their state of birth ranges from 0.91 to 0.97 and approaches one over time. The child sample is more selected: the share leaving the childhood commuting zone falls from 0.369 in the broader linked-child sample to about 0.260 among same-origin cousins. Appendix A1 reports the census comparisons, marriage-cohort coverage and descendant follow-up.

The empirical analysis therefore relies on destination-specific variation within linked families. The next section asks whether acquiring a connection to a particular destination through marriage changes the probability that the couple or its children subsequently use that destination.

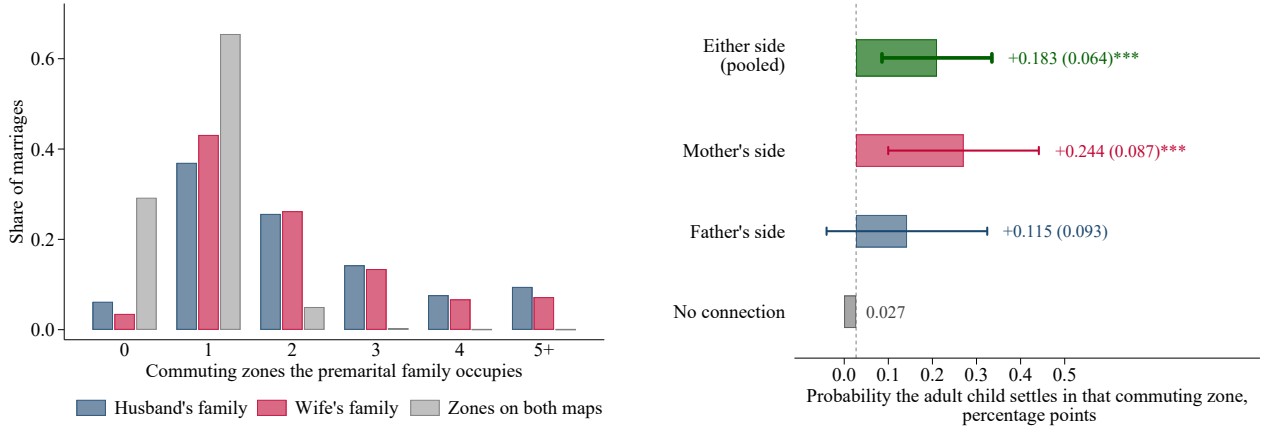
3 Marriage, family geography and location choice

Marriage combines the family geographies of two spouses and can thereby change where both the couple and their children live. I establish three facts. First, spouses bring only partly overlapping family geographies into marriage and where they come from determines how much new geography is added. Second, couples are more likely to settle in destinations where the spouse has relatives. Third, children later use destinations introduced through their parents' marriage, even when neither parent ever lived there.

Fact 1. *Marriage combines distinct family geographies and where spouses come from determines how much new geography is added.*

Before marriage, the husband's relatives occupy 2.21 commuting zones on average and the wife's 2.07. Only 0.77 zones are shared. Their union therefore spans 3.52 commuting zones, so adding the wife's family expands the husband's connected geography by 1.31 destinations, or 59%. The median marriage adds one commuting zone and 37.2% add none. Alternative measures give a similar picture: marriage raises the population living in connected places by roughly 56% and increases spatial extent from about 329 to 512 miles, while the rank of the best connected destination rises by only 0.065.

Figure 1: Marriage creates and transmits family geography



(a) Marriage merges two family geographies

(b) Children use destinations created by marriage

Notes: Panel (a) reports the distribution of the number of commuting zones occupied before marriage by the husband's family, the wife's family and both families. Values at five or more zones are grouped; means use the full distribution. Panel (b) reports the same-origin cousin estimates from Fact 3 with 95% confidence intervals. Bars show settlement-probability levels beginning from the 0.027-percentage-point probability of choosing an otherwise unconnected destination. The pooled specification combines the maternal and paternal samples with separate blood-family-by-destination fixed effects.

The amount of new geography depends strongly on where spouses are drawn from. I reassign wives within increasingly narrow groups and recompute the union of premarital maps. Reassignment within marriage cohort produces a union of 4.265 commuting zones, compared with 3.517 under observed matching. Restricting the comparison to the same state lowers the union to 4.105 zones. Restricting it to the same premarital commuting zone lowers the union to 3.564, only 1.4% above the observed union. The geographic concentration of marriage therefore accounts for most of the overlap between the two family geographies.

Socioeconomic status appears on a different margin. Conditional on cohort, the groom's and bride's premarital ranks have only small associations with the number of commuting zones in the union but strong positive associations with its quality. Higher-status households do not become connected to many more places through marriage. Instead, the places to which they become connected are systematically better. Marriage therefore changes not only how many places are connected through family but also which opportunities those places contain. I next ask whether couples use the destinations added through marriage.

Fact 2. *Couples are more likely to settle in destinations supplied by the spouse's family.*

Sibling comparison. I compare siblings from the same family who marry spouses with different premarital family geographies. The siblings share family background but acquire different destination connections through their spouses. Among brothers, this variation comes

Table 2: Marriage expands family reach

	Husband's map	Wife's map	Union
Connected commuting zones	2.213	2.070	3.517
Shared zones		0.765	
Gain relative to husband's map		1.305 zones (59.0%)	
Best connected place, rank	0.522		0.587
<i>Alternative spouse assignments</i>			
	Observed union	Random union	Shortfall
Cohort only	3.517	4.265	17.5%
+ five-year age band	3.517	4.264	17.5%
+ state	3.517	4.105	14.3%
+ premarital commuting zone	3.513	3.564	1.4%
<i>Premarital family-status gradients</i>			
	Groom rank	Bride rank	
Union size, zones	−0.0615*** (0.0123)	−0.0214* (0.0121)	
Union quality	0.1491*** (0.0010)	0.1513*** (0.0010)	

Notes: The sample contains 778,678 couples observed before and after marriage. Family maps are measured before marriage. Alternative unions use reassignments within increasingly narrow geographic groups and average 500 draws. The shortfall is the observed-minus-random gap divided by the random union. Status regressions use 632,073 couples with both premarital household-head ranks observed.

from their wives' families; among sisters, it comes from their husbands' families. At the couple-destination level, I estimate

$$Y_{id} = \alpha_{g(i)d} + \beta A_{id}^{spouse} + \varphi \tilde{A}_{id}^{spouse} + \gamma A_{id}^{own} + \delta S_{id} + X'_{id}\Gamma + \varepsilon_{id}, \quad (2)$$

where Y_{id} indicates that couple i settles in destination d and $\alpha_{g(i)d}$ is a sibling-family-by-destination fixed effect. A destination is acquired through the wife when her close relatives live there before marriage, his do not and neither spouse previously lived there; the husband-side definition is symmetric. The fixed effect absorbs the family's persistent attraction to the destination. The comparison is therefore between siblings considering the same place when only one sibling's spouse has relatives there.

A destination supplied only by the wife is 0.284 percentage points more likely to be chosen, while a destination supplied only by the husband is 0.332 points more likely to be chosen. The unconditional probability of settling in any particular one of the 709 commuting zones is 0.141%. More than 95% of the relevant family-destination groups contain siblings who differ in whether the spouse has a connection there, so the estimates come predominantly from within-family,

within-destination comparisons (Table 3).

The role of local geography. Spouses from the same premarital commuting zone may have similar family geographies. To distinguish geography specific to the realized spouse from geography common to that place, I assign each spouse the family map of another individual from the same premarital commuting zone, marriage cohort and five-year age group. Settlement remains strongly related to the actual spouse’s geography, while the reassigned geography shows little corresponding relationship. The difference between the two coefficients is 0.315 percentage points on the wife side and 0.655 points on the husband side. The baseline relationship therefore reflects geography specific to the realized spouse rather than connections generally available among spouses from the same commuting zone.

Timing of family connections. The timing of deaths among relatives provides a further test of whether a family connection matters when the couple chooses where to live. I restrict the sample to destinations linked to the spouse through a single eligible relative and estimate

$$Y_{id} = \alpha_{g(i)d} + \beta_{\text{pre}} D_{id}^{\text{pre}} + \beta_{\text{post}} D_{id}^{\text{post}} + X'_{id} \Gamma + \varepsilon_{id}, \quad (3)$$

where D_{id}^{pre} indicates that the sole relative connecting the spouse to destination d dies before marriage and D_{id}^{post} indicates that the relative dies only after the household location decision. The sibling-family-by-destination fixed effect $\alpha_{g(i)d}$ absorbs persistent differences in a family’s attraction to the destination. The coefficient difference $\beta_{\text{pre}} - \beta_{\text{post}}$ therefore compares otherwise analogous destination-specific connections that differ in whether they remain available when the couple chooses where to live.

A death before marriage reduces settlement in the connected destination by 4.1 percentage points, while a death after the location decision has essentially no effect. The resulting pre–post difference is -4.95 percentage points. The timing contrast therefore links settlement to the availability of the family connection at the time of choice rather than to persistent differences across destinations or families.

The broader migration evidence is consistent with the same mechanism. Larger spouse maps predict a greater probability of leaving the premarital location, while relative map quality primarily predicts which side of the family the couple follows. Appendix A2.2 reports these margins and household outcomes conditional on the realized destination. These results show that marriage-created geography shapes where couples settle. I next ask whether the same connections remain relevant for the next generation.

Fact 3. *Children use destinations that entered the family through their parents’ marriage.*

Next, I ask whether children later settle in destinations contributed by one parental lineage

Table 3: Couples use destinations added through marriage

	Through wife	Through husband
<i>A. Destinations added through the realized marriage</i>		
Realized spouse connection, p.p.	0.284*** (0.063)	0.332*** (0.067)
Reassigned spouse connection, p.p.	−0.031 (0.051)	−0.323*** (0.049)
Realized minus reassigned, p.p.	0.315***	0.655***
p , realized = reassigned	< 0.001	< 0.001
Family–destination groups with within-group variation	96.2%	95.4%
Generic destination probability		0.141%
Couple–destination observations	37,737,234	34,360,976
<i>B. Sole-bridge death timing</i>		
Relative dies before marriage	−0.0407*** (0.0121)	
Relative dies after location decision	+0.0088 (0.0164)	
Pre minus post	−0.0495 (0.0203)	
Observations	216,223	

Notes: Panel A estimates equation (2) with sibling-family-by-destination fixed effects. A destination is added through marriage when one spouse has close kin there, the other does not and neither spouse previously lived there. Reassigned spouses are drawn from the realized spouse’s premarital commuting zone, marriage cohort and five-year age band. Standard errors are two-way clustered on sibling family and destination. Panel B restricts the sample to destinations connected through a single eligible relative and compares deaths before marriage with deaths after the couple’s location decision.

even when neither parent ever lived there. A maternal option for child i and destination d is

$$A_{id}^M = \mathbf{1} \{ K_{d,t_0}^M > 0, K_{d,t_0}^F = 0, d \neq o_i, d \notin \mathcal{H}_i^P \}, \quad (4)$$

where o_i is the child’s childhood commuting zone and $\mathcal{H}_i^P$ is the set of all locations occupied by either parent before the child’s adult location decision. The paternal definition is symmetric. A treated destination was therefore connected to one parental lineage before marriage, absent from the other lineage, different from the child’s origin and never used by either parent before the child’s adult location choice.

Cousin comparison. To test whether the geography introduced through marriage is transmitted to the next generation, I compare cousins from the same blood family who grew up in the same commuting zone but inherit different family connections through the other parent. Among children of brothers, this variation comes from their mothers’ family geographies;

among children of sisters, it comes from their fathers' family geographies. I estimate

$$Y_{id} = \alpha_{g(i)d} + \beta A_{id}^{merge} + X_{id}'\Gamma + \varepsilon_{id}, \quad (5)$$

where $\alpha_{g(i)d}$ is a blood-family-by-destination fixed effect. The comparison therefore holds fixed the blood family, the destination and the childhood commuting zone. Identification comes from different branches of the same extended family whose parents married into different destination-specific geographies.

The main identifying concern is that extended families may have persistent ties to particular destinations. Family and destination fixed effects account for differences in overall mobility across families and in the attractiveness of destinations, but they do not account for a given family's persistent connection to a particular place. I therefore compare branches of the same blood family within the same destination by including blood-family-by-destination fixed effects. The maternal and paternal estimates fall from 0.744 and 0.682 percentage points with fixed effects to 0.244 and 0.115 under this specification.

Pooling the maternal and paternal samples, a marriage-created connection raises the probability that the adult child settles in that destination by 0.183 percentage points, relative to a generic destination probability of 0.027%. The maternal estimate is 0.244 percentage points and the paternal estimate is 0.115; the two estimates are not statistically distinguishable.

Same-location reassignment. I next compare the realized marriage-created geography with family geography available through another spouse from the same premarital commuting zone. Under the preferred interacted specification, the pooled coefficient is 0.184 percentage points for the realized marriage and -0.058 points for the same-location reassignment. The difference is 0.242 percentage points ($p = 0.003$). Children's adult locations are therefore more strongly related to the geography created by their parents' realized marriage than to geography available through another spouse from the same commuting zone.

Adult outcomes. Children who use inherited family connections also follow economically distinct paths. Relative to siblings who remain in their childhood CZ, those who move to family-connected destinations have higher occupational scores and wages and are more likely to live in urban areas. Their outcomes are nevertheless lower than those of siblings who move to destinations without family connections. These comparisons are descriptive because the destination ultimately chosen by the child is endogenous. They show that the use of inherited family geography is associated with economically meaningful mobility. Appendix [A2.3.4](#) reports these comparisons and the corresponding ex-ante outcomes.

Table 4: Children use destinations created by their parents' marriage

	Maternal	Paternal	Pooled
<i>A. Preferred: blood family $\times$ destination</i>			
Marriage-created destination, p.p.	0.244*** (0.087)	0.115 (0.093)	0.183*** (0.064)
p , maternal = paternal			0.314
<i>B. Realized and same-location reassigned family geography</i>			
Realized marriage, p.p.	0.244*** (0.087)	0.116 (0.093)	0.184*** (0.064)
Same-location reassignment, p.p.	-0.060 (0.067)	-0.056 (0.064)	-0.058 (0.047)
Realized minus reassigned, p.p.	0.304***	0.172	0.242***
p , realized = reassigned	0.009	0.129	0.003
Children	69,013	63,448	123,464
Cousin groups	13,975	12,579	26,554
Generic destination probability		0.027%	

Notes: The sample compares cousins from the same blood family who grew up in the same commuting zone. The treatment excludes the childhood CZ and every destination occupied by either parent before the child's adult location decision. Standard errors cluster on the blood family. The pooled column stacks the maternal and paternal samples with separate family-by-destination fixed effects. Same-location reassigned maps are drawn from spouses with the same premarital CZ, marriage cohort and five-year age band as the realized spouse and are subjected to the same leave-out rule.

4 Theoretical model

The empirical results show that marriage creates family connections to new destinations that shape where couples and their children later live. The model asks how these responses shape the geography inherited by the next generation. Marriage can matter directly by changing the destinations connected through family and indirectly by changing where the parents establish the household. The latter changes the child's origin, the value of staying and the migration costs to every future destination.

I use a two-generation location model. The observed marriage and destination fundamentals are taken as given. Parents choose where to establish the household and value both their own location and the spatial opportunities available to their children. Children grow up at the chosen location, inherit the merged family geography and later choose where to live as adults. The model therefore links marriage to both the family connections children inherit and the place where they grow up.

Timing and family geography. For couple i , I index time relative to marriage. Period 0 is the premarital period, Period 1 is the parental location choice and Period 2 is the linked child's adult location choice. Spouse origins and family maps are measured in Period 0.¹

Let $\mathcal{K}_i^H$ and $\mathcal{K}_i^W$ denote the commuting zones occupied by the husband's and wife's eligible blood relatives before marriage. The family geography available to the couple after marriage is their union,

$$\mathcal{K}_i^P = \mathcal{K}_i^H \cup \mathcal{K}_i^W, \quad Z_i^P(d) = \mathbf{1}\{d \in \mathcal{K}_i^P\}. \quad (6)$$

Thus, $Z_i^P(d)$ indicates whether either spouse has eligible premarital family in commuting zone d , regardless of which lineage supplies the connection.

The child inherits the same union,

$$\mathcal{K}_i^C = \mathcal{K}_i^H \cup \mathcal{K}_i^W, \quad Z_i^C(d) = \mathbf{1}\{d \in \mathcal{K}_i^C\}. \quad (7)$$

Thus, $Z_i^C(d)$ indicates whether the child has inherited family in commuting zone d , regardless of which parental lineage supplies the connection. The parents themselves are not counted in Z_i^C because the parental household enters the child's problem through the childhood origin.

The child's location choice. Let $D_{d,2}^C$ denote the common value of destination d in Period 2, including wages, amenities and other destination characteristics, and let $M^C(n, d)$ denote the migration cost from childhood origin n to destination d . If the parents establish the household in n , the value of adult destination d is

$$\begin{aligned} B_{ic}^C(d | n) = & D_{d,2}^C + \alpha_{\text{away}} \mathbf{1}\{d \neq n\} - M^C(n, d) \\ & + \phi_D Z_i^C(d) \mathbf{1}\{d \neq n\} + \phi_O Z_i^C(n) \mathbf{1}\{d = n\}. \end{aligned} \quad (8)$$

The parameter α_{away} captures the common relative value of leaving the childhood origin. The migration cost $M^C(n, d)$ captures the cost of reaching destination d from n . The coefficient ϕ_D captures the value of inherited family at an outside destination, while ϕ_O captures the value of inherited family at the childhood origin (Zerecero, 2021).

With Type-I extreme-value location shocks, the child chooses d with probability

$$P_{ic}^C(d | n) = \frac{\exp\{B_{ic}^C(d | n)/\sigma_N\}}{\sum_{r \in \mathcal{D}} \exp\{B_{ic}^C(r | n)/\sigma_N\}}. \quad (9)$$

¹Appendix A3 reports the timing structure and sample construction.

The corresponding value is

$$O_{ic}(n) = \sigma_N \log \left[\sum_{d \in \mathcal{D}} \exp\{B_{ic}^C(d | n)/\sigma_N\} \right]. \quad (10)$$

This is the child's spatial opportunity value associated with childhood origin n . It summarizes the value of staying, the attractiveness and migration cost of every outside destination and the inherited family geography.

The parents' location choice. The spouses enter marriage from premarital locations o_i^H and o_i^W . Let $M^H(o_i^H, n)$ and $M^W(o_i^W, n)$ denote the migration costs of establishing the household in n from the husband's and wife's premarital origins. The deterministic value of location n is

$$\begin{aligned} B_i^P(n) &= D_{n,1}^P - M^H(o_i^H, n) - M^W(o_i^W, n) \\ &\quad + \theta_P Z_i^P(n) + \beta O_{ic}(n). \end{aligned} \quad (11)$$

The first line of equation (11) captures common destination value and migration costs from each spouse's premarital origin. The second contains the family motives relevant for household location. The coefficient θ_P captures the value of having family in destination n , regardless of which premarital lineage supplies the connection, while $\beta O_{ic}(n)$ captures the value parents place on their children's future spatial opportunities. Parent choice probabilities take the same logit form,

$$P_i^P(n) = \frac{\exp\{B_i^P(n)/\sigma_N\}}{\sum_{r \in \mathcal{D}} \exp\{B_i^P(r)/\sigma_N\}}. \quad (12)$$

Two channels of spatial inheritance. Changing the family geography brought into marriage affects the next generation through two channels. Holding parental location fixed, it changes $Z_i^C(d)$ and therefore the destinations connected through inherited family. I call this the *direct inherited-access channel*. The same change can also alter the parental location distribution P_i^P . This changes the child's origin and hence the home alternative, migration costs to every adult destination and whether inherited family is located at home or elsewhere. I call this the *parental-location channel*.

5 Model quantification

I quantify the model in two blocks using linked U.S. families observed between 1900 and 1950. The model is not estimated at a single representative date. Instead, the parameters governing responses to migration costs and family connections are common across cohorts, while desti-

nation attractiveness varies by census year. I first estimate the child's adult location problem and use it to calculate the value of every possible childhood origin. I then use these values when estimating where the parents establish the household. The location likelihoods determine the migration-cost parameters and destination values, while the three family coefficients are disciplined by the corresponding empirical responses. Table 5 summarizes the resulting parameterization.

Samples and timing. The parent sample contains 272,873 couples in four ten-year marriage windows: 1900–1910, 1910–1920, 1920–1930 and 1930–1940. For each couple, premarital origins and family geography are measured in the census at the beginning of the window and household location in the census at its end. Parent choices are therefore observed in 1910, 1920, 1930 and 1940. The child block is estimated on 1,327,663 children observed as adults from 1910 through 1950. Each child is evaluated using destination values from that child's adult census year. Appendix Table A4.1.1 reports the complete calendar structure.

Parameterization. I use the migration dispersion parameter $\sigma_N = 0.51$ following Looser (2026) and set parental altruism to $\beta = 0.231$ following Eckert and Kleineberg (2024). For quantification, migration costs are proportional to log distance:

$$\begin{aligned} M^C(n, d) &= \kappa_C \log(1 + \text{dist}(n, d)), \\ M^H(o_i^H, n) &= \kappa_H \log(1 + \text{dist}(o_i^H, n)), \\ M^W(o_i^W, n) &= \kappa_W \log(1 + \text{dist}(o_i^W, n)). \end{aligned} \quad (13)$$

Child destination values $D_{d,y}^C$ are estimated separately for each of the 709 CZs and each adult census year $y \in \{1910, 1920, 1930, 1940, 1950\}$. Parent destination values $D_{n,y}^P$ are estimated for each CZ and parental census year $y \in \{1910, 1920, 1930, 1940\}$. These values absorb common differences in local attractiveness within each year.

Child location. Let n_{ic} denote child c 's observed childhood CZ, d_{ic} the observed adult destination and y_{ic} the adult census year. The deterministic value used in estimation is

$$\begin{aligned} B_{ic}^C(d \mid n_{ic}) &= D_{d,y_{ic}}^C + \alpha_{\text{away}} \mathbf{1}\{d \neq n_{ic}\} - \kappa_C \log(1 + \text{dist}(n_{ic}, d)) \\ &\quad + \phi_D Z_i^C(d) \mathbf{1}\{d \neq n_{ic}\} + \phi_O Z_i^C(n_{ic}) \mathbf{1}\{d = n_{ic}\}. \end{aligned} \quad (14)$$

The corresponding full-choice-set probability is

$$P_{ic}^C(d \mid n_{ic}) = \frac{\exp\{B_{ic}^C(d \mid n_{ic})/\sigma_N\}}{\sum_{r \in \mathcal{D}} \exp\{B_{ic}^C(r \mid n_{ic})/\sigma_N\}}. \quad (15)$$

Table 5: Parameterization of the quantitative model

Symbol	Description	Value	Discipline
<i>A. Fixed</i>			
σ_N	Location-shock scale	0.51	Looser (2026)
β	Parental altruism	0.231	Eckert and Kleineberg (2024)
<i>B. Child location</i>			
α_{away}	Value of leaving childhood CZ	5.586	Child location likelihood
κ_C	Child migration-cost gradient	2.113	Child location likelihood
ϕ_O	Family at childhood origin	0.365	Origin-family loss
ϕ_D	Family at outside destination	0.340	Fact 3
$D_{d,y}^C$	Child destination-year value	—	Child location likelihood
<i>C. Parent location</i>			
κ_H	Distance cost from husband’s origin	0.614	Parent location likelihood
κ_W	Distance cost from wife’s origin	0.563	Parent location likelihood
θ_P	Family at parental destination	1.686	Fact 2
$D_{n,y}^P$	Parent destination-year value	—	Parent location likelihood

Notes: All utility coefficients are in units of the normalized location shock and are common across cohorts. Destination values vary by census year. The child location likelihood is estimated on 1,327,663 children and the parent likelihood on 272,873 couples. The origin-family coefficient is calibrated using the separate origin-loss sample. For each candidate ϕ_D , the remaining child parameters are re-estimated; for each candidate θ_P , the remaining parent parameters are re-estimated. Appendix A4 gives the exact likelihoods, sampled-alternative correction and auxiliary-moment mappings.

Conditional on (ϕ_O, ϕ_D) , I estimate

$$\left(\hat{\alpha}_{\text{away}}, \hat{\kappa}_C, \left\{\hat{D}_{d,y}^C\right\}\right) = \arg \max \sum_{i,c} \log P_{ic}^C(d_{ic} \mid n_{ic}). \quad (16)$$

For computation, I estimate this likelihood using sampled alternatives with the standard inclusion-probability correction described in Appendix A4.2. It is important to note that all model moments and quantitative exercises use the full set of 709 CZs. The home-versus-away choices discipline α_{away} , variation in choices across distances disciplines κ_C and destination-by-year choice shares discipline $D_{d,y}^C$.

The two family coefficients are disciplined separately. I choose ϕ_O so that removing inherited family from the childhood origin reproduces the increase in out-migration after a loss of family at the origin, $m_O^{\text{model}}(\hat{\phi}_O) = \hat{m}_O^{\text{data}} = 0.070969$. This calibration is carried out on the separate origin-loss sample and $\hat{\phi}_O$ is then held fixed. I choose ϕ_D using the same-origin cousin comparison from Fact 3. Restricting that comparison to the origins, destinations and marriage windows represented in the quantitative model gives $\hat{m}_D^{\text{data}} = 0.002460832$. For each candidate ϕ_D , I re-estimate equation (16), evaluate probabilities over all 709 destina-

tions and apply the same cousin regression to the model probabilities. The estimate satisfies $m_D^{\text{model}}(\hat{\phi}_D) = \hat{m}_D^{\text{data}} = 0.002460832$.

With the child block estimated, I evaluate the location problem over all 709 destinations. For every child and every possible childhood origin n ,

$$O_{ic}(n) = \sigma_N \log \left[\sum_{d \in \mathcal{D}} \exp \{ B_{ic}^C(d | n) / \sigma_N \} \right]. \quad (17)$$

If couple i has C_i linked children, the continuation value entering the parents' problem is

$$\bar{O}_i(n) = \frac{1}{C_i} \sum_{c=1}^{C_i} O_{ic}(n). \quad (18)$$

Each child retains the destination values from that child's own adult census year before these values are averaged.

Parent location. I next estimate where couples establish the household. If o_i^H and o_i^W are the spouses' premarital CZs and y_i^P is the post-marriage census year, the deterministic value of location n is

$$\begin{aligned} B_i^P(n) &= D_{n,y_i^P}^P - \kappa_H \log(1 + \text{dist}(o_i^H, n)) - \kappa_W \log(1 + \text{dist}(o_i^W, n)) \\ &\quad + \theta_P Z_i^P(n) + \beta \bar{O}_i(n). \end{aligned} \quad (19)$$

Using the same sampled-logit correction as in the child block, let $\tilde{P}_i^P(n)$ denote the probability used in estimation. Conditional on θ_P and the fixed value of β , I estimate

$$\left(\hat{\kappa}_H, \hat{\kappa}_W, \left\{ \hat{D}_{n,y}^P \right\} \right) = \arg \max \sum_i \log \tilde{P}_i^P(n_i), \quad (20)$$

where n_i is the observed post-marriage CZ. The two migration-cost gradients are separately identified because, for couples entering marriage from different CZs, a candidate destination generally lies at different distances from the husband's and wife's origins. Destination-by-year choice shares discipline $D_{n,y}^P$.

I choose θ_P to reproduce the pooled spouse-supplied-destination response from Fact 2. Restricting the empirical specification to the population represented in the quantitative model gives $\hat{m}_P^{\text{data}} = 0.00161$. For each candidate θ_P , I re-estimate equation (20), evaluate location probabilities over all 709 destinations and apply the same pooled destination regression to the model probabilities. The estimate satisfies $m_P^{\text{model}}(\hat{\theta}_P) = \hat{m}_P^{\text{data}} = 0.00161$. The resulting estimates are $\hat{\theta}_P = 1.686$, $\hat{\kappa}_H = 0.614$ and $\hat{\kappa}_W = 0.563$. The standard error of $\hat{\theta}_P$ is 0.186. In migration-cost units, $\hat{\theta}_P / \hat{\kappa}_H$ implies that family at a destination offsets a 15.6-fold increase

in one plus distance from the husband’s premarital origin. For children, $\hat{\phi}_D/\hat{\kappa}_C$ corresponds to a 17% reduction in distance. Although only the pooled parental response is targeted, the husband-side response is 0.140 percentage points in the data and 0.134 in the model, while the wife-side response is 0.184 and 0.190.

Model fit. I compare the model with the data that are not matched directly. For children, 63.07% remain in their childhood CZ in the model, compared with 63.23% in the data. Among movers, mean log distance is 5.273 in the model and 5.280 in the data. The main discrepancy is settlement near inherited family: the model predicts that 52.79% of children settle in a CZ containing inherited family, compared with 60.72% in the data. The gap is concentrated outside the childhood origin, where the corresponding shares are 2.45% and 7.15%.

The parent block also reproduces the main settlement margins closely. The model predicts that 80.83% of couples settle in either spouse’s premarital CZ, compared with 80.17% in the data, and that 80.88% settle in a family-connected destination, compared with 81.56%. [Appendix A4.4](#) reports the full child and parent location diagnostics.

6 Family geography, persistence and mobility

The empirical results show that family connections affect where couples settle and where their children later live. I now use the model to ask how much this geography contributes to the persistence of families in places and to differences in the opportunities children can reach. I first remove family ties from location choices altogether. I then study who is most constrained by family geography, whether its geographic reach or the places it connects to matter more and how these effects are transmitted from parents to children.

I focus on three sets of outcomes. The first describes where families live: whether parents remain near their premarital locations and whether children stay in their childhood commuting zone. The second measures the quality of children’s adult destinations using the within-year wage rank of each commuting zone. The third measures persistence across generations by relating the quality of children’s destinations to parental socioeconomic rank.

6.1 Family ties and persistence across generations

I begin by asking how much family ties contribute to keeping families in the same places across generations. I remove family proximity from both generations’ location choices: parents no longer receive utility from living near their premarital relatives and children no longer receive utility from inherited relatives at home or elsewhere. Distance costs, destination values, the observed couples and their family maps remain unchanged. The comparison therefore isolates the contribution of family ties to spatial persistence.

Family ties strongly anchor parents to inherited locations. Without them, parents are about 20 percentage points less likely to settle in a spouse’s premarital commuting zone (Figure 2 and Table 6). This change in parental settlement carries forward to the next generation. Children become less likely to remain where they grew up and more likely to reach higher-opportunity places. As a result, the intergenerational persistence of place quality falls by about one third and the gap in access to top-quartile destinations between children of high- and low-status parents essentially disappears. Family ties therefore transmit place across generations rather than simply affecting where one generation chooses to live.

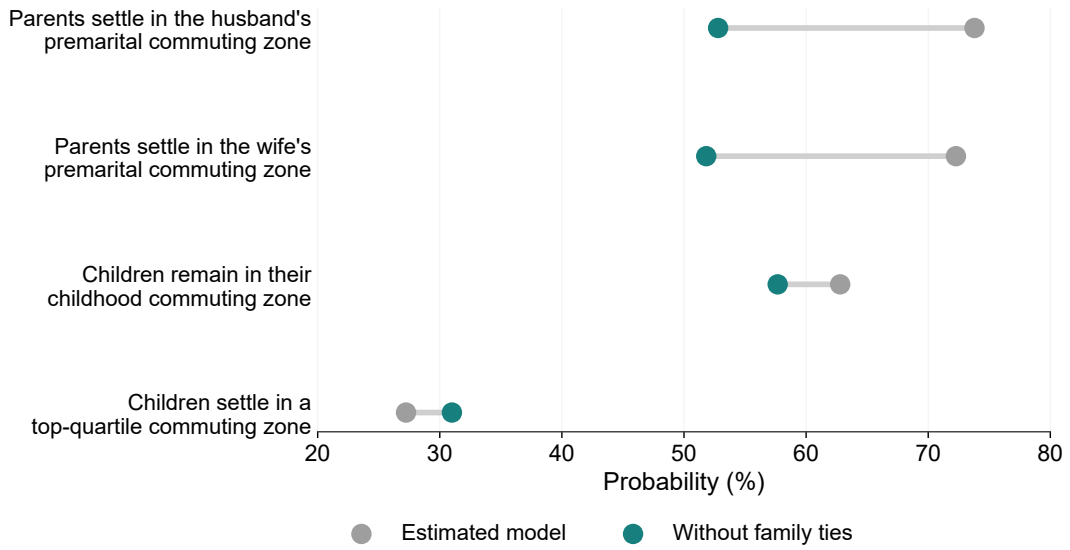


Figure 2: Family ties and the persistence of families in places

Notes: Grey markers report the estimated model with observed couples and family maps. Teal markers remove the value of family proximity in both generations, setting the parental family coefficient and the children’s destination and origin family coefficients to zero. Top-quartile commuting zones are defined on the within-year distribution of commuting-zone wage ranks.

Parents’ and children’s ties generate this persistence through different margins. I remove the parents’ family value and the children’s family values separately and allocate their interaction symmetrically using a Shapley decomposition. Parents’ ties account for more than four fifths of the effect on the quality of the destinations children reach because they shape where the household settles and hence where children grow up. Children’s own ties instead account for almost all of the additional tendency to remain in the childhood commuting zone. Family geography therefore keeps children close to home mainly through their own attachments, while it shapes the opportunities they reach mainly through the place inherited from their parents.

6.2 Family geography and the gap between rich and poor children

The effect of family geography is not evenly distributed across households. Children from lower-status families gain substantially more when family ties no longer constrain location choices.

Table 6: Family ties and persistence across generations

	Estimated	Without		Contribution	
Outcome	model	ties	Change	Parents	Children
<i>Panel A. Where parents settle (%)</i>					
Husband's premarital CZ	73.79	52.79	−21.00	−20.42	−0.58
Wife's premarital CZ	72.27	51.82	−20.45	−19.88	−0.57
Either spouse's family CZ	83.21	56.56	−26.65	−25.91	−0.75
<i>Panel B. Where children settle (%)</i>					
Childhood CZ	62.79	57.67	−5.12	−0.20	−4.91
Above-median CZ	60.46	63.21	2.76	2.25	0.50
Top-quartile CZ	27.23	31.00	3.76	3.06	0.70
<i>Panel C. Intergenerational persistence of place quality</i>					
Parental-rank slope	0.1076	0.0733	−0.0343	−0.0280	−0.0063
Q5–Q1 top-quartile gap (p.p.)	1.55	−0.64	−2.19	−1.74	−0.46

Notes: The estimated model uses observed couples and family maps. “Without ties” removes family proximity from both generations’ location choices. Parents’ and children’s contributions are Shapley values from the two single-generation removals and sum to the total change. The parental-rank slope regresses the wage rank of the child’s adult CZ on parental socioeconomic rank. Appendix Table A5.2.1 reports each specification separately.

Figure 3 shows a clear gradient: removing family ties raises access to above-median destinations by roughly 4–5 percentage points for children in the bottom three parental quintiles, but by less than 2 points for those near the top. The same pattern appears when family connections are geographically broadened rather than removed. Family geography therefore limits access to opportunity most for children whose parents begin with fewer economic resources.

I next ask how much of the difference between rich and poor children reflects the geography of the families their parents bring into marriage. I give couples in the bottom two quintiles the premarital family maps of couples in the top two quintiles from the same marriage cohort, while holding the spouses, their premarital origins and all other characteristics fixed. Access to top-quartile destinations rises by almost 5 percentage points and the relationship between parental rank and children’s destination quality falls by roughly one half. The original gap between children from the top and bottom of the parental distribution is more than eliminated. Differences in inherited family geography can therefore account for a meaningful part of the spatial opportunity gap between rich and poor children.

6.3 Reach and the places family connections lead

The previous exercises change family geography as a whole. I now ask which features of that geography matter. I begin conservatively by changing the family map brought into the household through only one spouse, drawing alternative maps from spouses of the same sex, marriage cohort, age group and socioeconomic-rank group. Expanding that map from the

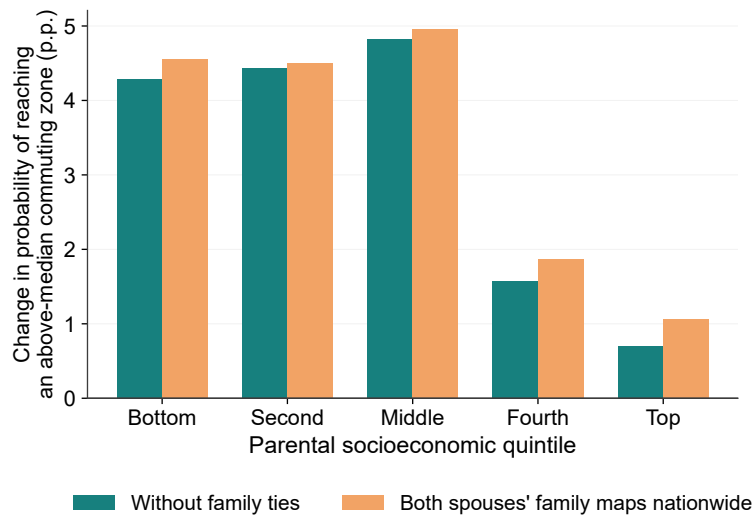


Figure 3: Family geography constrains poor families more

Notes: Each bar reports the change in the probability that a child settles in a commuting zone above the median of the within-year distribution of commuting-zone wage ranks, by quintile of parental socioeconomic rank. Teal bars remove the value of family proximity in both generations relative to the estimated model. Orange bars compare national with same-CZ assignment of both spouses' premarital family maps. Replacement maps preserve sex, marriage cohort, five-year age group and socioeconomic-rank group.

spouse's own CZ to anywhere in the country has only a modest average effect on children's access to top-quartile destinations. Most of this gain is already obtained within 250 kilometers (Appendix Figure A5.3.1). The small average reflects the fact that the other side of the family remains unchanged and continues to anchor the household near its original geography.

The effect becomes much larger when the household's full family geography changes. When both spouses bring family connections drawn from across the country rather than only from their own CZs, children's probability of reaching a top-quartile destination rises by almost 4 percentage points and the relationship between parental rank and destination quality falls by 30% (Figure 4). Moving the spouse's premarital origin together with the family map produces an even larger response. Family geography is therefore one part of a broader spatial bundle that marriage brings into the household: where a spouse comes from and where that spouse's relatives live jointly shape the locations available to the family.

Geographic reach matters most when the family starts in a place with few local opportunities. For couples whose two premarital CZs are both in the bottom quartile, allowing both spouses to bring nationally distributed family maps raises children's probability of reaching a top-quartile destination from about 11% to 18%. Broader family geography also makes parents more likely to establish the household outside the bottom of the place distribution. Connections to other places therefore matter most when the local alternative is poor.

Reach alone is not enough: the destinations to which relatives connect the household also matter. I rank family maps by the wage quality of the places where relatives live and compare

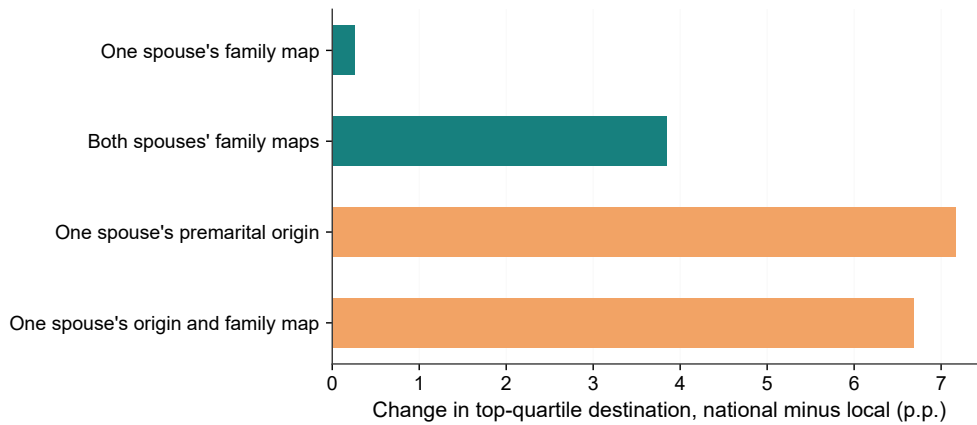


Figure 4: Family geography and the spouse's spatial bundle

Notes: Each bar compares national with same-CZ assignment of the stated component and reports the change in the probability that the child reaches a top-quartile destination. The first, third and fourth bars replace one spouse's component and average husband- and wife-side replacements. The second bar replaces both spouses' family maps simultaneously. Appendix Table A5.3.2 reports levels and the parental-rank slope.

maps within the same premarital CZ and marriage cohort. Children connected to higher-opportunity family geographies are more likely to reach top-quartile destinations, even though their parents begin in the same CZ. This pattern is not driven by socioeconomic assortative matching. Reassigning family maps by socioeconomic status within the same CZ has almost no effect on average outcomes or on the parental-rank gradient. What matters within a place is therefore where relatives live, not simply the socioeconomic characteristics of the spouse who brings those connections.

6.4 Family geography changes the child's origin

Family geography can affect the next generation in two ways. Children directly inherit connections to places where relatives live. But the same geography can also change where their parents establish the household, thereby changing the child's origin before the child makes any adult location choice. I separate these channels by first changing the child's inherited family map while holding the parental location distribution fixed and then allowing parents to adjust where they live.

Parental settlement is the larger channel. When one spouse's family map expands from the spouse's own CZ to nationwide reach, changes in where parents settle account for nearly three quarters of the resulting increase in children's access to top-quartile destinations. Direct inherited connections account for the remaining quarter. The same ordering appears when I compare low- and high-opportunity family maps within the same CZ: about two thirds of the effect operates through parental settlement (Table 7). This mirrors the earlier generation decomposition, where parents' ties account for most of the effect of family geography on the quality of the

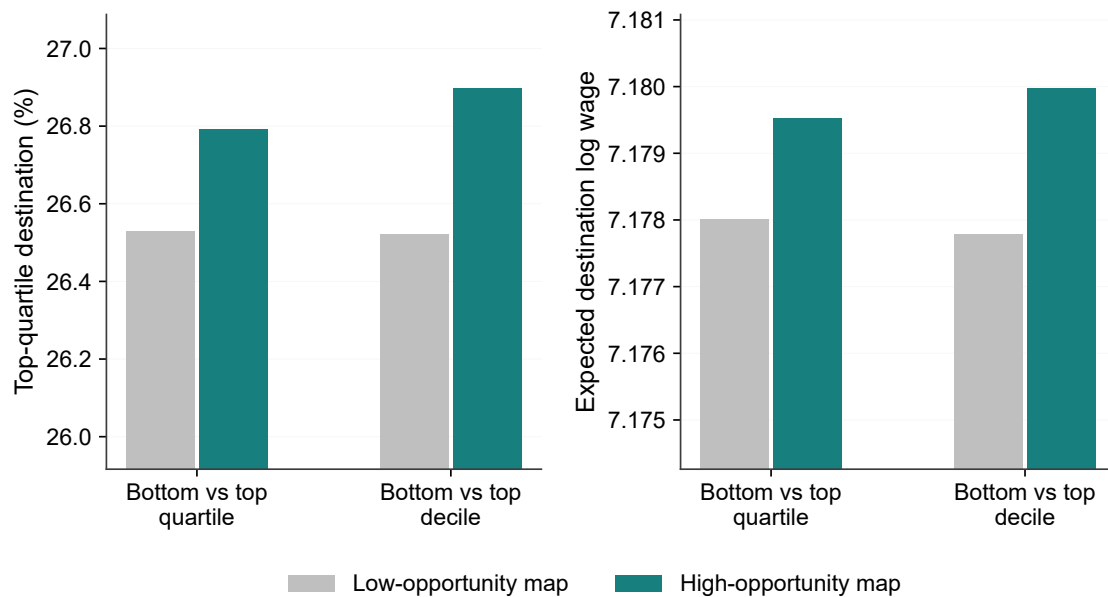


Figure 5: Family-map quality within the same commuting zone

Notes: The sample contains households in the bottom two quintiles of parental socioeconomic rank. The first comparison draws low- and high-opportunity maps from the bottom and top quartiles of the relative-count-weighted destination wage index within the same premarital commuting zone and marriage cohort. The second uses the bottom and top deciles.

places children reach.

Table 7: How family geography reaches the next generation

Channel	Share of total effect (%)
<i>Panel A. Same-CZ to national family maps</i>	
Direct inherited access	27.4
Parental settlement	72.6
Total	100.0
<i>Panel B. Low- to high-opportunity family map within CZ</i>	
Direct inherited access	33.1
Parental settlement	66.9
Total	100.0

Notes: The outcome is the probability that the child reaches a top-quartile destination. Panel A decomposes the effect of expanding one spouse's family map from the same CZ to nationwide reach. Panel B compares high- and low-opportunity family maps within the same premarital commuting zone for households in the bottom two quintiles of parental socioeconomic rank. The direct inherited-access component changes the child's family map while holding the full parental location distribution fixed. The parental-settlement component then allows the parental location policy to respond.

The strength of the parental channel follows from what changes when the child’s origin moves. A different household location changes the home alternative, the distance to every future destination and whether inherited relatives are nearby or far away. Changing the inherited family map while holding the origin fixed affects only the family value attached to particular places. The parental location choice therefore reshapes the child’s entire spatial choice problem rather than changing the attractiveness of only a few destinations.

Broader family geography also changes which side of the family anchors the household. When family maps extend beyond the spouses’ own CZs, parents become much more likely to settle in a place connected through only one side of the family. Marriage therefore does not simply add destinations to the child’s inherited geography. By changing which family connections parents follow, it determines where the child starts.

6.5 Parental motives for family proximity

The importance of parental settlement raises a final question: why do parents respond to family geography? In the model, a location can be attractive because parents value living near their own relatives or because they value the future spatial opportunities that the location creates for their children. I remove these motives separately and recompute parental and child choices. Parents respond primarily because they want to live near their own relatives. A Shapley decomposition attributes about two thirds of the effect of broader family reach to parents’ own-family proximity and about one fifth to the value they place on their children’s future opportunities (Table 8). The remainder arises even when both motives are removed because changing the family map also changes the child’s inherited connections directly. The main parental response is therefore contemporaneous, even though its consequences persist into the next generation.

Table 8: Parental motives for family proximity

Component	Share of total effect (%)
Effect absent both parental motives	12.5
Parents’ own-family proximity	68.1
Parental altruism	19.5
Total	100.0

Notes: The outcome is the change in the probability that the child reaches a top-quartile destination when one spouse’s family map expands from the same CZ to nationwide reach. The own-family and parental-altruism contributions are Shapley values, which allocate the interaction between the two motives symmetrically. Shares sum to 100 up to rounding.

This distinction clarifies how family geography becomes intergenerational. Parents do not

mainly move in response to the opportunities a location will eventually create for their children. They move because of their own family connections. But the location they choose becomes the child's home, changing the distances, family connections and destinations the child later faces. A contemporaneous preference for family proximity can therefore generate persistence in opportunity across generations.

Taken together, the quantitative results show that family ties are an important source of spatial persistence. They keep parents close to inherited locations, and those parental choices shape the geography faced by the next generation. The constraint is strongest for families with fewer economic resources and in low-opportunity places. Across the exercises, the main intergenerational channel is the same: family geography affects children's opportunities primarily by determining where they grow up.

7 Conclusion

This paper shows that marriage helps determine the geography transmitted across generations. Spouses bring relatives living in different places into the household, expanding family-connected geography by 59%. Couples respond to these connections when choosing where to live and children later use destinations introduced through the parental marriage. In the quantitative model, removing family proximity substantially weakens persistence in place and reduces the intergenerational relationship between parental status and destination quality by about one third. Most of this effect operates through where parents establish the household rather than through the destinations children inherit directly.

These results have implications for policies intended to increase geographic mobility and access to opportunity. Moving costs depend not only on distance, housing or information, but also on where a household's relatives live. Families with strong origin ties face a larger effective cost of leaving, while relatives at potential destinations can make relocation easier. The same relocation assistance can therefore generate different responses across households even when their economic incentives are otherwise similar. Targeting mobility programs only on the basis of destination wages or local economic conditions may miss this variation in the barriers households face.

The intergenerational channel also changes how the effects of mobility policy should be evaluated. A parental move does more than change the parents' current location: it changes where children grow up, their distance to future destinations and which relatives are nearby when they later make their own location choices. Policies that alter parental settlement can therefore affect the geography faced by the next generation even when children are not directly targeted. More broadly, childhood location is not simply an inherited initial condition. It is partly the outcome of earlier family formation and migration decisions, making family geography one mechanism through which differences in access to places persist across generations.

References

- Bergman, P., Chetty, R., DeLuca, S., Hendren, N., Katz, L. F., & Palmer, C. (2024). Creating moves to opportunity: Experimental evidence on barriers to neighborhood choice. *American Economic Review*, 114(5), 1281–1337. doi: 10.1257/aer.20200407
- Blumenstock, J. E., Chi, G., & Tan, X. (2025). Migration and the value of social networks. *Review of Economic Studies*, 92(1), 97–128. doi: 10.1093/restud/rdad113
- Bound, J., & Holzer, H. J. (2000). Demand shifts, population adjustments, and labor market outcomes during the 1980s. *Journal of Labor Economics*, 18(1), 20–54. doi: 10.1086/209949
- Chetty, R., & Hendren, N. (2018a). The impacts of neighborhoods on intergenerational mobility i: Childhood exposure effects. *Quarterly Journal of Economics*, 133(3), 1107–1162. doi: 10.1093/qje/qjy007
- Chetty, R., & Hendren, N. (2018b). The impacts of neighborhoods on intergenerational mobility ii: County-level estimates. *Quarterly Journal of Economics*, 133(3), 1163–1228. doi: 10.1093/qje/qjy006
- Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? the geography of intergenerational mobility in the united states. *Quarterly Journal of Economics*, 129(4), 1553–1623. doi: 10.1093/qje/qju022
- Chiappori, P.-A., Costa Dias, M., Meghir, C., & Zhang, H. (2025). *Changes in marital sorting: Theory and evidence from the us* (NBER Working Paper No. 33354). National Bureau of Economic Research. doi: 10.3386/w33354
- Compton, J., & Pollak, R. A. (2014). Family proximity, childcare, and women’s labor force attachment. *Journal of Urban Economics*, 79, 72–90. doi: 10.1016/j.jue.2013.03.007
- Costa, D. L., & Kahn, M. E. (2000). Power couples: Changes in the locational choice of the college educated, 1940–1990. *Quarterly Journal of Economics*, 115(4), 1287–1315. doi: 10.1162/003355300555079
- Eckert, F., Gvirtz, A., Liang, J., & Peters, M. (2020, February). *A method to construct geographical crosswalks with an application to US counties since 1790* (NBER Working Paper No. 26770). National Bureau of Economic Research. doi: 10.3386/w26770
- Eckert, F., & Kleineberg, T. (2024). *The geography of opportunity: Education, work, and intergenerational mobility across us counties* (mimeo). World Bank.
- Eika, L., Mogstad, M., & Zafar, B. (2019). Educational assortative mating and household income inequality. *Journal of Political Economy*, 127(6), 2795–2835. doi: 10.1086/702018
- Fernández, R., & Rogerson, R. (2001). Sorting and long-run inequality. *Quarterly Journal of Economics*, 116(4), 1305–1341. doi: 10.1162/003355301753265589
- Ganong, P., & Shoag, D. (2017). Why has regional income convergence in the U.S. declined?

- Journal of Urban Economics*, 102, 76–90. doi: 10.1016/j.jue.2017.07.002
- Greenwood, J., Guner, N., Kocharkov, G., & Santos, C. (2014). Marry your like: Assortative mating and income inequality. *American Economic Review*, 104(5), 348–353. doi: 10.1257/aer.104.5.348
- Guler, B., Guvenen, F., & Violante, G. L. (2012). Joint-search theory: New opportunities and new frictions. *Journal of Monetary Economics*, 59(4), 352–369. doi: 10.1016/j.jmoneco.2012.05.001
- Ilyés, V., Boza, I., Lórinicz, L., & Eriksson, R. H. (2023). How to enter high-opportunity places? the role of social contacts for residential mobility. *Journal of Economic Geography*, 23(2), 371–395. doi: 10.1093/jeg/lbac019
- Jia, N., Molloy, R., Smith, C., & Wozniak, A. (2023). The economics of internal migration: Advances and policy questions. *Journal of Economic Literature*, 61(1), 144–180. doi: 10.1257/jel.20211623
- Kaplan, G. (2012). Moving back home: Insurance against labor market risk. *Journal of Political Economy*, 120(3), 446–512.
- Konrad, K. A., Kunemund, H., Lommerud, K. E., & Robledo, J. R. (2002). Geography of the family. *American Economic Review*, 92(4), 981–998. doi: 10.1257/00028280260344551
- Kremer, M. (1997). How much does sorting increase inequality? *Quarterly Journal of Economics*, 112(1), 115–139. doi: 10.1162/003355397555145
- Looser, L. (2026, September). *Dynastic structural change: Kinship and economic reallocation*. (Working paper)
- Mincer, J. (1978). Family migration decisions. *Journal of Political Economy*, 86(5), 749–773. doi: 10.1086/260710
- Stuart, B. A., & Taylor, E. J. (2021). Migration networks and location decisions: Evidence from u.s. mass migration. *American Economic Journal: Applied Economics*, 13(3), 134–175. doi: 10.1257/app.20180294
- Zabek, M. (2024). Local ties in spatial equilibrium. *American Economic Journal: Macroeconomics*, 16(2), 287–317. doi: 10.1257/mac.20210326
- Zerecero, M. (2021). *The birthplace premium*. (Working paper)

APPENDIX

A1	Data construction and sample diagnostics	29
A2	Empirical evidence	33
A2.1	Marriage and family geography	33
A2.2	Couple location	35
A2.3	Child location	36
A3	Model details	38
A3.1	Timing and quantitative sample	40
A3.2	Family geography across generations	40
A3.3	Child location problem	41
A3.4	Parent location problem	42
A3.5	Normalization and interpretation	42
A3.6	Direct inheritance and parental settlement	43
A4	Model quantification and fit	43
A4.1	Samples, timing and parameter discipline	43
A4.2	Child location estimation	45
A4.3	Parent location estimation	47
A4.4	Model fit	48
A5	Quantitative exercises and sensitivity	50
A5.1	Construction of the quantitative exercises	50
A5.2	Family ties and persistence	51
A5.3	Geographic reach	52
A5.4	Family connections to high- and low-opportunity places	55
A5.5	Direct inherited access and parental settlement	56
A5.6	Parental motives for family proximity	58
A5.7	Additional decompositions and robustness	58
A5.8	Sampling uncertainty	59

A1 Data construction and sample diagnostics

This appendix describes the construction of the family lineages and marriage sample used in Section 2. I first describe how census households are linked into extended blood-family lineages, then document the marriage and descendant samples and finally report linkage selection and the construction of destination quality.

Reconstructing family lineages. This paper uses the extended family-lineage data from [Looser \(2026\)](#). The reconstruction begins with complete-count census households and extends blood-family identifiers across census waves using Census Tree person links. Within each census, relationship-to-head codes identify immediate blood and half-blood relatives. When an individual is linked to a later census, relatives observed in the new household are attached to the existing lineage when supported by the contemporaneous household structure. Repeating this procedure across census waves allows family identifiers to extend as households split, individuals move and new generations appear.

Pairwise relationships are recovered by tracing individuals to common ancestors. This identifies vertical relatives, such as parents and grandparents, and horizontal relatives, such as aunts, uncles and cousins. Relationships that cannot be inferred uniquely are treated as ambiguous and are not used to propagate a blood lineage. Spouses remain outside one another’s blood lineages. Children inherit both parental branches and can therefore belong to more than one reconstructed lineage.

Overlapping lineages can imply more than one genealogical path between the same two individuals. I de-duplicate these connections before constructing family maps and retain the relationship associated with the closest common ancestor. Each relative pair therefore contributes once to a given family measure. The marriage analysis uses parents, grandparents, siblings, children, aunts and uncles, nieces and nephews and first cousins.

By 1940, the reconstruction contains roughly 32 million individuals with at least one family lineage. Coverage is not uniform across the population or over time. The selection generated by cross-census linkage is documented below.

From lineages to premarital family maps. For a future spouse observed in premarital census t_0 , I retain eligible blood relatives alive and observed in t_0 , map their counties into harmonized commuting zones and collapse the resulting locations to a spouse-specific destination map. The future spouse is excluded from their own map. For spouse $s \in \{H, W\}$,

$$\mathcal{K}_i^s = \{d : \text{at least one eligible premarital relative of spouse } s \text{ lives in CZ } d\}.$$

The baseline map is extensive: a lineage is connected to a destination whenever at least one

eligible relative lives there. Keeping the two premarital lineages separate makes it possible to distinguish destinations connected to the husband, the wife or both. The child’s inherited family geography is the union of these two branches.

The descriptive analysis in Fact 1 also examines alternative measures of family reach, including the number of connected people, spatial extent and destination quality. Appendix A2.1 reports those exercises together with count-based family measures and the reweighting used to compare map completeness across marriage cohorts.

Marriage observations and census timing. A marriage observation requires both future spouses to be single in t_0 and married to one another in the next usable census t_1 . I exclude couples who share a reconstructed blood lineage, appear in the same premarital household or have ambiguous spouse links. Future spouses must be aged 10–40 at t_0 and 16–50 at t_1 .

The first cohort spans 1880–1900 because the 1890 census is unavailable. Subsequent cohorts use adjacent decennial censuses. Table A1.0.1 reports the resulting sample by marriage window.

Table A1.0.1: Marriage cohorts

Marriage window	Couples
1880–1900	135,856
1900–1910	151,043
1910–1920	167,418
1920–1930	156,940
1930–1940	128,173
1940–1950	39,248
Total	778,678

Notes: Both future spouses are observed single in the premarital census and married to one another in the next usable census. The first window spans twenty years because the 1890 census is unavailable.

Fact 1 uses all 778,678 marriages. Fact 2 additionally requires siblings from the same family whose spouses supply different destination-specific connections. Fact 3 requires linked children, reconstructed cousin relationships and a common childhood commuting zone within the comparison family group.

Cross-census person links. The lineage construction uses Census Tree person links at confidence setting 2. The marriage sample relies on the 1880–1900, 1900–1910, 1910–1920, 1920–1930, 1930–1940 and 1940–1950 transitions, together with earlier non-adjacent links when

needed to reconstruct family ancestry. Link counts increase with census population and coverage: the 1880–1900 transition contains 18.2 million matched pairs, the 1900–1910 transition 39.4 million, the 1910–1920 transition 47.7 million and the 1930–1940 transition 66.5 million. The final marriage cohort is smaller because linkage into 1950 is thinner. This affects cohort coverage rather than the definition of a premarital family map. Appendix A2.1 reports Fact 1 with the final cohort excluded.

Selection into reconstructed family trees. Cross-census linkage and multigenerational reconstruction do not produce a representative sample of the contemporaneous census. Table A1.0.2 compares individuals aged 15–40 who belong to a reconstructed family tree with the corresponding complete-count population.

Table A1.0.2: The linked family sample and the complete census

	1880		1900		1910	
	Census	Tree	Census	Tree	Census	Tree
Share in a tree		.330		.261		.240
Male	.506	.501	.508	.496	.515	.493
White	.871	.984	.881	.988	.885	.988
Black	.125	.016	.115	.012	.110	.011
Foreign-born	.163	.038	.162	.014	.175	.011
Urban	.307	.196	.426	.264	.498	.327
Farm	.402	.532	.324	.452	.281	.414
Outside birth state	.279	.271	.236	.220	.249	.233
	1920		1930		1940	
	Census	Tree	Census	Tree	Census	Tree
Share in a tree		.243		.244		.242
Male	.504	.494	.498	.500	.495	.505
White	.891	.991	.890	.991	.892	.992
Black	.105	.009	.105	.008	.103	.007
Foreign-born	.153	.009	.113	.007	.053	.006
Urban	.551	.387	.597	.449	.579	.439
Farm	.259	.375	.221	.310	.210	.291
Outside birth state	.257	.244	.266	.256	.240	.235

Notes: Individuals aged 15–40. “Census” denotes the complete-count population and “Tree” the subset observed in a reconstructed family lineage. Residence outside the state of birth is defined for the U.S.-born. Entries are un-weighted.

Selection is pronounced along race, nativity, urban residence and farming. The linked sample is more likely to be white, U.S.-born, rural and farm-resident throughout the period.

Selection on migration. Selection is smaller on migration than on several demographic characteristics. Table A1.0.3 compares tree-membership rates for U.S.-born individuals who remain in their birth state with those who live elsewhere.

Table A1.0.3: Tree membership by migration status

	1880	1900	1910	1920	1930	1940
Lives in state of birth	.384	.313	.294	.247	.244	.240
Lives elsewhere	.368	.285	.270	.230	.232	.233
Mover/stayer ratio	.957	.910	.918	.931	.950	.971
Mover coefficient, state FE	-.0197 (.0009)	-.0203 (.0007)	-.0136 (.0006)	-.0073 (.0005)	+.0029 (.0005)	+.0066 (.0004)

Notes: Tree-membership rates are reported separately by whether a U.S.-born individual resides in the state of birth. The final row reports the coefficient on living outside the birth state from a linear probability model for tree membership with state fixed effects and controls for sex, race, urban status, nativity, marital status and farm residence, estimated on a ten-percent random subsample of each census.

Movers are modestly less likely to enter a reconstructed tree early in the sample. The difference narrows over time: the mover-to-stayer tree-membership ratio rises from 0.910 in 1900 to 0.971 in 1940. This reduces concern that the lineage reconstruction mechanically excludes mobile individuals, but it does not address the identifying assumptions of the within-family comparisons.

Selection into the child sample. Following the next generation requires additional links. Of the 778,678 couples in Fact 1, 49.7% have at least one child observed at an adult census. Conditional on marriage cohort, broader union maps do not reduce child follow-up, while cross-CZ marriages are 3.2 percentage points less likely to contribute a followed child.

The same-origin cousin comparison is more restrictive. Roughly 1.33 million linked children are observed at an adult census, about 198,000 belong to a usable cousin group and 69,013 enter the preferred maternal same-origin comparison. The probability of leaving the childhood commuting zone is 0.369 in the broader linked-child sample and about 0.260 in the same-origin cousin sample.

Destination quality. The paper uses one destination-quality measure for historical description and another for the quantitative exercises. For the descriptive analysis, I calculate the mean male occupational-income score among individuals aged 20–60 for each commuting zone and census year and rank commuting zones within year using population weights. This measure is available consistently across the historical period and is used to characterize whether marriage adds access to economically better places.

The quantitative exercises instead use

$$Q_{dt} = \bar{s}_t \log w_{dt}^A + (1 - \bar{s}_t) \log w_{dt}^M,$$

where w_{dt}^A and w_{dt}^M are agricultural and manufacturing earnings in destination d and year t . The weight $\bar{s}_t$ is the mean predicted agricultural share across the two sectors in the linked-child sample. Sector probabilities are estimated by sex and premarital paternal-status group using observations outside the child’s own estimation fold.

Sectoral log wages are interpolated linearly within commuting zones across available census years and extrapolated when necessary. If a commuting zone has only one observed value, that value is carried across years. Remaining missing destination-year observations receive the corresponding year mean.

A2 Empirical evidence

This appendix provides additional evidence for the three empirical facts in Section 3. It documents the distribution of premarital family geography, shows where the within-family comparisons come from and reports additional robustness and outcome results for couples and children.

A2.1 Marriage and family geography

A2.1.1 Distribution of premarital family maps

The baseline family map includes parents, grandparents, siblings, children, aunts and uncles, nieces and nephews and first cousins observed before marriage. Table A2.1.1 shows how many commuting zones are connected through each spouse before marriage and how many are shared by both sides. Husbands and wives enter marriage with similarly sized family geographies, but the overlap between them is limited: 29.2% of marriages have no connected commuting zone in common and 65.5% share exactly one.

A2.1.2 Alternative measures of geographic reach

The expansion in family geography is not specific to counting connected commuting zones. Relative to the husband’s premarital map, marriage raises the population living in family-connected places from 0.903 to 1.410 million and increases the geographic span of connected places from 329 to 512 miles. Both measures increase by about 56%. The median marriage adds one commuting zone. By comparison, the rank of the highest-opportunity connected destination rises from 0.522 to 0.587. Marriage therefore expands the breadth and spatial reach of family connections much more than it changes the best destination already connected through family.

Table A2.1.1: Distribution of connected commuting zones before marriage

Connected zones	Husband	Wife	Shared
0	0.062	0.035	0.292
1	0.369	0.431	0.655
2	0.256	0.262	0.050
3	0.142	0.134	0.003
4	0.076	0.067	0.000
5 or more	0.095	0.071	0.000
Mean, full distribution	2.213	2.070	0.765

Notes: Entries are shares of marriages. The upper tail is grouped only for presentation; means use the full distribution. “Shared” counts commuting zones containing eligible relatives from both premarital lineages.

A2.1.3 Geographic concentration of marriage and map overlap

To assess how much of the overlap between spouses’ family geographies comes from where spouses originate, I compare the observed marriages with alternative pairings. Wives are re-assigned within progressively narrower sets of otherwise comparable women and the combined family geography is recomputed. I average each comparison over 500 reassignments. The broadest comparison preserves only the marriage cohort. I then additionally require similar age, the same state and finally the same premarital commuting zone.

The observed combined geography is 17.5% smaller than under reassignment within the same cohort and 14.3% smaller than under reassignment within the same state. Once the alternative wife is required to come from the same premarital commuting zone, the difference falls to 1.4%. Most of the overlap between spouses’ family geographies can therefore be accounted for by the geographic concentration of marriage.

A2.1.4 Tree completeness across marriage cohorts

The number of relatives observed in the reconstructed family trees varies across marriage cohorts and can mechanically affect both map size and overlap. I therefore reweight marriages so that cohorts are compared at a common distribution of observed close-kin counts and lineage size. The patterns in family-map breadth and overlap after 1900 are essentially unchanged. The 1880 cohort has weaker common support, consistent with the longer 1880–1900 linkage interval and more limited coverage of the earliest family trees.

This adjustment addresses a different issue from the population-representativeness weights used elsewhere in the paper. Those weights make linked individuals more comparable to the contemporaneous census population. Here, the purpose is instead to compare marriages with similar amounts of observed family information.

A2.1.5 Spatial diversification and status gradients

Marriage also changes the dispersion of the opportunities connected through family. The variance of the husband's premarital map-quality measure is 0.04777 and falls to 0.03916 after the two spouses' maps are combined, a decline of about 18%. Under reassignment within the same state, the variance falls further to 0.03609.

Family-map breadth and family-map quality differ in how much variation remains within the same premarital commuting zone. Marriage cohort and premarital-CZ effects explain about 12% of the variation in a wife's map breadth but about 87% of the variation in map quality. Most variation in the number of places connected through family therefore remains within a commuting zone, while the average quality of those places is much more geographically concentrated.

A2.2 Couple location

A2.2.1 Within-family comparisons and inference

The main comparison asks whether siblings from the same family are more likely to settle in a destination when that place is connected through one sibling's spouse but not the other's. I use the full set of 709 possible destinations and compare siblings within the same family–destination pair. These comparisons absorb any persistent tendency of the family to locate in a particular place.

Most relevant family–destination pairs contain within-family variation. Among brothers, 50,831 pairs contain at least one destination connected through a wife and 48,906 contain brothers who differ in whether that connection is present. The corresponding numbers among sisters are 53,358 and 50,911. Thus more than 95% of the relevant family–destination pairs contain within-family variation. Standard errors are two-way clustered by family and destination.

A2.2.2 Alternative spouse maps from the same commuting zone

A spouse's family geography partly reflects the place from which the spouse comes. To separate the realized spouse's family connections from geography that is generally available among spouses from the same area, I construct an alternative family map using another spouse of the same sex, premarital commuting zone, marriage cohort and five-year age group. The alternative spouse's premarital origin moves with that family map so that the same restrictions can be applied to the realized and alternative connections.

The realized and alternative maps are similar in size, and destinations in both maps are positively related to subsequent settlement when considered separately. When the two maps are included together, the realized spouse's connections remain more predictive of where the couple settles. Applying the recipient couple's geographic exclusions to the alternative map moves the

wife-side coefficient close to zero while leaving the difference between the realized and alternative maps essentially unchanged.

A2.2.3 Timing of the loss of a sole family connection

I also use changes in whether a spouse still has family at a destination when the couple chooses where to live. I focus on places connected through exactly one eligible relative. If that relative dies before marriage, the spouse no longer has an observed family connection to the destination when the household makes its location choice. If the death occurs after the couple has already chosen where to live, the connection was still present at the relevant time.

Losing the sole connection before the location decision lowers the probability of settling in that destination by 0.0407, with standard error 0.0121. The corresponding estimate for a death occurring after the decision is 0.0088, with standard error 0.0164. The difference between the two is -0.0495 . A broader measure based on the number of relatives gives a larger response before marriage and a small positive estimate after the location decision.

A2.2.4 Broader migration responses

Spouse geography also affects broader migration choices, not only the probability assigned to a particular destination. Among brothers, a one-log-unit increase in the geographic breadth of the wife's family raises the probability of leaving the husband's premarital commuting zone by 0.062 and of leaving both spouses' premarital commuting zones by 0.060. Among sisters, a wider husband's family map raises the corresponding probabilities by 0.106 and 0.049. Differences in the quality of the two family maps mainly predict which side of the family the couple follows.

A2.2.5 Household outcomes at the realized destination

Couples settling on the wife's side of the family are more likely to own their home and coreside with relatives, have higher female employment and lower male earnings. These comparisons condition on where the couple chooses to live and are therefore descriptive. They are consistent with family proximity being associated with household support and changes in how work is allocated within the household.

A2.3 Child location

A2.3.1 Within-family comparisons and destination controls

For children, I compare cousins from the same extended family who grow up in the same commuting zone but whose parents connect them to different destinations through marriage. The preferred comparison is within blood-family–destination pairs, so that cousins are compared

only when they share the same underlying family connection to a place but differ in whether the parental marriage adds that destination to their inherited family geography.

The generic probability of choosing a particular destination is 0.0002685 in the maternal sample and 0.000268 in the paternal sample. Among 17,916 blood-family–destination pairs with at least one marriage-created connection, 9,713 contain cousins who differ in whether they inherit that connection. Table A2.3.1 shows how the estimated relationship changes as increasingly detailed family-specific destination differences are absorbed.

Table A2.3.1: Child destination choice under alternative fixed effects

	Maternal	Paternal
Family + destination FE, p.p.	0.744*** (0.059)	0.682*** (0.058)
Family $\times$ destination FE, p.p.	0.244*** (0.087)	0.115 (0.093)
Variation absorbed	0.2% additive; 49.8% interacted	
Child–destination observations, interacted	46,840,484	43,092,386

Notes: Samples contain cousins who share the same childhood commuting zone. The interacted specification compares cousins within the same blood-family–destination pair and therefore absorbs persistent family-specific attraction to that destination. Singleton family–destination pairs are omitted because they do not contribute to the within-family comparison.

The more demanding comparison absorbs roughly half of the available variation but also removes persistent family-specific attraction to a destination. Combining the maternal and paternal estimates by inverse variance gives 0.184 with standard error 0.064. Stacking the two samples while allowing separate blood-family–destination effects gives 0.183 with standard error 0.064. Excluding the 8,997 children who appear in both samples gives 0.189 with standard error 0.069.

A2.3.2 Alternative parental maps from the same commuting zone

I next compare the family geography created by the realized parental marriage with geography that could have entered through another parent from the same premarital commuting zone. Under the preferred within-family–destination comparison, the coefficient on the realized marriage-created geography is 0.184, while the coefficient on the alternative geography is -0.058 . The difference is 0.242 percentage points.

I also repeat the reassignment one hundred times in the specification with separate family and destination effects. The average coefficient on the alternative maternal maps is 0.331, compared with 0.744 for the realized map. On the paternal side, the corresponding values are 0.330 and 0.682. None of the one hundred alternative assignments produces a coefficient as large as the

realized family geography.

A2.3.3 Overlap across maternal and paternal samples

The maternal and paternal samples are constructed in the same way but are not mutually exclusive. A child can appear in both when the required cousin structure is observed on both sides of the family. The preferred estimate combines the two samples while allowing separate blood-family–destination effects for each side. Excluding children who appear in both leaves the estimate essentially unchanged, from 0.183 to 0.189 percentage points.

A2.3.4 Economic outcomes

Outcomes by realized destination type. I first describe the adult outcomes associated with the different migration paths used in Fact 3. Within the same parental family, I distinguish children who remain in their childhood CZ, those who move to a destination connected through family and those who move to a destination without inherited family connections. I report family-connected destinations both for places introduced through the parental marriage and for the broader set of places containing inherited relatives.

Children who move to family-connected destinations have higher occupational scores, wages and urban residence than siblings who remain in the childhood CZ. Among movers, however, children who use a family connection have lower outcomes than those moving to destinations without inherited relatives. These comparisons do not identify the effect of the destination itself because children choose whether and where to move.

Outcomes associated with inherited family geography. I finally ask whether the geography created by the realized parental marriage predicts later economic outcomes beyond its effect on where the child lives. I compare the realized inherited geography with family geography obtained by reassigning the parent within the same premarital commuting zone. Unlike the preceding table, this comparison does not condition on the destination eventually chosen by the child. I group outcomes into four pre-specified families.

The realized and alternative family geographies do not differ systematically across these four broader outcome families after adjustment for multiple testing. The clearest reduced-form evidence therefore concerns where children locate rather than their subsequent economic outcomes.

A3 Model details

This appendix provides the timing, full choice problems and implementation details for the two-generation model in Section 4. The model is solved backward. I first solve the child’s adult

Table A2.3.2: Adult outcomes for children by realized destination type

Comparison group	Occ. score	Log wage	Urban	High school
<i>A. Relative to all other children</i>				
Marriage-created connection	1.156 (0.140)	0.098 (0.020)	0.092 (0.005)	0.005 (0.006)
Any family connection	0.815 (0.123)	0.081 (0.018)	0.069 (0.005)	−0.002 (0.005)
<i>B. Relative to stayers</i>				
Marriage-created connection	2.344 (0.163)	0.197 (0.025)	0.181 (0.006)	0.030 (0.007)
Any family connection	1.984 (0.142)	0.175 (0.022)	0.156 (0.005)	0.019 (0.006)
<i>C. Relative to movers to unconnected destinations</i>				
Marriage-created connection	−0.635 (0.213)	−0.035 (0.029)	−0.064 (0.008)	−0.038 (0.008)
Any family connection	−0.947 (0.191)	−0.053 (0.026)	−0.087 (0.007)	−0.039 (0.008)

Notes: Each entry comes from a separate regression of the adult outcome on an indicator for settling in a family-connected destination, with parental-couple fixed effects and controls for adult age and sex. Standard errors are clustered by parental couple. Marriage-created connections are destinations connected through exactly one parental lineage. “Any family connection” also includes destinations connected through both lineages. Connected destinations differ from the childhood CZ and exclude places previously occupied by either parent. Because these comparisons condition on the child’s realized destination, they are descriptive.

Table A2.3.3: Child outcomes under realized and alternative inherited geography

Outcome family	Realized	Alternative	Raw p	Holm p	N
Education	0.0029	−0.0060	0.322	1.000	66,515
Labor location	0.0328	0.0320	0.939	0.939	57,455
Household	0.0013	−0.0001	0.879	1.000	66,515
Marriage location	−0.0060	−0.0070	0.724	1.000	66,515

Notes: Children are compared within the same blood-family–destination pair. “Alternative” replaces the realized parental family geography with that of another parent from the same premarital commuting zone. The p -value tests equality of the realized and alternative coefficients. Holm adjustment is applied across the four pre-specified outcome families.

location problem for every possible childhood origin and then use the resulting continuation values in the parents’ household-location problem.

A3.1 Timing and quantitative sample

For couple i , t_i^0 is the premarital census and t_i^1 the first post-marriage census used for parental location. For child c , t_{ic}^2 is the adult census in which the child's destination is observed. Premarital origins and family maps are measured at t_i^0 . The parent chooses the household location at t_i^1 , which becomes the child's origin in the quantitative exercises. The child then chooses an adult destination at t_{ic}^2 .

The model preserves the observed census timing. The parent sample starts with the 1900–1910 marriage window; the 1880–1900 window is excluded because it spans twenty years. The child choice sample is larger because estimating the terminal child problem does not require observing the parental decision over the common ten-year window.

Table A3.1.1: Calendar coverage of the quantitative child sample

Adult child census	Children
1910	112,507
1920	246,149
1930	294,120
1940	392,424
1950	282,463
Total	1,327,663

Notes: Adult census year for linked children entering the quantitative child-location sample. Parent destination values are indexed by the corresponding observed parent period. Coefficients are pooled across cohorts.

A3.2 Family geography across generations

The parents and children use the same premarital lineages differently. For spouse $s \in \{H, W\}$,

$$Z_i^{P,s}(d) = \mathbf{1}\{d \in \mathcal{K}_i^s\},$$

and

$$F_i^P(d) = Z_i^{P,H}(d) + Z_i^{P,W}(d).$$

The parental measure can therefore take values zero, one or two. A destination connected to both lineages supplies two sources of direct family proximity to the couple.

The child inherits the union,

$$Z_i^C(d) = \mathbf{1}\{d \in \mathcal{K}_i^H \cup \mathcal{K}_i^W\},$$

so a destination enters the extensive child map once even when both parental lineages are present. The preferred child specification separately retains the total inherited relative count

$$N_i(d) = N_i^H(d) + N_i^W(d).$$

Alternative assignments move the extensive map and the corresponding count map together. The parents themselves are excluded from the inherited extended-family map because the parental household is represented by the child's origin n .

A3.3 Child location problem

Conditional on childhood origin n , the child's deterministic value of adult destination d is

$$\begin{aligned} B_{ic}^C(d | n) &= D_{d,t_{ic}^2}^C + \alpha_{\text{away}} \mathbf{1}\{d \neq n\} - \kappa_C \log(1 + \text{dist}(n, d)) \\ &\quad + [\phi_D Z_i^C(d) + \phi_2 \log(1 + N_i(d))] \mathbf{1}\{d \neq n\} + \phi_O Z_i^C(n) \mathbf{1}\{d = n\}. \end{aligned}$$

The home and outside alternatives therefore enter the logit denominator differently. For a general shock scale σ_N , define

$$\begin{aligned} H_{ic}(n) &\equiv \exp \left\{ \frac{D_{n,t_{ic}^2}^C + \phi_O Z_i^C(n)}{\sigma_N} \right\} \\ &\quad + \sum_{d \neq n} \exp \left\{ \frac{D_{d,t_{ic}^2}^C + \alpha_{\text{away}} - \kappa_C \log(1 + \text{dist}(n, d)) + \phi_D Z_i^C(d) + \phi_2 \log(1 + N_i(d))}{\sigma_N} \right\}. \end{aligned} \tag{A3.3.1}$$

Then

$$P_{ic}^C(d | n) = \frac{\exp\{B_{ic}^C(d | n)/\sigma_N\}}{H_{ic}(n)}$$

and

$$O_{ic}(n) = \sigma_N \log H_{ic}(n).$$

The inclusive value is computed for every candidate childhood origin $n \in \mathcal{D}$. Changing n affects the child problem through three objects simultaneously: which alternative is home, the distance from home to every other destination and whether inherited family is located at the home alternative rather than at an outside destination.

A3.4 Parent location problem

The couple enters marriage from premarital locations o_i^H and o_i^W . Define the distance component of the parental location cost as

$$M_i^P(n) = \kappa_H \log(1 + \text{dist}(o_i^H, n)) + \kappa_W \log(1 + \text{dist}(o_i^W, n)). \quad (\text{A3.4.1})$$

The value of household location n can then be written

$$B_i^P(n) = D_{n,t_i^1}^P - M_i^P(n) + \theta_P F_i^P(n) + \beta \bar{O}_i(n).$$

where

$$\bar{O}_i(n) = \frac{1}{C_i} \sum_{c=1}^{C_i} O_{ic}(n).$$

Averaging over linked children prevents couples with more observed children from receiving mechanically greater continuation value. Each child's continuation value is evaluated at that child's observed adult census t_{ic}^2 , so the model retains differences in the destination environment faced by siblings observed in different years.

Parent location probabilities are

$$P_i^P(n) = \frac{\exp\{B_i^P(n)/\sigma_N\}}{\sum_{r \in \mathcal{D}} \exp\{B_i^P(r)/\sigma_N\}}.$$

The parental problem is solved only after all child continuation values have been computed.

A3.5 Normalization and interpretation

Only utility differences are identified in each location problem. I impose the same Type-I extreme-value shock scale across generations and normalize $\sigma_N = 1$. This places parent and child location values in common units. Under this normalization, β is the relative weight parents place on the child's continuation value. I set $\beta = 0.231$ following [Eckert and Kleineberg \(2024\)](#); quantitative sensitivity to this externally supplied parameter is reported with the quantitative analysis.

The destination values D^P and D^C are common location values rather than causal place effects. They absorb wages, amenities and other destination-period determinants shared by decision makers in a generation. The baseline quantitative exercises hold these destination values fixed and reassign family geography.

A3.6 Direct inheritance and parental settlement

Let Z collect the child's inherited family map and relative counts and let P denote the parental location distribution induced by the corresponding parent problem. For any child outcome $T(P, Z)$, changing family geography from baseline state 0 to alternative state 1 can be written

$$T(P^1, Z^1) - T(P^0, Z^0) = \underbrace{T(P^0, Z^1) - T(P^0, Z^0)}_{\text{direct inherited access}} + \underbrace{T(P^1, Z^1) - T(P^0, Z^1)}_{\text{parental location}}. \quad (\text{A3.6.1})$$

The first term changes the inherited child geography while holding the entire parental location distribution fixed. The second allows parental settlement to respond and therefore changes the child's origin before the adult location choice. Under this ordering, the interaction between the new map and the new parental policy is assigned to the parental-location component. The quantitative section reports both terms separately.

A4 Model quantification and fit

This appendix provides the implementation details behind Section 5. I first describe the calendar structure, samples and weighting. I then give the sampled-choice likelihoods and the mappings from the empirical family responses to ϕ_O , ϕ_D and θ_P . I conclude with additional fit diagnostics.

A4.1 Samples, timing and parameter discipline

Families enter the model at different calendar dates but face the same sequence of decisions. Period 0 is the last census before marriage, Period 1 is the census in which the married couple's location is observed and Period 2 is the linked child's adult census. These periods therefore index stages of the family life cycle rather than common calendar years. For a couple marrying between 1900 and 1910, Periods 0 and 1 correspond to 1900 and 1910; for a couple marrying between 1920 and 1930, they correspond to 1920 and 1930.

The coefficients α_{away} , κ_C , ϕ_O , ϕ_D , κ_H , κ_W and θ_P are common across cohorts while destination values are not. A child observed as an adult in year y faces $D_{d,y}^C$, while a couple observed after marriage in year y faces $D_{n,y}^P$. Within each year, destination values are identified only up to a common additive constant, so I normalize their mean across the 709 CZs to zero.

Weighting. Every likelihood is unweighted at the level of the decision maker. Each child contributes once to the child likelihood and each couple once to the parent likelihood. Couples with more linked children therefore do not receive more weight in parent estimation. Instead, if couple i has C_i linked children, the continuation value entering the parents' problem is the

Table A4.1.1: Calendar structure of the quantitative samples

<i>Panel A. Parent sample</i>			
	Period 0	Period 1	
Marriage window	Premarital census	Parent census	Couples
1900–1910	1900	1910	109,685
1910–1920	1910	1920	105,073
1920–1930	1920	1930	55,509
1930–1940	1930	1940	2,606
Total			272,873
<i>Panel B. Child samples</i>			
Adult child census	Child block estimation	Children of parent sample	
1910	112,507	–	
1920	246,149	5,780	
1930	294,120	178,890	
1940	392,424	365,718	
1950	282,463	279,186	
Total	1,327,663	829,574	

Notes: Panel A shows the census mapping and size of the four ten-year marriage windows used in parent estimation. Periods 0 and 1 are relative to each family. Panel B reports the adult census years of children. The child block is estimated on all children with an observed childhood CZ. The children of couples in the parent sample enter parental continuation values and the quantitative exercises. Child destination values correspond to the child’s adult census year and parent destination values to the post-marriage census year.

simple average $C_i^{-1} \sum_c O_{ic}(n)$. Each child’s continuation value is first evaluated using the destination values from that child’s own adult census year. The model counterparts of the empirical family responses use the same observations and regression weights as the corresponding empirical specifications. Child outcomes in the quantitative exercises are averaged over children and parent outcomes over couples.

Fixed parameters and estimation sequence. I set $\sigma_N = 0.51$ following [Looser \(2026\)](#) and $\beta = 0.231$ following [Eckert and Kleineberg \(2024\)](#). The remaining parameters are estimated backward. First, I use the origin-loss response to obtain ϕ_O . Holding this value fixed, I profile over ϕ_D ; for every candidate ϕ_D , I re-estimate α_{away} , κ_C and $\{D_{d,y}^C\}$ from the child likelihood before evaluating the Fact 3 auxiliary moment. The resulting child block determines $\bar{O}_i(n)$ for every couple and candidate parental location. I then profile over θ_P ; for every candidate θ_P , I re-estimate κ_H , κ_W and $\{D_{n,y}^P\}$ from the parent likelihood before evaluating the Fact 2 auxiliary moment. All auxiliary moments, fit statistics and quantitative exercises are evaluated

over the complete set of 709 CZs.

A4.2 Child location estimation

Let n_{ic} , d_{ic} and y_{ic} denote the observed childhood CZ, adult CZ and adult census year of child c in family i . Conditional on childhood origin n_{ic} , the deterministic value of destination d is

$$B_{ic}^C(d | n_{ic}) = D_{d,y_{ic}}^C + \alpha_{\text{away}} \mathbf{1}\{d \neq n_{ic}\} - \kappa_C \log(1 + \text{dist}(n_{ic}, d)) \\ + \phi_D Z_i^C(d) \mathbf{1}\{d \neq n_{ic}\} + \phi_O Z_i^C(n_{ic}) \mathbf{1}\{d = n_{ic}\}. \quad (\text{A4.2.1})$$

The destination-period values absorb common differences in destination attractiveness within each adult census year. The coefficient α_{away} governs the home-versus-away margin, κ_C governs the distance gradient and ϕ_D and ϕ_O govern inherited family outside and at the childhood origin.

Sampled alternatives and likelihood. Each child chooses among 709 CZs. To reduce the size of the estimation problem, the sampled choice set $\mathcal{S}_{ic}$ always contains the observed adult destination, the childhood origin and every family-connected destination. I then draw at most 50 alternatives uniformly from the remaining nonmandatory destinations. Let U_{ic} denote the number of nonmandatory alternatives and $K_{ic} = \min\{50, U_{ic}\}$ the number drawn. Mandatory alternatives have inclusion probability one, while each sampled nonmandatory destination has inclusion probability

$$\pi_{ic}^S = \frac{K_{ic}}{U_{ic}}. \quad (\text{A4.2.2})$$

The McFadden correction therefore adds $-\log \pi_{ic}^S$ to the index of a sampled nonmandatory destination. The resulting probability is

$$\tilde{P}_{ic}^C(d | n_{ic}) = \frac{\exp\{(B_{ic}^C(d | n_{ic}) - \log \pi_{ic}^S(d)) / \sigma_N\}}{\sum_{r \in \mathcal{S}_{ic}} \exp\{(B_{ic}^C(r | n_{ic}) - \log \pi_{ic}^S(r)) / \sigma_N\}}, \quad (\text{A4.2.3})$$

where $\pi_{ic}^S(d) = 1$ for mandatory alternatives. Conditional on (ϕ_O, ϕ_D) , I solve

$$\left(\hat{\alpha}_{\text{away}}, \hat{\kappa}_C, \left\{ \hat{D}_{d,y}^C \right\} \right) = \arg \max \sum_{i,c} \log \tilde{P}_{ic}^C(d_{ic} | n_{ic}), \quad (\text{A4.2.4})$$

subject to the normalization that destination values average zero across CZs within each census year. The home-versus-away choices discipline α_{away} , relative destination choices across distances discipline κ_C and destination-by-year choice shares discipline $D_{d,y}^C$.

Family at the childhood origin. The coefficient ϕ_O is calibrated from the increase in out-migration after losing family at the childhood origin. The empirical response is $\hat{m}_O^{\text{data}} = 0.070969$. For child c in family i , let

$$P_{ic}^{\text{leave}}(z; \phi_O) = 1 - P_{ic}^C(n_{ic} \mid n_{ic}; Z_i^C(n_{ic}) = z, \phi_O), \quad (\text{A4.2.5})$$

where $z \in \{0, 1\}$ indicates whether inherited family is present at the childhood origin and the probability is evaluated over all 709 CZs. The model response is

$$\hat{m}_O^{\text{model}}(\phi_O) = \frac{1}{N_O} \sum_{(i,c) \in \mathcal{I}_O} [P_{ic}^{\text{leave}}(0; \phi_O) - P_{ic}^{\text{leave}}(1; \phi_O)], \quad (\text{A4.2.6})$$

where $\mathcal{I}_O$ is the separate sample underlying the origin-loss estimate. I choose

$$\hat{\phi}_O : \quad \hat{m}_O^{\text{model}}(\hat{\phi}_O) = 0.070969. \quad (\text{A4.2.7})$$

This gives $\hat{\phi}_O = 0.365$. I hold this value fixed when estimating the remaining child parameters. Because the origin-loss sample differs from the full child likelihood sample, the corresponding removal exercise in the final child block need not equal the calibration target observation for observation. In the final block, removing family from the childhood origin raises out-migration by 0.076 among children whose childhood CZ contains inherited family.

Family at an outside destination. I discipline ϕ_D using the pooled same-origin cousin comparison from Fact 3. For each candidate value, I first solve equation (A4.2.4), giving

$$\hat{v}_C(\phi_D) = \left(\hat{\alpha}_{\text{away}}(\phi_D), \hat{\kappa}_C(\phi_D), \left\{ \hat{D}_{d,y}^C(\phi_D) \right\} \right).$$

I then evaluate full-choice-set probabilities for the child-destination observations in the empirical cousin comparison. Let A_j denote the marriage-created-destination indicator and $\tilde{A}_j$ its residual after removing the same fixed effects and controls used in Fact 3. The model-implied coefficient is

$$\hat{m}_D^{\text{model}}(\phi_D) = \frac{\sum_j \tilde{A}_j P_j^C(\phi_D)}{\sum_j \tilde{A}_j^2}. \quad (\text{A4.2.8})$$

The full empirical sample gives 0.001833758. Restricting the same specification to the origins, destinations and marriage windows represented in the quantitative model gives $\hat{m}_D^{\text{data}} = 0.002460832$, which is the structural target. I choose

$$\hat{\phi}_D : \quad \hat{m}_D^{\text{model}}(\hat{\phi}_D) = 0.002460832. \quad (\text{A4.2.9})$$

The resulting value is $\hat{\phi}_D = 0.340$. Because α_{away} , κ_C and the destination-period values are re-estimated for every candidate ϕ_D , the auxiliary moment determines the family coefficient jointly with the child location likelihood rather than holding the other child parameters fixed.

Continuation values. After estimating the child block, I evaluate the model over all 709 destinations. For every linked child and every candidate parental location n ,

$$O_{ic}(n) = \sigma_N \log \left[\sum_{d \in \mathcal{D}} \exp \{ B_{ic}^C(d | n) / \sigma_N \} \right]. \quad (\text{A4.2.10})$$

If couple i has C_i linked children, the object entering the parents' problem is

$$\bar{O}_i(n) = \frac{1}{C_i} \sum_{c=1}^{C_i} O_{ic}(n). \quad (\text{A4.2.11})$$

Each child retains the destination values associated with that child's adult census year before these continuation values are averaged.

A4.3 Parent location estimation

Let o_i^H and o_i^W denote the husband's and wife's premarital CZs and let y_i^P denote the census year in which the married couple's location is observed. The deterministic value of establishing the household in n is

$$\begin{aligned} B_i^P(n) = & D_{n, y_i^P}^P - \kappa_H \log(1 + \text{dist}(o_i^H, n)) - \kappa_W \log(1 + \text{dist}(o_i^W, n)) \\ & + \theta_P Z_i^P(n) + \beta \bar{O}_i(n). \end{aligned} \quad (\text{A4.3.1})$$

With $\sigma_N = 1$, the full-choice-set probability is

$$P_i^P(n) = \frac{\exp \{ B_i^P(n) / \sigma_N \}}{\sum_{r \in \mathcal{D}} \exp \{ B_i^P(r) / \sigma_N \}}. \quad (\text{A4.3.2})$$

Sampled alternatives and likelihood. The parent likelihood uses the same alternative-sampling correction as the child block. The sampled set $\mathcal{S}_i$ always contains the observed post-marriage CZ and every CZ containing eligible premarital family of either spouse. At most 50 destinations are drawn uniformly from the remaining alternatives. If U_i is the number of nonmandatory destinations and $K_i = \min\{50, U_i\}$, each sampled nonmandatory destination has inclusion probability K_i/U_i and receives the corresponding $-\log(K_i/U_i)$ correction. Let

$\tilde{P}_i^P(n)$ denote this corrected sampled-set probability. For a candidate value of θ_P , I solve

$$\left(\hat{\kappa}_H, \hat{\kappa}_W, \left\{\hat{D}_{n,y}^P\right\}\right) = \arg \max \sum_i \log \tilde{P}_i^P(n_i), \quad (\text{A4.3.3})$$

subject to mean-zero destination values within each census year. The parent estimation sample contains 272,873 couples and 14,672,610 sampled couple–destination rows. Couples whose spouses enter marriage from different CZs separately discipline κ_H and κ_W , because the same candidate destination then generally lies at different distances from the two premarital origins.

Parental family proximity. I discipline θ_P using the pooled spouse-supplied-destination response from Fact 2. Restricting the empirical specification to the couples and destinations represented in the quantitative model gives $\hat{m}_P^{\text{data}} = 0.00161$. For every candidate θ_P , I first re-estimate κ_H , κ_W and $\{D_{n,y}^P\}$ from equation (A4.3.3). I then evaluate the model over all 709 CZs and apply the same pooled destination regression used in the data. Let S_j denote the spouse-supplied-destination indicator and $\tilde{S}_j$ its residual after removing the same design-specific fixed effects and controls. The model-implied coefficient is

$$\hat{m}_P^{\text{model}}(\theta_P) = \frac{\sum_j \tilde{S}_j P_j^P(\theta_P)}{\sum_j \tilde{S}_j^2}. \quad (\text{A4.3.4})$$

The stacked auxiliary regression contains 28,606,732 couple–destination rows. The residualized treatment and denominator are computed once because they do not depend on θ_P . I choose

$$\hat{\theta}_P : \quad \hat{m}_P^{\text{model}}(\hat{\theta}_P) = 0.00161. \quad (\text{A4.3.5})$$

The resulting estimates are

$$\hat{\theta}_P = 1.6860, \quad \hat{\kappa}_H = 0.6137, \quad \hat{\kappa}_W = 0.5626.$$

The standard error of $\hat{\theta}_P$ is 0.1856. The model uses a single parental family coefficient regardless of which spouse supplies the connection. As an untargeted check, the husband-side coefficient is 0.00140 in the data and 0.00134 in the model, while the wife-side coefficient is 0.00184 in the data and 0.00190 in the model.

A4.4 Model fit

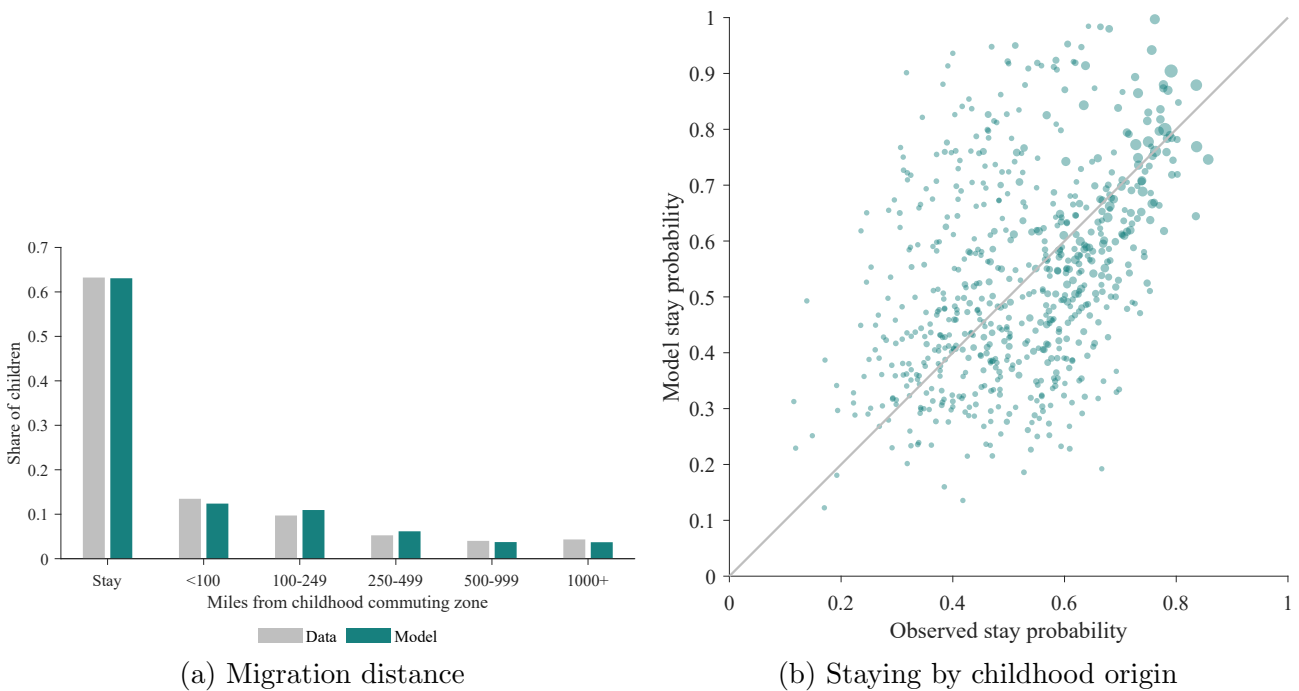
The family responses used to discipline ϕ_O , ϕ_D and θ_P are targeted auxiliary moments. I also evaluate the estimated model on aggregate margins of child and parent location choice that are

not matched directly as auxiliary moments.

Child location. The model closely reproduces the aggregate home-versus-away margin and migration distance. In the data, 63.23% of children remain in their childhood CZ, compared with 63.07% in the model. Among movers, mean log distance from the childhood origin is 5.280 in the data and 5.273 in the model. The main discrepancy concerns inherited family destinations. In the data, 60.72% of children settle in a CZ containing inherited family, compared with 52.79% in the model. The gap is concentrated outside the childhood origin, where the corresponding shares are 7.15% and 2.45%.

Migration distance and staying across origins. Figure A4.4.1 compares observed and model-implied child location choices over the full set of 709 CZs. Panel (a) reports the migration-distance distribution, including the probability of remaining in the childhood CZ. Panel (b) compares observed and predicted variation in staying across childhood origins.

Figure A4.4.1: Untargeted fit



Notes: Panel (a) compares the observed and model-implied distribution of children across migration-distance categories, including staying in the childhood CZ. Panel (b) compares observed and model-implied probabilities of remaining in the childhood CZ across origins with at least 100 linked children; marker size reflects the number of children raised in each origin. Model predictions integrate over all 709 CZs from the observed childhood origin.

Across the 650 childhood CZs represented in Panel (b), the weighted correlation between observed and model-implied stay probabilities is 0.581, with an RMSE of 0.141. Regressing the observed stay probability on its model-implied counterpart gives a slope of 0.429.

Parent location. The parent model also reproduces the main settlement margins closely. In the data, 80.17% of couples settle in either spouse’s premarital CZ, compared with 80.83% in the model. The share settling in a destination connected to either spouse’s premarital family is 81.56% in the data and 80.88% in the model. Mean log distance from the husband’s premarital origin is 1.510 in the data and 1.512 in the model; the corresponding values for the wife’s origin are 1.562 and 1.574. The model places somewhat more probability on very long moves: mean distance from the husband’s origin is 89.6 miles in the data and 119.2 in the model, while mean distance from the wife’s origin is 86.1 and 121.9 miles.

A5 Quantitative exercises and sensitivity

This appendix describes the quantitative exercises in Section 6 and reports additional results and sensitivity checks.

A5.1 Construction of the quantitative exercises

All exercises keep the observed couple, both spouses’ premarital commuting zones, household characteristics, destination values and estimated parameters fixed, except where stated. After each change, I recompute parental and child location choices using the estimated model. Parents face destination values from the family’s post-marriage census and children face destination values from the census in which they are observed as adults. All reported outcomes use the full set of 709 commuting zones. Parent outcomes average over couples and child outcomes over children.

Removing family ties. The family-ties exercises set family coefficients to zero while keeping observed family maps. Removing parents’ ties sets the parental family coefficient θ_P to zero. Removing children’s ties sets the destination and origin family coefficients ϕ_D and ϕ_O to zero, which also changes the continuation value parents attach to each location. Removing both sets all three to zero. The contribution of each generation is the Shapley value computed from the two single-generation removals and the joint removal.

Reassigning family maps. The map exercises replace a spouse’s premarital family map with the map of another spouse of the same sex, marriage cohort, five-year age group and socioeconomic-rank group. In the one-spouse exercises, I carry out the replacement separately for husbands and wives and average the two effects. The both-spouse exercise replaces both maps simultaneously. The geographic restriction on the replacement map ranges from the spouse’s own commuting zone to anywhere in the country. The socioeconomic-rank restriction is kept fixed throughout, except when I explicitly study socioeconomic matching.

Equalizing family geography. The equalization exercise ranks couples by the average premarital socioeconomic rank of the two spouses. Each of the 119,816 couples in the bottom two quintiles receives the husband’s and wife’s premarital family maps of one couple drawn at random from the top two quintiles of the same marriage cohort, anywhere in the country. All other couples keep their observed maps. Quintiles of parental rank used to report child outcomes are defined over children and therefore do not coincide exactly with the couple-level quintiles used to select recipients.

A5.2 Family ties and persistence

Table A5.2.1 reports the estimated model and the three family-tie removals separately. Removing parents’ ties alone accounts for almost all of the change in parental settlement and most of the change in the persistence of place quality. Removing children’s ties alone accounts for most of the decline in the probability that children remain in their childhood commuting zone. For parental settlement and the persistence of place quality, removing both is close to the sum of the two separate removals, so the Shapley allocation in Table 6 is not sensitive to the treatment of the interaction. For children’s probability of remaining in the childhood commuting zone the two removals overlap substantially, and the Shapley allocation divides the overlap equally between the two generations.

Table A5.2.1: Family ties and persistence: separate removals

	Estimated model	Parents’ ties removed	Children’s ties removed	Both removed
<i>Panel A. Where parents settle (%)</i>				
Husband’s premarital CZ	73.79	53.58	73.42	52.79
Wife’s premarital CZ	72.27	52.59	71.91	51.82
CZ with either spouse’s relatives	83.21	57.50	82.66	56.56
CZ with exactly one lineage	17.57	8.64	17.30	8.38
Top-quartile CZ	21.32	24.11	21.36	24.19
<i>Panel B. Where children settle (%)</i>				
Childhood CZ	62.79	61.62	56.91	57.67
Above-median CZ	60.46	62.84	61.09	63.21
Top-quartile CZ	27.23	30.42	28.06	31.00
<i>Panel C. Intergenerational persistence of place quality</i>				
Parental-rank slope	0.1076	0.0787	0.1003	0.0733
Q5–Q1 gap in top-quartile access (p.p.)	1.55	–0.26	1.02	–0.64

Notes: Parents’ ties removed sets $\theta_P = 0$. Children’s ties removed sets $\phi_D = \phi_O = 0$. Both removed sets all three to zero. Observed couples, family maps, premarital origins, distance coefficients and destination values are unchanged in every column.

Equalizing family geography. Table A5.2.2 reports the equalization exercise. Replacing the family maps of low-status couples moves a large share of them away from family-connected destinations: the probability that parents settle where either spouse has premarital relatives falls from 83.2% to 48.0%, because the donated relatives typically live far from the recipients' own premarital origins. Children of low-status parents gain most, and the parental-rank slope falls by almost half.

Table A5.2.2: Equalizing family geography between low- and high-status couples

	Estimated model	Equalized geography	Change	95% interval
Parents settle in husband's premarital CZ (%)	73.79	63.35	−10.44	[−10.50, −10.39]
Parents settle in CZ with relatives (%)	83.21	48.03	−35.18	[−35.37, −35.04]
Children remain in childhood CZ (%)	62.79	60.85	−1.94	[−1.96, −1.93]
Children reach top-quartile CZ (%)	27.23	29.59	2.36	[2.33, 2.38]
Children of bottom two quintiles (%)	26.60	31.52	4.92	
Parental-rank slope	0.1076	0.0560	−0.0516	[−0.0523, −0.0509]
Q5–Q1 gap in top-quartile access (p.p.)	1.55	−3.81	−5.36	[−5.45, −5.26]

Notes: Couples in the bottom two quintiles of average premarital socioeconomic rank receive both premarital family maps of a randomly drawn couple from the top two quintiles of the same marriage cohort, drawn from anywhere in the country. Spouses, premarital origins and all other characteristics are unchanged. Intervals are based on 200 paired resamples clustered at the root-family level.

A5.3 Geographic reach

The reach exercise asks how much outcomes change when one spouse's family connections may come from progressively farther away. For the same-CZ case, the replacement map must come from the observed spouse's premarital commuting zone. I then expand this radius to 100, 250, 500 and 1,000 kilometers and finally to the entire country. If no alternative map is available within a given radius, the household keeps its observed map.

Most of the effect is realized at relatively short distances. Extending family reach to 250 kilometers produces 94% of the increase in top-quartile access obtained when maps can be drawn from anywhere in the country. The parental-rank slope continues to decline at longer distances, so the reduction in the persistence of place quality is spread more evenly across the distance profile. The average effect also conceals considerable heterogeneity: between the same-CZ and nationwide assignments, the mean absolute change in a household's top-quartile probability is 1.49 percentage points and 58% of households experience an increase.

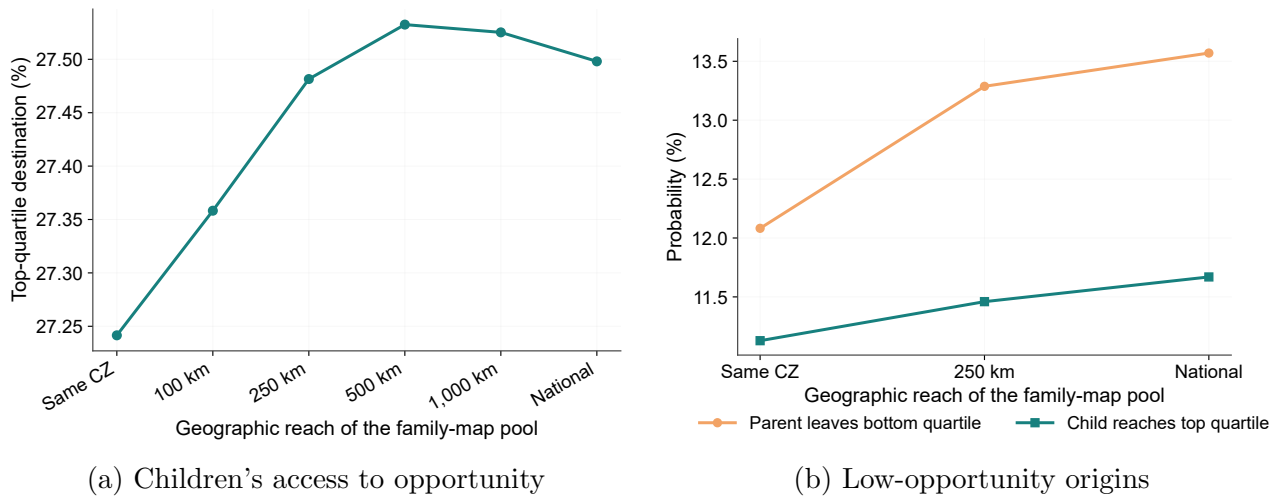
Varying the full household family geography. The one-spouse exercise changes the family map of only one spouse. I therefore also consider broader changes to the household's geog-

Table A5.3.1: Child outcomes by the geographic reach of family connections

Reach	Top-quartile CZ (%)	Remains in childhood CZ (%)	Parental-rank slope	Q5–Q1 gap (p.p.)	Fallback share (%)
Same CZ	27.24	62.79	0.1076	1.54	0.00
100 km	27.36	62.71	0.1073	1.55	1.05
250 km	27.48	62.51	0.1068	1.54	0.15
500 km	27.53	62.48	0.1063	1.51	0.04
1,000 km	27.53	62.48	0.1060	1.49	0.01
National	27.50	62.46	0.1056	1.44	0.00

Notes: Each row replaces one spouse's premarital family map with a map from another spouse satisfying the same sex, marriage-cohort, age and socioeconomic-rank restrictions. The replacement spouse's premarital CZ must lie within the stated distance. If no alternative map is available, the observed map is retained and the household is counted in the fallback column. Husband- and wife-side replacements are averaged. The Q5–Q1 gap is the difference in top-quartile access between children of top- and bottom-quintile parents.

Figure A5.3.1: Family reach and access to opportunity



Notes: Panel (a) expands the geographic pool from which one spouse's premarital family map is reassigned. Panel (b) restricts the sample to couples whose two premarital commuting zones both lie in the bottom quartile of the destination wage-rank distribution. In both panels, husband- and wife-side reassignments are averaged.

raphy. The first replaces both spouses' family maps simultaneously. The second changes only one spouse's premarital origin. The third changes that spouse's premarital origin together with the associated family map. The latter two exercises capture a broader change in the geography brought into the marriage and are not interpreted as pure family-map effects.

Changing both family maps has a much larger effect than changing only one side of the family: top-quartile access rises by 3.84 percentage points, the probability that children remain in their childhood commuting zone falls from 62.8% to 58.2% and the parental-rank slope falls by 30%. Changing the spouse's premarital origin produces a still larger response, showing that the

Table A5.3.2: Changes in family geography and premarital origin

Component varied	Top-quartile destination (%)			Parental-rank slope	
	Same CZ	National	Change	Same CZ	National
One spouse's family map	27.24	27.50	0.26	0.1076	0.1056
Both spouses' family maps	27.25	31.09	3.84	0.1075	0.0752
One spouse's premarital origin	27.24	34.41	7.16	0.1076	0.0461
One spouse's origin and family map	27.24	33.92	6.68	0.1076	0.0532

Notes: Replacement pools preserve sex, marriage cohort, five-year age group and socioeconomic-rank group. Same-CZ replacements are drawn from the spouse's premarital commuting zone, while national replacements may come from anywhere in the country. Rows 1, 3 and 4 vary one spouse at a time and average husband- and wife-side replacements. Row 2 replaces both family maps simultaneously. Premarital origins remain fixed except in rows 3 and 4.

geography associated with marriage operates through both family connections and the place from which the spouse enters the household.

Low-opportunity origins. I next repeat the reach exercise for families starting in low-opportunity places. I restrict the sample to couples whose two premarital CZs both lie in the bottom quartile of the within-year destination wage-rank distribution. Because this restriction is based on premarital locations, the sample is determined before parents choose where to establish the household. It contains 183,051 children. Defining low opportunity instead by the average rank of the two premarital origins yields 193,888 children.

Table A5.3.3: Families starting in low-opportunity commuting zones

Outcome	Same CZ	250 km	National	National – same CZ
Parent leaves a bottom-quartile CZ	0.1208	0.1329	0.1357	0.0149
Parent reaches an above-median CZ	0.0757	0.0825	0.0863	0.0106
Parent reaches a top-quartile CZ	0.0356	0.0387	0.0405	0.0049
Expected parent destination rank	0.2178	0.2226	0.2245	0.0067
Child reaches an above-median CZ	0.2465	0.2522	0.2556	0.0091
Child reaches a top-quartile CZ	0.1113	0.1146	0.1167	0.0054
Expected child destination rank	0.3314	0.3354	0.3373	0.0059

Notes: The sample contains couples whose two premarital commuting zones lie in the bottom quartile of the pooled destination wage-rank distribution. It contains 183,051 children. The sample is defined using premarital origins, before the post-marriage location choice. One spouse's family map is reassigned and husband- and wife-side reassignments are averaged. The 250-kilometer specification delivers 61.2% of the nationwide change in the child's top-quartile probability. The conditional 95% interval for the nationwide change is [0.0053, 0.0055].

Broader family reach has larger effects for these households. Reassigning one spouse's map raises the probability that parents leave the bottom quartile by 12.3% relative to the same-CZ baseline and children's probability of reaching a top-quartile destination by 4.9%. Reassigning

both spouses' maps raises children's probability of reaching a top-quartile destination from 11.1% to 18.2% and of reaching an above-median destination from 24.6% to 36.7%.

A5.4 Family connections to high- and low-opportunity places

Geographic reach does not capture where family connections lead. Two spouses from the same premarital CZ may have relatives in very different parts of the country. To summarize these differences, I assign each premarital family map an opportunity measure equal to the average destination wage of the places where relatives live, weighted by the number of relatives in each destination. For spouse j ,

$$A_j = \frac{\sum_{d \in \mathcal{K}_j} N_j(d) Q_{d,t_j}}{\sum_{d \in \mathcal{K}_j} N_j(d)}, \quad (\text{A5.4.1})$$

where $\mathcal{K}_j$ is the set of commuting zones containing eligible relatives, $N_j(d)$ is the number of those relatives in d and Q_{d,t_j} is the destination-year wage index.

I then compare family maps from the bottom and top quartiles of this measure while keeping the spouse's premarital CZ and marriage cohort fixed. The comparison therefore changes where the spouse's relatives live without changing the spouse's geographic origin.

Table A5.4.1: High- and low-opportunity family maps within the same commuting zone

Outcome	Low map (Q1–Q2)	High map (Q1–Q2)	Difference (Q1–Q2)	Difference (Q4–Q5)
Top-quartile destination (%)	26.53	26.79	0.264	0.345
Expected destination rank	0.5386	0.5405	0.00199	0.00209
Expected destination log wage	7.1780	7.1795	0.00151	0.00187
<i>Top-quartile effect, lower-status households (p.p.)</i>				
Total			0.264	
Direct inherited access			0.088	
Parental settlement			0.177	

Notes: Map opportunity is the relative-count-weighted mean wage index of the destinations containing premarital relatives. Low- and high-opportunity maps are drawn from the bottom and top quartiles of this measure within the spouse's premarital CZ and marriage cohort. The Q1–Q2 columns refer to households in the bottom two quintiles of parental socioeconomic rank. The final column reports the corresponding difference for households in the top two quintiles. Components may not sum exactly because of rounding.

Family maps pointing toward higher-wage destinations improve children's subsequent destination outcomes, and the effect is somewhat larger among higher-status households. Comparing the bottom and top deciles instead of quartiles raises the effect on top-quartile access from 0.264 to 0.378 percentage points for lower-status households (Figure 5). Measured per within-

cell standard deviation of map quality, the two comparisons imply nearly the same effect, about 0.26 to 0.27 percentage points.

Socioeconomic matching within CZ. The previous comparison raises the possibility that higher-status spouses simply bring systematically better family maps. To separate socioeconomic matching from family geography, I redraw family maps within the same premarital CZ without requiring the replacement spouse to have a similar socioeconomic rank. I compare these assignments with the baseline specification that preserves socioeconomic rank.

Table A5.4.2: Socioeconomic matching and family-map assignment

Diagnostic	Status matched	Status random
Within-CZ slope of raw map wage rank	0.00404	−0.00032
Within-CZ slope of model-relevant access	−0.51552	−0.51470
Q5–Q1 raw map wage rank	0.10646	0.10350
Median absolute status-rank gap	0.0000	0.0979
<i>Household-level change, random minus matched</i>		
Mean absolute change in map access		0.08883
Mean absolute change in parent CZ wage rank		0.00142
Mean absolute change in child log wage		0.00149
Signed mean change in child log wage		0.000005
Jaccard overlap of assigned maps		0.313

Notes: Status matched preserves the spouse’s socioeconomic-rank group when assigning a replacement family map. Status random drops this restriction while continuing to draw maps from the same premarital CZ. The first panel describes how the assignments change map composition. The second reports household-level differences between the two assignments.

Removing socioeconomic matching changes the maps assigned to individual households but has essentially no effect on the average child outcome. The small signed effect therefore reflects offsetting gains and losses. A stronger version of the exercise gives every spouse in the bottom two quintiles of own socioeconomic rank the family map of a randomly drawn spouse from the top two quintiles of the same premarital CZ and marriage cohort. Children of recipients gain only 0.02 percentage points in top-quartile access and the parental-rank slope falls by 0.3%. Within a commuting zone, the family geographies of high- and low-status spouses are too similar for socioeconomic sorting on family maps to matter for the persistence of place quality.

A5.5 Direct inherited access and parental settlement

Family geography can affect children in two ways. It changes the destinations to which children are directly connected through relatives and it can change where their parents establish the household. To separate these channels, I compute the effect in two steps. First, I give the

child the broader family map while keeping the parental location distribution exactly as it was under local family reach. This is the direct inherited-access effect. I then allow parental location choices to adjust to the new family geography. The additional change is the parental-settlement effect.

Formally, let $T(G, Z)$ denote a child outcome when G is the parental location distribution and Z the inherited family map. Let L denote the same-CZ family geography and N nationwide family reach. Then

$$T(G^N, Z^N) - T(G^L, Z^L) = \underbrace{T(G^L, Z^N) - T(G^L, Z^L)}_{\text{direct inherited access}} + \underbrace{T(G^N, Z^N) - T(G^L, Z^N)}_{\text{parental settlement}}. \quad (\text{A5.5.1})$$

The second term includes the interaction between the broader family map and the new parental location.

Table A5.5.1: Direct inherited access and parental settlement

Outcome	Total	Direct access	Parental settlement
Top-quartile destination, p.p.	0.257	0.070	0.186
Adult destination percentile	0.124	0.040	0.084
Stays in childhood CZ, p.p.	-0.332	-0.465	0.133
Mean log migration distance	0.021	0.027	-0.006
Parental-rank slope	-0.0020	-0.0003	-0.0017
Q5-Q1 top-quartile gap, p.p.	-0.094	-0.013	-0.082

Notes: The exercise expands one spouse's family map from the same CZ to nationwide reach. The direct-access column changes the child's inherited family map while keeping the full parental location distribution at its same-CZ value. The parental-settlement column then allows parental location choices to adjust. Components may not sum exactly because of rounding.

The parental-settlement channel accounts for most of the change across several measures of destination quality. It contributes 72.6% of the increase in top-quartile access, 68.1% of the increase in the expected destination percentile and 85% of the decline in the parental-rank slope.

Broader family geography also changes where parents locate relative to the two family lineages. The probability that parents choose a destination connected through exactly one spouse's pre-marital family rises from 19.44% to 76.23%.

The same decomposition for high- versus low-opportunity family maps assigns 33.1% of the top-quartile effect to direct inherited access and 66.9% to parental settlement.

A5.6 Parental motives for family proximity

Parental location choices respond to family geography for two reasons in the model. Parents value living near their own relatives and they value the future location opportunities available to their children. I quantify the contribution of each motive by recomputing location choices after removing them one at a time and then removing both together.

Table A5.6.1: Parental motives for family proximity

Specification	Same-CZ top-Q	National top-Q	Difference, p.p.
Both motives retained	27.24	27.50	0.257
Child-continuation motive removed	26.55	26.71	0.162
Own-family proximity removed	30.42	30.45	0.036
Both motives removed	28.98	29.02	0.032
Own-family proximity, Shapley contribution			0.175
Child continuation, Shapley contribution			0.050

Notes: The outcome is the child's probability of reaching a top-quartile destination. Parental and child location choices are recomputed in each specification. The Shapley contributions divide the interaction between the two parental motives symmetrically.

Most of the response comes from parents' own preference for family proximity. The Shapley decomposition assigns about 68% of the total effect to this motive and about 20% to the value parents place on their children's future opportunities.

A5.7 Additional decompositions and robustness

Husband- and wife-side variation. The results are similar depending on which spouse's family map is changed. Broadening husband-side family connections raises top-quartile access by 0.266 percentage points, compared with 0.247 points for wife-side connections.

A single family tie in a high-opportunity place. I add one relative in the nearest top-quartile commuting zone to the family maps of the 55,788 couples whose two premarital commuting zones both lie in the bottom quartile. The added tie is on average 331 miles from the husband's premarital origin and is visible to both parents and children. Parents settle in the tied commuting zone with probability 0.37%, and children's probability of reaching a top-quartile destination rises from 11.14% to 11.23%. A single distant tie does little against the distance cost and the remaining family geography at the origin, which is why the exercises in the main text change the household's family geography as a whole.

Family geography across marriage cohorts. The reach of realized family maps rises over the sample period. The mean log distance from the husband's premarital origin to the

commuting zones containing premarital relatives increases from 3.06 for couples married in the 1900s to 3.36 in the 1910s and 3.30 in the 1920s, and the number of family-connected commuting zones rises from 3.5 to 3.9. Over the same cohorts, the model-implied parental-rank slope of destination quality falls from 0.179 to 0.105. To ask whether the change in family maps contributes to this decline, I give each household a family map drawn from couples married after or before the median marriage year within the same premarital commuting zone and socioeconomic-rank group. The two assignments yield nearly identical outcomes: top-quartile access differs by 0.002 percentage points and the parental-rank slope by 0.0002. Conditional on the premarital origin, later family maps are not better positioned, so the decline in the persistence of place quality across cohorts reflects changes in destination values rather than in family geography.

Adverse childhood origin. I also consider whether broader inherited family geography provides alternatives when the child’s origin becomes less attractive. After parental location has been determined, I impose an unanticipated one-standard-deviation reduction in utility at the realized childhood CZ. I then measure the probability that the child chooses an unaffected top-quartile destination outside that CZ. Changing inherited family connections alone raises this probability from 22.68% to 22.86%. Allowing parental location to respond raises it further to 22.94%.

A5.8 Sampling uncertainty

I assess sampling uncertainty by drawing 200 paired resamples of families, clustered at the root-family level. For each resample, I keep the estimated structural parameters and assigned family maps fixed and recompute the quantitative outcomes. The resulting intervals therefore capture sampling variation across families.

Table A5.8.1: Sampling uncertainty in the quantitative estimates

Outcome	Estimate	95% interval
<i>Without family ties</i>		
Parents settle in husband's premarital CZ, p.p.	−21.00	[−21.03, −20.96]
Children reach top-quartile CZ, p.p.	3.76	[3.73, 3.80]
Parental-rank slope	−0.0343	[−0.0352, −0.0335]
<i>Equalized family geography</i>		
Parental-rank slope	−0.0516	[−0.0523, −0.0509]
<i>Family-map reassignment</i>		
Same CZ to national: top-quartile destination, p.p.	0.257	[0.245, 0.266]
Same CZ to national, $\theta_P \pm 1.96$ s.e., p.p.	0.257	[0.227, 0.283]
Same CZ to 250 km: top-quartile destination, p.p.	0.240	[0.232, 0.247]
High minus low map: expected destination log wage	0.00151	[0.00147, 0.00155]
Low-opportunity origins: top-quartile destination, p.p.	0.541	[0.530, 0.550]

Notes: Intervals are based on 200 paired resamples clustered at the root-family level. Structural parameters and assigned family maps are held fixed across resamples. The row with $\theta_P \pm 1.96$ s.e. instead re-estimates the parent block with the parental family coefficient fixed at the bounds of its 95% confidence interval, 1.32 and 2.05, and reports the resulting range of the effect.