The Great Migration and Educational Opportunity
Cavit Baran, Eric Chyn, and Bryan A. Stuart
Online Appendix

A Derivation of Selection Correction

Section 3 follows Finkelstein, Gentzkow and Williams (2021) and introduces the place effect selection correction as equation (5). This appendix reviews how this selection correction equation is derived from Assumptions 1 and 2 in Section 3.

Finkelstein, Gentzkow and Williams (2021) provide the following proposition:

Proposition 1 *Assumption 1 is equivalent to*

$$\eta_j^{dest} = \frac{SD\left(\eta_j^{\text{dest}}\right)}{SD\left(h_j^{\text{dest}}\right)} h_j^{\text{dest}}.$$

Proof: Note that h_j^{dest} and η_j^{dest} are normalized to have mean zero, which implies that $Cov(T_{ij}, h_j^{\text{dest}}) = \frac{N}{N'} h_j^{\text{dest}} p(1-p)$ and $Cov(T_{ij}, \eta_j^{\text{dest}}) = \frac{N}{N'} \eta_j^{\text{dest}} p(1-p)$, where N is the total number of migrants, N' is the number of migrants with $T_{ij} = 0$, and $p = Pr(T_{ij} = 1)$. Assumption 1 is then equivalent to:

$$\frac{\frac{N}{N'} h_j^{\text{dest}} p(1-p)}{SD(T_{ij}) SD(h_j^{\text{dest}})} = \frac{\frac{N}{N'} \eta_j^{\text{dest}} p(1-p)}{SD(T_{ij}) SD(\eta_j^{\text{dest}})}.$$

Proposition 1 follows immediately after canceling terms in the above expression. □

This proposition is helpful since it can be combined with Assumption 2 to obtain equation (5) in Section 3, which is the desired selection-correction expression.

B Model of Selective Migration

This appendix describes a simple model of selective migration and describes how the selection correction procedure used in the paper adjusts observed patterns in the data to identify causal place effects.

In broad terms, the model features parents who care about their own income and the later-life income of their children. Places differ in the returns to parent human capital and the translation of child schooling capital into educational attainment. When the returns to parent human capital or child schooling capital are increasing in the underlying levels of human capital, there is selective migration in the sense that parents and children with higher levels of human capital move to locations with higher returns.

B.1 Formal Model

Let i index a family, which consists of a parent and a child. The parent chooses the location that maximizes his or her utility. The parent's indirect utility from moving to location j is

$$u_{i,j} = W^p_{i,j} + \delta W^c_{i,j} - \kappa_j - e_{i,j}, \tag{B1}$$

where $W^p_{i,j}$ is the present discounted value of the parent's earnings if they live in location j, $W^c_{i,j}$ is the present discounted value of the child's (later-life) earnings if their parent moves to location j, $\delta \in [0, 1]$ is a parameter describing how much weight the parent places on the earnings of their child, $\kappa_j \geq 0$ is a moving cost, and $e_{i,j}$ is an idiosyncratic preference term. Parents can choose to remain in their origin location, $o(i)$, in which case the moving cost equals 0.

The parent has human capital $\theta^p_i \geq 0$, which is assumed to be fixed. The earnings of a parent that lives in location j are:

$$W^p_{i,j} = f(\theta^p_i; \boldsymbol{\phi_j}), \tag{B2}$$

where $\boldsymbol{\phi_j}$ is a vector that parametrizes earnings in location j, normalized so that $\partial f/\partial \boldsymbol{\phi_j} > 0$. Parent human capital consists of years of schooling, $S^p_i \geq 0$, and an independent component, $\eta^p_i \geq 0$:

$$\theta^p_i = S^p_i + \eta^p_i. \tag{B3}$$

The child's years of schooling if the parent moves to location j is $S^c_{i,j} \geq 0$. We assume that the parent considers the lifetime earnings of the child to be:

$$W^c_{i,j} = g(S^c_{i,j}; \boldsymbol{\psi}), \tag{B4}$$

where $\boldsymbol{\psi}$ is a vector that parametrizes earnings, normalized so that $\partial g/\partial \boldsymbol{\psi} > 0$. Equation (B4) makes the simplifying assumption that parents focus on the first-order effect of how their location choice affects their child's schooling but not the second-order effect of how their location choice affects the return to schooling that their child will earn in the future. The years of schooling attained by a child are:

$$S^c_{i,j} = \theta^c_i + \gamma_j, \tag{B5}$$

where $\theta^c_i \geq 0$ summarizes the determinants of a child's educational attainment that do not depend on where the child lives, and γ_j is a place effect on years of schooling. We refer to θ^c_i as the child's level of schooling capital. Equation (B5) is a direct analog of equation (1).

Child schooling capital can be correlated with parent human capital. We summarize this relationship as:

$$\theta^c_i = \rho \theta^p_i + \nu_i, \tag{B6}$$

where ρ describes the degree of intergenerational persistence in human capital and $\nu_i \geq 0$ is an independent component of child schooling capital. Equations (B3) and (B6) imply that child schooling capital can be correlated with parent schooling, which is a key feature for our empirical approach.

B.2 Selective Migration in the Model

At this point, we can use the model to discuss selective migration. Plugging equations (B2), (B4), and (B5) into the indirect utility function in equation (B1) yields:

$$u_{i,j} = f(\theta_i^p; \boldsymbol{\phi_j}) + \delta g(\theta_i^c + \gamma_j; \boldsymbol{\psi}) - \kappa_j - e_{i,j}. \tag{B7}$$

Equation (B7) summarizes the determinants of the parent's location decision. The presence of selective migration depends on the shape of the functions $f(\cdot)$ and $g(\cdot)$. If the cross-partial derivative $\partial^2 f/(\partial\theta_i^p \partial\boldsymbol{\phi_j}) > 0$, then a parent with a higher level of human capital gains more by moving to a destination with a better labor market (i.e., one with a higher $\boldsymbol{\phi_j}$) than a parent with a lower level of human capital. As a result, there will tend to be selective migration in terms of parent human capital. Selective migration in terms of child schooling capital—so that children with higher levels of schooling capital, θ_i^c, tend to live in places with higher place effects on their schooling, γ_j—similarly requires that $\partial^2 g/(\partial\theta_i^c \partial\gamma_j) > 0$. Given the functional form in equation (B5), this cross-partial derivative will be positive if $g(\cdot)$ is a convex function. The value of δ is likely to be less than 1, both because parents are not perfectly altruistic and because children's earnings are discounted in present value terms.

To more concretely demonstrate selection into migration, let us briefly consider a version of the model that has three simplifying assumptions. First, we specify functional forms for the earnings of parents and children. Motivated by the large literature which models log earnings as a linear function of years of schooling following Mincer (1958), we assume that the earnings of the parent and child are:

$$W_{i,j}^p = \exp(\phi_j^0 + \phi_j^1 \theta_i^p), \tag{B8}$$
$$W_{i,j}^c = \exp(\psi^0 + \psi^1 S_i^c). \tag{B9}$$

Equation (B8) allows for the possibility that locations differ both in the earnings received by all individuals, as captured by the intercept term $\phi_j^0 \geq 0$, and the possibility that locations differ in the return to parent human capital, as captured by the slope term $\phi_j^1 \geq 0$. These functional forms generate selective migration in terms of both parent human capital and child schooling capital. Second, we assume there are just two locations: a single destination j and a single origin o. Third, we assume that parents do not care about their child's educational attainment (i.e., $\delta = 0$).

In this simplified version of our model, a parent prefers migrating to location j over staying in the origin if

$$\exp(\phi_j^0 + \phi_j^1 \theta_i^p) - \kappa_j + e_{i,j} > \exp(\phi_o^0 + \phi_o^1 \theta_i^p) + e_{i,o}, \tag{B10}$$

which says that the return to migration in terms of higher earnings outweighs the moving cost and idiosyncratic preference difference. If the labor market returns in location j are higher than those in the origin ($\phi_j^0 > \phi_o^0$ and $\phi_j^1 > \phi_o^1$), then all else equal there will be a threshold level of parent human capital such that parents with θ_i^p above this threshold will prefer location j.[56] Intuitively,

[56] This can be seen by rearranging equation (B10) so that the left-hand side is an increasing function of θ_i^p and

parents with a higher level of human capital are willing to pay the moving cost because they benefit by more from the higher return in labor market j. In a more general scenario where parents also care about their child's educational attainment, there could be selective migration in terms of both parent human capital and child schooling capital. We discuss the implications of selection in terms of child schooling capital in detail below.

B.3 Adjusting for Selective Migration to Estimate Place Effects

We use this model to describe how the selection correction procedure used in the text adjusts observed patterns in the data to estimate causal place effects.

For simplicity, we assume that there are two destinations, $j \in \{L, H\}$, with location H having a higher place effect on children's schooling: $\gamma_H > \gamma_L$. Using equation (B5), the difference in average education levels between migrant children that live in the two destinations depends on both selection and place effects:

$$\underbrace{E[S^c_{i,j}|j(i) = H] - E[S^c_{i,j}|j(i) = L]}_{\text{Observed difference in child schooling}} = \underbrace{E[\theta^c_i|j(i) = H] - E[\theta^c_i|j(i) = L]}_{\text{Selection on unobserved child schooling capital}} + \underbrace{\gamma_H - \gamma_L}_{\text{Place effects}}. \tag{B11}$$

The most natural concern in our setting is that the average level of children's unobserved schooling capital in destination H is higher than in destination L: $E[\theta^c_i|j(i) = H] > [\theta^c_i|j(i) = L]$. This implies that a simple comparison of education levels would overstate the difference in place effects.

There are two distinct reasons why children's unobserved schooling capital might be higher in the destination with a higher place effect in this model. The first reason is that parents care about their child's long-run outcomes when deciding where to move (i.e., $\delta > 0$) *and* the long-run earnings return for a child of moving to a destination with a higher schooling place effect is increasing in the level of child schooling capital (i.e., $\partial^2 g/(\partial\theta^c_i\partial\gamma_j) > 0$). The second reason is that parents care about their own earnings when deciding where to move, a parent with a higher level of human capital will gain more earnings by moving to a destination with a better labor market (i.e., $\partial^2 f/(\partial\theta^p_i\partial\boldsymbol{\phi_j}) > 0$), destinations that have higher place effects on child schooling also have higher wage returns to parents (i.e., $\phi^0_H > \phi^0_L$ and $\phi^1_H > \phi^1_L$), and there is a positive correlation between parent human capital and child schooling capital (i.e., $\rho > 0$). The functional forms in equations (B8) and (B9) generate both forms of selective migration.

These same reasons suggest that parent human capital will also be higher in destination H: $E[\theta^p_i|j(i) = H] > E[\theta^p_i|j(i) = L]$. Although parent human capital is not observed, we can observe parents' education level. As long as parents' education displays the same pattern of selection as parents' human capital, we have a situation where destination H features both higher levels of child schooling capital (which is not observed) and higher levels of parent education (which is observed). In this simple example, this implies that

$$\underbrace{E[\theta^c_i|j(i) = H] - E[\theta^c_i|j(i) = L]}_{\text{Selection on unobserved child schooling capital}} = \alpha \times \underbrace{E[S^p_i|j(i) = H] - E[S^p_i|j(i) = L]}_{\text{Selection on observed parent schooling}}, \tag{B12}$$

the right-hand side does not depend on θ^p_i: $\exp(\phi^0_j + \phi^1_j\theta^p_i) - \exp(\phi^0_o + \phi^1_o\theta^p_i) > \kappa_j - (e_{i,j} - e_{i,o})$. The threshold level of parent human capital is a function of location-specific earnings returns ($\phi^0_j, \phi^1_j, \phi^0_o, \phi^1_o$), moving costs ($\kappa_j$), and idiosyncratic preferences ($e_{i,j}, e_{i,o}$).

for some positive constant α. Given a value for α, we could use equations (B11) and (B12) to identify place effects in this simple example. This example describes the basic intuition behind approaches that use selection on observed variables to address selection on unobserved variables (Altonji, Elder and Taber, 2005; Oster, 2019; Finkelstein, Gentzkow and Williams, 2021). A key benefit of the approach of Finkelstein, Gentzkow and Williams (2021) is that it uses additional moments of the data to quantify how selection on observables translates into selection on unobservables. Because we study a setting where there is potential selection on both parent and child human capital, it is not straightforward to derive further analytical expressions for the estimator developed by Finkelstein, Gentzkow and Williams (2021). In Section 4.6, we probe the robustness of our results to different assumptions about this scaling factor.

C Empirical Bayes Adjustment

When reporting county-specific place effects, we use an empirical Bayes adjustment to account for finite sample bias. This section provides details on the adjustment, following Chetty and Hendren (2018*b*) and Finkelstein, Gentzkow and Williams (2021).

Let γ_j be the true education place effect, which is normalized to have mean 0. Let M be the average causal place effect (which is 0 by construction). We assume that γ_j is a normally distributed random variable:

$$\gamma_j = M + \eta_j, \tag{C1}$$

with $\eta_j \sim N(0, \chi^2)$.

We assume that estimates of γ_j are measured with idiosyncratic error:

$$\hat{\gamma}_j = \gamma_j + \nu_j, \tag{C2}$$

where the estimation error is $\nu_j \sim N(0, s_j^2)$ and s_j is the standard error of $\hat{\gamma}_j$.

Based on equations (C1) and (C2), we have:

$$\hat{\gamma}_j = M + \eta_j + \nu_j, \tag{C3}$$

where η_j is assumed to be independent of ν_j. This implies that

$$\chi^2 = var(\eta_j) = var(\hat{\gamma}_j) - var(\nu_j) \tag{C4}$$

$$= var(\hat{\gamma}_j) - E(s_j^2). \tag{C5}$$

Equation (C5) yields an estimate of χ^2 using the variance of estimated place effects and the average standard error of place effects. We calculate standard errors for each place effect using the bootstrap.

In this framework, we can compute optimal predictions γ_j^{EB} for each county j by minimizing the mean squared prediction error:

$$\sum_{j=1}^{J} (\gamma_j^{EB} - \gamma_j)^2. \tag{C6}$$

We write the true causal effect of moving to j as:

$$\gamma_j = \beta_{1j} M + \beta_{2j} \hat{\gamma}_j. \tag{C7}$$

We cannot estimate this regression (since γ_j is unobserved), but the hypothetical regression would allow us to recover J-many β_{1j} and β_{2j} coefficients that would allow us to compute optimal predictions γ_j^{EB} that minimize the mean squared prediction error:

$$\gamma_j^{EB} = \hat{\beta}_{1j} M + \hat{\beta}_{2j} \hat{\gamma}_j. \tag{C8}$$

While β_{1j} and β_{2j} cannot be estimated directly, we can construct them using the assumptions in equations (C1) and (C2). Specifically, we have

$$\gamma_j^{EB} = \left(\frac{\chi^2}{\chi^2 + s_j^2} \right) \hat{\gamma}_j + \left(\frac{s_j^2}{\chi^2 + s_j^2} \right) M. \tag{C9}$$

Since we assume $M = 0$, this simplifies to:

$$\gamma_j^{EB} = \left(\frac{\chi^2}{\chi^2 + s_j^2} \right) \hat{\gamma}_j. \tag{C10}$$

We use equation (C10) to construct empirical-Bayes-adjusted place effects. We estimate χ^2 using equation (C5), s_j^2 using the bootstrap, and $\hat{\gamma}_j$ using the selection correction described in Section 3. The empirical Bayes adjustment shrinks place effects with larger standard errors towards the grand mean, which is zero.

D Matched Sample Details

We create our matched sample by linking men from the 1920 and 1940 full count Censuses, provided by IPUMS and accessed through the NBER server. We start with U.S.-born Black men age 3–52 in the 1920 full count Census. Following Abramitzky et al. (2021*a*), we link these men to themselves in 1940 using the following procedure[57]:

1. In each dataset, we clean first and last names to remove any non-alphabetic characters and standardize nicknames.

2. We link individuals from 1920 to 1940 in the following way:

   (a) For each 1920 record, we look for records in the 1940 data that match on cleaned first and last name, race, birth state, and exact birth year. If the match is unique, then we call this pair a match. If there is more than one observation, then we drop the 1920 record from the search and call it unmatched.

   (b) For the remaining records for which we did not find matches in the previous step, we search for a unique match within $+/-1$ year of the birth year in 1920. We only accept unique matches.

[57] In our linking procedure, we also download and use command files from Abramitzky et al. (2021*b*).

(c) We repeat the previous step by looking for a unique match within $+/-2$ years. If the record still has no unique match, then we call it unmatched.

3. We repeat the same procedure in (2), but this time we link individuals from 1940 to 1920. In the 1940 Census, we restrict the sample to U.S.-born Black men age 25–70.

4. We take the intersection of the two linked samples.

For our analysis, we focus on U.S.-born Black men who were age 3–52 and living in their birth state in 1920.[58] Our match rate for this group is 14.1 percent. This number is similar to match rates for African Americans in the literature. Eriksson (2019), who links Black men in 1940 to the 1900, 1910, or 1920 Censuses, obtains a match rate of 18.6 percent. This is slightly higher than our match rate because some men have two chances to be matched.

Appendix Table 4 shows that, relative to all U.S.-born Black men who were age 10–50 and living in their birth state in 1920 (Column 1), the subset of men in the matched sample (Column 2) are more likely to be literate and have higher likelihood of living in urban areas in the 1920 Census. The fathers of men in our matched sample also have higher socioeconomic status. These differences are statistically significant at the one-percent level, which is not surprising given the very large samples. However, the differences are quite small in magnitude (e.g., the literacy gap is 4.3 percentage points, which is about 6 percent of the average literacy rate).

Our analysis of place effects focuses on children age 14–18 observed in the 1940 Census. To study selection for this sample, we first examine Southern-born Black men who are age 25–70 in the 1940 Census. These are potential fathers of children in our sample. Column 1 of Appendix Table 5 reports statistics for this group, and column 2 reports statistics for the matched sample subset age 25–70. Relative to all Black household heads, men in the matched sample have 0.3 more years of schooling (6 percent) and $27 more earnings (6 percent). Columns 4 and 5 report statistics for children of the samples in columns 1 and 2. The matched subset has 0.2 more years of schooling (3 percent). Although these differences are statistically significant, the magnitude of the differences are substantively small.

E Alternative Approaches to Estimating the Effect of Moving North

Section 3.3 compares our approach to estimating the effect of moving North—averaging county-specific place effects—and a common approach in prior work—estimating the coefficient on a North indicator. This section describes results from these approaches.

Appendix Table 6 reports results from regressions where the dependent variable is years of schooling for Black children age 14–18. We use the matched sample for this analysis and begin by focusing on children whose parents are born in the South to be consistent with past work. Column 1 shows that children whose parents moved to the North have 2.2 additional years of schooling than children whose parents remained in the South. In column 2, we control for the child's age and sex, plus indicators for the father's and mother's years of schooling and the head of household's 1920 state of residence. Including these variables reduces the difference considerably, to 1.4 years of schooling, indicating the presence of selection. In column 3, we limit the sample to individuals residing in a county with at least 10 migrants, to be consistent with the restriction used in our

[58] Allowing for the $+/-$ two year difference, this produces a matched sample ages 25–70 in 1940.

preferred approach. This restriction mainly eliminates counties in the South with few migrants. The estimated North-South difference falls to 1.0 years, consistent with these counties having less-educated children.

The regression in column 3 controls for a limited number of covariates, which raises concern about selection driving these results. In column 4, we add fixed effects for the head of household's 1920 county of residence. The estimated difference remains large at 1.0 years. In column 5, we focus on a sample of children for whom a cousin is observed (by virtue of their fathers being in the same 1920 household). This sample restriction does not change the estimated difference. Column 6 adds fixed effects for the father's 1920 household. These fixed effects absorb all differences in children's education that are common across 1920 family lines. In this case, the point estimate falls to 0.37, and the standard error nearly triples. The sensitivity of the results for children contrasts to estimates for adults, where household fixed effects generally have little impact (Collins and Wanamaker, 2014; Boustan, 2017). The specification with household fixed effects is quite demanding of the data, and the results in column 6 do not provide conclusive evidence on the effects of moving North on children's education.

To focus on specifications that are more comparable to our selection-correction approach, the model used for column 7 allows state of origin fixed effects to differ by migrant status and includes county of residence fixed effects for non-migrants (which follows equation (3)). This leads to an increase in the coefficient, to 1.2 years. In column 8, we use the same sample and control variables, but we calculate the implied North-South difference by estimating a regression that replaces the North indicator with 1940 county of residence fixed effects for migrants and constructing averages, as in equation (6). The implied North-South difference is identical from this approach. This highlights the fact that it is possible to aggregate fixed effects to recover the previously-estimated moving North parameter.[59] Finally, columns 9 and 10 do not restrict the sample to children whose parents are born in the South, yielding extremely similar results when estimating a North indicator or aggregating fixed effects. The results in column 10 are the main point of interest because the associated model relies on cross-state migration as a source of identifying variation. Recall that the selection-correction approach from Section 3 also relies on these types of comparisons.

Notably, the implied North-South difference in column 10 is 1.2 years, which is considerably larger than the 0.8 year difference reported in Figure 4. To explain this gap, we turn to Appendix Table 7, where column 1 repeats the estimate from column 10 of Appendix Table 6. The implied North-South difference depends on estimated place effects and weights that reflect the number of observations in each county, as shown in equation (6). The table shows that results are extremely similar when using the full sample to construct observation weights and place effects for counties in the matched sample (columns 2 and 3). In contrast, adjusting for selection on unobservables matters considerably: the North-South difference in column 4 falls by 39 percent, from 1.2 years to 0.7 years. The similarity of the column 4 estimate to the full sample estimate in Figure 4 implies that our focus on the matched sample does not explain the discrepancy in the North-South difference.[60] The empirical Bayes adjustment, shown in column 5, leads to only a slight decrease in this difference.

In sum, Appendix Tables 6 and 7 highlight the importance of adjusting for selection on unob-

[59]In a regression without covariates, the North coefficient is equal to the difference in average place effects. With covariates, the equality need not be exact.

[60]The estimate in column 4 of 0.7 years is slightly smaller than the estimate of 0.8 years in Figure 4. This is because the full sample contains more counties than the matched sample.

servables. This is possible with our approach, which also allows us to examine heterogeneity in place effects across counties.

F Bounding Exercise to Account for Potential Mortality Effects

This section conducts a bounding analysis to account for selective survival of children. The motivation for this exercise is based on prior research that highlights the potential for migration from the rural South to the urban North during the early 20th century to increase Black infant mortality (Eriksson and Niemesh, 2016). Our approach computes bounds on county-level place effects that account for the fact that we can only estimate place effects on children who survive.

We begin by writing the true place effect as the weighted average between children who do and do not survive:

$$\gamma_j^* = p_j^{Die}\gamma_j^{Die} + (1 - p_j^{Die})\gamma_j, \tag{F1}$$

where γ_j^* is the true, unobserved place effect, and p_j^{Die} is the share of children whose parents moved to county j but died before aging into our sample, which contains individuals ages 14–18. The place effect among this group is γ_j^{Die}, while γ_j is the place effect in our sample of children observed in the 1940 Census, defined in Section 3.

The key challenge to evaluating equation (F1) is that we cannot estimate place effects for individuals who die before reaching our sample age criteria. However, we can construct an upper and lower bound for γ_j^* by making extreme assumptions about γ_j^{Die}. In particular, we assume that γ_j^{Die} is bounded from above by the maximum estimate of γ_j in our sample. We also assume that γ_j^{Die} is bounded from below by the minimum estimate of γ_j in our sample. This leads to upper and lower bounds:

$$\gamma_j^{UB} = p_j^{Die}\gamma^{UB} + (1 - p_j^{Die})\gamma_j \tag{F2}$$
$$\gamma_j^{LB} = p_j^{Die}\gamma^{LB} + (1 - p_j^{Die})\gamma_j. \tag{F3}$$

The ideal estimate of p_j^{Die} is the share of children whose parents move to county j and do not survive to age 14. Unfortunately, data to construct this estimate do not exist. Instead, we use infant mortality data from Bailey et al. (2018).[61] Infant mortality rates were considerably higher than child mortality rates, which could lead this approach to overstate the potential importance of mortality.

As an initial examination of the nature of selective mortality, Panel A of Appendix Figure 12 plots the infant mortality rate and our main place effect estimates. The infant mortality rate is lower in counties with higher place effects. However, this relationship is modest in size, as a one-year increase in place effects is associated with a 0.4 percentage point decrease in the infant mortality rate (whose average is 6.2 percent in our sample of counties). This correlation provides little reason to worry that our estimates of positive place effects stem mainly from higher mortality rates.

We summarize our results by calculating the average upper and lower bounds of place effects for counties in the South and North. Panel B of Appendix Figure 12 shows that the average

[61] We use county-level infant mortality rates calculated from 1933–1937. Note that 1933 is the first year where we can observe infant mortality rates for all counties in our sample.

bounds are relatively narrow. In the South, the migrant-weighted average upper and lower bounds are -0.36 and -0.66, respectively. In the North, the migrant-weighted average upper and lower bounds are 0.39 and 0.15, respectively. These estimates suggest that the effect of moving North is at least a 0.51-year increase in schooling, and no more than a 1.05-year increase. Given the conservative nature of these bounds, we view the similarity of our main estimate—a 0.83-year increase in schooling—as reassuring.

G Sources and Details for County-Level Measures

This appendix provides definitions and sources for the county-level measures used in our analysis.

G.1 Schooling

We create measures of historical schooling and school quality using a variety of sources. For 1940, we compute average years of schooling for non-migrant Black individuals (i.e., individuals in a household where the head still lives in his/her state of birth in 1940) ages 14–18 using the complete count Census.

We measure school quality for African Americans in 1940 using two different school resource data sets. For ten Southern states with segregated schools, we construct race-specific school quality variables using county-level data for the year 1939–1940 from Carruthers and Wanamaker (2019). For other states, we construct county-level school quality variables (for all races) for the year 1939–1940 using city-level data from Biennial Surveys of Education (U.S. Office of Education, 1947). These surveys contain data for cities with at least 10,000 residents, and we aggregate cities within a county.[62] As a result, we do not have data on counties where there is no city with at least 10,000 residents in 1940, and the data do not represent school quality in rural areas. We believe that this is a minor limitation, as 88 percent of African Americans in the North lived in urban areas in 1940. More Black students attended rural schools in the South, but the Carruthers and Wanamaker (2019) data cover these schools. We compute county-level averages for teacher salary, number of teachers per pupil, and term length. We impute variables using nearby years when necessary. Unfortunately, school resource data are not available for Florida for the relevant years. In addition to the states covered in Carruthers and Wanamaker (2019), schools were segregated in Delaware, the District of Columbia, Maryland, Missouri, Oklahoma, Virginia, and West Virginia. Since race-specific data are not available, we use Biennial Survey data for these states. We also construct an indicator variable for school segregation being required by law in 1940 by using Jim Crow laws by state from Sutherland (1955).

We also measure teachers per pupil in modern times. County-level data on teachers per pupil are not available in 1990. We use the NCES Common Core Data on teachers per pupil for 2000, the earliest academic year that is available and features nearly complete coverage for our sample of counties.[63]

[62] We use geographic crosswalks from U.S. Cities Database (n.d.) to match cities with counties by city and state name.

[63] For Massachusetts, Tennessee, and New York City boroughs, we use teacher pupil ratios from 2003, 2004 and 2005, respectively. Teacher pupil ratios were not reported for Buffalo, SD; Issaquena, MS; and Winkler, TX. For each of these counties, we use information available on the adjacent counties that were served by the same school district. Additionally, we use teacher pupil ratios from different years for several Virginia counties and independent

Finally, we also use the 1940 Census to create measures of high grade 9 enrollment at the county-level. This is defined based on the ratio of ninth to eighth grade enrollment for Black children ages 12 to 17. We create an indicator for high grade 9 enrollment based on whether the ratio of ninth to eighth grade enrollment is at least 0.5; our results are similar if we also define the threshold to be 0.25. This measure proxies for access to secondary schools. To the best of our knowledge, county-level data on the availability of secondary schools for Black children are not available.

G.2 Local Economic Conditions

We use the complete count 1940 Census from Ruggles et al. (2020) and summary files from the 1990 Census and 2005–2009 American Community Survey from Manson et al. (2019) to create measures of median household income, average earnings, manufacturing employment, the Gini index of income inequality, and the poverty rate.

We calculate median household income in 1940 as follows. We begin with the complete count Census data and remove all individuals in group quarters. We impute earnings for individuals who are self-employed.[64] We sum up all earned income at the household level and construct the county-level median. For 1990, we use Census summary files.

We construct average wage and salary income for non-migrant Black men in 1940 (i.e., individuals in a household where the head still lives in his/her state of birth in 1940). We restrict our sample to all non-migrant Black men ages 25–64 and drop individuals with missing income.

To calculate the manufacturing employment share in 1940, we first remove anyone who reports an industry that is "N/A," "Housework at home," "School response (students, etc.)," "Retired," or "Non-industrial response." We then classify as employed in manufacturing anyone with an industry code (1950 basis) that takes a value between 300 and 500. For 1990, we create the same measure using the Census summary files.

We construct poverty rates in the 1940 Census following Barrington (1997). We measure poverty at the family level instead of the household level.[65] We calculate family income using wage and salary income for wage earners and imputed income for the self-employed (as described above). We assign a 1939 poverty threshold to each family based on gender of household head, farm status, family size and number of children (Barrington, 1997, Table 1). We remove families with more than nine members as no poverty line was defined for larger families. We compute county-level poverty rates as the share of families whose income is below the corresponding

cities: Staunton City (1998); Charles City, King William, Lancaster, Williamsburg City (2001); Alleghany, Bedford City, Emporia City, Fairfax City (2002); and Clifton Forge City (2006).

[64]The 1940 Census contains only wage and salary earnings. To impute income for the self-employed, we use 1960 Census data from Ruggles et al. (2021) on individuals age 18–64 who are not currently enrolled in school, not in group quarters or on active military duty, and for whom occupation is not missing. We measure median earned income for each race (Black or White), region (of which there are four), and occupation (1950 basis) cell. If there are fewer than 10 observations in a region-race-occupation cell, we use median earned income by region and occupation. Then, we calculate where in the distribution of 1960 wage and salary income each median earnings value falls. Our earned income imputation equals the appropriate percentile of the 1940 wage and salary income distribution.

[65]We use the "famunit" variable for this purpose. For instance, we count each hired hand or servant and his/her family as a separate unit if they are not related to the head. We ensure that every member is related to each other in a family unit by using the "relate" and "sfrelate" variables. One exception is that we assign any individual in a single-member family unit to the primary family in the household if that individual is 14 or younger.

poverty line.[66]

To measure income inequality, we compute Gini coefficients. We use the complete count Census data for 1940. We begin with all family units that we defined to calculate poverty rates above. We restrict our sample to family units with children (defined as having at least one member age 14 or younger). Following Chetty et al. (2014), we compute the Gini coefficient in county j as:

$$\text{Gini}_j = \frac{2\text{Cov}(X_{ij}, P_{ij})}{\overline{X}_j}$$

where $\overline{X}_j$ is the mean family income in county j and $\text{Cov}(X_{ij}, P_{ij})$ is the covariance between family income (X_{ij}) and the percentile rank (P_{ij}) of family i in county j. We estimate $\text{Cov}(X_{ij}, P_{ij})$ by regressing percentile rank (P_{ij}) on family income (X_{ij}) for each county and multiplying the estimated coefficient by the variance of family income in county j. To measure income inequality for the later period, we use Census-produced Gini coefficients from the 2005–2009 American Community Survey.

G.3 Crime and Social Capital Measures

Ideally, we would measure homicides in 1940 to align with our other variables, but these data are not available.[67] Instead, we use annual homicide counts from the Vital Statistics of the United States for the years 1947–1950 (National Office of Vital Statistics, 1949, 1950, 1951, 1952).[68] Because homicides are rare in some counties, we construct the average homicide count over all available years for each county. In the denominator of homicides per capita, we use 1950 county population from Haines and Inter-university Consortium for Political and Social Research (2010). To measure homicide rates in 1990, we use the 1990 FBI Uniform Crime Reports (Federal Bureau of Investigation, 2016), which contain murders reported to police.[69]

We measure the number of lynchings per capita using data from Bailey et al. (2008), which contains information on all known lynchings for several Southern states (AL, AR, FL, GA, KY, LA, MS, NC and TN). We compute the county-level sum of lynchings during the period 1882–1929. We construct the final measure using total population as recorded in the 1940 Census.

We construct the incarceration rate in 1940 following Eriksson (2019). We start by classifying as incarcerated anyone reporting correctional institutions as their group quarter type (i.e., when the group quarters variable “gqtype” is equal to 2). We require any inmate to report a relationship to the household head that is either “institutional inmate” or “boarder/lodger.” To account for inconsistent reporting of group quarter type, we also keep any “institutional inmate” with a group quarter that is not a correctional institution. For 1990, we use the Incarceration Trends dataset from the Vera Institute of Justice (2015). We add up the jail and prison admissions that originate in each county, dividing by 1990 population.

[66]Ross, Danziger and Smolensky (1987) describe an alternative approach to measuring poverty in 1939 (see also Barrington (1997)). Our results are extremely similar when we use this approach.

[67]Homicide counts from FBI Uniform Crime Reports are available in 1940, but these data cover only large cities during this period.

[68]Here we also use U.S. County and City Data Book Consolidated File (U.S. Bureau of the Census, 2012) as a crosswalk between county names and county fips codes.

[69]Since the crime data is available at the agency level, we also use data from National Archive of Criminal Justice Data (2007), which provides a crosswalk between agencies and counties.

To measure residential segregation in 1940, we use the segregation index developed by Logan and Parman (2017*b*). They use information on the race of next-door neighbors to assess the amount of residential segregation relative to scenarios with complete segregation and no segregation. A key advantage of their approach is that it can be used in rural areas. For segregation in 1990, we construct a Theil index using tract-level data on the share of population that is White, Black, Hispanic, and another race.

We also use county-level data on the presence of National Association for the Advancement of Colored People (NAACP) chapters that come from Gregory and Estrada (2019). Compiled from NAACP annual reports and the branch bulletins, this database shows the spread of NAACP branches between 1912 and 1964. We use the year a local branch was first mentioned in the database to create a measure of whether a county had a local NAACP chapter by 1940, 1950, and 1960.

G.4 Demographic Measures

We use the complete count 1940 Census from Ruggles et al. (2020) and summary files from the 1990 Census (Manson et al., 2019) to create measures of population density, percent in an urban area, percent on a farm, and percent of the population that is Black.

H Place Effects on Earnings for Black Adult Migrants

In this appendix, we explore place effects for adults using our selection correction approach from Section 3. Our analysis is motivated by the fact that place effects for parental income could be a key channel that drives place effects on schooling for children. We create a sample of Black men ages 25–64 from the 1940 Census and estimate impacts on log earnings.[70] For the selection correction approach, we include fixed effects for age in the vector of demographic variables X_i. The key vector that measures selection on observables, H_i, contains fixed effects for an individual's years of schooling, marital status, and number of own children in the household.

Our main finding is that our selection-corrected estimates imply that there were notable labor market benefits for Black men who moved to locations in the North. Similar to our main analysis for children, Appendix Figure 13 displays separate densities for county-level place effect estimates on men's log earnings in the North and South. We estimate that there was a 42 percent increase in earnings from moving North. As noted in Section 5, the place effects on adult earnings are strongly related to place effects on children's education (correlation: 0.59), which suggests that much of the relationship with median Black family income is driven by earnings gains available to adult migrants.

While we find that migration North led to substantial increases in adults' earnings, our estimate of a 42 percent increase in earnings is smaller than the evidence presented in Collins and Wanamaker (2014) and Boustan (2017), who find gains of 80–130 percent. We differ from this prior work by estimating selection-corrected county-level place effects and by examining a broader age range. By comparison, these previous papers estimate regressions of log earnings on an indicator

[70] The 1940 Census measures wage and salary income, but not total earnings (which also includes self-employment income). We impute earned income for self-employed individuals based on their race, region, and occupation, as detailed in Appendix G.

for living in the North and various controls. Collins and Wanamaker (2014) examine a sample of men ages 21–40 in the 1930 Census, which does not contain direct measures of earnings and so requires the use of earnings imputed by occupation. Boustan (2017) uses a sample of men ages 18–38 in the 1940 Census. We focus on a broader group of 25–64-year-old men than these papers to obtain a larger sample size (which is helpful for our estimation of county-level place effects) and to better represent the fathers of children in our main analysis sample.

We examine potential explanations for why we find a smaller effect of moving North in Appendix Tables 8 and 9. Column 1 of Appendix Table 8 shows that earnings of Black men in our adult sample were 98 percent larger in the North, consistent with the results in Collins and Wanamaker (2014) and Boustan (2017). The North-South earnings gap narrows to 75 percent when controlling for age, education, marital status, number of children, and origin state, which provides initial evidence of the potential for selective migration. When additionally limiting the sample to destination counties with at least 10 migrants (which we use to increase the reliability of our place effect estimates), the North-South gap falls to 61 percent. The remaining columns of Appendix Table 8 show that estimates are similar when controlling for origin county fixed effects (column 4), limiting the sample to brothers and including 1920 household fixed effects (columns 5–6), and using the same controls for observed variables as in the selection correction approach (columns 7–10).

Appendix Table 9 examines the importance of adjusting for selection on unobserved characteristics. Column 3 shows that the average effect of moving to the North based on county-level place effects when not adjusting for selection on unobservables is 56 percent. Adjusting for selection lowers this earnings gain to 42 percent.Thus, adjusting for selection on unobservables leads to a 25 percent decrease in the effect of moving North ($= (0.42 - 0.56)/0.56$).

In sum, a substantial amount of the difference between our bottom-line estimate of a 42 percent earnings gain from moving North and the 80–130 percent estimate from prior work appears to be explained by controlling for observed variables (in particular, education) and focusing on a subset of counties for which there is a sufficiently large sample of migrants that we can feasibly estimate place effects. A smaller, but still significant, share of the difference is explained by adjusting for selection on unobserved variables.

Appendix Table 1: Correlates of 1940 Place Effects on White Children's Education

	Dependent Variable: Place effect, children's education		
	Bivariate Regressions	Multivariate Regressions	
	(1)	(2)	(3)
Teachers per pupil	0.253***	0.219***	0.210***
	(0.0475)	(0.0476)	(0.0497)
Median White household income	-0.0615*	-0.210***	-0.213***
	(0.0338)	(0.0400)	(0.0403)
Homicide rate	-0.218***	-0.191***	-0.179***
	(0.0434)	(0.0434)	(0.0488)
Incarceration rate	-0.0505	-0.0550	-0.0551
	(0.0447)	(0.0444)	(0.0446)
NAACP chapter	0.0807**	0.0902**	0.0827*
	(0.0368)	(0.0405)	(0.0423)
South indicator			-0.0655
			(0.0997)
Observations (counties)	715	715	715
R-squared	–	0.108	0.108

Notes: This table reports correlates of 1940 place effects for White children's education. Sample is limited to the counties for which we estimate place effects for both Black and White children. We normalize all variables to have mean zero and standard deviation one. All regressions include a series of indicators for whether variables are missing. Heteroskedasticity-robust standard errors in parentheses. Column 1 reports estimates of separate bivariate regressions for each explanatory variable. Columns 2–3 report estimates of multivariate regressions. See Appendix G for details on variable construction and sources. Statistical significance is denoted by: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.
Source: Authors' calculations using 1940 Census (Ruggles et al., 2020)

Appendix Table 2: Correlation of Upward Mobility in 1990s and Place Effects in 1940

	DV: Upward mobility, 1990s		
	Black upward mobility (1)	Pooled upward mobility (2)	Exposure effects (3)
Place effect, 1940	0.210***	0.432***	0.304***
	(0.0340)	(0.0330)	(0.0348)
Observations (counties)	728	728	728
R-squared	0.045	0.186	0.093

Notes: Table reports correlations between measures of upward mobility from the 1990s and place effects from the 1940 Census. Columns 1 and 2 use the mean household income rank for children whose parents were at the 25th percentile of the national income distribution from Chetty et al. (2020). Column 1 uses upward mobility for Black children, and column 2 uses upward mobility for children of all races. Column 3 uses exposure effects from Chetty and Hendren (2018*b*). We standardize place effect estimates and the upward mobility measure so that normalized measures have a mean of 0 and a standard deviation of 1. As a result, point estimates in this table are correlation coefficients. Heteroskedasticity-robust standard errors in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.
Source: Authors' calculation using the 1940 Census (Ruggles et al., 2020), Chetty and Hendren (2018*b*), and Chetty et al. (2020)

Appendix Table 3: Place Effects and Mechanisms, Within-Place Estimates, Robustness Using Exposure Effect Measure

	Dependent Variable: Δ Opportunity Measure (1990s vs 1940)			
	Bivariate Regressions	Multivariate Regressions		
	(1)	(2)	(3)	(4)
Δ Teachers per pupil	0.156***	0.142***	0.111***	0.107***
	(0.0322)	(0.0321)	(0.0295)	(0.0338)
Δ Median Black household income	0.229***	0.201***	0.158***	0.156***
	(0.0456)	(0.0410)	(0.0388)	(0.0412)
Δ Homicide rate	-0.208***	-0.179***	-0.0717**	-0.0701**
	(0.0353)	(0.0345)	(0.0346)	(0.0354)
Δ Incarceration rate	-0.106***	-0.112***	-0.0991***	-0.101***
	(0.0278)	(0.0276)	(0.0263)	(0.0285)
Δ NAACP chapter	0.153	0.00857	-0.0287	-0.0288
	(0.0943)	(0.0919)	(0.0904)	(0.0906)
Δ Percent Black	-0.915***		-0.725***	-0.721***
	(0.0758)		(0.0805)	(0.0832)
South indicator				0.0216
				(0.111)
Observations (counties)	728	728	728	728
R-squared	–	0.147	0.230	0.230

Notes: Separately for each year, we normalize all variables to have mean zero and standard deviation one. We then construct the change from 1940 to the 1990s, except for the change in the presence of a NAACP chapter, which is from 1940 to 1960. The dependent variable is the difference between exposure effects from Chetty and Hendren (2018*b*) and place effects in 1940. All regressions include a series of indicators for whether variables are missing. Heteroskedasticity robust standard errors in parentheses. Column 1 reports estimates of separate bivariate regressions for each explanatory variable. Columns 2–4 report estimates of multivariate regressions. See Appendix G for details on variable construction and sources. Statistical significance is denoted by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Source: Authors' calculations using 1940 Census (Ruggles et al., 2020) and Chetty et al. (2020)

Appendix Table 4: Comparing the Full Population and the Matched Sample of Adults in the 1920 Census

	1920 U.S.-born Black men age 3–52	Matched sample subset	Difference
	(1)	(2)	(3)
Age	21.366	19.537	1.830***
Urban status	0.250	0.267	-0.016***
Farm status	0.554	0.548	0.006***
Number of siblings	1.980	2.215	-0.235***
Literate	0.765	0.808	-0.043***
School attendance	0.447	0.460	-0.013***
Father's Duncan Index	13.782	14.036	-0.254***
Father's literacy	0.686	0.724	-0.038***
Father's farmer status	0.689	0.651	0.038***
North	0.130	0.178	-0.048***
South	0.870	0.822	0.048***
Observations	3,425,187	501,284	–

Notes: Column 1 reports summary statistics for all U.S.-born Black men who were age 3–52 in 1920 and living in their birth state. Column 2 contains the subset of these men in the matched sample. Column 3 reports the difference between these columns, with stars indicating statistical significance based on heteroskedasticity-robust standard errors. * $p < 0.10$ ** $p < 0.05$, *** $p < 0.01$

Source: Authors' calculation using the 1920 and 1940 Census (Ruggles et al., 2020)

Appendix Table 5: Comparing the Full Population and Matched Sample of Children in the 1940 Census

	1940 Southern-born Black men age 25–70 who have children age 14–18 (1)	Matched sample subset, age 25–70 (2)	Difference (3)	1940 children age 14–18 of Southern-born Black men (4)	Matched sample subset of children (5)	Difference (6)
Years of schooling	4.509	4.782	-0.273***	6.671	6.888	-0.216***
Earnings	476.186	503.369	-27.184***	97.605	93.372	4.233**
Age	46.475	46.822	-0.347***	15.888	15.895	-0.007
Urban status	0.444	0.449	-0.005	0.419	0.424	-0.004
Farm status	0.369	0.366	0.003	0.396	0.394	0.002
Married	0.912	0.921	-0.009***	0.007	0.007	0.000
North	0.286	0.314	-0.028***	0.276	0.303	-0.027***
South	0.714	0.686	0.028***	0.724	0.697	0.027***
Observations	116,366	24,148	–	179,335	37,623	–

Notes: Column 1 reports summary statistics for all Southern-born Black men age 25–70 in the 1940 Census who have children between age 14 and 18. Column 2 contains the subset of these men in the matched sample. Column 4 contains children age 14–18 of Southern-born Black men in column 1, and column 5 contains the matched sample subset. Columns 3 and 6 report the difference between these columns, with stars indicating statistical significance based on heteroskedasticity-robust standard errors.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' calculation using the 1920 and 1940 Census (Ruggles et al., 2020)

Appendix Table 6: Comparison of North Indicator to Place Effects, Black Children's Educational Attainment

	DV: Years of schooling									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
North indicator	2.180***	1.411***	1.008***	1.007***	1.064***	0.369	1.198***		1.210***	
	(0.098)	(0.087)	(0.120)	(0.109)	(0.123)	(0.335)	(0.124)		(0.124)	
Implied North-South difference from place effects								1.198		1.207
Observations	105,347	105,347	41,092	41,092	26,082	26,082	41,092	41,092	46,867	46,867
Sample										
Household (HH) head born in South	X	X	X	X	X	X	X	X		
In destination county with at least 10 migrants			X	X	X	X	X	X	X	X
At least 2 children with same 1920 family					X	X				
Controls										
Age, sex, parents' education		X	X	X	X	X	X	X	X	X
HH head 1920 state FE		X	X	X	X	X	X	X	X	X
HH head 1920 county FE				X						
HH head 1920 family FE						X				
HH head 1920 state FE $\times$ mover indicator							X	X	X	X
1940 county FE $\times$ non-mover indicator							X	X	X	X

Notes: The first row reports results from regressing years of schooling on an indicator for living in the North and controls. Standard errors are clustered by 1940 county of residence. The implied North-South difference from place effects comes from a regression that replaces the North indicator with county fixed effects for migrants. We calculate the difference between the average fixed effects in the North and the average fixed effects in the South, where each average is constructed using weights equal to the number of migrants in each county, as in equation (6). Sample contains African Americans age 14–18. Statistical significance is denoted by: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Source: Authors' calculation using the 1920 and 1940 Census (Ruggles et al., 2020)

Appendix Table 7: Comparison of Estimated North-South Differences Across Samples and Adjustments, Black Children's Educational Attainment

	DV: Years of schooling				
	(1)	(2)	(3)	(4)	(5)
Implied North-South difference from place effects	1.207	1.207	1.205	0.736	0.718
Place effects sample	Matched	Matched	Full	Full	Full
Observation weight sample	Matched	Full	Full	Full	Full
Selection correction				X	X
Empirical Bayes adjustment					X

Notes: Table reports the difference between the average fixed effects in the North and the average fixed effects in the South, where each average is constructed using weights equal to the number of migrants in each county as in equation (6). We estimate place effects and measure the number of migrants using the matched sample and full sample (for the latter, focusing on counties for which fixed effects are estimated in the matched sample). Column 4 uses selection-corrected place effects, and column 5 further uses empirical-Bayes-adjusted effects. For all columns, we use the same specification as in column 10 of Appendix Table 6.

Source: Authors' calculation using the 1920 and 1940 Census (Ruggles et al., 2020)

Appendix Table 8: Comparison of North Indicator to Place Effects, Black Adult Log Earnings

	DV: Log wage and salary earnings									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
North indicator	0.683***	0.559***	0.477***	0.473***	0.484***	0.506***	0.425***		0.424***	
	(0.030)	(0.029)	(0.030)	(0.028)	(0.031)	(0.028)	(0.032)		(0.032)	
Implied North-South difference from place effects								0.427		0.425
Observations	292,978	288,993	204,122	204,122	49,939	49,939	204,122	204,122	239,704	239,704
Sample										
Born in South	X	X	X	X	X	X	X	X		
In destination county with at least 10 migrants			X	X	X	X	X	X	X	X
At least 2 adult men with same 1920 family					X	X				
Controls										
Age, education, marital status, children		X	X	X	X	X	X	X	X	X
1920 state FE		X	X	X	X	X	X	X	X	X
1920 county FE				X						
1920 family FE						X				
1920 state FE $\times$ mover indicator							X	X	X	X
1940 county FE $\times$ non-mover indicator							X	X	X	X

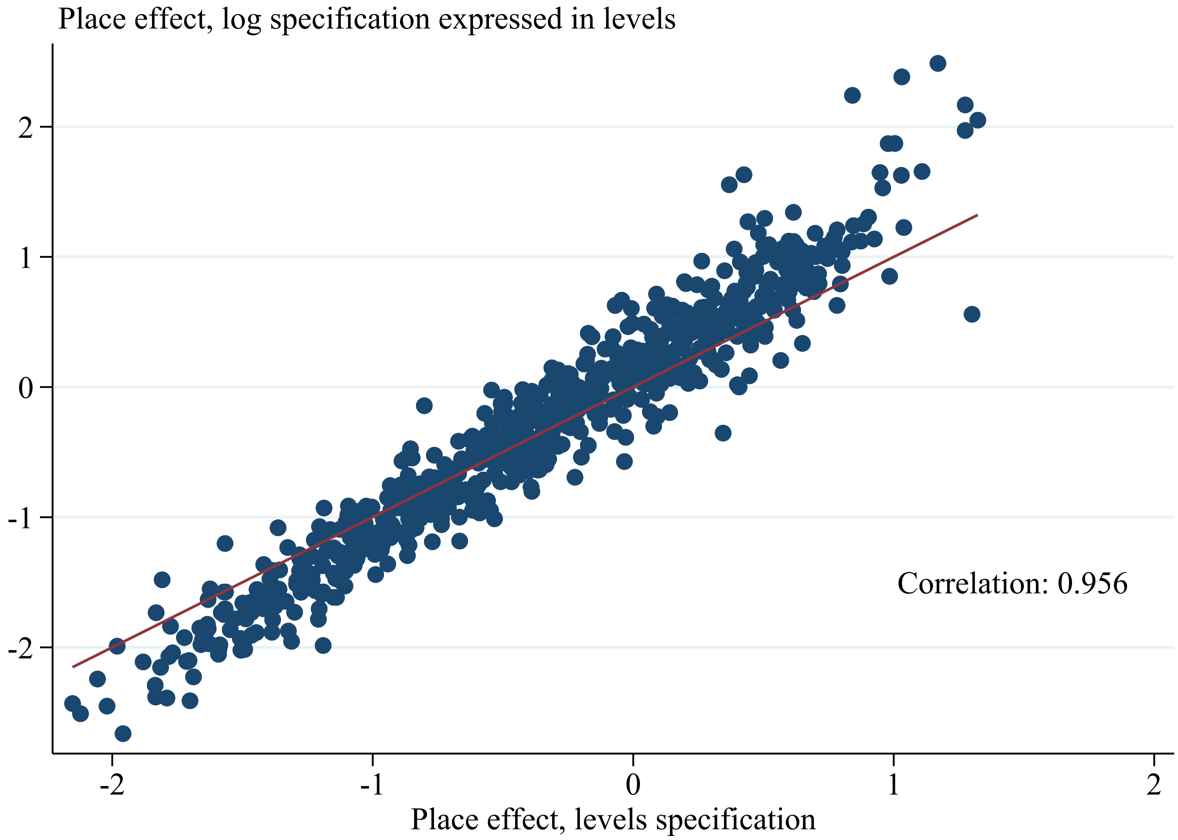
Notes: The first row reports results from regressing log earnings on an indicator for living in the North and controls. Standard errors are clustered by 1940 county of residence. The implied North-South difference from place effects comes from a regression that replaces the North indicator with county fixed effects for migrants. We calculate the difference between the average fixed effects in the North and the average fixed effects in the South, where each average is constructed using weights equal to the number of migrants in each county, as in equation (6). Sample contains African American men age 25–64.
Source: Authors' calculation using the 1920 and 1940 Census (Ruggles et al., 2020)

Appendix Table 9: Comparison of Estimated North-South Differences Across Samples and Adjustments, Black Adult Log Earnings

	DV: Log wage and salary earnings				
	(1)	(2)	(3)	(4)	(5)
Implied North-South difference from place effects	0.425	0.425	0.442	0.348	0.346
Place effects sample	Matched	Matched	Full	Full	Full
Observation weight sample	Matched	Full	Full	Full	Full
Selection correction				X	X
Empirical Bayes adjustment					X

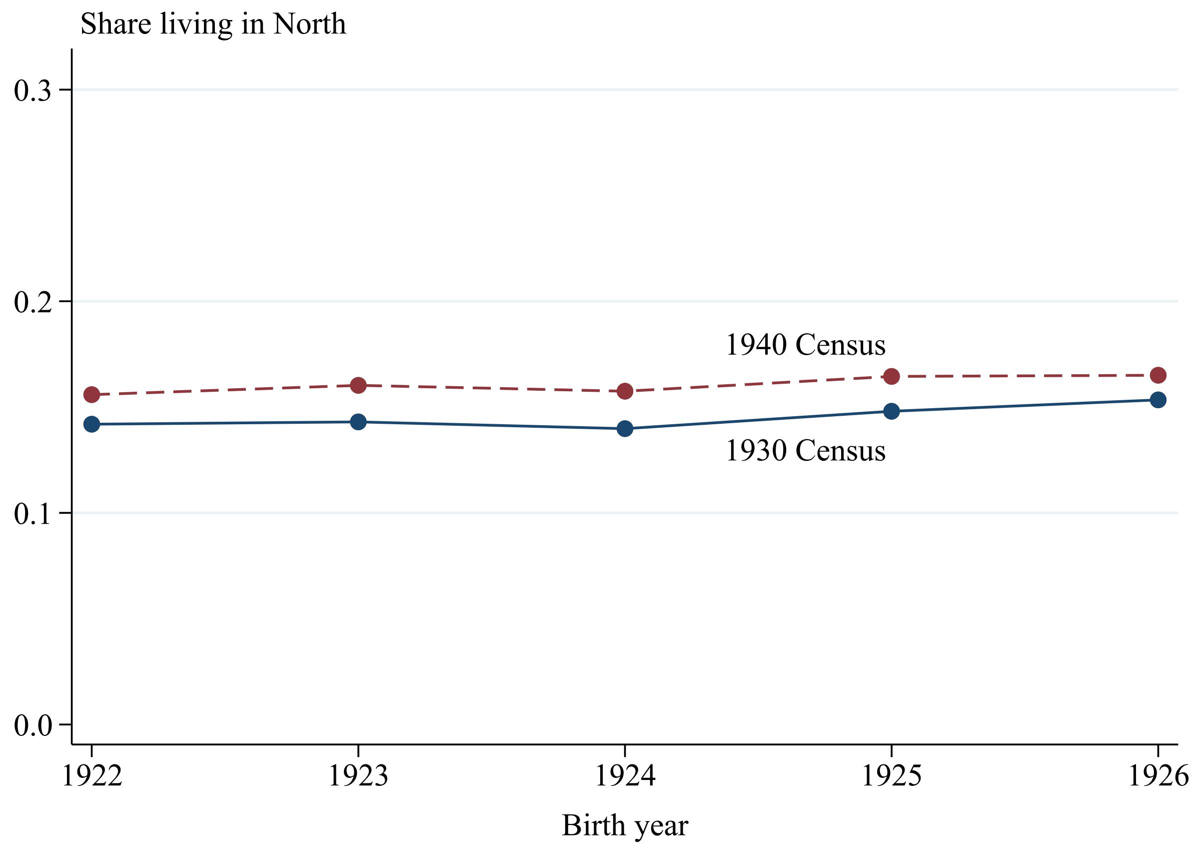
Notes: Table reports the difference between the average fixed effects in the North and the average fixed effects in the South, where each average is constructed using weights equal to the number of migrants in each county as in equation (6). We estimate place effects and measure the number of migrants using the matched sample and full sample (for the latter, focusing on counties for which fixed effects are estimated in the matched sample). Column 4 uses selection-corrected place effects, and column 5 further uses empirical-Bayes-adjusted effects. For all columns, we use the same specification as in column 10 of Appendix Table 8.
Source: Authors' calculation using the 1920 and 1940 Census (Ruggles et al., 2020)

Appendix Figure 1: Comparison of Place Effects from Levels vs. Log Specification



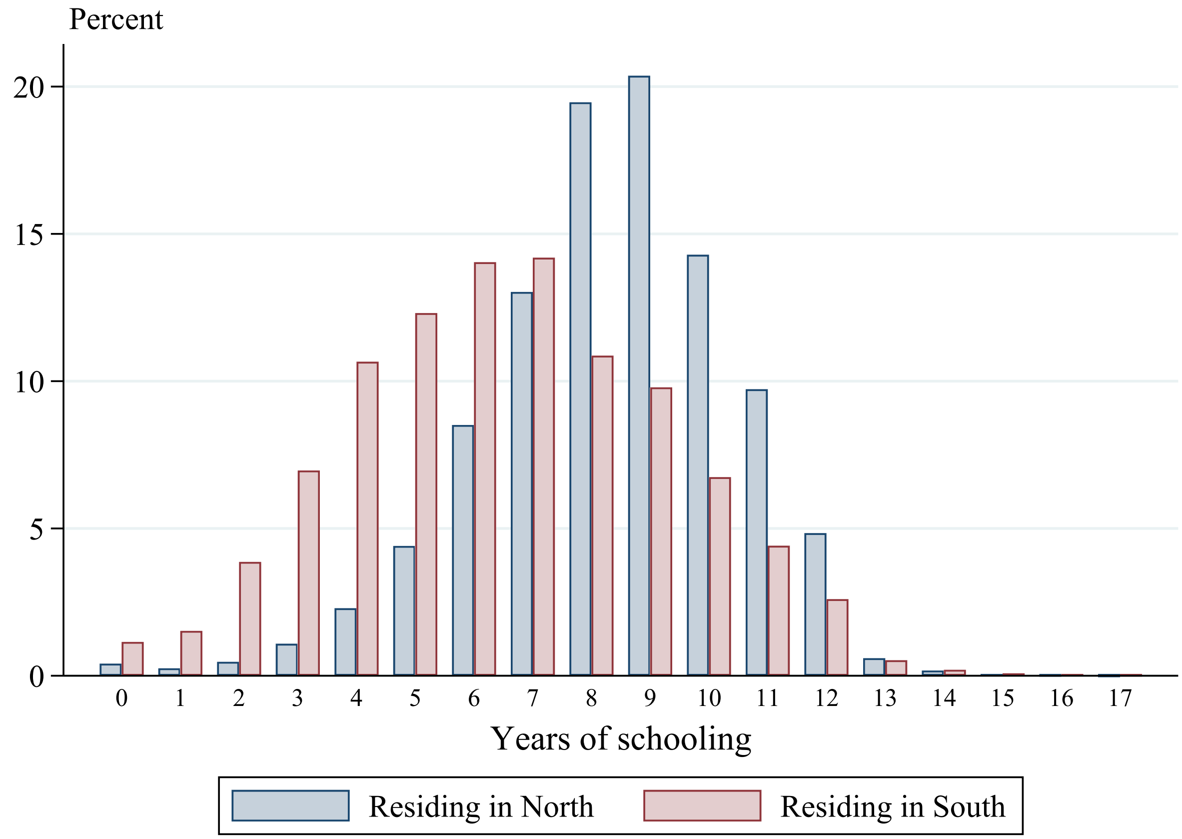
Notes: Figure displays empirical-Bayes-adjusted place effects from our baseline specification (x-axis) against an alternative specification (y-axis) that estimates place effects on log years of schooling and converts to levels using the formula: place effect in levels $= [\exp(\text{place effect in logs}) - 1] \times$ mean. For the alternative specification, we estimate the empirical-Bayes-adjusted place effects using the log place effects, and then convert both the adjusted and unadjusted log place effects separately into level place effects. We calculate correlations using non-empirical-Bayes-adjusted place effects.
Source: Authors' calculations using 1940 Census (Minnesota Population Center and Ancestry.com, 2013).

Appendix Figure 2: Summary Statistics on Northern Migration in the 1930 and 1940 Censuses



Notes: Figure displays the share of each birth cohort (x-axis) that is living in the North in the 1930 (navy, solid) and 1940 (maroon, dashed) Censuses. In the 1930 Census, the overall average share living in the North is about 15 percent. In the 1940 Census, the overall average share living in the North is 16 percent. Samples contain Black children born from 1922 to 1926 who are living in a household where the head was born in the South.
Source: Authors' calculations using the 1930 and 1940 Censuses (Ruggles et al., 2020)

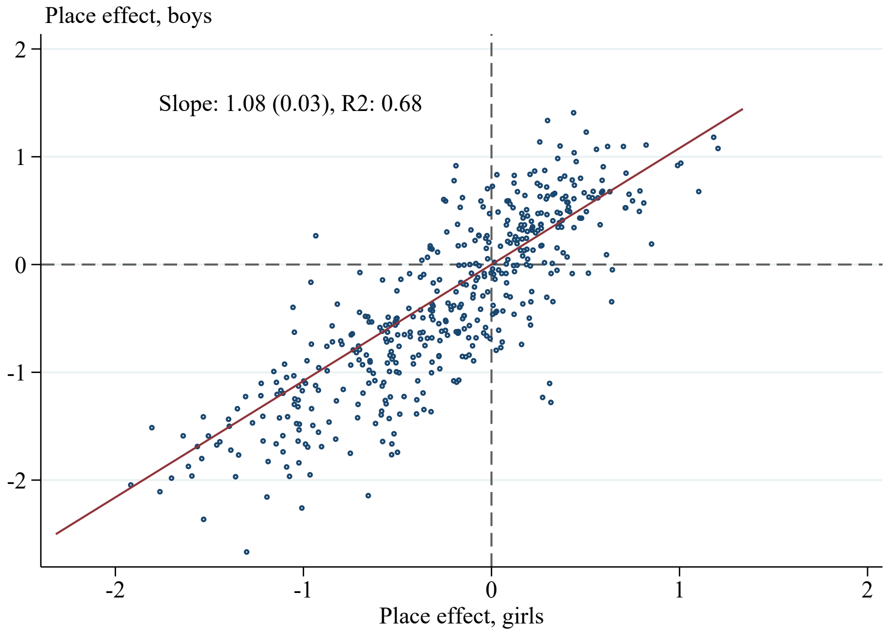
Appendix Figure 3: Educational Attainment by 1940 Place of Residence, Black Children Age 14–18 with Migrant Parents



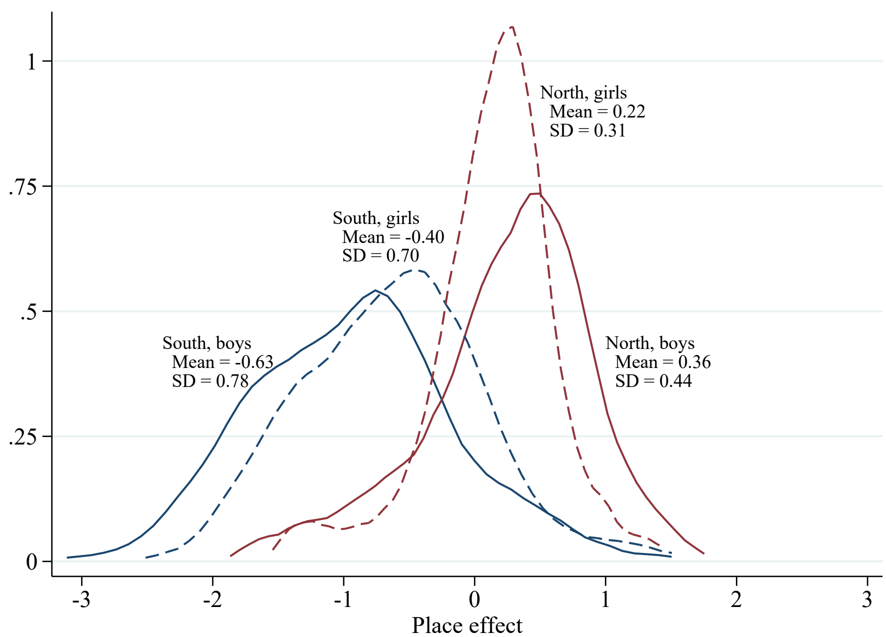
Notes: Sample contains Black children age 14–18 whose household head was born in the South and lives outside the head's birth state in 1940.
Source: Authors' calculations using 1940 Census (Ruggles et al., 2020)

Appendix Figure 4: Comparison of Place Effects by Sex, Black Children Age 14–18

(a) Bivariate Relationship

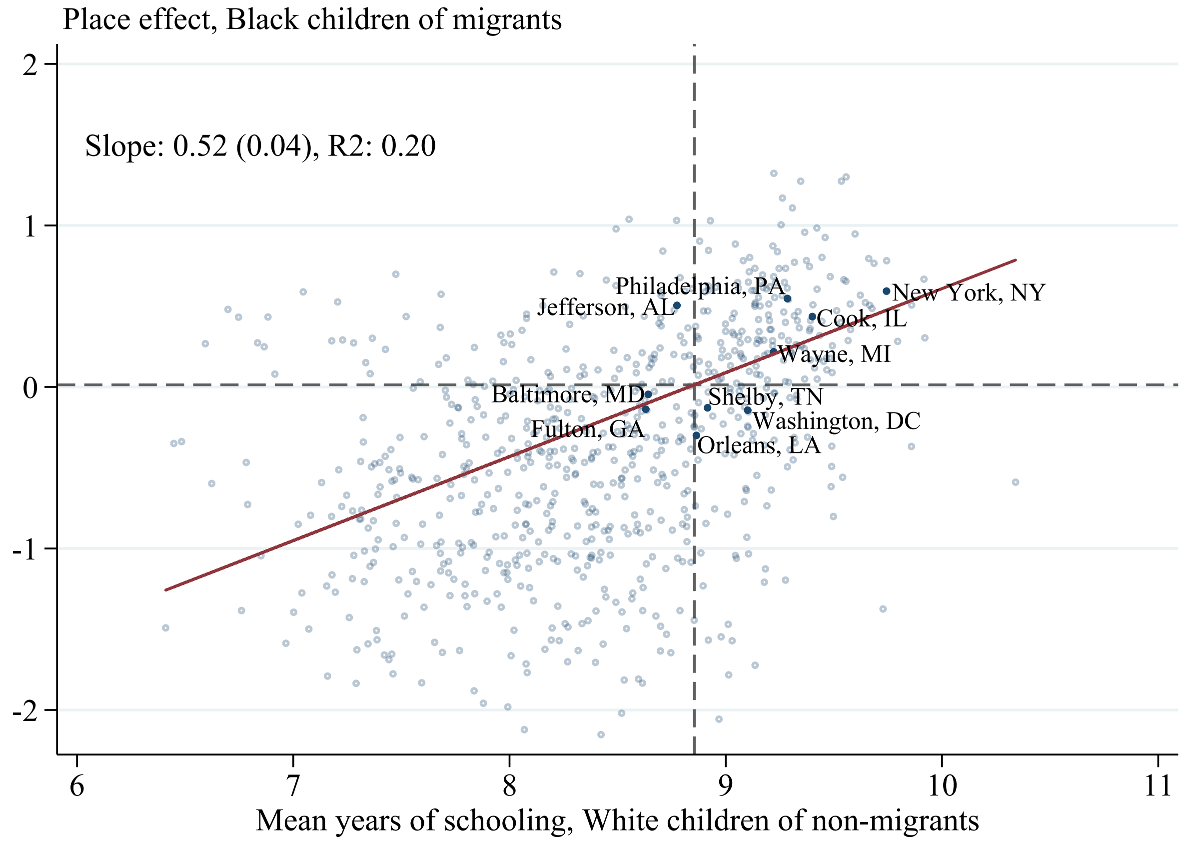


(b) Density of Place Effects, by Region



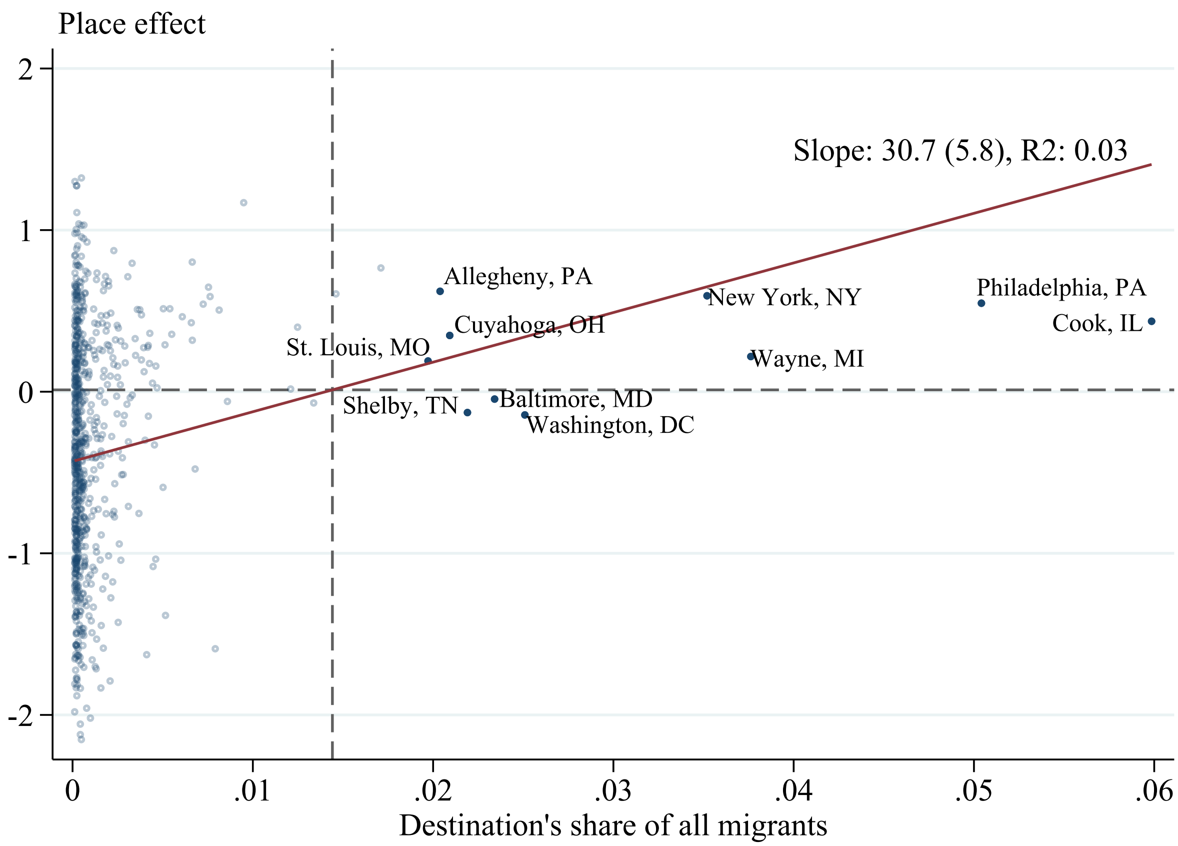
Notes: Panel A displays empirical-Bayes-adjusted place effects for boys and girls age 14–18 in 1940. Dashed lines are migrant-weighted averages. To estimate the line of best fit, we use non-empirical-Bayes-adjusted place effects. Panel B shows the density of place effects in the South and North, alongside migrant-weighted averages and standard deviations.
Source: Authors' calculations using 1940 Census (Ruggles et al., 2020)

Appendix Figure 5: Place Effects for Black Children versus Average Years of Schooling for White Non-Migrants, Ages 14–18



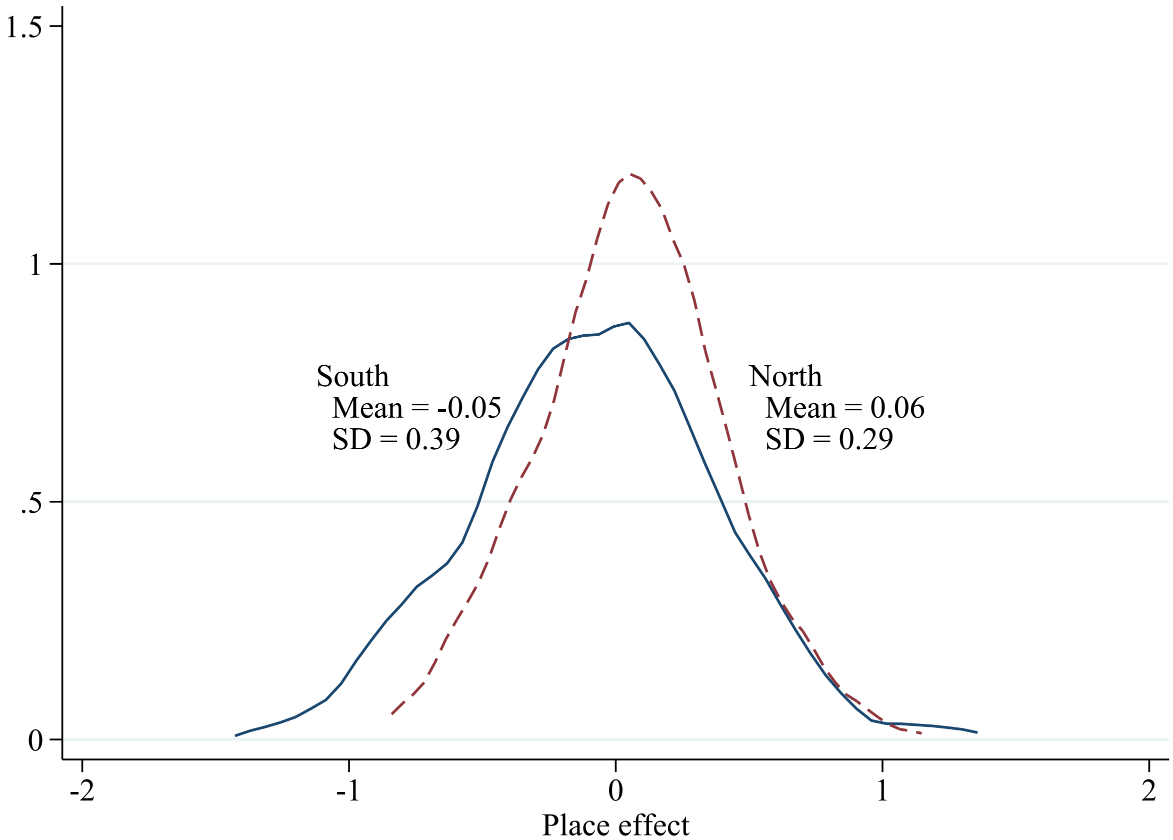
Notes: Figure displays empirical-Bayes-adjusted place effects for Black children against average years of schooling for White non-migrants. Dashed lines are migrant-weighted averages (0.00 and 8.86). The ten largest counties in terms of 1940 Black population are labeled. To estimate the line of best fit, we use non-empirical-Bayes-adjusted place effects as the dependent variable.
Source: Authors' calculations using the 1940 Census (Minnesota Population Center and Ancestry.com, 2013).

Appendix Figure 6: Place Effects versus Share of Migrants in Destination, Black Children Age 14–18



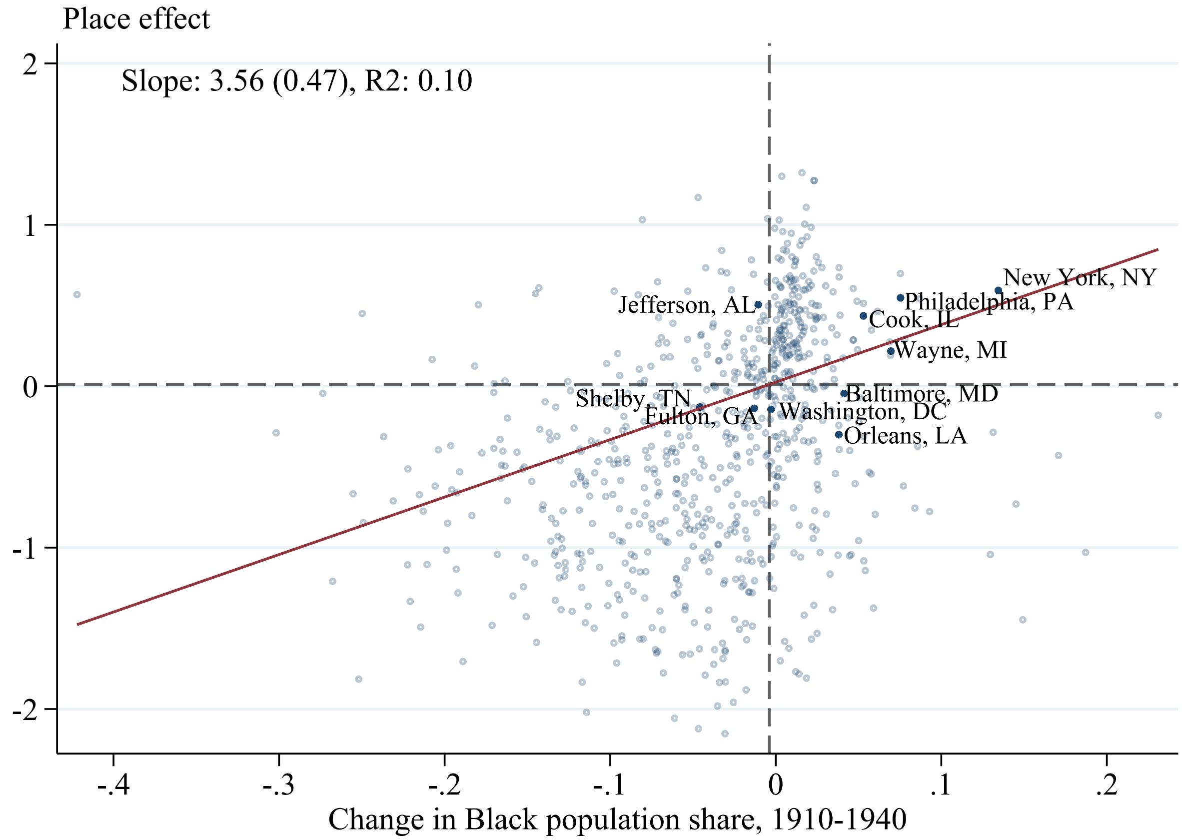
Notes: Figure displays empirical-Bayes-adjusted place effects against the share of children of migrants in each destination. Dashed lines are migrant-weighted averages (0.00 and 0.01). The ten largest counties in terms of migrant child share are labeled; these counties contain 31.5 percent of all migrant children.
Source: Authors' calculations using 1940 Census (Ruggles et al., 2020)

Appendix Figure 7: Distribution of Place Effects on Years of Schooling in South and North, White Children Age 14–18



Notes: Figure shows density of place effect estimates in the South and North for White children age 14–18 whose household head was born in the South. Migrant-weighted averages and standard deviations are reported.
Source: Authors' calculations using 1940 Census (Ruggles et al., 2020)

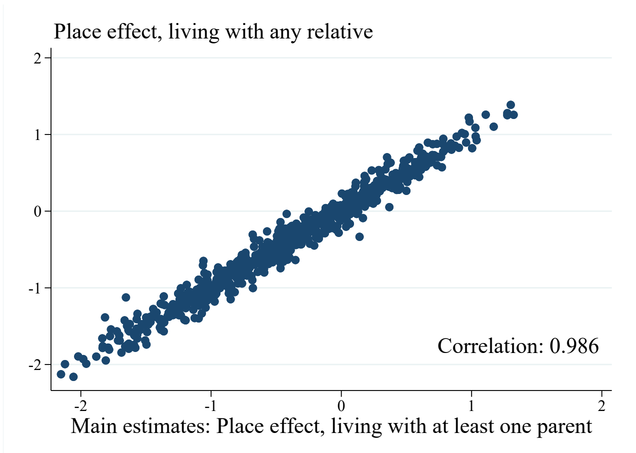
Appendix Figure 8: Place Effects versus Change in Black Population Share from 1910–1940



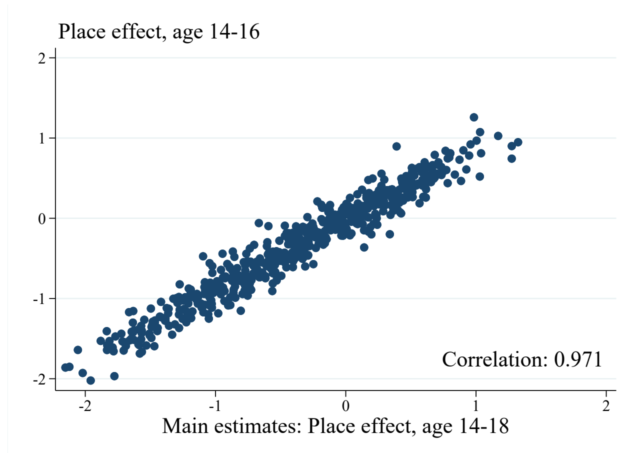
Notes: Figure displays empirical-Bayes-adjusted place effects against the change in the Black population share from 1910–1940. Dashed lines are migrant-weighted averages (0.00 and -0.004). The ten largest counties in terms of 1940 Black population are labeled.
Source: Authors' calculations using 1940 Census (Ruggles et al., 2020)

Appendix Figure 9: Robustness to Sample Selection and Censoring

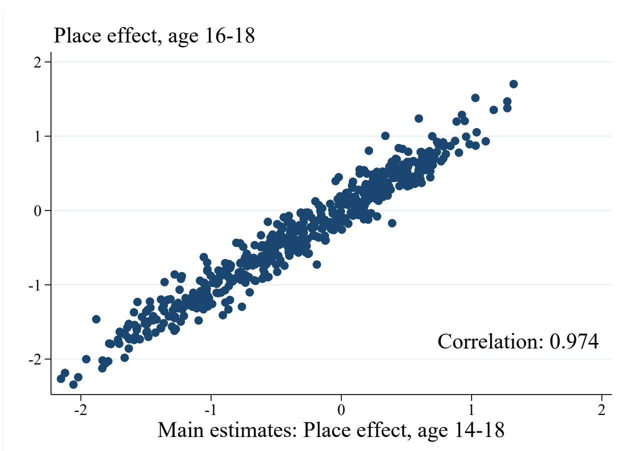
(a) Robustness to Sample Selection: Parental Co-Residence



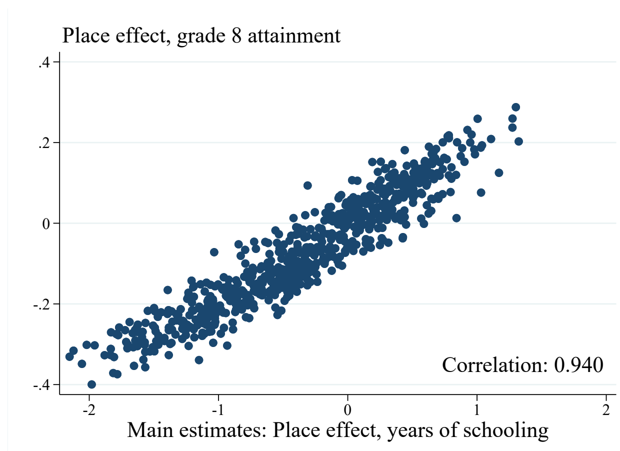
(b) Robustness to Sample Selection: Age



(c) Robustness to Censoring: Age

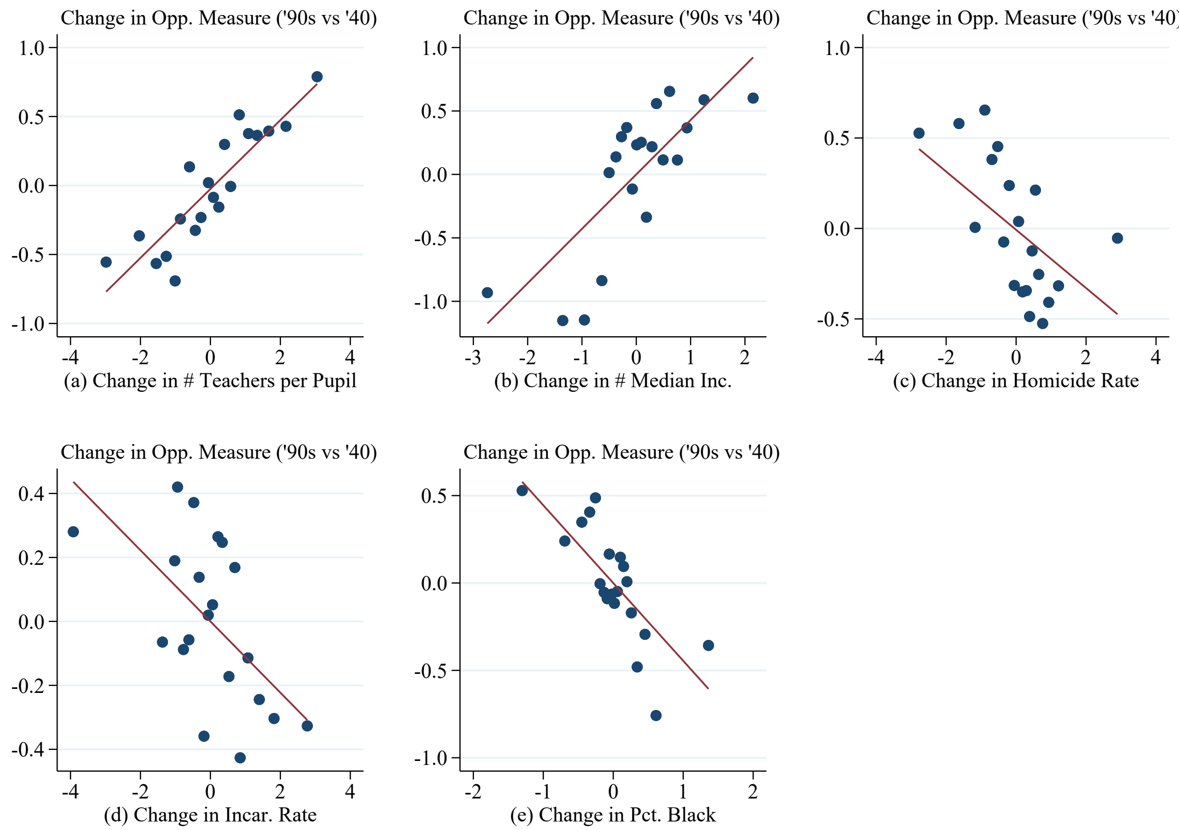


(d) Robustness to Censoring: Dependent Variable



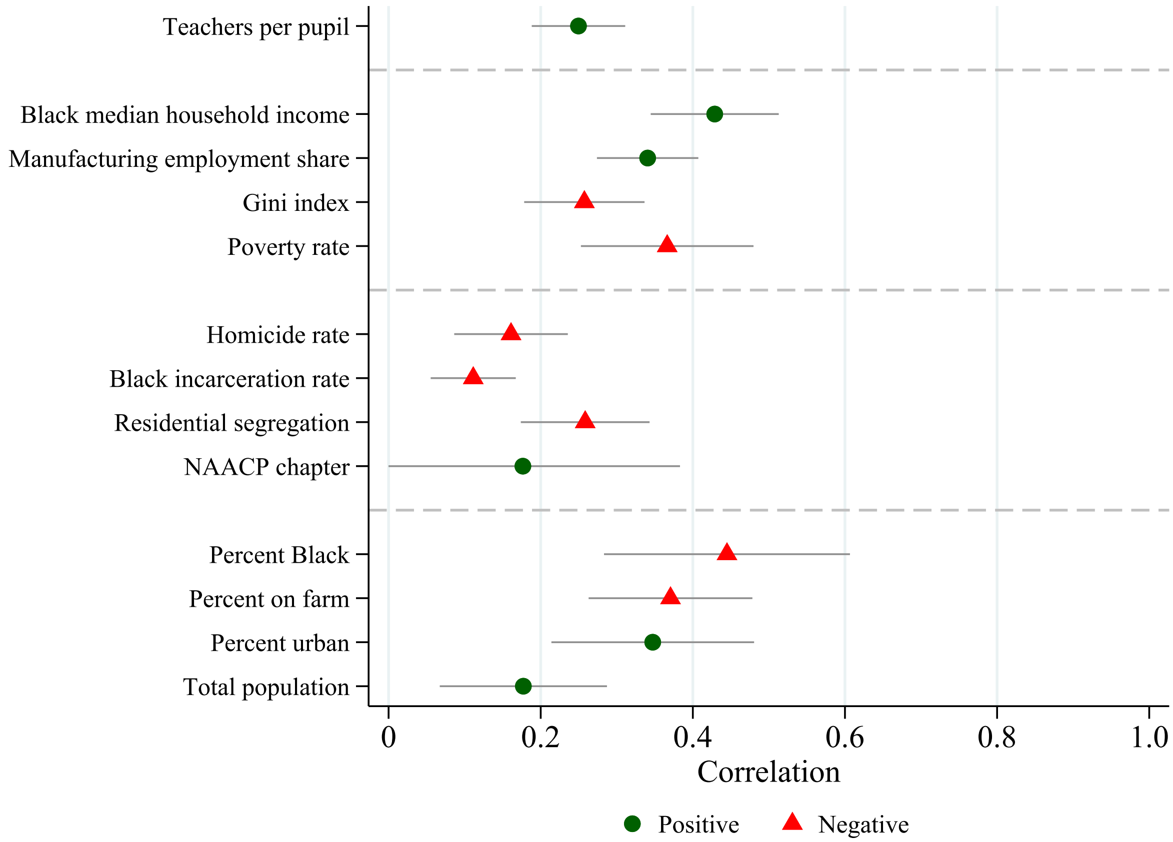
Notes: Figure displays empirical-Bayes-adjusted place effects for Black children. Across all panels, our main estimates (for years of education of children ages 14–18 that live with at least one parent) are shown on the x-axis. The y-axis in Panel A displays place effects for children ages 14–18 that live with any relative. Panel B shows results for children ages 14–16 that live with at least one parent. Panel C shows results for children ages 16–18 that live with at least one parent. Panel D plots place effects for grade 8 attainment among our main sample.
Source: Authors' calculations using 1940 Census (Ruggles et al., 2020)

Appendix Figure 10: Place Effect Mechanisms, Within-Place Estimates, Binned Scatterplot



Notes: Figure displays the relationship between the change in opportunity measures and the change in place characteristics. Each observation in the plot represents the average change within binned values of the x and y axis. We group the data into 20 equally-sized bins. The 1940 measure of place effects is based on our analysis of the 1940 Census. For a contemporary opportunity measure, we use the upward mobility measure from Chetty et al. (2020). Upward mobility is the mean household income rank for children whose parents were at the 25th percentile of the national income distribution. This statistic is calculated for children born between 1978 and 1983, who grew up during the 1990s. Both measures of opportunity are empirical-Bayes-adjusted. We normalize all variables to have a standard deviation of one and a mean of zero. We compute the change for each standardized variable between the contemporary and historical periods. The construction of measures of place characteristics is described in Appendix G.
Source: Authors' calculations using 1940 Census (Ruggles et al., 2020) and Chetty et al. (2020)

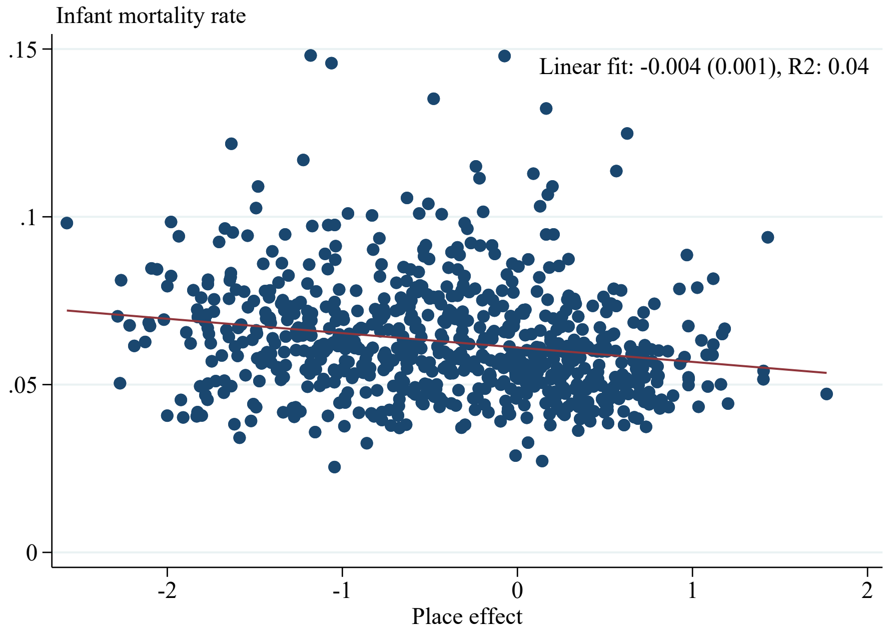
Appendix Figure 11: Place Effect Mechanisms, Within-Place Estimates, Bivariate Results



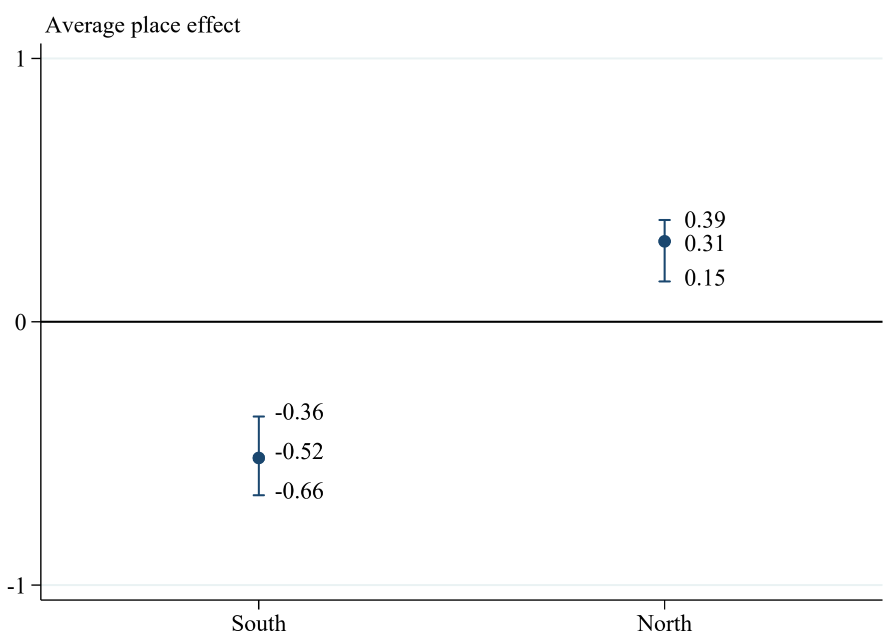
Notes: Figure displays correlations based on an analysis of the change in opportunity measures and the change in place characteristics. The 1940 measure of place effects is based on our analysis of the 1940 Census. For a contemporary opportunity measure, we use the upward mobility measure from Chetty et al. (2020). Upward mobility is the mean household income rank for children whose parents were at the 25th percentile of the national income distribution. This statistic is calculated for children born between 1978 and 1983. The construction of measures of place characteristics is described in Appendix G. We normalize all variables to have a standard deviation of one and a mean of zero. We compute the change for each standardized variable between the contemporary and historical periods. Correlations are based on the change in normalized measures.
Source: Authors' calculations using 1940 Census (Ruggles et al., 2020) and Chetty et al. (2020)

Appendix Figure 12: Sensitivity of Results to Child Mortality Differences

(a) Relationship between Infant Mortality Rates and Place Effect Estimates

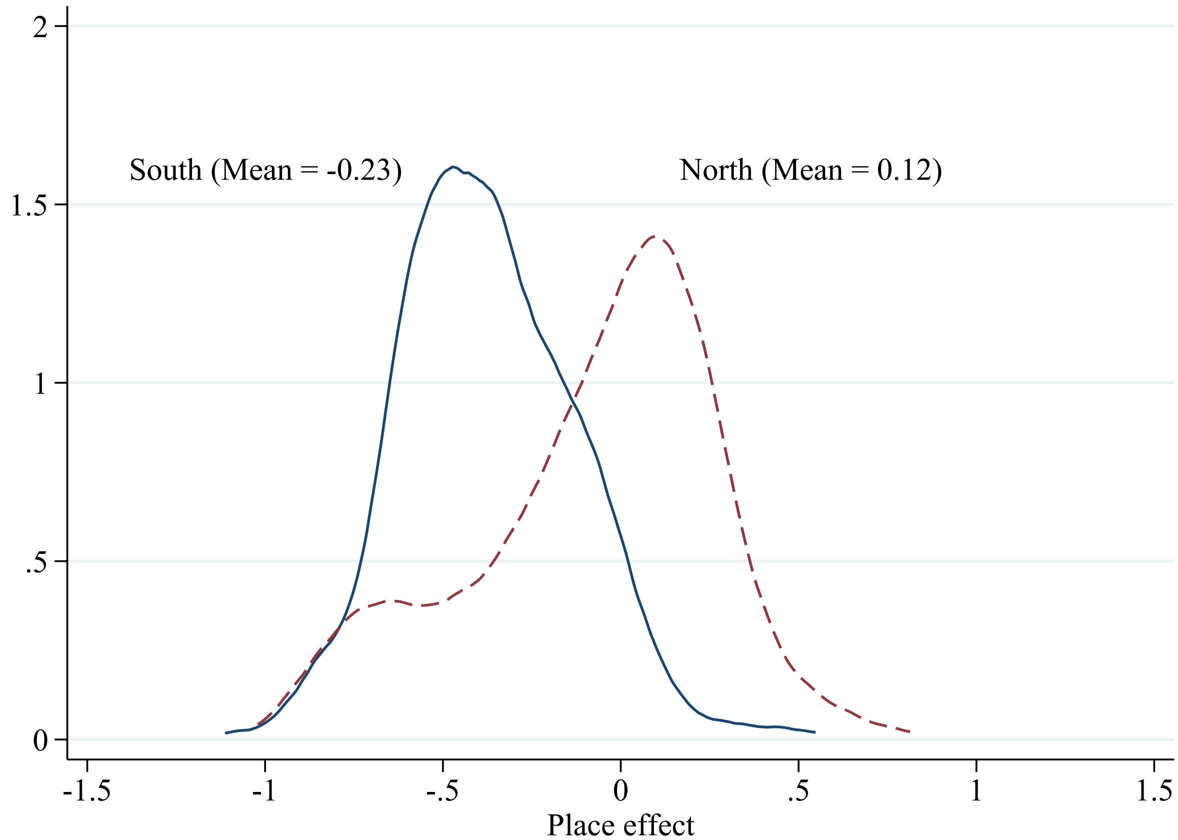


(b) Upper and Lower Bounds of Place Effects in the South and North



Notes: Panel A displays the relationship between infant mortality rates from 1933–1937 and our baseline place effect estimates. Panel B displays the average upper and lower bound for county place effects in the South and North, respectively. Section F provides details on the constructions of the bounds.
Source: Authors' calculations using 1940 Census (Ruggles et al., 2020) and infant mortality records (Bailey et al., 2008)

Appendix Figure 13: Density of Place Effects on Adult Earnings in South versus North, Black Men Age 25–64



Notes: Figure shows density of place effect estimates in the South and North. Migrant-weighted averages are reported.
Source: 1940 Complete Count Census