Online Appendix

"Migration Networks and Location Decisions: Evidence from U.S. Mass Migration"

Bryan A. Stuart and Evan J. Taylor

A Derivation of Network Index

Appendix A derives the expression for the network index in equation (4).

First, recall the definition of the network index, $\Delta_{j,k} \equiv \mathbb{E}[N_{-i,j,k}|D_{i,j,k}=1] - \mathbb{E}[N_{-i,j,k}|D_{i,j,k}=0]$. Because $\mathbb{E}[N_{-i,j,k}|\cdot] = (N_j - 1)\,\mathbb{E}[D_{i',j,k}|\cdot]$ for $i' \neq i$, we can rewrite this as

$$\Delta_{j,k} = (N_j - 1)\left(\mathbb{E}[D_{i',j,k}|D_{i,j,k}=1] - \mathbb{E}[D_{i',j,k}|D_{i,j,k}=0]\right),\ i' \neq i. \tag{A.1}$$

The law of iterated expectations implies that the probability of moving from birth town j to destination k can be written

$$P_{j,k} = \mathbb{E}[D_{i',j,k}|D_{i,j,k}=1]P_{j,k} + \mathbb{E}[D_{i',j,k}|D_{i,j,k}=0](1-P_{j,k}). \tag{A.2}$$

Using the definition $\mu_{j,k} \equiv \mathbb{E}[D_{i',j,k}|D_{i,j,k}=1]$ and rearranging equation (A.2) yields

$$\mathbb{E}[D_{i',j,k}|D_{i,j,k}=0] = \frac{P_{j,k}(1-\mu_{j,k})}{1-P_{j,k}}. \tag{A.3}$$

Hence, we have

$$\mathbb{E}[D_{i',j,k}|D_{i,j,k}=1] - \mathbb{E}[D_{i',j,k}|D_{i,j,k}=0] = \mu_{j,k} - \frac{P_{j,k}(1-\mu_{j,k})}{1-P_{j,k}} \tag{A.4}$$

$$= \frac{\mu_{j,k} - P_{j,k}}{1-P_{j,k}}. \tag{A.5}$$

Substituting equation (A.5) into equation (A.1) yields

$$\Delta_{j,k} = (N_j - 1)\left(\frac{\mu_{j,k} - P_{j,k}}{1-P_{j,k}}\right). \tag{A.6}$$

Applying the law of iterated expectations to the first term of the covariance of location decisions, $C_{j,k}$, yields

$$C_{j,k} \equiv \mathbb{E}[D_{i',j,k}D_{i,j,k}] - \mathbb{E}[D_{i',j,k}]\,\mathbb{E}[D_{i,j,k}] \tag{A.7}$$

$$= \mathbb{E}[D_{i',j,k}|D_{i,j,k}=1]P_{j,k} - (P_{j,k})^2 \tag{A.8}$$

Using the definition of $\mu_{j,k}$ and rearranging yields $\mu_{j,k} - P_{j,k} = C_{j,k}/P_{j,k}$. Substituting this ex-

pression into (A.6), and noting that Assumption 1 implies that $P_{j,k} = P_{g,k}$, yields equation (4).

B Generalized Method of Moments Formulation

B.1 Basic Model

As described in the text, we can derive the destination-level network index, Δ_k, in two ways: as a weighted average of $\Delta_{j,k}$ or by assuming that for each destination $\Delta_{j,k}$ is constant across birth towns within a birth state. Both approaches lead to the same point estimate of the destination-level network index, but the latter approach allows us to use GMM to estimate standard errors.

If we assume that the network index, $\Delta_{j,k}$, is constant across birth towns within a birth state, the destination-level network index, Δ_k, can be written

$$\Delta_k = \Delta_{j,k} = \frac{C_{j,k}(N_j - 1)}{P_{j,k} - P_{j,k}^2}. \tag{A.9}$$

It is useful to rewrite this as

$$\Delta_k \left(P_{j,k} - P_{j,k}^2\right) - C_{j,k}(N_j - 1) = 0. \tag{A.10}$$

To conduct inference, we treat the birth town group as the unit of observation. Aggregating across towns within a birth town group yields

$$\Delta_k Y_{g,k} - X_{g,k} = 0, \tag{A.11}$$

where

$$X_{g,k} \equiv \sum_{j \in g} C_{j,k}(N_j - 1) \tag{A.12}$$

$$Y_{g,k} \equiv \sum_{j \in g} P_{j,k} - P_{j,k}^2. \tag{A.13}$$

In the text, we describe how we construct our estimates $\widehat{P_{j,k}}$, $\widehat{P_{j,k}^2}$, and $\widehat{C_{j,k}}$. These estimates immediately lead to estimates $\widehat{X_{g,k}}$ and $\widehat{Y_{g,k}}$, which can be written as deviations from the underlying parameters,

$$\widehat{X_{g,k}} = X_{g,k} + u_{g,k}^X \tag{A.14}$$

$$\widehat{Y_{g,k}} = Y_{g,k} + u_{g,k}^Y. \tag{A.15}$$

This allows us to rewrite equation (A.11),

$$\Delta_k \widehat{Y_{g,k}} - \widehat{X_{g,k}} + (\Delta_k u_{g,k}^Y - u_{g,k}^X) = 0. \tag{A.16}$$

Because we have unbiased estimates of $P_{j,k}$, $P_{j,k}^2$, and $C_{j,k}$, we have unbiased estimates of $X_{g,k}$

and $Y_{g,k}$. This implies that

$$\mathbb{E}\left[\Delta_k \widehat{Y_{g,k}} - \widehat{X_{g,k}}\right] = 0. \tag{A.17}$$

Equation (A.17) is the basis of our GMM estimator. The sample analog is

$$\frac{1}{G}\sum_g \left(\widehat{\Delta_k}\widehat{Y_{g,k}} - \widehat{X_{g,k}}\right) = 0, \tag{A.18}$$

where G is the number of birth town groups in a state. This can be rewritten

$$\widehat{\Delta_k} = \frac{\sum_j \widehat{C_{j,k}}(N_j - 1)}{\sum_{j'} \widehat{P_{j',k}} - \widehat{P^2_{j',k}}}. \tag{A.19}$$

Equation (A.19) is identical to equation (8).

The above derivation is for a single destination-level network index, but can easily be expanded to consider all K destination-level network index parameters. The aggregated moment condition is

$$\mathbb{E}\begin{bmatrix} \Delta_1\widehat{Y_{g,1}} - \widehat{X_{g,1}} \\ \vdots \\ \Delta_K\widehat{Y_{g,K}} - \widehat{X_{g,K}} \end{bmatrix} \equiv \mathbb{E}\left[\boldsymbol{f}(\boldsymbol{w}_g, \boldsymbol{\Delta})\right] = \mathbf{0}, \tag{A.20}$$

where $\boldsymbol{w}_g$ is observed data used to construct $\widehat{\boldsymbol{X}_{\boldsymbol{g}}}$ and $\widehat{\boldsymbol{Y}_{\boldsymbol{g}}}$ and $\boldsymbol{\Delta} \equiv (\Delta_1, \ldots, \Delta_K)'$ is a $K \times 1$ vector of destination-level network index parameters.

Under standard conditions (e.g., Cameron and Trivedi, 2005), the asymptotic distribution of $\boldsymbol{\Delta}$ is

$$\sqrt{G}(\hat{\boldsymbol{\Delta}} - \boldsymbol{\Delta}) \xrightarrow{d} \mathcal{N}\left[\mathbf{0}, \hat{\boldsymbol{F}}^{-1}\hat{\boldsymbol{S}}(\hat{\boldsymbol{F}}')^{-1}\right], \tag{A.21}$$

where

$$\hat{\boldsymbol{F}} = \frac{1}{G}\sum_g \left.\frac{\partial \boldsymbol{f}_g}{\partial \boldsymbol{\Delta}'}\right|_{\hat{\boldsymbol{\Delta}}} \tag{A.22}$$

$$= \frac{1}{G}\sum_g \begin{bmatrix} \widehat{Y_{g,1}} & 0 & 0 & \cdots & 0 \\ 0 & \widehat{Y_{g,2}} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & \widehat{Y_{g,K}} \end{bmatrix} \tag{A.23}$$

and

$$\hat{\boldsymbol{S}} = \frac{1}{G}\sum_g \boldsymbol{f}(\boldsymbol{W}_g, \hat{\boldsymbol{\Delta}})\boldsymbol{f}(\boldsymbol{W}_g, \hat{\boldsymbol{\Delta}})'. \tag{A.24}$$

While it is convenient to describe the asymptotic properties when grouping all destinations together into $\mathbf{\Delta}$, we estimate each destination-level network index parameter Δ_k independently.

B.2 Comparing Estimates from Two Models

The GMM framework facilitates a comparison of estimates from different models. Under the null hypothesis we wish to test, we have two unbiased estimates for $X_{g,k}$ and $Y_{g,k}$:

$$\widehat{X^1_{g,k}} = X_{g,k} + u^X_{g,k} \tag{A.25}$$

$$\widehat{Y^1_{g,k}} = Y_{g,k} + u^Y_{g,k} \tag{A.26}$$

$$\widehat{X^2_{g,k}} = X_{g,k} + v^X_{g,k} \tag{A.27}$$

$$\widehat{Y^2_{g,k}} = Y_{g,k} + v^Y_{g,k}. \tag{A.28}$$

We estimate the unrestricted version of the model using GMM, for which the sample analog of the moment condition is

$$\frac{1}{G}\sum_g \begin{pmatrix} \widehat{\Delta^1_k}\widehat{Y^1_{g,k}} - \widehat{X^1_{g,k}} \\ \widehat{\Delta^2_k}\widehat{Y^2_{g,k}} - \widehat{X^2_{g,k}} \end{pmatrix} \tag{A.29}$$

This simply stacks the two estimates of the destination-level network index, Δ_k into a single, exactly-identified system.

Let $\Delta^1 \equiv N^{-1}\sum_k N_k\Delta_k$ be the migrant-weighted average of the destination-level network index parameters, where $N \equiv \sum_k N_k$ is the total number of migrants from a birth state. We are interested in testing whether $\Delta^1 = \Delta^2$. To test this hypothesis, we form the test statistic

$$\hat{t} = \frac{\widehat{\Delta^1} - \widehat{\Delta^2}}{\left(\widehat{\text{Var}}[\widehat{\Delta^1} - \widehat{\Delta^2}]\right)^{1/2}}. \tag{A.30}$$

Given destination-level network index estimates $\widehat{\Delta^1_k}$ and $\widehat{\Delta^2_k}$, it is straightforward to construct the averages $\widehat{\Delta^1}$ and $\widehat{\Delta^2}$. To estimate the variance in the denominator of the test statistic, we assume that destination-level network index estimates are independent of each other. Given the large number of sending birth towns, and the large number of destinations, we believe that the covariance between two destination-level network index estimates is likely small. Furthermore, we are not confident in our ability to reliably estimate the covariance of the covariances of location decisions, as would be necessary if we did not assume independence. Under the independence assumption, we can estimate $\widehat{\text{Var}}[\widehat{\Delta^1} - \widehat{\Delta^2}]$ as the appropriately weighted sum of

$$\widehat{\text{Var}}[\widehat{\Delta^1_k} - \widehat{\Delta^2_k}] = \widehat{\text{Var}}[\widehat{\Delta^1_k}] + \widehat{\text{Var}}[\widehat{\Delta^2_k}] - 2\widehat{\text{Cov}}[\widehat{\Delta^1_k}, \widehat{\Delta^2_k}] \tag{A.31}$$

which we obtain from the GMM variance estimate.

One issue with calculating this test statistic is that, when estimating the variance of our network index estimates under the extension in Section 3.4, we ignore the variance that arises because $\widetilde{P_{j,k}}$

and $\widetilde{P^2_{j,k}}$ rely on OLS estimates. We could account for this variance using a bootstrap, but the computational cost is very high, as it takes about 48 hours to construct the estimates in Column 3 of Table 4. Not accounting for this variance means that the p-values are too low. When pooling states for Southern black migrants, the p-value for the comparison between columns 1 and 2 of Table 4 is 0.33; it is 0.44 for columns 1 and 3. For Great Plains white migrants, the p-values are both 0.00.

C Estimating Cross-Group Network Indices

When estimating cross-group network indices, we are interested in the expected increase in the number of type b people from birth town j that move to destination county k when an arbitrarily chosen person i of type w is observed to make the same move,

$$\Delta^{b|w}_{j,k} \equiv \mathbb{E}[N^b_{j,k}|D^w_{i,j,k}=1] - \mathbb{E}[N^b_{j,k}|D^w_{i,j,k}=0]. \tag{A.32}$$

The steps described in Appendix A yield

$$\Delta^{b|w}_{j,k} = \frac{C^{b,w}_{j,k} N^b_j}{P^w_{j,k}(1-P^w_{j,k})}, \tag{A.33}$$

where $C^{b,w}_{j,k}$ is the covariance of location decisions between migrants of type b and w, N^b_j is the number of type b migrants born in j, and $P^w_{j,k}$ is the probability that a migrant of type w moves from j to k.

We estimate $P^w_{j,k}$ as described in the text. To estimate $C^{b,w}_{j,k}$, consider the model

$$D^b_{i,j(i),k} D^w_{i',j(i'),k} = \alpha_{g,k} + \sum_{j\in g} \beta^{b,w}_{j,k} 1[j(i)=j(i')=j] + \epsilon_{i,i',k}. \tag{A.34}$$

This model is analogous to equation (1) in the text and yields the following covariance estimator,

$$\widehat{C^{b,w}_{j,k}} = \frac{N^b_{j,k}N^w_{j,k}}{N^b_j N^w_j} - \frac{\sum_{j\in g}\sum_{j'\neq j\in g} N^b_{j,k}N^w_{j',k}}{\sum_{j\in g}\sum_{j'\neq j\in g} N^b_j N^w_{j'}}. \tag{A.35}$$

We estimate the destination-level network index as

$$\hat{\Delta}^{b|w}_k = \sum_j \left(\frac{\widehat{P^w_{j,k}} - \widehat{(P^w_{j,k})^2}}{\sum_{j'} \widehat{P^w_{j',k}} - \widehat{(P^w_{j',k})^2}} \right) \hat{\Delta}^{b|w}_{j,k}. \tag{A.36}$$

We only estimate network indices for destinations which received at least ten black and white migrants from a given state. When calculating weighted averages of $\hat{\Delta}^{b|w}_k$, we use the number of type w individuals who moved to each destination.

D Addressing Measurement Error due to Incomplete Migration Data

Network index estimates depend on population flows observed in the Duke SSA/Medicare data, which is incomplete because some individuals die before enrolling in Medicare and some individuals' birth town information is unavailable. We first address the consequences of measurement error due to incomplete migration data under a missing at random assumption. If we observe a random sample of migration flows for each birth town-destination pair, then measurement error does not bias estimates of the covariance of location decisions, $C_{j,k}$, or moving probabilities, $P_{g,k}$. As a result, equation (4) implies that network index estimates will be attenuated because we undercount the number of migrants from each town, N_j.

More specifically, let N_j^* be the true number of migrants from birth town j that live to age 65, let α be the coverage rate, and assume that

$$N_j = \alpha N_j^*. \tag{A.37}$$

We approximate the coverage rate by dividing the number of individuals in the Duke SSA/Medicare data by the number of individuals in decennial census data who are born from 1916-1936 and survive to age 65.[1] The overall coverage rate is 64.9 percent for African Americans from the South and 82.2 percent for whites from the Great Plains (Appendix Table A.12), which implies that $N_j^* \approx 1.54 N_j$ for Southern African Americans and $N_j^* \approx 1.22 N_j$ for Great Plains whites. As an approximate measurement error correction, network index estimates should be multiplied by factors of 1.54 and 1.22 for Southern black and Great Plains white migrants. Appendix Table A.13 presents results that reflect state-specific coverage rate adjustments. The weighted average of destination-level network index estimates is 3.06 for Southern African Americans and 0.46 for Great Plains whites. Adjusting for incomplete data under a missing at random assumption increases both the magnitude of network index estimates and the black-white gap.

An alternative approach is to define N_j^* as the true number of migrants that live to a younger age, such as 40. Under this benchmark, coverage rates would be lower, and the estimates that adjust for measurement error would be larger. We do not focus on this alternative because, as described in the text, our data are best-suited for measuring long-run location decisions for individuals who survive to age 65.

Without making a missing at random (MAR) assumption, we can derive a lower bound on the network index and show that estimates of this lower bound still reveal sizable migration networks. As described in the text, the network index, $\Delta_{j,k}$, depends on the covariance of location decisions for migrants from birth town j to destination k, $C_{j,k}$, the probability of moving from birth town group g to destination k, $P_{g,k}$, and the number of migrants from town j, N_j. To focus on the key issues, suppose that we have an unbiased estimate of $P_{g,k}$ and consider the consequences of measurement error in $C_{j,k}$ and N_j. Let $\Delta_{j,k}^*$ and $C_{j,k}^*$ be the true values of the network index and covariance of location decisions. The true parameters are connected through the equation

$$\Delta_{j,k}^* = \frac{C_{j,k}^*(N_j^* - 1)}{P_{g,k} - P_{g,k}^2}. \tag{A.38}$$

[1] We use the 1990 Census to construct coverage rates for individuals born from 1916-1925 and the 2000 Census for individuals born from 1926-1935.

Using the definition of covariance, it is straightforward to show that

$$C^*_{j,k} = \alpha^2 C_{j,k} + 2\alpha(1-\alpha)C^{\text{in, out}}_{j,k} + (1-\alpha)^2 C^{\text{out, out}}_{j,k}, \tag{A.39}$$

where $C_{j,k}$ is the covariance of location decisions between migrants who are in our data, $C^{\text{in, out}}_{j,k}$ is the average covariance of location decisions between a migrant who is in our data and a migrant who is not, and $C^{\text{out, out}}_{j,k}$ is the average covariance of location decisions between migrants who are not in our data.

When not assuming that data are MAR, the covariance of location decisions among migrants not in our data ($C^{\text{in, out}}_{j,k}$ and $C^{\text{out, out}}_{j,k}$) could differ from the covariance of location decisions between migrants who are in our data ($C_{j,k}$). As a result, the network index based on our data, $\Delta_{j,k}$, might not simply be attenuated, as implied by the MAR assumption. In general, we cannot point identify the network index under this more general measurement error model. However, we can construct a lower bound for the strength of migration networks. In particular, we make the extreme assumptions that there are no interactions between migrants in and out of our data, so that $C^{\text{in, out}}_{j,k} = 0$, and that there are no interactions between migrants out of our data, so that $C^{\text{out, out}}_{j,k} = 0$. In this case, equations (A.37), (A.38), and (A.39) imply that

$$\Delta^*_{j,k} \geq \alpha \Delta_{j,k}, \tag{A.40}$$

so that we can estimate a lower bound on the true network index by multiplying the estimated network index by the coverage rate.[2] Combining the average coverage rates (64.9 and 82.2 percent) with the average destination-level network index estimates from Table 3, we estimate a lower bound for the network index of 1.26 for African Americans and 0.31 for whites. These lower bounds, which depend on extremely conservative assumptions about the migration of individuals not in our data, still reveal sizable networks, especially among African Americans.

E Differences in Family Size and the Black-White Gap

Appendix E provides a more detailed discussion of whether differences in family size explain the black-white network index gap.

To explore this issue, we decompose the network index into a component for migrants from the same family, Δ^{fam}, and a component for migrants not from the same family, Δ^{not}. To examine the importance of differences in family size, we assume that black and white network indices differ

[2]Proof: If $C^{\text{in, out}}_{j,k} = C^{\text{out, out}}_{j,k} = 0$, equations (A.37), (A.38), and (A.39) imply

$$\begin{aligned}\Delta^*_{j,k} &= \frac{\alpha^2 C_{j,k}\left(\frac{N_j}{\alpha} - 1\right)}{P_{g,k} - P^2_{g,k}} \\ &\geq \frac{\alpha^2 C_{j,k}\left(\frac{N_j}{\alpha} - \frac{1}{\alpha}\right)}{P_{g,k} - P^2_{g,k}} = \alpha \Delta_{j,k},\end{aligned}$$

where the inequality comes from noting that $\alpha \in [0,1]$ and assuming $C_{j,k} \geq 0$, and the final equality comes from equation (4) in the text. One could also construct upper bounds, but these are not particularly informative.

only because of differences in family size. Then we have

$$\Delta_b = \Delta^{fam} P_b^{fam} + \Delta^{not}(1 - P_b^{fam}) \tag{A.41}$$

$$\Delta_w = \Delta^{fam} P_w^{fam} + \Delta^{not}(1 - P_w^{fam}), \tag{A.42}$$

where Δ_b is the network index among black migrants, and P_b^{fam} is the probability that two randomly chosen black migrants are from the same family. Δ_w and P_w^{fam} are defined analogously. The black-white network index gap is

$$\Delta_b - \Delta_w = (\Delta^{fam} - \Delta^{not})(P_b^{fam} - P_w^{fam}). \tag{A.43}$$

Our data do not allow us to estimate $\Delta^{fam} - \Delta^{not}$, but we can use equation (A.43), along with estimates of $\Delta_b - \Delta_w$ and $P_b^{fam} - P_w^{fam}$ to explore whether it is reasonable to conclude that differences in family size explain the black-white gap. As described in the text, our average network indices for black and white migrants are 1.94 and 0.38. In the 1940 Census, the average within-household family size for individuals born from 1916-1936 is 6.16 for African Americans from the South and 5.25 for whites from the Great Plains. In the Duke SSA/Medicare dataset, there are 142 black migrants and 181 white migrants per town. However, as discussed in the text, the Duke data undercount the total number of migrants. If we inflate the migrant counts by 1.54 and 1.22, then we estimate 219 black migrants and 221 white migrants per town. Combining the Census family size estimates with the adjusted Duke migrant estimates, we have P_b^{fam} = 6.16/219 = 0.028 and P_w^{fam} = 5.25/221 = 0.024. With these estimates, $\Delta^{fam} - \Delta^{not}$ would have to equal 520 (= 1.56/0.003) people for differences in family size to fully explain the black-white gap. This is clearly implausible.

To construct an upper bound on the probability that two randomly chosen migrants are from the same family, we use the 100 percent sample of the 1940 Census to count the number of individuals in a county born from 1916-1936 with the same last name (Minnesota Population Center and Ancestry.com, 2013). On average, there are 54.5 African Americans with the same last name and 14.7 whites with the same last name. Using these numbers in the numerator leads to estimates of P_b^{fam} = 54.5 / 219 = 0.249 and P_w^{fam} = 14.7/221 = 0.067. In this case, $\Delta^{fam} - \Delta^{not}$ would have to equal 8.57 (=1.56/0.182) people for differences in family size to fully explain the black-white gap. This approach considerably overestimates the extent of family connections, because many individuals with the same last name are not related, and we use counties, instead of towns, as the geographic unit in the numerator of P^{fam}. Even still, this gap seems too large to us. In sum, differences in family size might explain some, but not all, of the differences in migration networks between black and white migrants.

F Calculating County-Specific Relative Wages

Appendix F provides details on how we calculate county-specific relative wages, which we use as a correlate of destination network strength.

Consider the following model for the log hourly wage of individual i,

$$\ln(w_i) = X_i\theta_r + \phi_{r,c} + \epsilon_i, \tag{A.44}$$

where X_i is a vector of observed covariates—a constant term plus indicators for detailed educational attainment (of which there are 23), age, marital status (married or not), birth state, and birth state-by-age—and $\phi_{r,c}$ is a race-specific county fixed effect. We define the black-white relative wage in county c as

$$RW_{bw,c} = \phi_{b,c} - \phi_{w,c} + \bar{X}_b(\theta_b - \theta_w) \tag{A.45}$$

where $\bar{X}_b$ is the mode of X_i among black individuals.[3] A higher value of the black-white relative wage indicates less discrimination in county c. To construct the black-white relative wage, we estimate equation (A.44) on the sample of U.S. born men age 16–64 who are a wage/salary worker and have at least 26 weeks of work in the prior year in the 1940 complete count Census data. To study discrimination against white migrants from the Great Plains (which existed, albeit less severely), we construct a similar relative wage for white men who are born in the five Great Plains states or outside the border region shown in Figure 2.

[3]We use the mode instead of the mean because the variables in X_i are categorical. This only affects the mean of the relative wage, and so our results would be identical from any other choice of $\bar{X}$.

Table A.1: Industry of Migrants and Non-Migrants, Southern Blacks and Great Plains Whites, 1950

	Percent of Group Working in Industry			
	Southern Black		Great Plains White	
	Migrants (1)	Non-Migrants (2)	Migrants (3)	Non-Migrants (4)
Agriculture, Forestry, and Fishing	1.23%	35.92%	9.38%	31.60%
Mining	1.33%	1.20%	2.02%	3.65%
Construction	10.19%	8.12%	11.98%	9.14%
Manufacturing	37.88%	22.09%	23.80%	10.98%
Transportation, Communication, and Other Utilities	11.80%	7.89%	9.58%	9.59%
Wholesale and Retail Trade	13.61%	10.46%	16.47%	16.87%
Finance, Insurance, and Real Estate	2.21%	0.78%	2.39%	2.20%
Business and Repair Services	2.99%	1.67%	4.11%	3.49%
Personal Services	6.30%	5.24%	2.16%	1.83%
Entertainment and Recreation Services	1.03%	0.63%	1.15%	0.76%
Professional and Related Services	3.95%	3.31%	5.67%	4.27%
Public Administration	6.57%	2.33%	11.08%	5.17%
Other	0.92%	0.35%	0.22%	0.43%

Note: Sample contains currently employed males, age 20-60 in the 1950 Census.
Source: Authors' calculations using Ruggles et al. (2019)

Table A.2: Size of Birth Town Groups Chosen by Cross Validation

Birth State	(1)
Panel A: Southern Black Migrants	
Alabama	52
Florida	138
Georgia	40
Louisiana	48
Mississippi	42
North Carolina	52
South Carolina	30
Panel B: Great Plains White Migrants	
Kansas	128
Nebraska	128
North Dakota	84
Oklahoma	68
South Dakota	112
Panel C: Southern White Migrants	
Alabama	156
Florida	270
Georgia	168
Louisiana	136
Mississippi	170
North Carolina	50
South Carolina	266

Notes: Table displays the results of a cross validation procedure that chooses the length of the square grid used to define birth town groups. See text for details.
Source: Authors' calculations using Duke SSA/Medicare data

Table A.3: Number of Birth Towns and Migrants, by Birth State

Birth State	Birth Towns (1)	Migrants (2)	Migrants Per Town (3)
Panel A: Southern Black Migrants			
Alabama	693	96,269	138.9
Florida	203	19,158	94.4
Georgia	566	77,038	136.1
Louisiana	460	55,974	121.7
Mississippi	660	120,454	182.5
North Carolina	586	78,420	133.8
South Carolina	461	69,399	150.5
All States	3,629	516,712	142.4
Panel B: Great Plains White Migrants			
Kansas	883	139,374	157.8
Nebraska	643	134,011	208.4
North Dakota	592	92,205	155.8
Oklahoma	966	200,392	207.4
South Dakota	474	78,541	165.7
All States	3,558	644,523	181.1

Notes: Sample limited to towns with at least 10 migrants in the data.
Source: Authors' calculations using Duke SSA/Medicare data

Table A.4: Characteristics of Migrants and Non-Migrants from Linked Census Data, African Americans from South

	1920–1930 Censuses		1930–1940 Censuses		p-value, column differences		
	Non-Migrants (1)	Migrants (2)	Non-Migrants (3)	Migrants (4)	(1) – (2) (5)	(3) – (4) (6)	(2) – (4) (7)
Individual and family characteristics							
Attending school (age 6-17)	64.6	63.9	73.8	69.8	0.10	0.00	0.00
Literate (age 10+)	71.0	77.3	80.4	83.6	0.00	0.00	0.00
Age in 1920/1930	12.9	14.4	13.5	14.8	0.00	0.00	0.00
Father's age in 1920/1930	45.7	46.7	45.5	46.9	0.00	0.00	0.06
Mother's age in 1920/1930	38.7	39.9	39.3	40.6	0.00	0.00	0.00
Parent present	87.8	84.2	85.6	80.8	0.00	0.00	0.00
Parent literate	71.3	73.8	81.5	81.9	0.00	0.26	0.00
Owner-occupied housing	23.3	25.4	25.2	29.0	0.00	0.00	0.00
Number of siblings	3.6	3.3	3.5	3.0	0.00	0.00	0.00
Father occupation: professional	1.0	1.7	1.5	2.5	0.00	0.00	0.00
Father occupation: farmer	74.4	71.2	66.8	61.8	0.00	0.00	0.00
Father occupation: clerical or sales	0.2	0.5	0.5	0.8	0.00	0.00	0.00
Father occupation: craftsmen or operative	4.8	7.4	7.0	9.5	0.00	0.00	0.00
Father occupation: laborer or service worker	11.2	12.6	16.0	18.2	0.00	0.00	0.00
Father occupation: farm laborer	8.4	6.6	8.1	7.2	0.00	0.00	0.07
Town/city characteristics:							
Not in city	81.4	74.1	75.4	68.5	0.00	0.00	0.00
City population $\leq$ 25,000	13.1	17.4	14.6	19.1	0.00	0.00	0.00
City population $>$ 25,000	5.5	8.5	10.0	12.4	0.00	0.00	0.00

1910 County characteristics:							
Percent black	53.5	53.2	52.0	52.7	0.02	0.00	0.01
Percent of farmers who are black	50.3	50.3	48.4	49.5	0.78	0.00	0.00
Percent of black farmers who are owners	25.3	24.4	26.7	25.7	0.00	0.00	0.00
Percent of white farmers who are owners	58.3	57.2	58.9	58.2	0.00	0.00	0.00
Percent of farm acres in cotton	17.0	17.7	15.9	16.9	0.00	0.00	0.00
Percent of crop value in cotton	44.4	45.8	42.1	44.3	0.00	0.00	0.00
Black literacy rate (age 10+)	62.1	63.7	62.7	64.2	0.00	0.00	0.00
White literacy rate (age 10+)	92.0	92.5	91.8	92.5	0.00	0.00	0.37
Black school attendance rate (age 6-14)	55.3	57.3	55.6	57.7	0.00	0.00	0.00
White school attendance rate (age 6-14)	75.7	75.9	75.5	75.8	0.01	0.00	0.41
Black population density	34.7	37.7	39.4	39.1	0.00	0.62	0.02
White population density	44.0	46.0	57.3	50.0	0.07	0.00	0.01
Number of individuals	109,851	19,881	139,363	15,391			

Notes: Table reports averages of indicated variables. The 1920–1930 sample contains black men age 18-30 in 1930 who are born in Alabama, Florida, Georgia, Louisiana, Mississippi, North Carolina, or South Carolina and can be matched to the 1920 Census. Migrants are individuals who live outside the former Confederate states in 1930. The 1930–1940 sample uses the same age range and definitions. Source: Authors' calculations using Minnesota Population Center and Ancestry.com (2013), Haines and ICPSR (2010)

Table A.5: Characteristics of Migrants and Non-Migrants from Linked Census Data, Whites from Great Plains

	1920–1930 Censuses		1930–1940 Censuses		p-value, column differences		
	Non-Migrants (1)	Migrants (2)	Non-Migrants (3)	Migrants (4)	(1) – (2) (5)	(3) – (4) (6)	(2) – (4) (7)
Individual and family characteristics							
Attending school (age 6-17)	86.0	85.4	89.0	87.9	0.04	0.00	0.00
Literate (age 10+)	99.3	99.6	99.5	99.6	0.00	0.05	0.66
Age in 1920/1930	12.9	13.8	13.5	14.3	0.00	0.00	0.00
Father's age in 1920/1930	46.2	47.0	46.1	46.5	0.00	0.00	0.00
Mother's age in 1920/1930	41.0	41.9	41.5	41.9	0.00	0.00	0.44
Parent present	96.4	93.9	95.2	92.7	0.00	0.00	0.00
Parent literate	98.0	97.8	98.2	98.2	0.03	0.42	0.00
Owner-occupied housing	67.1	62.6	57.5	51.5	0.00	0.00	0.00
Number of siblings	3.4	3.0	3.0	2.9	0.00	0.00	0.00
Father occupation: professional	8.7	14.5	10.1	11.5	0.00	0.00	0.00
Father occupation: farmer	68.5	52.2	59.9	53.4	0.00	0.00	0.03
Father occupation: clerical or sales	3.9	6.2	5.3	6.5	0.00	0.00	0.33
Father occupation: craftsmen or operative	10.9	17.8	13.9	16.4	0.00	0.00	0.00
Father occupation: laborer or service worker	5.3	6.4	7.9	8.7	0.00	0.00	0.00
Father occupation: farm laborer	2.6	2.7	2.9	3.4	0.26	0.00	0.00
Town/city characteristics:							
Not in city	64.8	48.9	61.1	55.4	0.00	0.00	0.00
City population $\leq$ 25,000	28.4	41.7	28.8	34.7	0.00	0.00	0.00
City population $>$ 25,000	6.7	9.4	10.1	9.8	0.00	0.09	0.12

1910 County characteristics:							
Percent black	2.7	2.9	3.0	3.2	0.00	0.00	0.00
Percent of white farmers who are owners	58.7	59.1	58.9	58.5	0.00	0.00	0.00
Percent of farm acres in cotton	2.0	1.9	2.1	2.5	0.34	0.00	0.00
Percent of crop value in cotton	6.3	6.3	6.7	8.1	0.81	0.00	0.00
White literacy rate (age 10+)	97.7	97.6	97.7	97.6	0.00	0.00	0.92
White school attendance rate (age 6-14)	86.2	86.0	85.8	85.5	0.00	0.00	0.00
Black population density	2.6	2.7	2.8	2.4	0.05	0.00	0.00
White population density	45.0	50.1	47.2	41.6	0.00	0.00	0.00
Number of individuals	188,700	19,620	215,457	34,893			

Notes: Table reports averages of indicated variables. The 1920–1930 sample contains white men age 18–30 in 1930 who are born in Kansas, Nebraska, Oklahoma, North Dakota, or South Dakota and can be matched to the 1920 Census. Migrants are individuals who live outside these states plus the light grey border region in Figure 2. The 1930–1940 sample uses the same age range and definitions.
Source: Authors' calculations using Minnesota Population Center and Ancestry.com (2013), Haines and ICPSR (2010)

Table A.6: Characteristics of Migrants and Non-Migrants from Linked Census Data, Whites from South

	1920–1930 Censuses		1930–1940 Censuses		p-value, column differences		
	Non-Migrants (1)	Migrants (2)	Non-Migrants (3)	Migrants (4)	(1) – (2) (5)	(3) – (4) (6)	(2) – (4) (7)
Individual and family characteristics							
Attending school (age 6-17)	82.0	82.2	84.5	85.1	0.68	0.12	0.00
Literate (age 10+)	95.0	97.3	97.0	98.3	0.00	0.00	0.00
Age in 1920/1930	13.1	14.2	13.6	14.4	0.00	0.00	0.00
Father's age in 1920/1930	45.5	46.8	45.8	46.8	0.00	0.00	0.95
Mother's age in 1920/1930	40.4	41.3	40.9	41.6	0.00	0.00	0.01
Parent present	94.9	92.5	93.8	91.4	0.00	0.00	0.00
Parent literate	93.5	95.3	95.4	96.2	0.00	0.00	0.00
Owner-occupied housing	56.6	58.4	49.7	51.8	0.00	0.00	0.00
Number of siblings	3.6	3.4	3.3	3.0	0.00	0.00	0.00
Father occupation: professional	9.0	12.9	10.7	15.5	0.00	0.00	0.00
Father occupation: farmer	66.0	57.1	54.8	46.6	0.00	0.00	0.00
Father occupation: clerical or sales	4.1	5.8	6.1	8.5	0.00	0.00	0.00
Father occupation: craftsmen or operative	12.8	16.8	18.1	20.3	0.00	0.00	0.00
Father occupation: laborer or service worker	5.3	5.1	7.3	7.0	0.43	0.34	0.00
Father occupation: farm laborer	2.8	2.3	3.0	2.1	0.01	0.00	0.25
Town/city characteristics:							
Not in city	70.1	60.5	66.0	56.5	0.00	0.00	0.00
City population $\leq$ 25,000	20.7	26.6	21.2	26.1	0.00	0.00	0.33
City population $>$ 25,000	9.2	12.9	12.8	17.5	0.00	0.00	0.00

1910 County characteristics:							
Percent black	35.7	35.1	35.7	35.8	0.00	0.88	0.00
Percent of farmers who are black	30.1	29.6	29.8	30.0	0.01	0.49	0.23
Percent of black farmers who are owners	33.7	36.1	34.6	36.7	0.00	0.00	0.04
Percent of white farmers who are owners	60.7	61.8	61.2	62.6	0.00	0.00	0.00
Percent of farm acres in cotton	11.6	11.0	11.1	10.9	0.00	0.00	0.41
Percent of crop value in cotton	34.8	32.6	33.5	32.4	0.00	0.00	0.45
Black literacy rate (age 10+)	65.3	66.8	65.7	67.0	0.00	0.00	0.06
White literacy rate (age 10+)	89.9	90.6	90.0	90.9	0.00	0.00	0.00
Black school attendance rate (age 6-14)	56.0	57.8	56.2	57.9	0.00	0.00	0.34
White school attendance rate (age 6-14)	74.3	74.9	74.1	74.8	0.00	0.00	0.53
Black population density	37.7	39.7	39.4	38.7	0.01	0.36	0.29
White population density	80.4	81.4	83.8	77.4	0.62	0.00	0.13
Number of individuals	293,678	14,167	355,197	12,004			

Notes: Table reports averages of indicated variables. The 1920–1930 sample contains white men age 18–30 in 1930 who are born in Alabama, Florida, Georgia, Louisiana, Mississippi, North Carolina, or South Carolina and can be matched to the 1920 Census. Migrants are individuals who live outside the former Confederate states in 1930. The 1930–1940 sample uses the same age range and definitions.
Source: Authors' calculations using Minnesota Population Center and Ancestry.com (2013), Haines and ICPSR (2010)

Table A.7: Average Network Index Estimates, Southern White Migrants

Birth State	Number of Migrants (1)	Unweighted Average (2)	Weighted Average (3)
Alabama	43,157	0.204	0.516
		(0.014)	(0.052)
Florida	27,426	0.046	0.072
		(0.006)	(0.100)
Georgia	31,299	0.082	0.117
		(0.007)	(0.021)
Louisiana	31,303	0.122	0.269
		(0.011)	(0.071)
Mississippi	28,001	0.118	0.186
		(0.010)	(0.021)
North Carolina	47,146	0.179	0.412
		(0.012)	(0.040)
South Carolina	14,605	0.068	0.094
		(0.005)	(0.029)
All States	222,937	0.131	0.280
		(0.004)	(0.021)

Notes: Column 2 is an unweighted average of destination-level network index estimates, $\hat{\Delta}_k$. Column 3 is a weighted average, where the weights are the number of people who move from each state to destination k. Birth town groups are defined by cross validation. Standard errors in parentheses.
Source: Authors' calculations using Duke SSA/Medicare data

Table A.8: Average Network Index Estimates, By Size of Birth Town and Destination, Southern White Migrants

Exclude Largest Birth Towns:	No	Yes	No	Yes
Exclude Largest Destinations:	No	No	Yes	Yes
Birth State	(1)	(2)	(3)	(4)
Alabama	0.516	0.456	0.531	0.479
	(0.052)	(0.045)	(0.071)	(0.062)
Florida	0.072	0.071	0.134	0.034
	(0.100)	(0.013)	(0.082)	(0.009)
Georgia	0.117	0.104	0.119	0.095
	(0.021)	(0.012)	(0.019)	(0.015)
Louisiana	0.269	0.208	0.198	0.150
	(0.071)	(0.022)	(0.035)	(0.017)
Mississippi	0.186	0.184	0.135	0.133
	(0.021)	(0.021)	(0.013)	(0.013)
North Carolina	0.412	0.397	0.337	0.317
	(0.040)	(0.036)	(0.040)	(0.034)
South Carolina	0.094	0.088	0.058	0.054
	(0.029)	(0.022)	(0.013)	(0.012)
All States	0.280	0.253	0.262	0.225
	(0.021)	(0.013)	(0.021)	(0.015)

Notes: All columns contain weighted averages of destination-level network index estimates, $\hat{\Delta}_k$, where the weights are the number of people who move from each state to destination k. Column 1 includes all birth towns and destinations. Column 2 excludes birth towns with 1920 population greater than 20,000 when estimating each $\hat{\Delta}_k$. Column 3 excludes all destination counties which intersect in 2000 with the ten largest non-South CMSAs as of 1950: New York, Chicago, Los Angeles, Philadelphia, Boston, Detroit, Washington D.C., San Francisco, Pittsburgh, and St. Louis, in addition to counties which received fewer than 10 migrants. Column 4 excludes large birth towns and large destinations. Birth town groups are defined by cross validation. Standard errors are in parentheses.
Source: Authors' calculations using Duke SSA/Medicare data

Table A.9: Average Cross-Race Network Index Estimates, Southern Black and White Migrants

Birth State	All Counties (1)	Excluding Largest CMSAs (2)
Panel A: Black Migrants Induced to Location by White Migrant		
Alabama	0.188	0.130
	(0.106)	(0.150)
Florida	0.026	0.005
	(0.059)	(0.036)
Georgia	-0.028	0.040
	(0.039)	(0.044)
Louisiana	-0.066	0.068
	(0.196)	(0.038)
Mississippi	0.246	0.049
	(0.185)	(0.033)
North Carolina	-0.010	-0.005
	(0.062)	(0.011)
South Carolina	0.197	-0.025
	(0.161)	(0.027)
All States	0.071	0.050
	(0.048)	(0.033)
Panel B: White Migrants Induced to Location by Black Migrant		
Alabama	0.052	0.038
	(0.048)	(0.042)
Florida	0.047	-0.018
	(0.064)	(0.036)
Georgia	-0.020	0.004
	(0.014)	(0.014)
Louisiana	-0.137	0.016
	(0.066)	(0.017)
Mississippi	-0.056	0.020
	(0.030)	(0.011)
North Carolina	0.021	-0.002
	(0.029)	(0.022)
South Carolina	-0.019	0.020
	(0.013)	(0.018)
All States	-0.019	0.019
	(0.015)	(0.013)

Notes: Table A.9 contains weighted averages of cross-race destination-level network index estimates. Birth town groups are defined by cross validation. Standard errors in parentheses.
Source: Authors' calculations using Duke SSA/Medicare data

Table A.10: Average Network Index Estimates, Birth Town Groups Defined by Cross Validation and Counties

	Cross Validation		Counties	
Birth State	Unweighted (1)	Weighted (2)	Unweighted (3)	Weighted (4)
Panel A: Southern Black Migrants				
Alabama	0.770	1.888	0.616	1.393
	(0.049)	(0.195)	(0.034)	(0.170)
Florida	0.536	0.813	0.597	0.811
	(0.052)	(0.117)	(0.087)	(0.317)
Georgia	0.735	1.657	0.544	0.887
	(0.048)	(0.177)	(0.039)	(0.279)
Louisiana	0.462	1.723	0.399	2.209
	(0.039)	(0.478)	(0.039)	(0.920)
Mississippi	0.901	2.303	0.742	2.166
	(0.050)	(0.313)	(0.051)	(0.401)
North Carolina	0.566	1.539	0.402	1.022
	(0.039)	(0.130)	(0.028)	(0.123)
South Carolina	0.874	2.618	0.774	2.132
	(0.054)	(0.301)	(0.049)	(0.224)
All States	0.736	1.938	0.599	1.608
	(0.020)	(0.110)	(0.017)	(0.151)
Panel B: Great Plains White Migrants				
Kansas	0.128	0.255	0.106	0.194
	(0.007)	(0.024)	(0.008)	(0.028)
Nebraska	0.141	0.361	0.121	0.399
	(0.008)	(0.082)	(0.009)	(0.117)
North Dakota	0.174	0.464	0.156	0.385
	(0.012)	(0.036)	(0.010)	(0.029)
Oklahoma	0.112	0.453	0.102	0.372
	(0.008)	(0.036)	(0.007)	(0.036)
South Dakota	0.163	0.350	0.135	0.273
	(0.009)	(0.026)	(0.008)	(0.027)
All States	0.137	0.380	0.119	0.329
	(0.004)	(0.022)	(0.004)	(0.028)

Notes: Columns 1 and 3 are unweighted averages of destination-level network index estimates, $\hat{\Delta}_k$. Columns 2 and 4 are weighted averages, where the weights are the number of people who move from each state to destination k. In columns 1 and 2, we define birth town groups using cross validation, as described in the text. In columns 3 and 4, we use counties. Standard errors in parentheses.
Source: Authors' calculations using Duke SSA/Medicare data

Table A.11: Average Network Index Estimates, Birth Town Groups Based on Different Grid Sizes

	Weighted Average			Unweighted Average		
Grid Size:	50	100	200	50	100	200
Birth State	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Southern Black Migrants						
Alabama	1.869	2.256	2.398	0.759	0.846	0.913
	(0.203)	(0.198)	(0.196)	(0.046)	(0.046)	(0.045)
Florida	0.919	0.856	0.944	0.595	0.553	0.560
	(0.196)	(0.117)	(0.117)	(0.158)	(0.087)	(0.055)
Georgia	1.760	2.190	2.421	0.780	0.859	0.916
	(0.163)	(0.185)	(0.168)	(0.055)	(0.053)	(0.049)
Louisiana	1.887	2.097	2.660	0.469	0.508	0.549
	(0.542)	(0.507)	(0.717)	(0.038)	(0.034)	(0.035)
Mississippi	2.432	2.778	3.216	0.910	1.001	1.056
	(0.327)	(0.270)	(0.217)	(0.049)	(0.048)	(0.042)
North Carolina	1.557	1.719	1.877	0.566	0.629	0.678
	(0.133)	(0.149)	(0.139)	(0.041)	(0.043)	(0.037)
South Carolina	3.255	3.620	4.080	0.982	1.074	1.156
	(0.380)	(0.348)	(0.280)	(0.054)	(0.052)	(0.045)
All States	2.090	2.401	2.713	0.761	0.834	0.891
	(0.120)	(0.109)	(0.112)	(0.020)	(0.019)	(0.017)
Panel B: Great Plains White Migrants						
Kansas	0.256	0.256	0.253	0.122	0.127	0.130
	(0.028)	(0.026)	(0.019)	(0.010)	(0.008)	(0.006)
Nebraska	0.366	0.373	0.379	0.130	0.142	0.146
	(0.090)	(0.079)	(0.062)	(0.008)	(0.008)	(0.008)
North Dakota	0.424	0.490	0.529	0.164	0.177	0.186
	(0.032)	(0.037)	(0.038)	(0.011)	(0.011)	(0.011)
Oklahoma	0.425	0.488	0.514	0.107	0.115	0.119
	(0.038)	(0.038)	(0.034)	(0.008)	(0.008)	(0.007)
South Dakota	0.291	0.343	0.365	0.149	0.162	0.169
	(0.024)	(0.024)	(0.026)	(0.011)	(0.010)	(0.009)
All States	0.360	0.396	0.413	0.128	0.138	0.143
	(0.024)	(0.022)	(0.018)	(0.004)	(0.004)	(0.004)

Notes: Columns 1–3 are weighted averages of destination-level network index estimates, $\widehat{\Delta}_k$, where the weights are the number of people who move from each state to destination k. Columns 4–6 are unweighted averages. We define birth town groups as square grids, with the length of each square varying from 50 to 200 miles. Standard errors in parentheses.
Source: Authors' calculations using Duke SSA/Medicare data

Table A.12: Coverage Rates, Duke SSA/Medicare Dataset

Sample:	All	All	All	Men	Women	Cohort 1916–25	Cohort 1926–36
Birth State	Duke/SSA coverage rate, all (1)	Duke/SSA percent with town identified (2)	Duke/SSA coverage rate, town identified (3)	Duke/SSA coverage rate, town identified (4)	Duke/SSA coverage rate, town identified (5)	Duke/SSA coverage rate, town identified (6)	Duke/SSA coverage rate, town identified (7)
Panel A: Southern Black Individuals							
Alabama	86.4%	78.7%	68.0%	73.1%	64.6%	65.1%	70.4%
Florida	82.7%	83.6%	69.2%	72.3%	66.9%	65.5%	72.1%
Georgia	85.0%	73.1%	62.2%	65.2%	60.2%	57.0%	67.5%
Louisiana	85.2%	84.5%	72.0%	74.3%	70.3%	67.5%	76.0%
Mississippi	88.9%	74.7%	66.4%	69.7%	64.1%	63.9%	68.6%
North Carolina	88.5%	72.5%	64.2%	64.6%	63.9%	61.5%	66.5%
South Carolina	90.8%	61.9%	56.2%	57.3%	55.5%	53.6%	58.7%
All States	87.2%	74.4%	64.9%	67.6%	63.1%	61.3%	68.1%
Panel B: Great Plains White Individuals							
Kansas	88.1%	92.5%	81.5%	84.8%	78.6%	78.6%	84.4%
Nebraska	89.2%	93.3%	83.2%	87.5%	79.6%	80.8%	85.7%
North Dakota	88.1%	89.7%	79.0%	81.9%	76.7%	74.3%	84.0%
Oklahoma	93.1%	89.9%	83.7%	86.0%	81.8%	79.4%	87.7%
South Dakota	88.9%	91.2%	81.1%	82.6%	79.8%	78.7%	83.5%
All States	90.1%	91.3%	82.2%	85.2%	79.8%	78.8%	85.6%

Notes: Column 1 reports the number of individuals in the Duke SSA/Medicare dataset divided by the number of individuals in the 1990/2000 Census. Column 2 reports the share of individuals in the Duke SSA/Medicare dataset for whom birth town and destination county are identified. Columns 3–7 reports the number of individuals in the Duke SSA/Medicare dataset for whom birth town and destination county are identified divided by the number of individuals in the 1990/2000 Census. We use the 1990 Census for individuals born from 1916–1925 and the 2000 Census for individuals born from 1926–1936. The sample includes individuals living inside and outside their birth region.

Source: Authors' calculations using Duke SSA/Medicare data and Ruggles et al. (2019)

Table A.13: Average Network Index Estimates, Adjusted for Incomplete Migration Data

Sample: Birth State	All (1)	Men (2)	Women (3)	1916–25 Cohort (4)	1926–36 Cohort (5)
Panel A: Southern Black Migrants					
Alabama	2.775	1.160	1.621	1.273	1.671
	(0.286)	(0.120)	(0.175)	(0.144)	(0.165)
Florida	1.175	0.533	0.633	0.454	0.780
	(0.170)	(0.085)	(0.128)	(0.102)	(0.123)
Georgia	2.665	0.959	1.722	1.569	1.287
	(0.284)	(0.111)	(0.205)	(0.213)	(0.118)
Louisiana	2.394	1.184	0.992	0.973	1.676
	(0.664)	(0.316)	(0.389)	(0.225)	(0.573)
Mississippi	3.467	1.457	2.042	1.397	2.237
	(0.471)	(0.202)	(0.297)	(0.211)	(0.307)
North Carolina	2.398	1.023	1.403	1.216	1.326
	(0.203)	(0.099)	(0.121)	(0.115)	(0.109)
South Carolina	4.659	1.936	2.762	2.361	2.480
	(0.535)	(0.199)	(0.381)	(0.310)	(0.260)
All States	3.057	1.271	1.790	1.463	1.792
	(0.167)	(0.071)	(0.109)	(0.086)	(0.108)
Panel B: Great Plains White Migrants					
Kansas	0.313	0.151	0.178	0.204	0.167
	(0.029)	(0.016)	(0.017)	(0.020)	(0.014)
Nebraska	0.433	0.176	0.256	0.270	0.234
	(0.098)	(0.050)	(0.050)	(0.057)	(0.046)
North Dakota	0.587	0.250	0.338	0.387	0.277
	(0.046)	(0.021)	(0.031)	(0.032)	(0.021)
Oklahoma	0.541	0.250	0.291	0.285	0.319
	(0.043)	(0.023)	(0.023)	(0.025)	(0.026)
South Dakota	0.431	0.187	0.248	0.266	0.225
	(0.032)	(0.017)	(0.021)	(0.022)	(0.018)
All States	0.463	0.205	0.262	0.278	0.252
	(0.026)	(0.014)	(0.014)	(0.016)	(0.013)

Notes: Table A.13 reports weighted averages of destination-level network index estimates, adjusted for incomplete migration data using the coverage rates in Appendix Table A.12. Birth town groups are defined by cross validation. Standard errors in parentheses.
Source: Authors' calculations using Duke SSA/Medicare data and Ruggles et al. (2019)

Table A.14: Summary Statistics, Destination County Characteristics

Variable	Mean	S.D.	N
Panel A: Southern Black Migrants			
Network index estimate, $\widehat{\Delta}_k$	0.722	1.358	1515
Manufacturing employment share, 1910	0.178	0.112	1515
Agriculture employment share, 1910	0.242	0.173	1515
Log distance from birth state	6.688	0.517	1515
Direct railroad connection from birth state	0.092	0.290	1515
One-stop railroad connection from birth state	0.547	0.498	1515
Log population, 1910	11.230	1.155	1515
Percent African-American, 1910	0.043	0.077	1515
Percent rural, 1910	0.464	0.296	1515
Black church members per capita, 1916	0.190	0.195	1515
Small destination indicator	0.615	0.487	1515
Black-white relative wage, 1940	-0.132	0.194	1408
Logan-Parman segregation measure, 1940	0.493	0.213	1408
Panel B: Great Plains White Migrants			
Network index estimate, $\widehat{\Delta}_k$	0.140	0.438	4104
Manufacturing employment share, 1910	0.123	0.101	4104
Agriculture employment share, 1910	0.413	0.215	4104
Log distance from birth state	6.799	0.353	4104
Direct railroad connection from birth state	0.107	0.309	4104
One-stop railroad connection from birth state	0.486	0.500	4104
Log population, 1910	10.262	1.059	4104
Percent African-American, 1910	0.117	0.188	4104
Percent rural, 1910	0.701	0.284	4104
White church members per capita, 1916	0.422	0.185	4104
Small destination indicator	0.858	0.349	4104
Plains-not Plains relative wage, 1940	0.044	0.178	2311
Panel C: Southern White Migrants			
Network index estimate, $\widehat{\Delta}_k$	0.127	0.552	3357
Manufacturing employment share, 1910	0.139	0.111	3357
Agriculture employment share, 1910	0.340	0.198	3357
Log distance from birth state	6.765	0.594	3357
Direct railroad connection from birth state	0.083	0.276	3357
One-stop railroad connection from birth state	0.481	0.500	3357
Log population, 1910	10.596	1.137	3357
Percent African-American, 1910	0.037	0.073	3357
Percent rural, 1910	0.598	0.293	3357
White church members per capita, 1916	0.413	0.165	3357
Small destination indicator	0.766	0.424	3357

Notes: The unit of observation is a birth state-destination county pair. Sample includes destination counties for which we estimate a network index.

Source: Authors' calculations using Duke SSA/Medicare data, Black et al. (2015) data, Census (1992), Haines and ICPSR (2010), Logan and Parman (2017) data, Minnesota Population Center and Ancestry.com (2013), Ruggles et al. (2019)

Table A.15: Network Index Estimates and Destination County Characteristics, Birth Town Groups Defined by Counties

	Dependent variable: Destination-level network index estimate					
	Southern Black Migrants			Great Plains White Migrants		
	(1)	(2)	(3)	(4)	(5)	(6)
Manufacturing employment share, 1910	1.629	-0.091	-0.139	-0.068	-0.381	-0.383
	(0.731)	(0.520)	(0.519)	(0.091)	(0.166)	(0.165)
Manufacturing employment share by small destination indicator		3.036	3.024		0.415	0.421
		(1.224)	(1.168)		(0.190)	(0.189)
Agriculture employment share, 1910	0.087	-0.563	-0.610	0.055	0.077	0.073
	(0.255)	(0.401)	(0.404)	(0.036)	(0.125)	(0.125)
Agriculture employment share by small destination indicator		0.964	0.902		-0.010	-0.005
		(0.497)	(0.509)		(0.128)	(0.128)
Small destination indicator		-0.668	-0.661		-0.031	-0.035
		(0.294)	(0.287)		(0.069)	(0.069)
Log distance from birth state	-0.325	-0.288	-0.317	0.041	0.047	0.039
	(0.067)	(0.077)	(0.061)	(0.026)	(0.027)	(0.028)
Direct railroad connection from birth state	0.348	0.352	0.355	0.159	0.155	0.146
	(0.114)	(0.115)	(0.140)	(0.034)	(0.034)	(0.035)
One-stop railroad connection from birth state	0.227	0.216	0.189	0.058	0.052	0.051
	(0.092)	(0.088)	(0.097)	(0.014)	(0.013)	(0.013)
Log population, 1910	0.055	0.051	0.049	0.022	0.028	0.027
	(0.059)	(0.060)	(0.066)	(0.009)	(0.009)	(0.009)
Percent African-American, 1910	-1.566	-1.362	-1.384	-0.185	-0.201	-0.198
	(0.331)	(0.362)	(0.320)	(0.031)	(0.032)	(0.032)
Percent rural, 1910	-0.265	-0.264	-0.249	-0.033	-0.034	-0.034
	(0.195)	(0.198)	(0.215)	(0.036)	(0.037)	(0.037)
Black/white church members per capita, 1916	-0.400	-0.300	-0.315	-0.073	-0.064	-0.064
	(0.153)	(0.154)	(0.163)	(0.033)	(0.032)	(0.032)
Birth state fixed effects			x			x
R-squared	0.057	0.066	0.074	0.031	0.034	0.034
N (birth state-destination county pairs)	1,515	1,515	1,515	4,104	4,104	4,104
Destination counties	382	382	382	1230	1230	1230

Notes: See notes to Table 6. Birth town groups are defined by counties.
Source: Authors' calculations using Duke SSA/Medicare data, Black et al. (2015) data, Census (1992), Haines and ICPSR (2010), Ruggles et al. (2019)

Table A.16: Network Index Estimates and Destination County Characteristics, Southern White Migrants

Dependent variable: Destination-level network index estimate			
	(1)	(2)	(3)
Manufacturing employment share, 1910	0.465	-0.015	0.020
	(0.186)	(0.155)	(0.158)
Manufacturing employment share by small destination indicator		0.700	0.706
		(0.317)	(0.319)
Agriculture employment share, 1910	0.068	0.191	0.226
	(0.054)	(0.158)	(0.159)
Agriculture employment share by small destination indicator		-0.081	-0.099
		(0.169)	(0.167)
Small destination indicator		-0.115	-0.113
		(0.076)	(0.075)
Log distance from birth state	-0.033	-0.036	-0.004
	(0.018)	(0.018)	(0.018)
Direct railroad connection from birth state	0.058	0.060	0.074
	(0.038)	(0.038)	(0.037)
One-stop railroad connection from birth state	0.054	0.047	0.055
	(0.021)	(0.020)	(0.020)
Log population, 1910	0.016	0.015	0.022
	(0.016)	(0.016)	(0.015)
Percent African-American, 1910	-0.247	-0.327	-0.249
	(0.102)	(0.099)	(0.097)
Percent rural, 1910	0.043	0.012	0.010
	(0.063)	(0.064)	(0.063)
White church members per capita, 1916	-0.083	-0.089	-0.075
	(0.046)	(0.049)	(0.049)
Birth state fixed effects			x
R-squared	0.012	0.017	0.027
N (birth state-destination county pairs)	3,357	3,357	3,357
Destination counties	784	784	784

Notes: See notes to Table 6.
Source: Authors' calculations using Duke SSA/Medicare data, Black et al. (2015) data, Census (1992), Haines and ICPSR (2010), Ruggles et al. (2019)

Table A.17: Network Index Estimates and Destination County Characteristics, Additional Explanatory Variables

	Dependent variable: Destination-level network index estimate					
	Southern Black Migrants			Great Plains White Migrants		
	(1)	(2)	(3)	(4)	(5)	(6)
Manufacturing employment share, 1910	1.670	0.289	0.228	-0.085	-0.281	-0.283
	(0.585)	(0.703)	(0.709)	(0.129)	(0.162)	(0.161)
Manufacturing employment share by small destination indicator		2.539	2.601		0.283	0.288
		(1.010)	(1.001)		(0.230)	(0.228)
Agriculture employment share, 1910	0.322	-0.181	-0.248	0.121	0.224	0.217
	(0.331)	(0.488)	(0.491)	(0.080)	(0.148)	(0.146)
Agriculture employment share by small destination indicator		0.773	0.716		-0.109	-0.103
		(0.506)	(0.516)		(0.141)	(0.140)
Small destination indicator		-0.543	-0.552		0.039	0.035
		(0.273)	(0.272)		(0.076)	(0.075)
Log distance from birth state	-0.408	-0.370	-0.363	0.069	0.084	0.076
	(0.075)	(0.079)	(0.078)	(0.048)	(0.051)	(0.056)
Direct railroad connection from birth state	0.316	0.321	0.302	0.246	0.249	0.245
	(0.123)	(0.124)	(0.142)	(0.044)	(0.046)	(0.050)
One-stop railroad connection from birth state	0.215	0.204	0.160	0.103	0.097	0.098
	(0.082)	(0.080)	(0.084)	(0.024)	(0.024)	(0.023)
Log population, 1910	0.058	0.058	0.067	0.015	0.028	0.027
	(0.058)	(0.061)	(0.062)	(0.014)	(0.014)	(0.014)
Percent African-American, 1910	-2.193	-2.006	-1.923	-0.224	-0.242	-0.240
	(0.446)	(0.429)	(0.432)	(0.040)	(0.043)	(0.043)
Percent rural, 1910	-0.361	-0.345	-0.298	-0.090	-0.077	-0.076
	(0.209)	(0.227)	(0.231)	(0.057)	(0.054)	(0.054)
Black/white church members per capita, 1916	-0.438	-0.357	-0.338	-0.144	-0.126	-0.127
	(0.191)	(0.201)	(0.202)	(0.051)	(0.050)	(0.050)
Black-white / Plains-not Plains Relative wage, 1940	0.401	0.391	0.402	-0.120	-0.115	-0.115
	(0.183)	(0.187)	(0.191)	(0.043)	(0.043)	(0.042)
Logan-Parman segregation measure, 1940	0.274	0.281	0.276			
	(0.222)	(0.214)	(0.217)			
Birth state fixed effects			x			x
R-squared	0.102	0.109	0.123	0.048	0.052	0.053
N (birth state-destination county pairs)	1,408	1,408	1,408	2,311	2,311	2,311
Destination counties	335	335	335	642	642	642

Notes: See notes to Table 6. Columns 1–3 contain the black-white relative wage, and columns 4–6 contain the Plains-not Plains relative wage. See Appendix F for details on how these are constructed. Standard errors, clustered by destination county, are in parentheses.

Source: Authors' calculations using Duke SSA/Medicare data, Black et al. (2015) data, Census (1992), Haines and ICPSR (2010), Logan and Parman (2017) data, Minnesota Population Center and Ancestry.com (2013), Ruggles et al. (2019)

Table A.18: Summary Statistics, Birth County Characteristics

Variable	Mean	S.D.	N
Panel A: Southern Black Migrants			
Network index estimate, $\widehat{\Delta_c}$	1.729	3.555	546
Percent of black farmers who are owners, 1910	0.324	0.254	546
Percent of black individuals in owner-occupied housing, 1910	0.264	0.145	546
Percent of black workers in agriculture, 1910	0.665	0.224	546
Percent of black workers in manufacturing, 1910	0.071	0.083	546
Percent of farm acreage in cotton, 1910	0.135	0.106	546
Log black population density, 1910	2.581	1.005	546
Black church members per capita, 1916	0.391	0.188	546
Rosenwald school exposure	0.202	0.214	546
Black literacy rate (10+), 1910	0.632	0.100	546
Black school attendance rate (6–14), 1910	0.544	0.131	546
Railroad exposure	0.546	0.404	546
Percent African-American, 1910	0.443	0.213	546
Percent rural, 1910	0.899	0.168	546
Percent voting for Strom Thurmond, 1948	0.457	0.335	546
Panel B: Great Plains White Migrants			
Network index estimate, $\widehat{\Delta_c}$	0.354	0.642	383
Percent of white farmers who are owners, 1910	0.646	0.212	383
Percent of white individuals in owner-occupied housing, 1910	0.642	0.164	383
Percent of white workers in agriculture, 1910	0.643	0.140	383
Percent of white workers in manufacturing, 1910	0.021	0.028	383
Percent of farm acreage in cotton, 1910	0.016	0.045	383
Log white population density, 1910	2.539	0.926	383
White church members per capita, 1916	0.307	0.116	383
White literacy rate (10+), 1910	0.978	0.021	383
White school attendance rate (6–14), 1910	0.851	0.072	383
Railroad exposure	0.539	0.391	383
Percent African-American, 1910	0.020	0.05	383
Percent rural, 1910	0.892	0.177	383
Panel C: Southern White Migrants			
Network index estimate, $\widehat{\Delta_c}$	0.212	0.772	576
Percent of white farmers who are owners, 1910	0.618	0.147	576
Percent of white individuals in owner-occupied housing, 1910	0.554	0.107	576
Percent of white workers in agriculture, 1910	0.645	0.185	576
Percent of white workers in manufacturing, 1910	0.084	0.074	576
Percent of farm acreage in cotton, 1910	0.127	0.106	576
Log white population density, 1910	2.862	0.802	576
White church members per capita, 1916	0.469	0.197	576
Rosenwald school exposure	0.198	0.218	576
White literacy rate (10+), 1910	0.907	0.064	576
White school attendance rate (6–14), 1910	0.745	0.085	576
Railroad exposure	0.533	0.415	576
Percent African-American, 1910	0.424	0.219	576
Percent rural, 1910	0.904	0.165	576
Percent voting for Strom Thurmond, 1948	0.437	0.335	576

Notes: Sample includes birth counties containing at least one town with at least 10 migrants in the Duke data. Railroad exposure is the share of migrants in a county that lived along a railroad. Rosenwald school exposure is the average Rosenwald coverage experienced over ages 7–13.
Source: Authors' calculations using Duke SSA/Medicare data, Aaronson and Mazumder (2011) data, Black et al. (2015) data, Census (1992), Haines and ICPSR (2010), ICPSR (1999), Ruggles et al. (2019)

Table A.19: Network Index Estimates and Birth County Characteristics, Southern White Migrants

Dependent variable: Birth county-level network index estimate	(1)	(2)
Percent of white farmers who are owners, 1910	-1.587	-1.514
	(0.764)	(0.725)
Percent of white individuals in owner-occupied housing, 1910	1.907	1.966
	(0.988)	(0.958)
Percent of white workers in agriculture, 1910	-0.588	-0.477
	(0.366)	(0.427)
Percent of white workers in manufacturing, 1910	-0.014	0.143
	(0.950)	(1.002)
Percent of farm acreage in cotton, 1910	-0.263	-0.427
	(0.501)	(0.561)
Log white population density, 1910	-0.052	-0.052
	(0.105)	(0.116)
White church members per capita, 1916	-0.212	-0.216
	(0.186)	(0.195)
Rosenwald school exposure	0.379	0.431
	(0.166)	(0.210)
White literacy rate (10+), 1910	-0.479	-0.608
	(0.774)	(0.839)
White school attendance rate (6–14), 1910	-0.334	-0.126
	(0.442)	(0.447)
Railroad exposure	0.075	0.082
	(0.067)	(0.067)
Percent African-American, 1910	-0.851	-0.633
	(0.253)	(0.287)
Percent rural, 1910	0.910	0.880
	(0.370)	(0.366)
Percent voting for Strom Thurmond, 1948	0.108	-0.352
	(0.103)	(0.195)
Birth state fixed effects		x
R-squared	0.131	0.138
N (birth counties)	576	576

Notes: See notes to Table 7.
Source: Authors' calculations using Duke SSA/Medicare data, Aaronson and Mazumder (2011) data, Black et al. (2015) data, Census (1992), Haines and ICPSR (2010), ICPSR (1999), Ruggles et al. (2019)

Table A.20: Average Network Index Estimates, by Destination Region

	Destination Region			
	Northeast (1)	Midwest (2)	West (3)	South (4)
Panel A: Southern Black Migrants				
Alabama	1.237	2.356	0.813	-
	(0.161)	(0.295)	(0.272)	-
Florida	0.978	0.793	0.264	-
	(0.172)	(0.169)	(0.107)	-
Georgia	1.546	2.067	0.410	-
	(0.243)	(0.310)	(0.205)	-
Louisiana	0.282	1.138	2.169	-
	(0.101)	(0.206)	(0.734)	-
Mississippi	0.924	2.662	1.036	-
	(0.105)	(0.396)	(0.130)	-
North Carolina	1.678	0.908	0.185	-
	(0.149)	(0.176)	(0.040)	-
South Carolina	2.907	1.223	0.211	-
	(0.351)	(0.167)	(0.055)	-
All States	1.860	2.259	1.402	-
	(0.120)	(0.195)	(0.345)	-
Panel B: Great Plains White Migrants				
Kansas	0.079	0.452	0.281	0.051
	(0.019)	(0.095)	(0.031)	(0.006)
Nebraska	0.080	0.439	0.420	0.063
	(0.014)	(0.096)	(0.109)	(0.009)
North Dakota	0.107	0.405	0.524	0.047
	(0.027)	(0.057)	(0.046)	(0.009)
Oklahoma	0.051	0.390	0.542	0.074
	(0.007)	(0.091)	(0.047)	(0.007)
South Dakota	0.061	0.485	0.381	0.058
	(0.013)	(0.069)	(0.034)	(0.011)
All States	0.073	0.434	0.442	0.062
	(0.007)	(0.039)	(0.029)	(0.004)

Notes: All columns contain weighted averages of destination-level network index estimates, $\hat{\Delta}_k$, where the weights are the number of people who move from each state to destination k. We define destination regions slightly differently than the Census Bureau because we treat the former Confederate states as the South. The Census South region includes Delaware, the District of Columbia, Maryland, West Virginia, Kentucky, and Oklahoma. We include the first four states in the Northeast and the latter two in the Midwest. We do not estimate network indices for African Americans who move to the South. Birth town groups are defined by cross validation. Standard errors are in parentheses.
Source: Authors' calculations using Duke SSA/Medicare data

Table A.21: Average Network Index Estimates, by Destination Region, Southern White Migrants

	Destination Region			
	Northeast (1)	Midwest (2)	West (3)	South (4)
Alabama	0.140	1.048	0.208	-
	(0.021)	(0.123)	(0.034)	-
Florida	0.090	0.070	0.277	-
	(0.017)	(0.020)	(0.104)	-
Georgia	0.104	0.307	0.082	-
	(0.013)	(0.049)	(0.023)	-
Louisiana	0.159	0.450	0.331	-
	(0.027)	(0.100)	(0.100)	-
Mississippi	0.067	0.301	0.127	-
	(0.014)	(0.052)	(0.014)	-
North Carolina	0.549	0.489	0.302	-
	(0.063)	(0.122)	(0.048)	-
South Carolina	0.111	0.081	0.073	-
	(0.011)	(0.012)	(0.022)	-
All States	0.275	0.534	0.220	-
	(0.024)	(0.044)	(0.026)	-

Notes: See note to Appendix Table A.20.
Source: Authors' calculations using Duke SSA/Medicare data

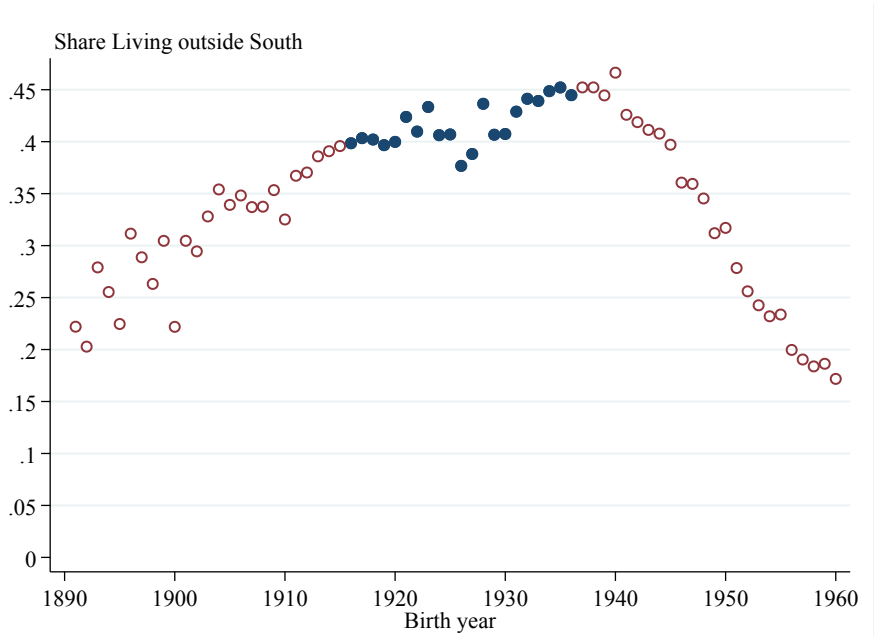
Table A.22: Changes in Regional Migration Patterns in a Counterfactual without Migration Networks

	Destination Region			
Birth State	Northeast (1)	Midwest (2)	West (3)	South (4)
Panel A: Southern Black Migrants				
Alabama	4,354	-6,706	2,353	
	[19.2%]	[-10.8%]	[20.5%]	
Florida	-224	-68	291	
	[-1.8%]	[-1.7%]	[12.2%]	
Georgia	621	-2,147	1,526	
	[1.6%]	[-6.6%]	[25.4%]	
Louisiana	1,267	1,809	-3,076	
	[31.2%]	[11.7%]	[-8.4%]	
Mississippi	2,951	-7,303	4,352	
	[38.9%]	[-7.6%]	[25.5%]	
North Carolina	-2,252	1,033	1,220	
	[-3.3%]	[14.9%]	[36.9%]	
South Carolina	-2,175	1,056	1,119	
	[-3.6%]	[14.8%]	[43.5%]	
All States	4,541	-12,325	7,785	
	[2.1%]	[-5.5%]	[9.8%]	
Panel B: Great Plains White Migrants				
Kansas	485	-179	-1,645	1,339
	[6.9%]	[-1.1%]	[-1.7%]	[6.6%]
Nebraska	673	284	-2,462	1,505
	[10.7%]	[2.0%]	[-2.5%]	[10.9%]
North Dakota	305	527	-1,508	676
	[10.0%]	[4.6%]	[-2.1%]	[10.4%]
Oklahoma	819	995	-5,000	3,186
	[12.2%]	[7.4%]	[-3.3%]	[11.3%]
South Dakota	292	171	-1,126	662
	[9.7%]	[1.5%]	[-2.0%]	[9.4%]
All States	2,574	1,799	-11,740	7,368
	[9.9%]	[2.7%]	[-2.5%]	[9.7%]

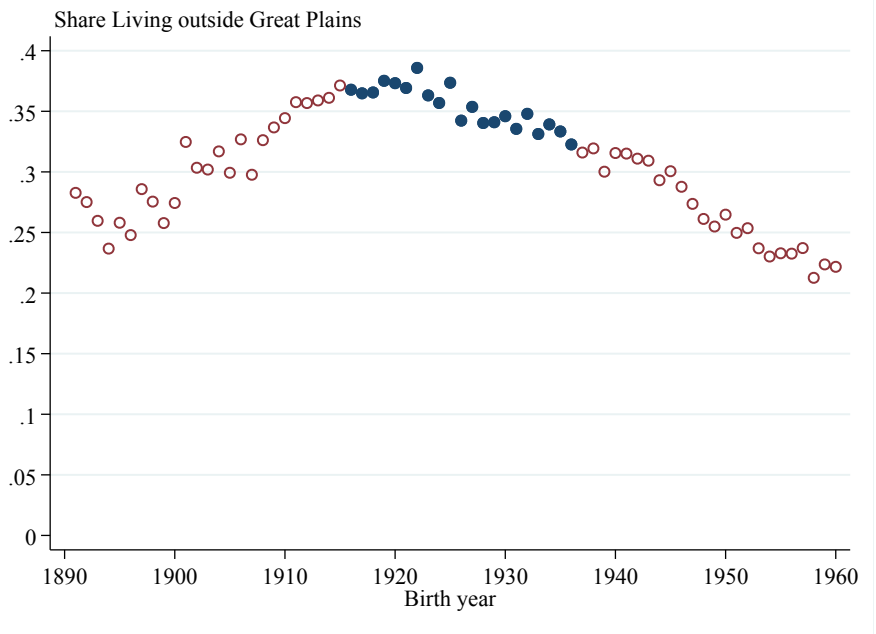
Notes: Table contains estimates of N_k^{cf}, the number of migrants that would have chosen destination county k in the absence of migration network, aggregated over all counties in each region. Percent changes of the number of migrants in the counterfactual are in brackets. See the text for details.
Source: Authors' calculations using Duke SSA/Medicare data

Figure A.1: Migration Rates Around Ages 40–49

(a) Southern Black Migrants

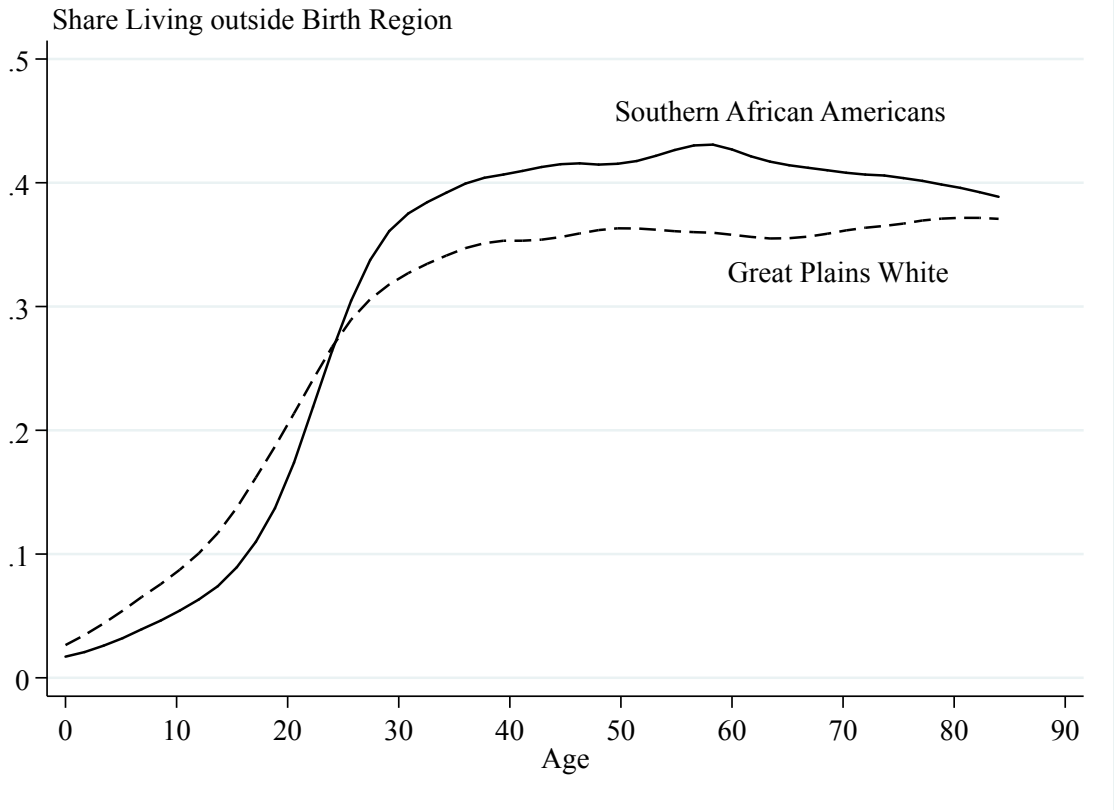


(b) Great Plains White Migrants



Notes: Panel A reports the share of African Americans born in AL, FL, GA, LA, MS, NC, and SC living outside of the former Confederate States. Panel B reports the share of whites born in KS, NE, ND, OK, and SD living outside of the Great Plains and border area shaded in light grey in Figure 2. For individuals born from 1891–1900, we measure their location using the 1900 Census. For individuals born from 1901–1910, we use the 1910 Census, and so forth. The shaded circles correspond to individuals born from 1916–1936, who comprise our sample from the Duke SSA/Medicare data.
Source: Authors' calculations using 1940–2000 Census data from Ruggles et al. (2019)

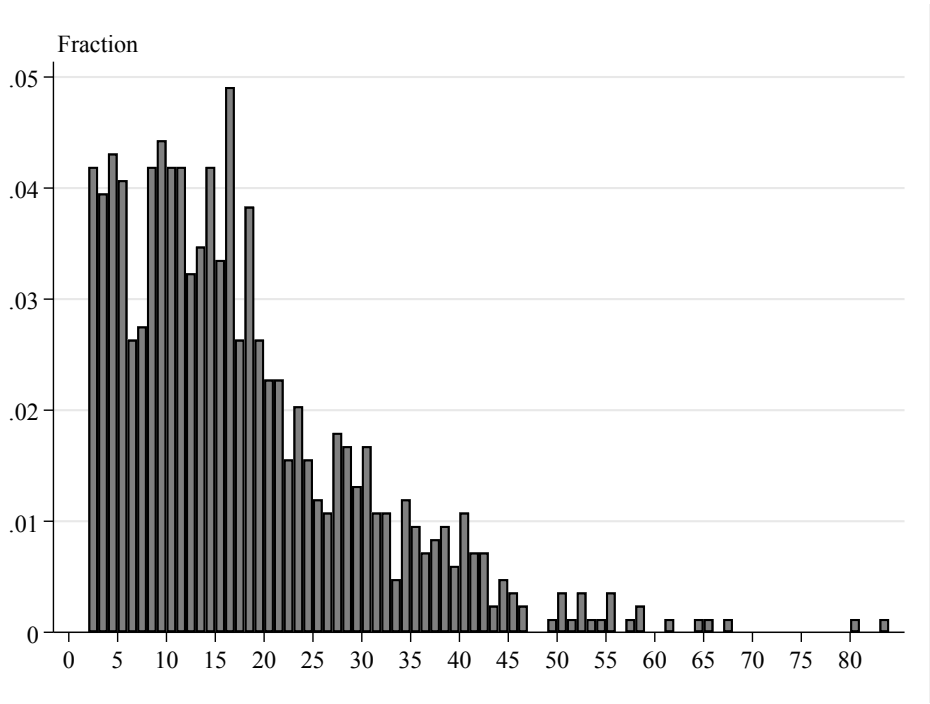
Figure A.2: Share Living Outside Birth Region, 1916–1936 Cohorts, by Age



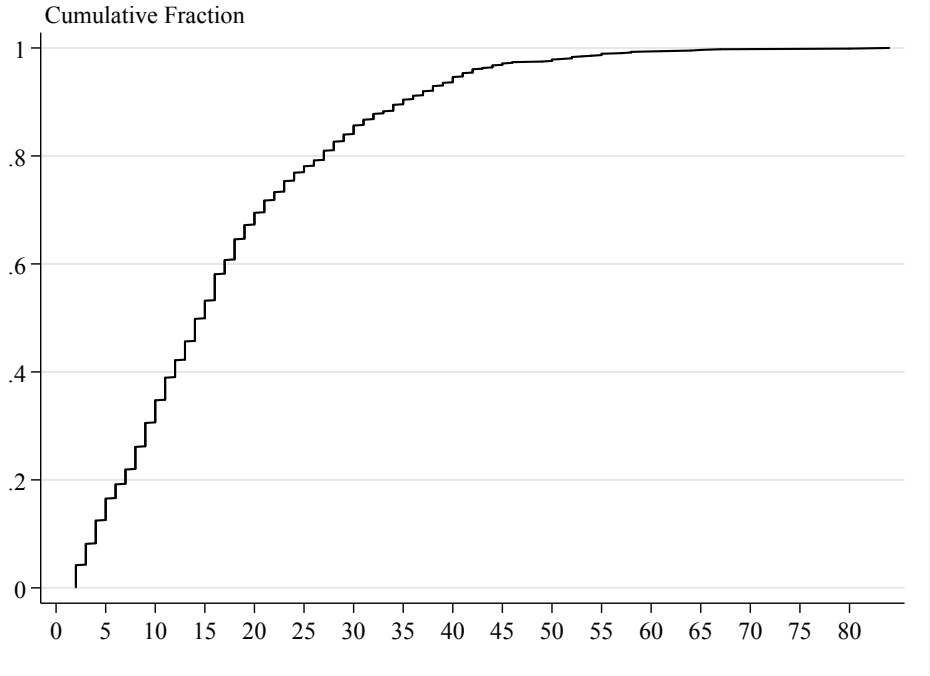
Notes: The solid line shows the percent of African Americans born from 1916–1936 in the seven Southern birth states we analyze (dark grey states in Figure 2) living outside the South (light and dark grey states) at the time of Census enumeration. The dashed line shows the percent of whites born from 1916–1936 from the Great Plains states living outside the Great Plains or Border States. Both lines are locally mean-smoothed relationships of the underlying observations.
Source: Authors' calculations using Ruggles et al. (2019)

Figure A.3: Number of Towns per Birth Town Group, Cross Validation, Southern Black Migrants

(a) Histogram



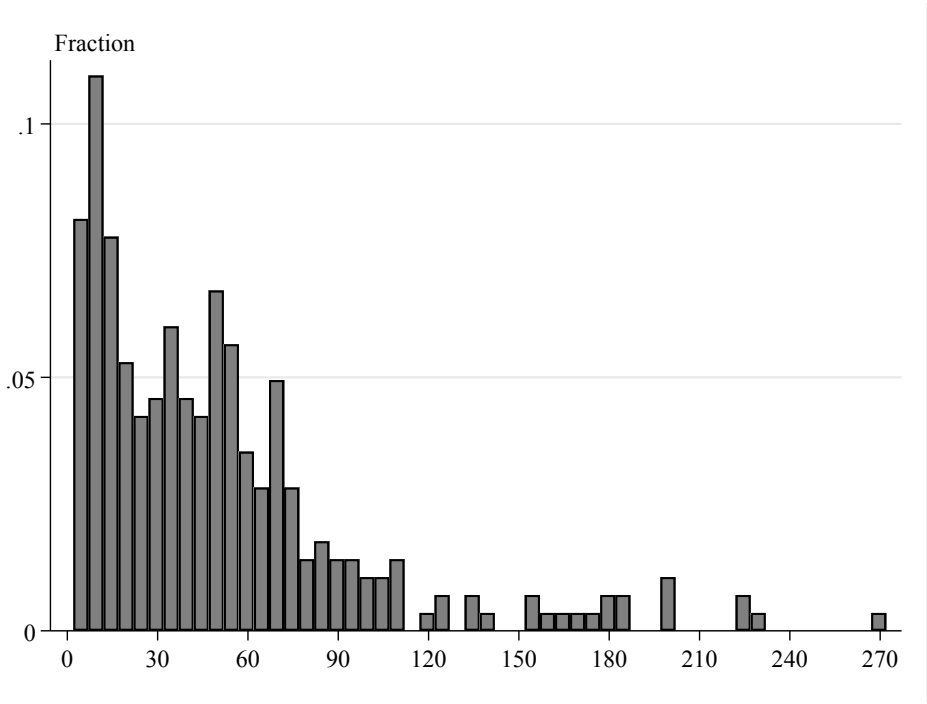
(b) Cumulative Distribution



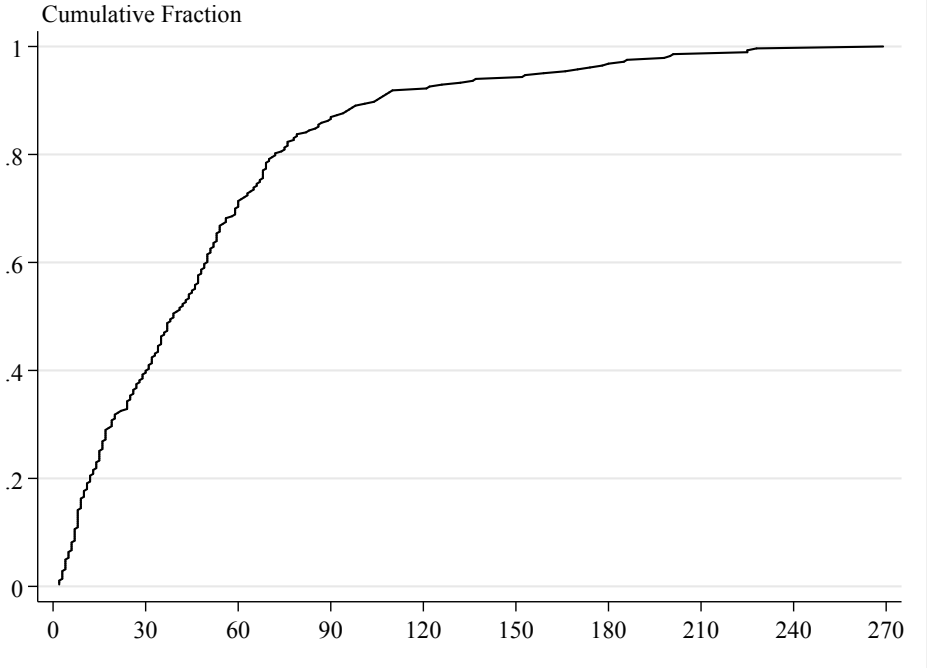
Notes: Figure excludes groups with a single town, as these are not used in the analysis. Bin width in panel (a) is 1.
Source: Authors' calculations using Duke SSA/Medicare data

Figure A.4: Number of Towns per Birth Town Group, Cross Validation, Great Plains White Migrants

(a) Histogram



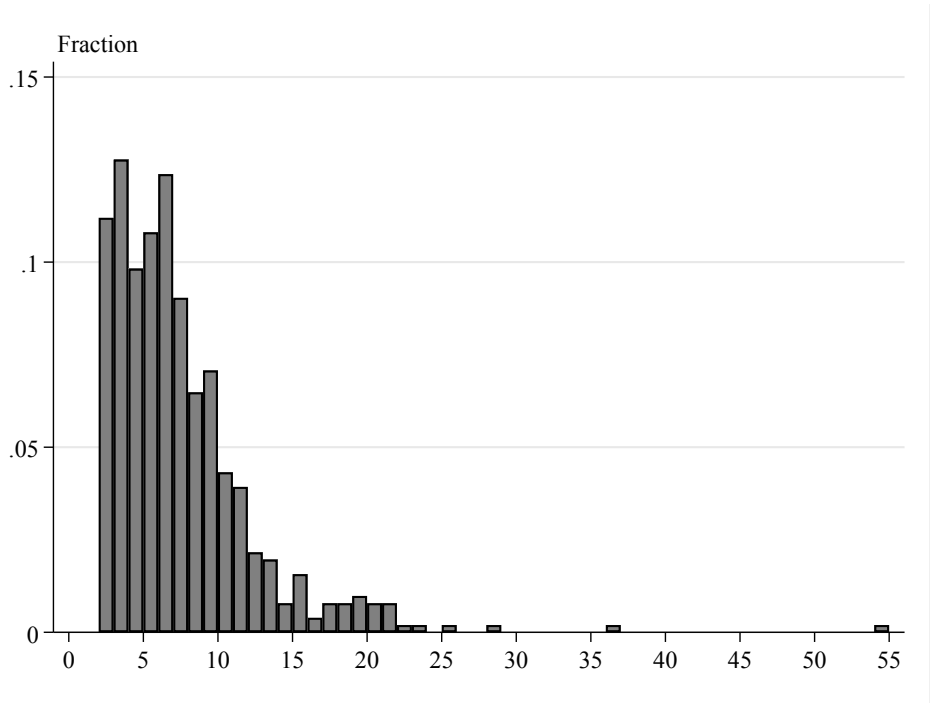
(b) Cumulative Distribution



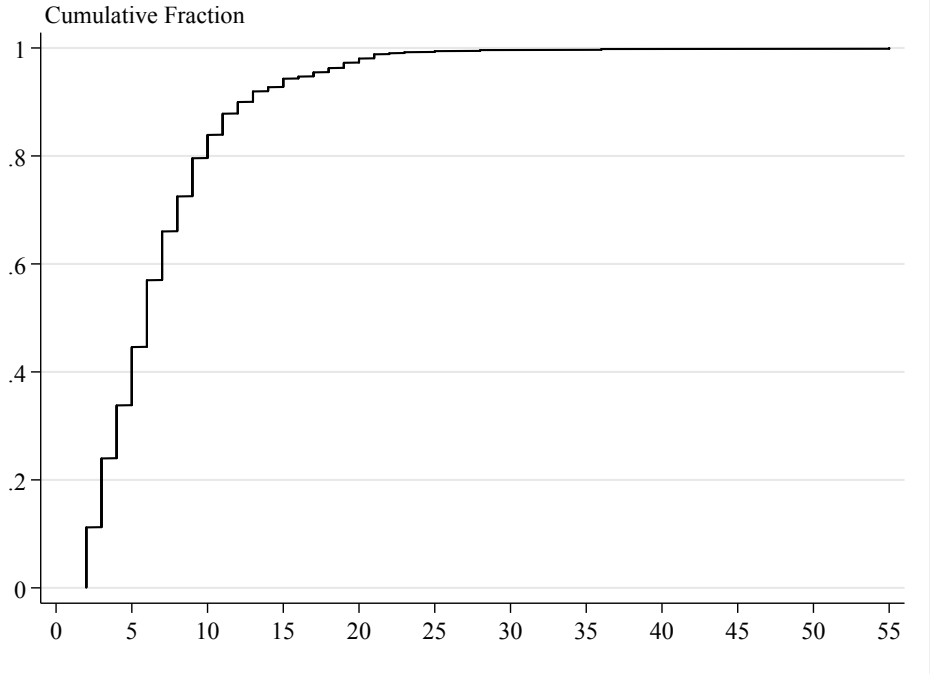
Notes: Figure excludes groups with a single town, as these are not used in the analysis. Bin width in panel (a) is 5.
Source: Authors' calculations using Duke SSA/Medicare data

Figure A.5: Number of Towns per Birth County, Southern Black Migrants

(a) Histogram



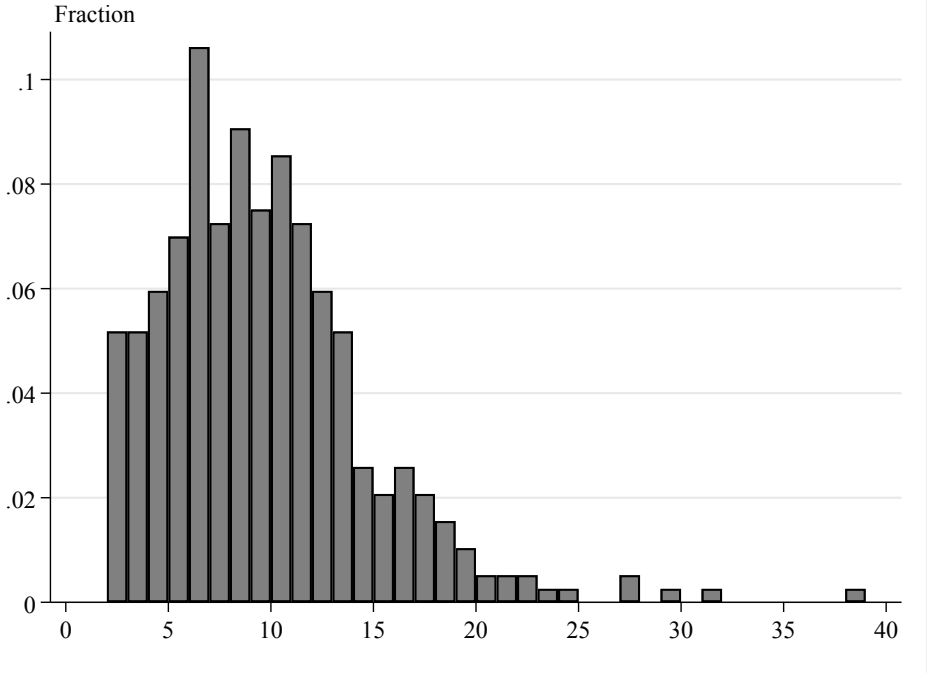
(b) Cumulative Distribution



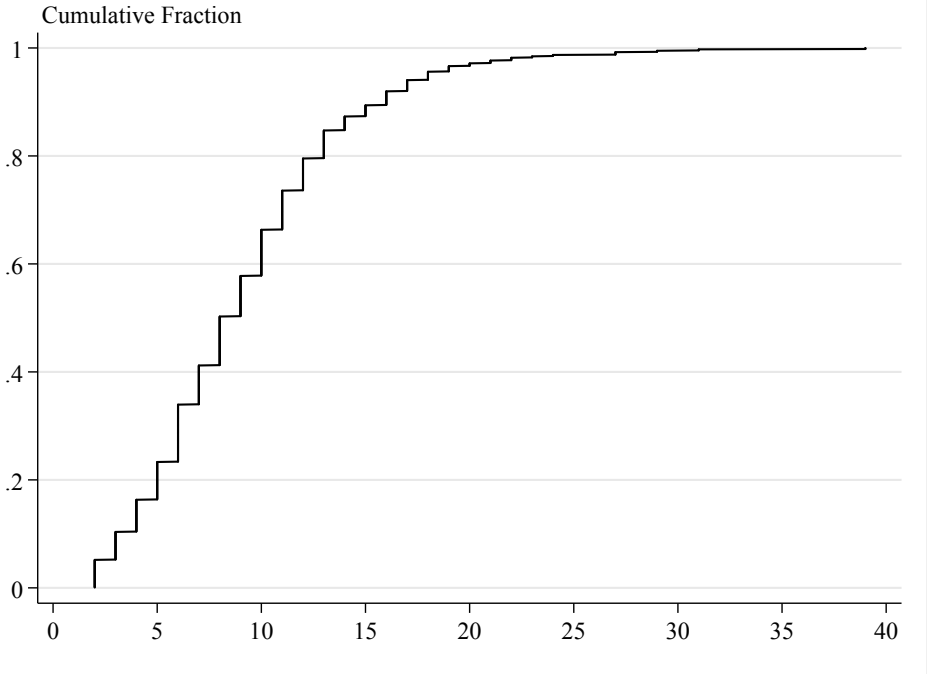
Notes: Figure excludes groups with a single town, as these are not used in the analysis. Bin width in panel (a) is 1.
Source: Authors' calculations using Duke SSA/Medicare data

Figure A.6: Number of Towns per Birth County, Great Plains White Migrants

(a) Histogram



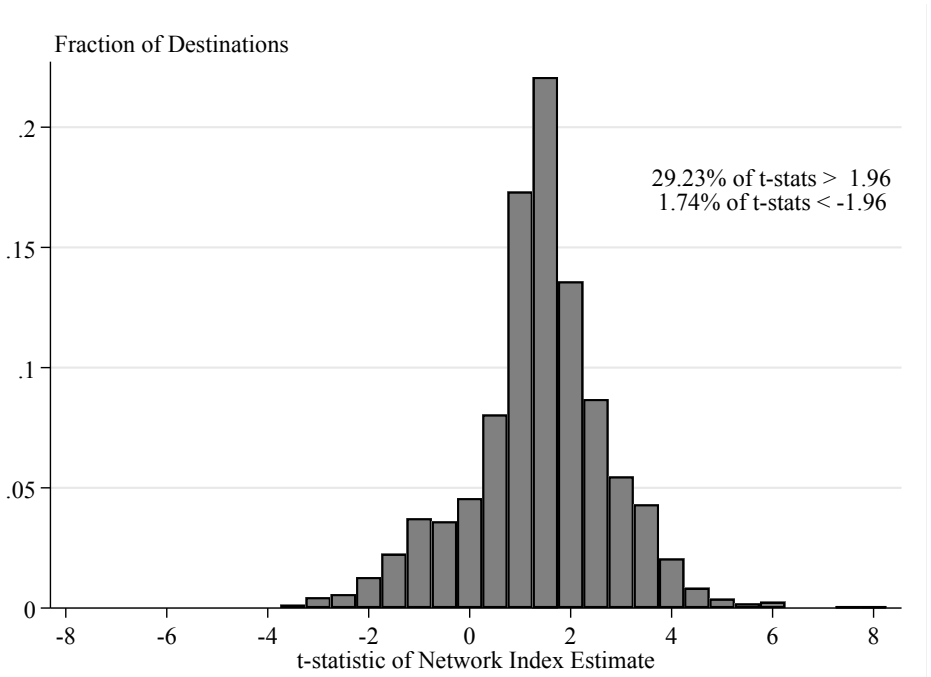
(b) Cumulative Distribution



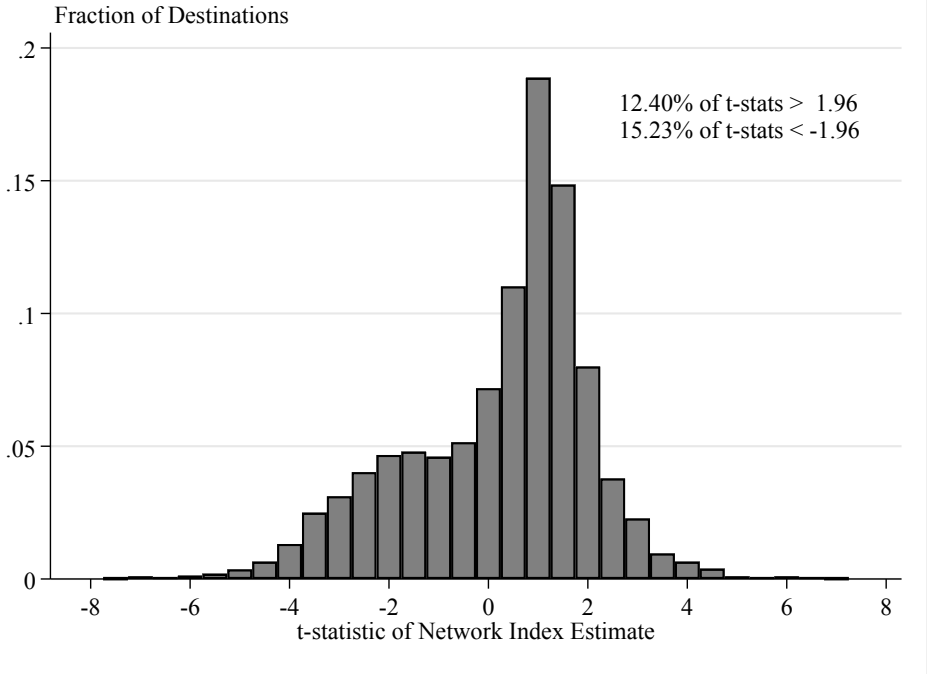
Notes: Figure excludes groups with a single town, as these are not used in the analysis. Bin width in panel (a) is 1.
Source: Authors' calculations using Duke SSA/Medicare data

Figure A.7: Distribution of Destination-Level Network Index t-statistics

(a) Southern Black Migrants

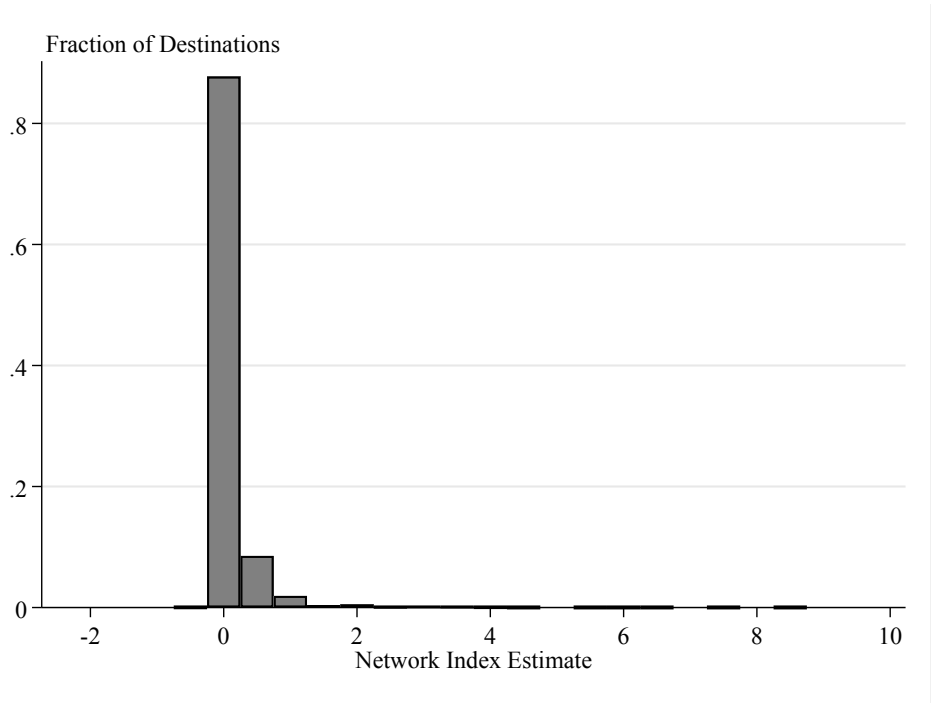


(b) Great Plains White Migrants



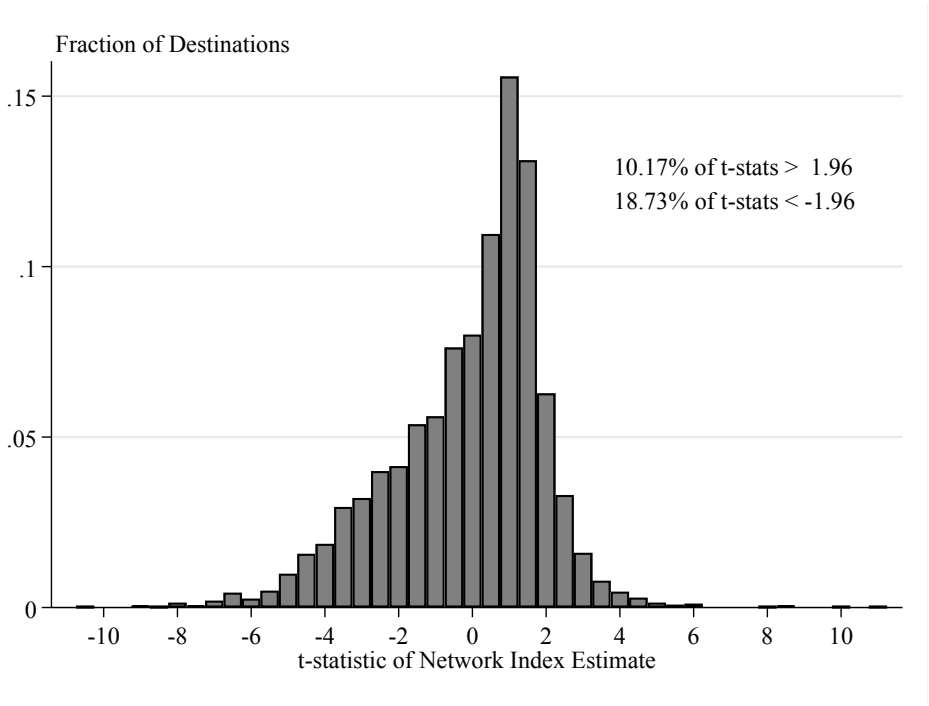
Notes: Bin width is 1/2. Birth town groups are defined by cross validation. Panel (a) omits the t-statistic of 13.7 from South Carolina to Hancock, WV.
Source: Authors' calculations using Duke SSA/Medicare data

Figure A.8: Distribution of Destination-Level Network Index Estimates, Southern White Migrants



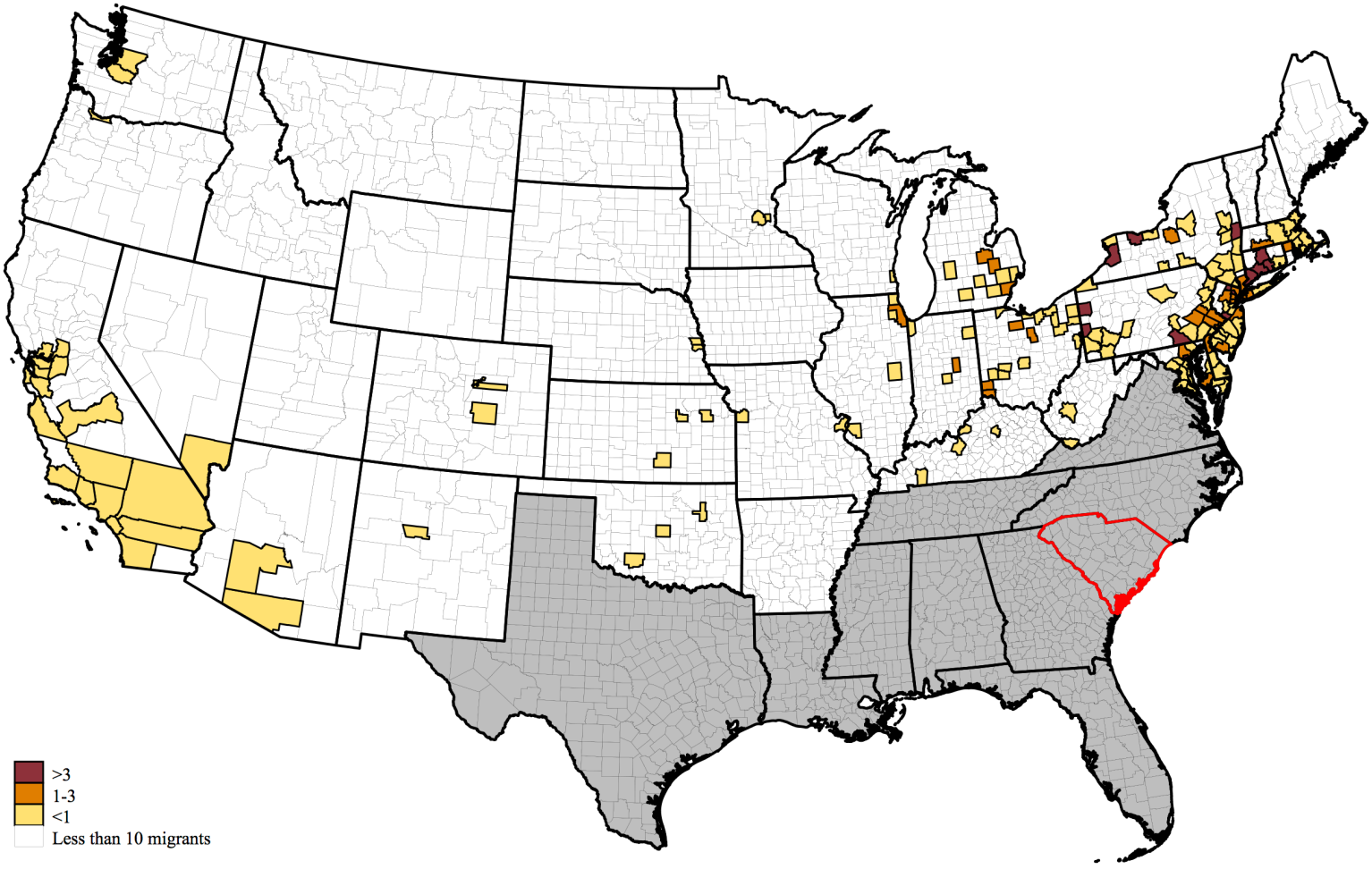
Notes: Bin width is 1/2. Figure omits estimate of $\hat{\Delta}_k = 19.3$ from Alabama to St. Joseph County, IN.
Source: Authors' calculations using Duke SSA/Medicare data

Figure A.9: Distribution of Destination-Level Network Index t-statistics, Southern White Migrants



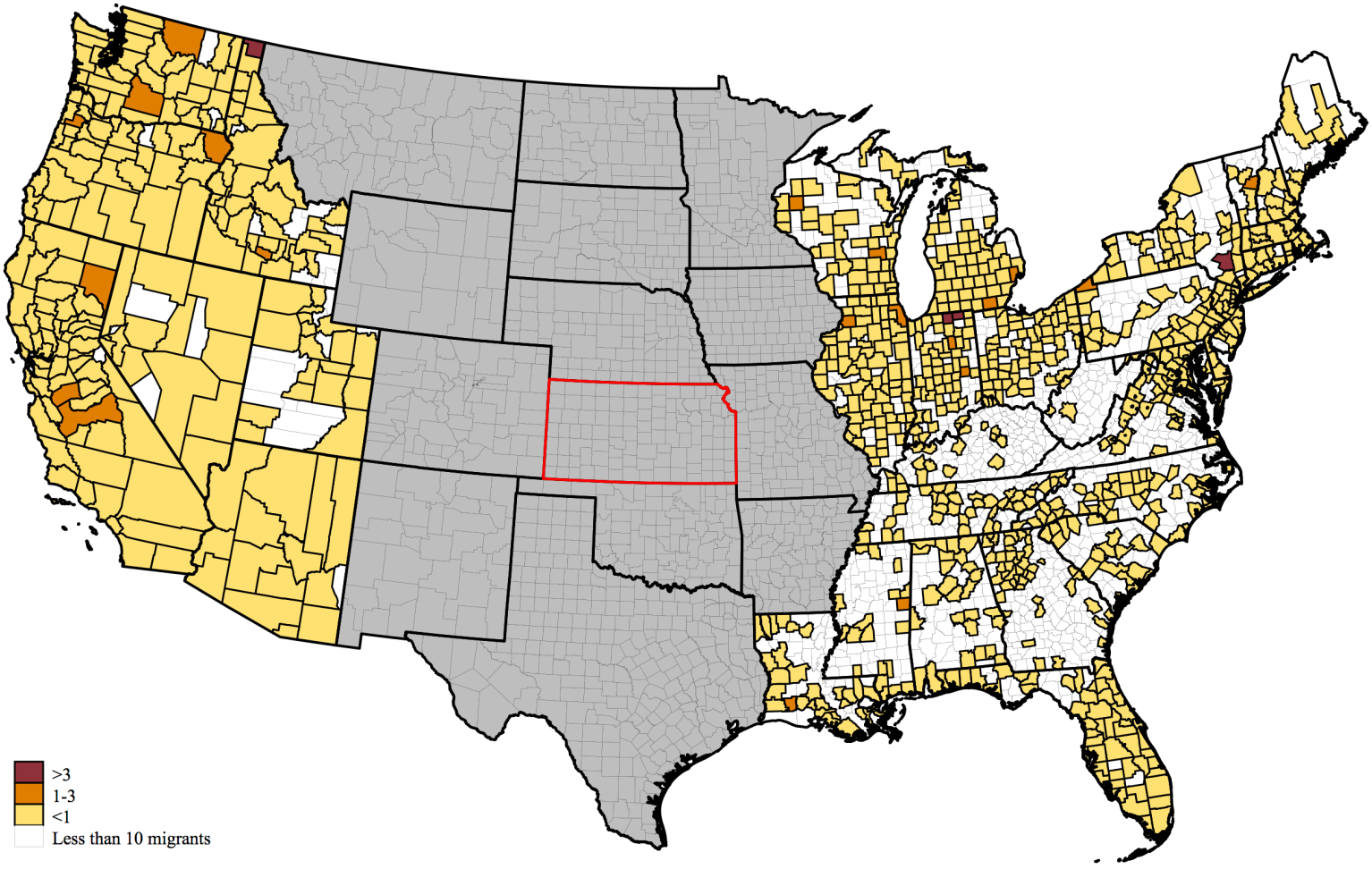
Note: Bin width is 1/2.
Source: Authors' calculations using Duke SSA/Medicare data

Figure A.10: Spatial Distribution of Destination-Level Network Index Estimates, South Carolina-born Black Migrants



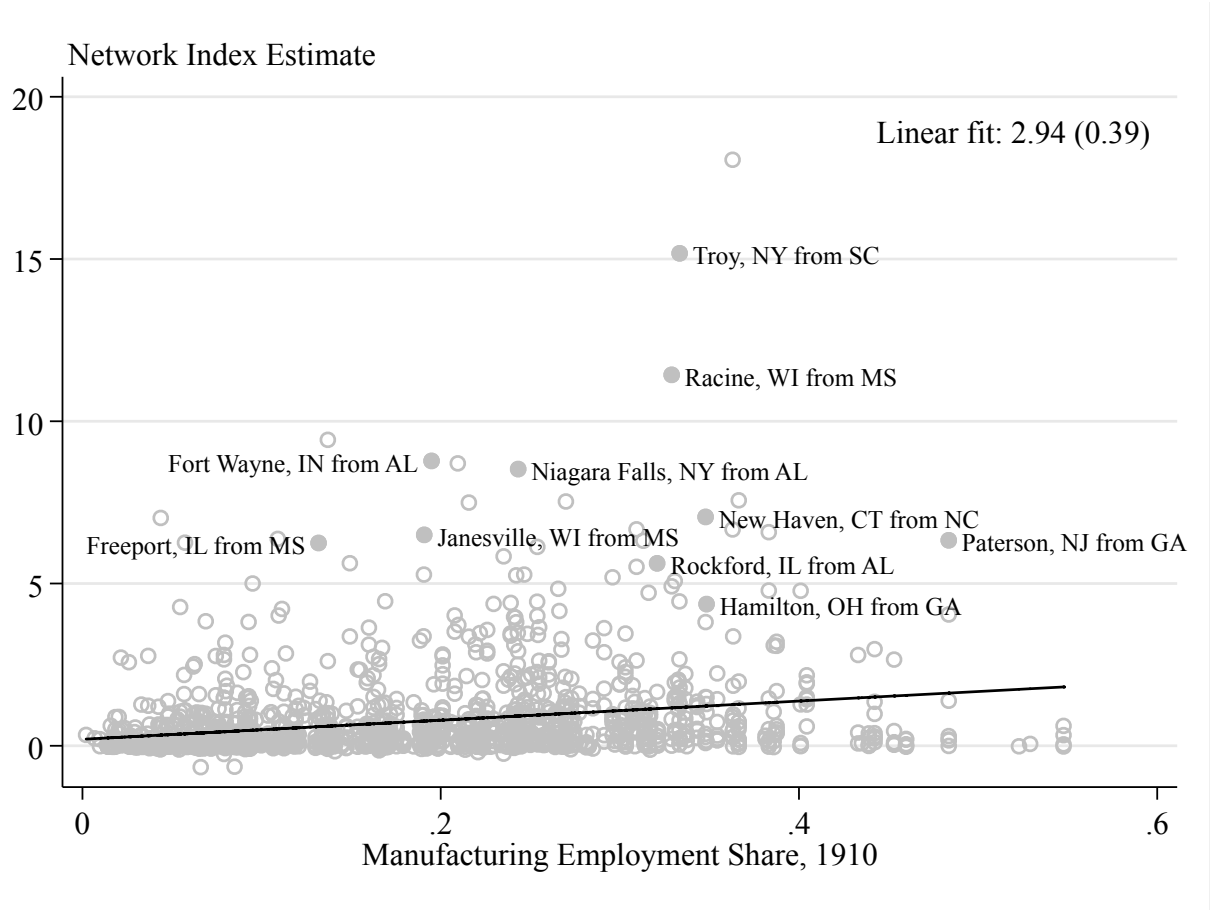
Notes: See note to Figure 4.

Figure A.11: Spatial Distribution of Destination-Level Network Index Estimates, Kansas-born White Migrants



Notes: See note to Figure 5.

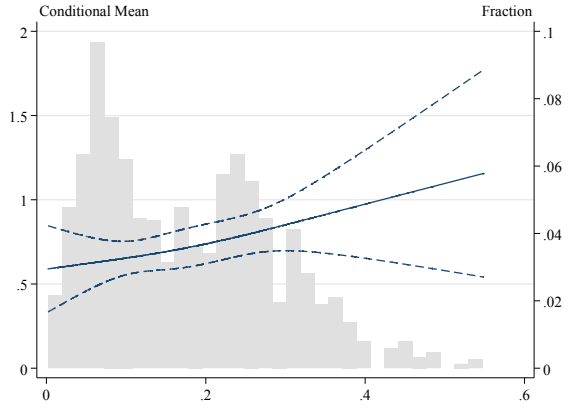
Figure A.12: Relationship between Southern Black Destination-Level Network Index Estimates and 1910 Manufacturing Employment Share



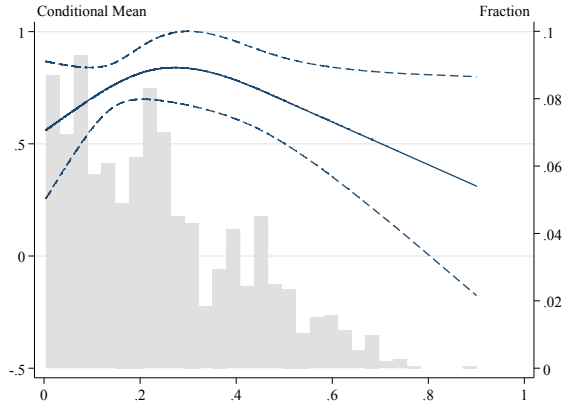
Notes: Linear prediction comes from an OLS regression that includes a constant and 1910 manufacturing employment share. Listed are the cities in Table 2.
Source: Authors' calculations using Duke SSA/Medicare data and Haines and ICPSR (2010)

Figure A.13: Nonlinear Relationship between Covariates and Destination County Network Index Estimates, Southern Black Migrants

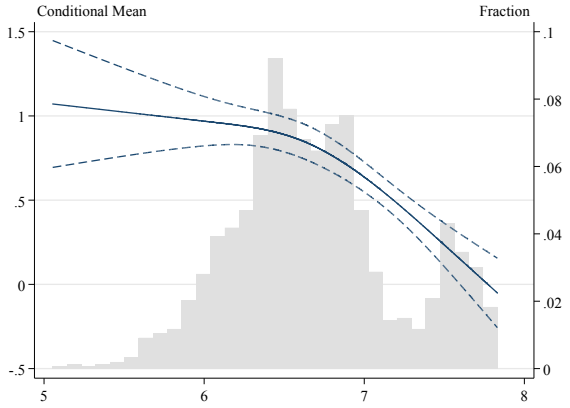
(a) Manufacturing employment share, 1910



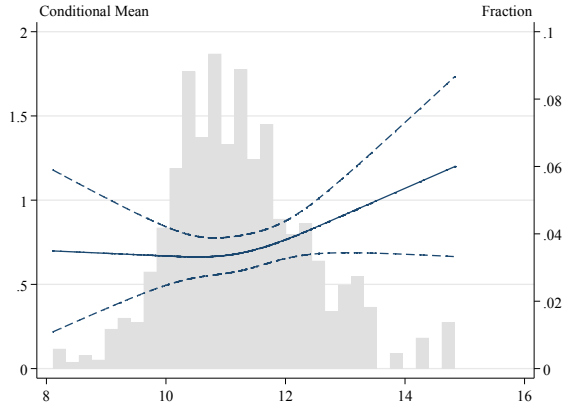
(b) Agriculture employment share, 1910



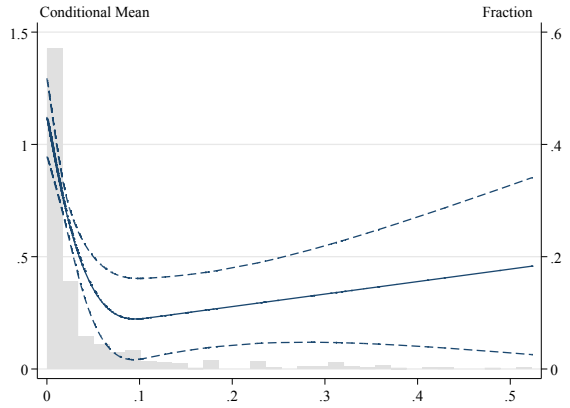
(c) Log distance from birth state



(d) Log population, 1910



(e) Percent African-American, 1910



(f) Percent rural, 1910

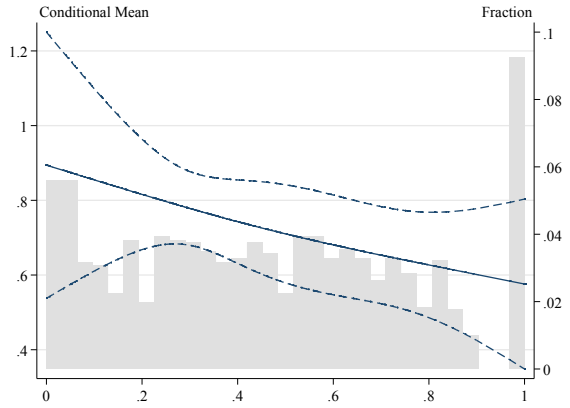
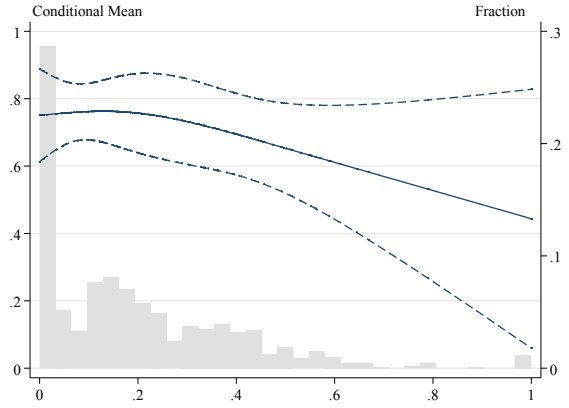
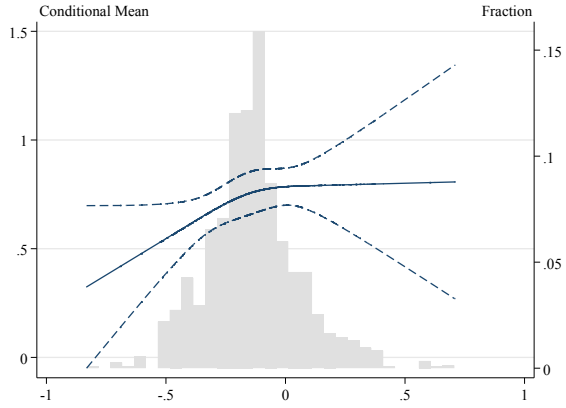


Figure A.13: Nonlinear Relationship between Covariates and Destination County Network Index Estimates, Southern Black Migrants

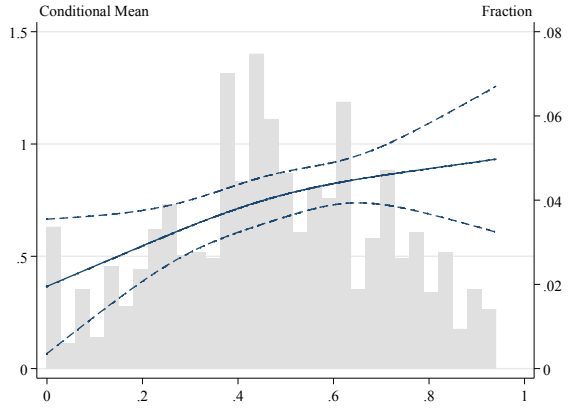
(g) Black church members per capita, 1916



(h) Logan-Parman segregation measure, 1940



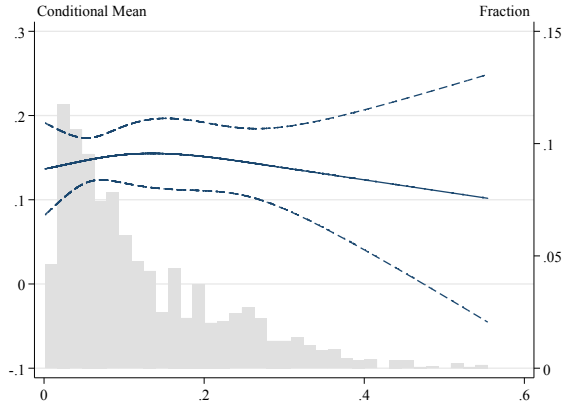
(i) Black-white relative wage, 1940



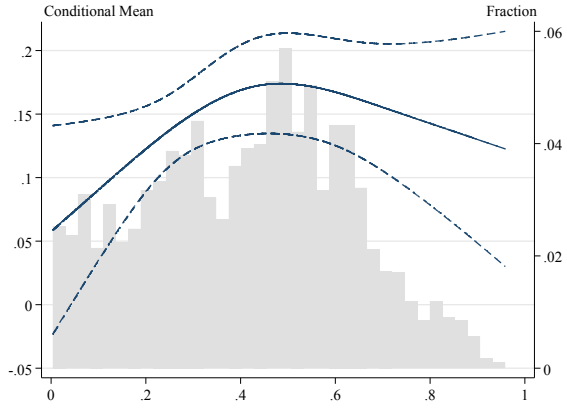
Notes: The solid blue line is the conditional mean of the birth county network index as a function of the indicated independent variable. Dashed lines are 95 percent confidence intervals. Results come from regressing destination county network index estimates against restricted cubic splines in the nine indicated variables, plus indicators for whether the destination has a direct or one-stop connection from the birth state. Grey bars are histograms of the underlying independent variable (right scale). See notes to Table 6.
Source: Authors' calculations using Duke SSA/Medicare data, Black et al. (2015) data, Census (1992), Haines and ICPSR (2010), Logan and Parman (2017), Minnesota Population Center and Ancestry.com (2013)

Figure A.14: Nonlinear Relationship between Covariates and Destination County Network Index Estimates, Great Plains White Migrants

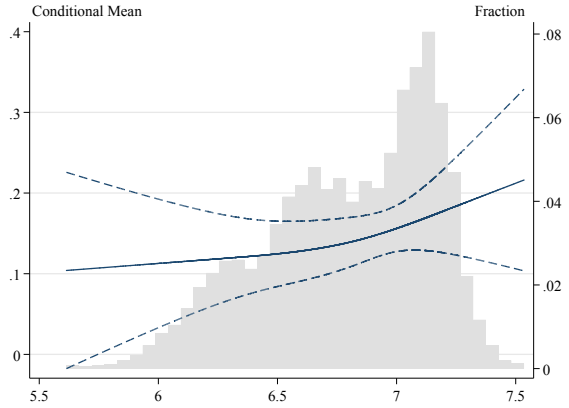
(a) Manufacturing employment share, 1910



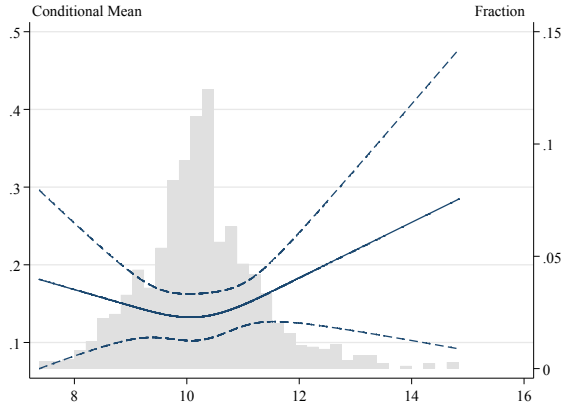
(b) Agriculture employment share, 1910



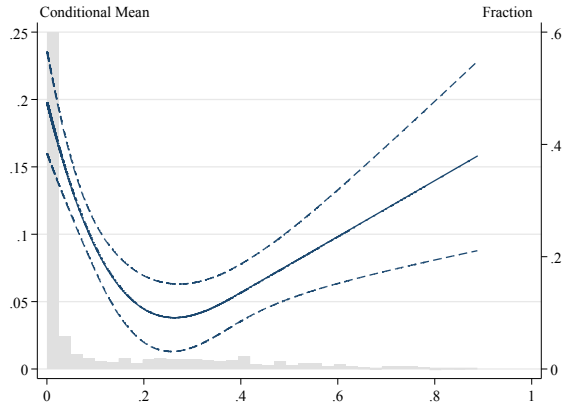
(c) Log distance from birth state



(d) Log population, 1910



(e) Percent African-American, 1910



(f) Percent rural, 1910

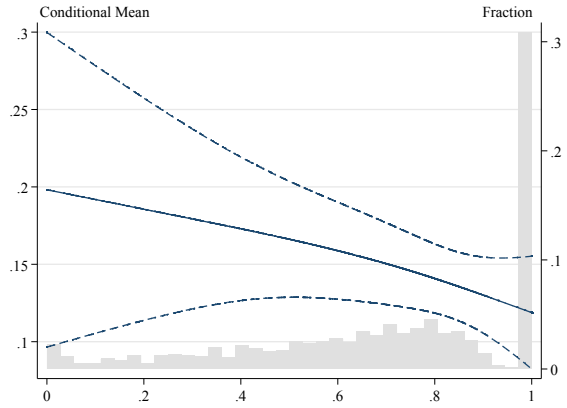
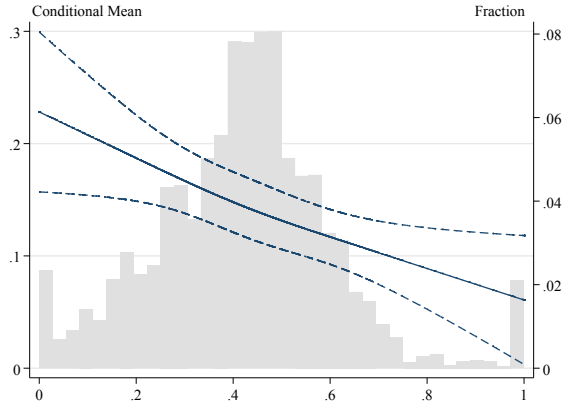
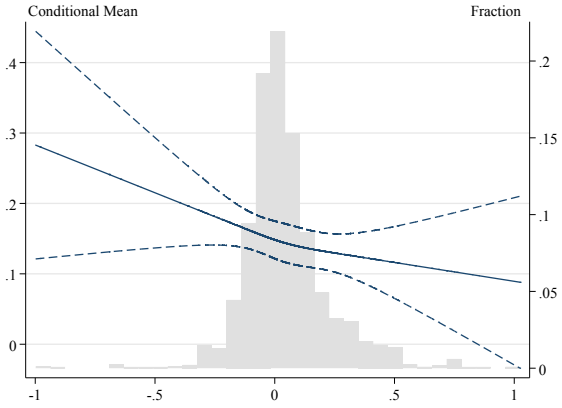


Figure A.14: Nonlinear Relationship between Covariates and Destination County Network Index Estimates, Great Plains White Migrants

(g) White church members per capita, 1916



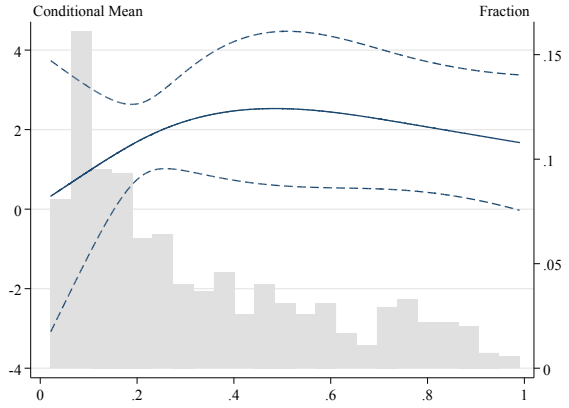
(h) Plains-not Plains relative wage, 1940



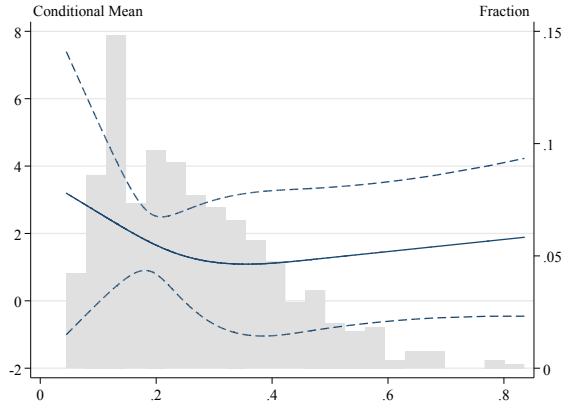
Notes: See notes to Figure A.13.
Source: Authors' calculations using Duke SSA/Medicare data, Black et al. (2015) data, Census (1992), Haines and ICPSR (2010), Logan and Parman (2017), Minnesota Population Center and Ancestry.com (2013)

Figure A.15: Nonlinear Relationship between Covariates and Birth County Network Index Estimates, Southern Black Migrants

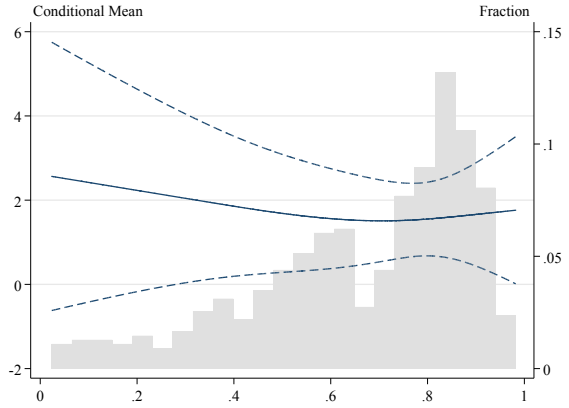
(a) Percent of farmers who are owners, 1910



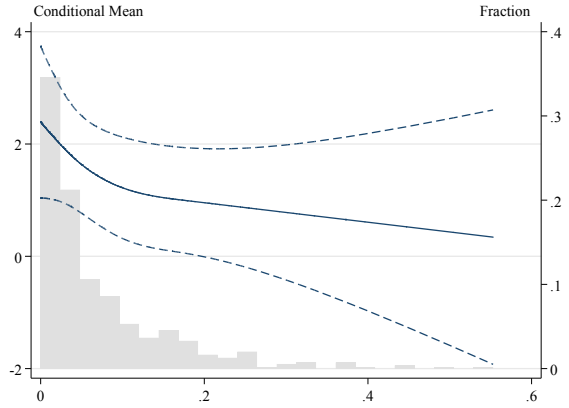
(b) Percent in owner-occupied housing, 1910



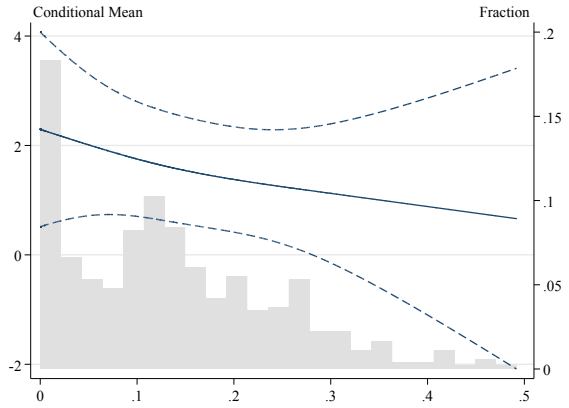
(c) Agriculture employment share, 1910



(d) Manufacturing employment share, 1910



(e) Percent of farm acreage in cotton, 1910



(f) Log population density, 1910

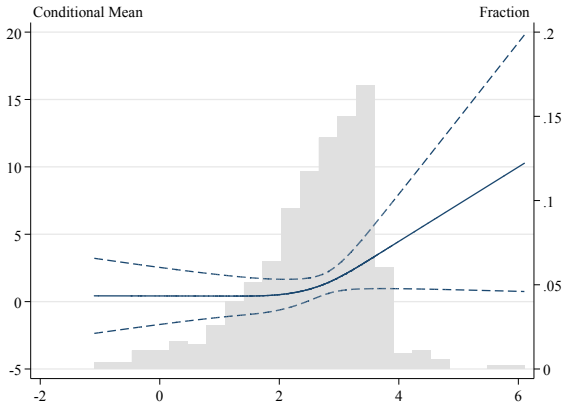
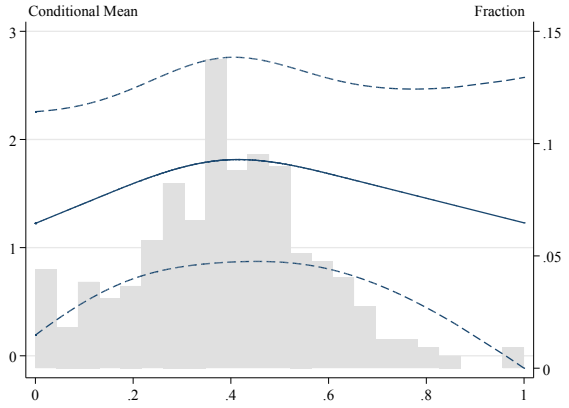
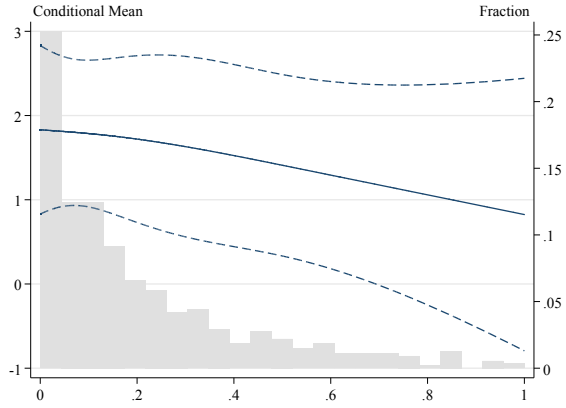


Figure A.15: Nonlinear Relationship between Covariates and Birth County Network Index Estimates, Southern Black Migrants

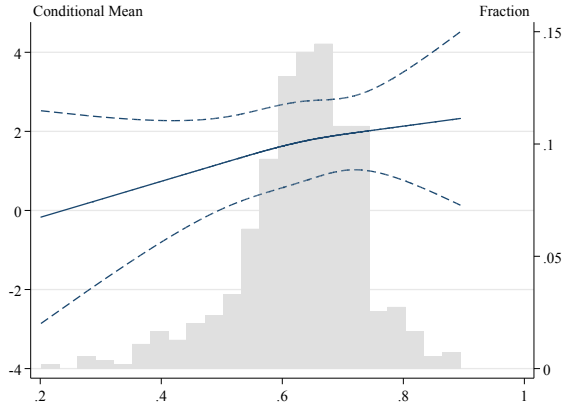
(g) Church members per capita, 1916



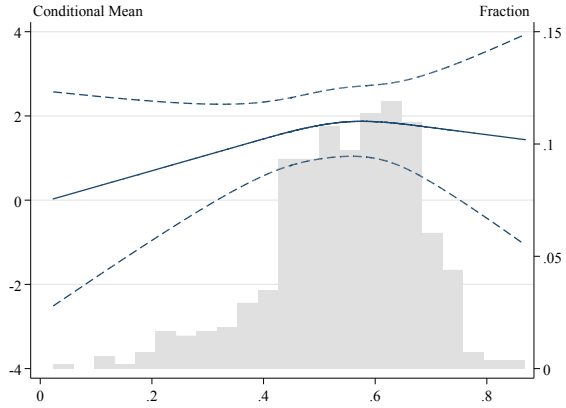
(h) Rosenwald school exposure



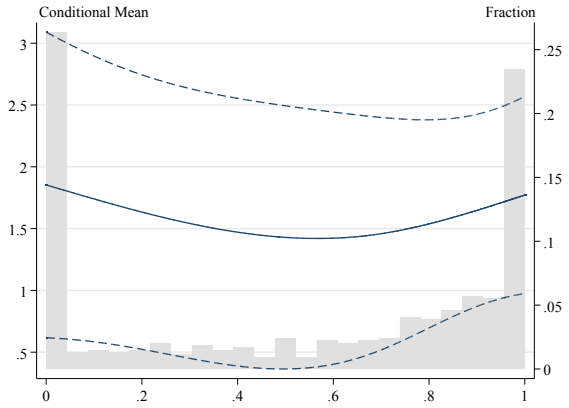
(i) Literacy rate (10+)



(j) School attendance rate (6–14)



(k) Railroad exposure



(l) Percent African-American, 1910

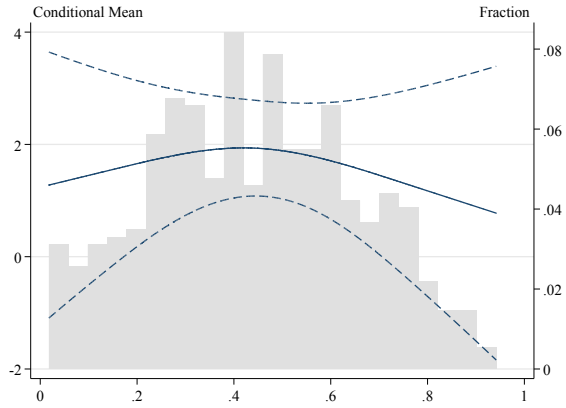
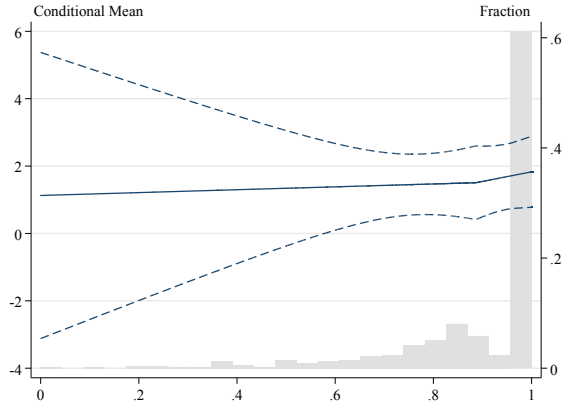
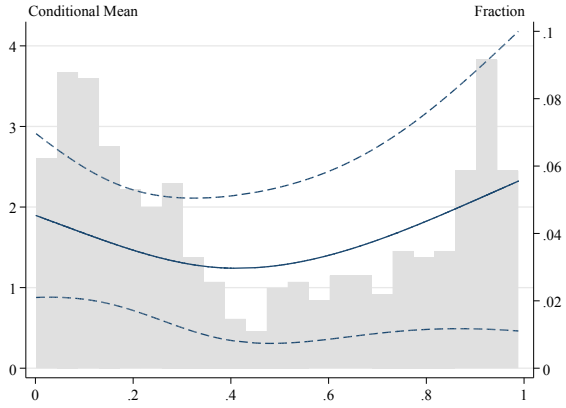


Figure A.15: Nonlinear Relationship between Covariates and Birth County Network Index Estimates, Southern Black Migrants

(m) Percent rural, 1910



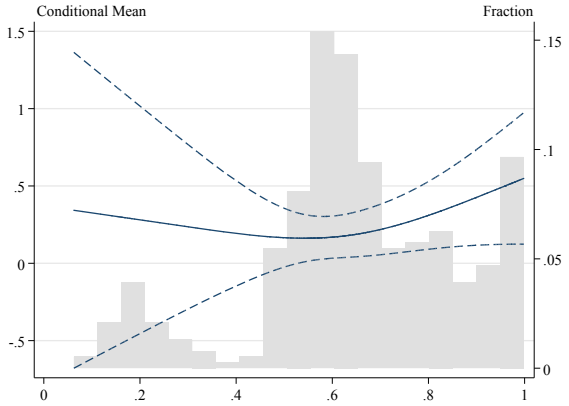
(n) Percent voting for Strom Thurmond, 1948



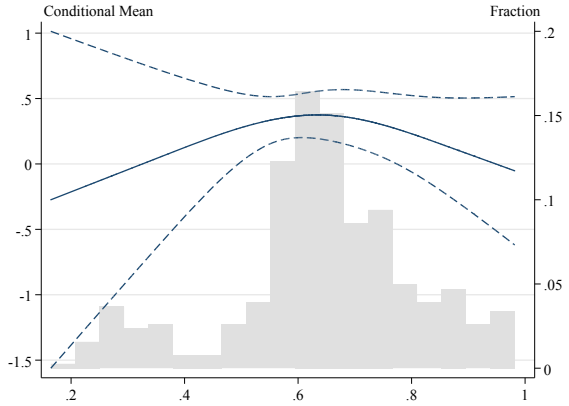
Notes: The solid blue line is the conditional mean of the birth county network index as a function of the indicated independent variable. Dashed lines are 95 percent confidence intervals. These results come from regressing birth county network index estimates against restricted cubic splines in the 14 indicated variables. Grey bars are histograms of the underlying independent variable (right scale). For panels A, B, C, D, F, G, I, and J, the explanatory variables are measured for African Americans.
Source: Authors' calculations using Duke SSA/Medicare data, Aaronson and Mazumder (2011) data, Black et al. (2015) data, Census (1992), Haines and ICPSR (2010), ICPSR (1999)

Figure A.16: Nonlinear Relationship between Covariates and Birth County Network Index Estimates, Great Plains White Migrants

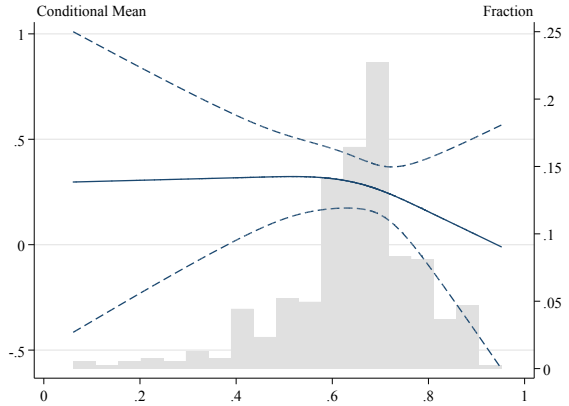
(a) Percent of farmers who are owners, 1910



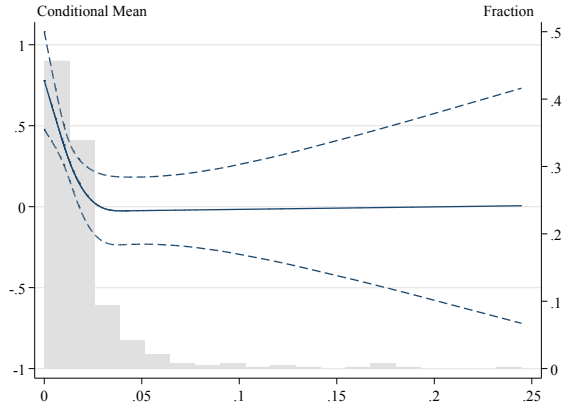
(b) Percent in owner-occupied housing, 1910



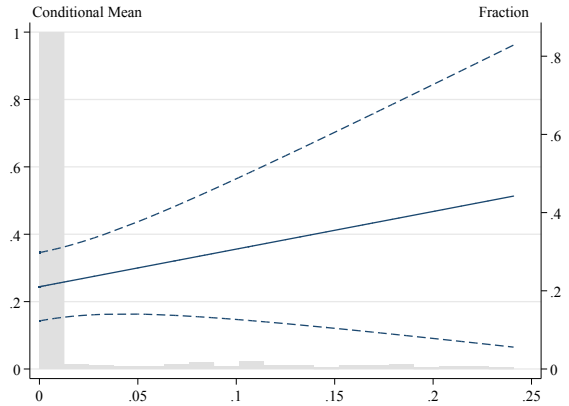
(c) Agriculture employment share, 1910



(d) Manufacturing employment share, 1910



(e) Percent of farm acreage in cotton, 1910



(f) Log population density, 1910

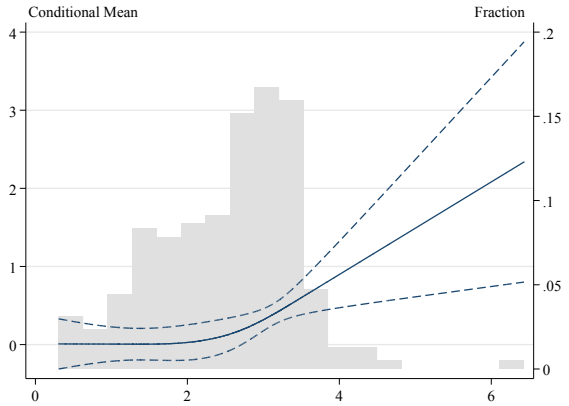
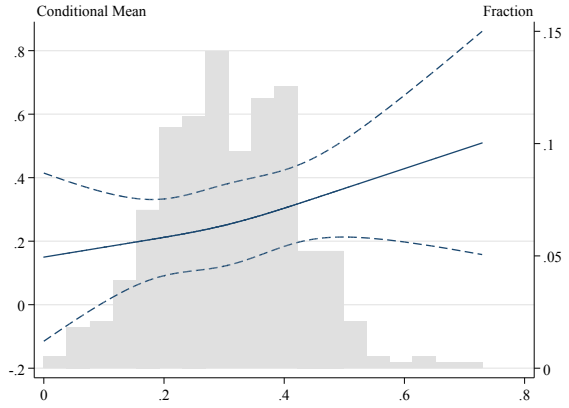
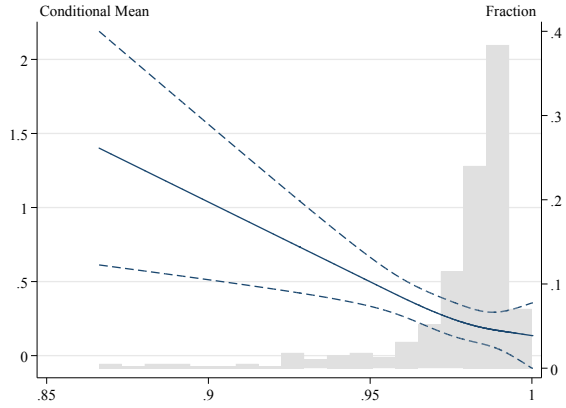


Figure A.16: Nonlinear Relationship between Covariates and Birth County Network Index Estimates, Great Plains White Migrants

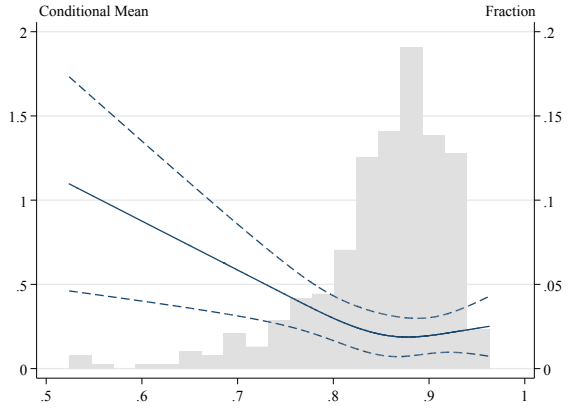
(g) Church members per capita, 1916



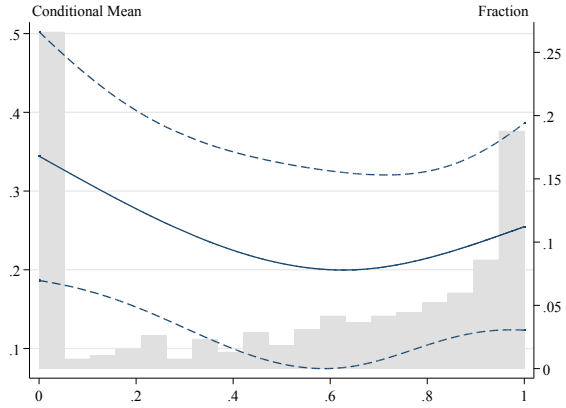
(h) Literacy rate (10+)



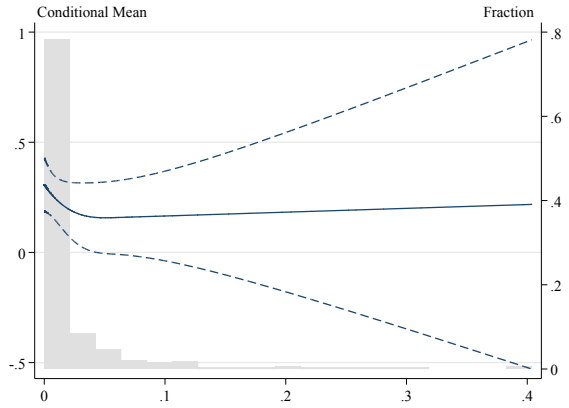
(i) School attendance rate (6–14)



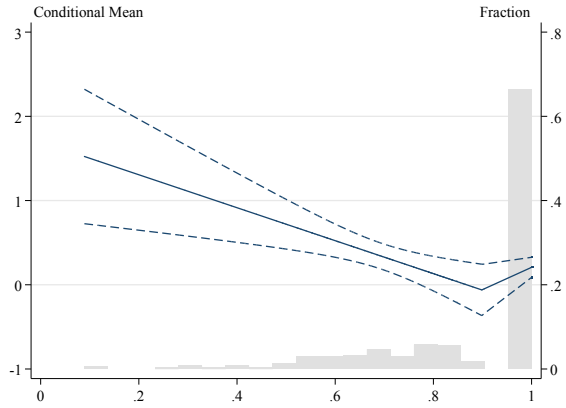
(j) Railroad exposure



(k) Percent African-American, 1910



(l) Percent rural, 1910



Notes: See notes to Figure A.15. For panels A, B, C, D, F, G, I, and J, the explanatory variables are measured for whites.

Source: Authors' calculations using Duke SSA/Medicare data, Aaronson and Mazumder (2011) data, Black et al. (2015) data, Census (1992), Haines and ICPSR (2010), ICPSR (1999)

References

Aaronson, Daniel, and Bhashkar Mazumder. 2011. "The Impact of Rosenwald Schools on Black Achievement." *Journal of Political Economy*, 119(5): 821–888.

Black, Dan A., Seth G. Sanders, Evan J. Taylor, and Lowell J. Taylor. 2015. "The Impact of the Great Migration on Mortality of African Americans: Evidence from the Deep South." *American Economic Review*, 105(2): 477–503.

Cameron, A. Colin, and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications.* New York: Cambridge University Press.

Census, United States Bureau of the. 1992. "Censuses of Religious Bodies, 1906-1936. Ann Arbor, MI: ICPSR [distributor]."

Haines, Michael R., and ICPSR. 2010. "Historical, Demographic, Economic, and Social Data: The United States, 1790-2002. Ann Arbor, MI: ICPSR [distributor]."

ICPSR. 1999. "United States Historical Election Returns, 1824-1968. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]."

Logan, Trevon D., and John M. Parman. 2017. "The National Rise in Residential Segregation." *The Journal of Economic History*, 77(1): 127–170.

Minnesota Population Center, and Ancestry.com. 2013. "IPUMS Restricted Complete Count Data: Version 1.0 [Machine-readable database]."

Minnesota Population Center and Ancestry.com. 2013. "IPUMS Restricted Complete Count Data: Version 1.0 [Machine-readable database]. Minneapolis: University of Minnesota."

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2019. "IPUMS USA: Version 9.0 [dataset]."