Essays on Human and Social Capital

by

Evan J. Taylor

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2017

Doctoral Committee:

Professor Jeffrey A. Smith, Chair
Professor Martha J. Bailey
Professor John E. Dinardo
Assistant Professor Andreas Hagemann

Evan J. Taylor
evanjt@umich.edu
ORCID iD: 0000-0002-6210-1005

DEDICATION

To my wife Alyssa.

ACKNOWLEDGEMENTS

I would like thank my committee: Martha Bailey, John Dinardo, Andreas Hagemann, and Jeff Smith. I am grateful also for comments and help from Dan Black, Clint Carter, Seth Sanders, Bryan Stuart and Lowell Taylor, and participants in seminars at the University of Michigan. Finally, I wish to acknowledge the generous financial support of the NSF Graduate Research Fellowship Program.

This dissertation uses confidential data from the U.S. Census. Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

TABLE OF CONTENTS

List of Figures

List of Tables

List of Appendices

ABSTRACT

Chapter 1: The Impact of College Education on Mortality: A Study of Marginal Treatment Effects.

With a newly constructed dataset that links the 2000 U.S. Census long-form to Social Security Administration records, I estimate the effect of college education on mortality. Using the proximity to college from birthplace as an instrument, I estimate the marginal treatment effect (MTE) of college education on 10-year mortality rate for adults aged 60-99 in the United States from 2000-2010.

The OLS results show a strong association between college education and lower mortality, consistent with the previous findings in the literature. The MTE results show that individuals that have unobserved characteristics that make them least likely to attend college have the largest effects of education in reducing mortality. This suggests that the individuals who would benefit most from receiving college education in terms of health are those do not attend college. The positive effects on reducing mortality are solely concentrated among men. For women, I find no evidence of an effect of education on old-age mortality. The MTE results, combined with evidence from the literature, also provide suggestive evidence that income is not the mechanism through which education reduces mortality for men.

Chapter 2: Social Interactions and Location Decisions: Evidence from U.S. Mass Migration. (with Bryan Stuart)

This paper estimates the strength through which social interactions influenced location decisions during two large scale migrations in the United States during early to mid 1900s. We examine the Great Migration of African-Americans out of the Southern United States and the Dust Bowl Migration of whites out of the Midwestern U.S. Using long-run data on migration patterns from the Social Security NUMIDENT file linked to Medicare Part B records for individuals born 1916-1936, we estimate the effect of social interactions on influencing where individuals decided to

migrate.

We find that social interactions were very important for blacks during the Great Migration in affecting location decisions. Our results suggest that 47-69 percent of blacks chose their destination city in the North because of influence from other people that were from their hometown. For whites, we estimate much smaller effects; only 14-24 percent of whites chose their destination city because of social interactions. For blacks, social interactions played a large role in influencing migration to destinations with large shares of manufacturing employment, suggesting the importance of access to information and labor demand conditions in facilitating social influence.

Chapter 3: The Effect of Social Migration on Crime: Evidence from the Great Migration. (with Bryan Stuart)

Using results from the second chapter of the dissertation, which shows that social interactions were very influential in guiding migration patterns during the Great Migration, this paper estimates the effect these patterns had on crime in U.S. cities from 1960-2009. We document the large variations in the connectedness of migrants from the South that moved to different Northern cities. For example, some cities received almost one-third of their migrants from only one origin town in the South, where other comparable cities received no more than 3% of migrants from any one place.

We find that, controlling for other economic characteristics, cities which received more connected migrants had lower crime rates from 1970-2000, which suggests an important role of social connectedness on crime during these periods. The results are largely driven by cities with a high population share of African Americans, and through crime increases among black juveniles. Cities that had more connected migrants had smaller increases in crime rates during the 1970s and 1980s.

CHAPTER 1 THE IMPACT OF COLLEGE EDUCATION ON OLD-AGE MORTALITY: A STUDY OF MARGINAL TREATMENT EFFECTS

1.1 INTRODUCTION

Well-educated people live longer and healthier lives than do their poorly-educated counterparts. This generalization holds across many countries, and has been true for more than a century (David M. Cutler and Adriana Lleras-Muney, 2010). The education-mortality gradient in the U.S. is dramatic; an analysis of 2000 mortality rates by Ellen R. Meara, Seth Richards and David M. Cutler (2008) shows that the life expectancy for a 25-year-old with at least some college education was approximately seven years longer than for an individual with a high school degree or less. It appears that the education-mortality gradient is, if anything, increasing over time. However, while the relationship between education and mortality is well documented, we have only modest evidence about the extent to which it is causal.

My contribution is an analysis evaluating the impact of education on mortality in the U.S. Like Meara, Richards and Cutler (2008), I focus attention on mortality among those who have attended at least some college, in comparison to those with high school education or less. This is a highly relevant margin of schooling progression; among cohorts I study over 30% of women and 40% of men attained education above the high school level.[1] In the U.S., we are currently engaged in an active debate concerning policies that promote college education. Benefits in terms of improved health and increased longevity may represent a substantial portion of the total gains of college education, so an understanding of causal impacts is crucial for this debate.

My empirical approach is designed to take on two important empirical challenges that arise when analyzing causal effects of education on such outcomes as income, health, or mortality: First, individuals who receive higher education differ along many unobserved characteristics that may impact the outcomes; put differently, there is selection into college education. Second, the effect of education on outcomes may vary across individuals, and, in particular, may differ for those who have unobserved characteristics which make them more or less likely to obtain higher

[1] Among individuals now in their prime working years, more than 60% have at least some college.

education. This type of heterogeneity in the treatment effect is called "essential heterogeneity" by James J. Heckman, Sergio Urzua and Edward Vytlacil (2006).

I use an instrumental variable approach that relies on distance to the nearest college or university (at age 17) from an individual's birth county. This instrument was used by Card (1995) to estimate the effect of education on earnings, and proximity to a nearby college has served as an instrument in many subsequent empirical analyses, including Cameron and Taber (2004) and Carneiro, Heckman and Vytlacil (2011).

To calculate treatment effects which may be heterogeneous in both unobserved and observed variables, I employ estimators of marginal treatment effects (MTEs) developed by James J. Heckman and Edward Vytlacil (1999, 2005, 2007). The estimation of MTEs provides improved conceptual clarity about impacts of education on mortality. As noted by Pedro Carneiro, James J. Heckman and Edward J. Vytlacil (2011), understanding how treatment effects vary across individuals is crucial in evaluating the marginal benefits of any policy which may increase access to college education. For example, it may be that gains are highest among individuals who already would attend college, i.e., the average benefits realized by individuals induced to attend college by the expansion policy are lower than the average benefits of those who would have attended anyway.

I estimate the MTE in two different ways—an approach using local instrumental variables, which evaluates the derivative of the expected outcome with respect to the estimated propensity score, and the separate estimation approach discussed in James J. Heckman and Edward Vytlacil (2007), and implemented in the recent work of Christian N. Brinch, Magne Mogstad and Matthew Wiswall (forthcoming, 2016). The separate approach involves the use of stronger parametric assumptions, with a corresponding benefit in terms of increased precision on estimation. This proves valuable in my application.

My analysis uses a newly constructed dataset that links the 2000 U.S. Census to NUMIDENT records from the U.S. Social Security Administration.[2] These data include records for almost 5 million individuals born in the United States, 1911–1940. The data contain educational attainment (recorded in the 2000 Census) and place of birth, date of birth, and date of death for the deceased (from the NUMIDENT file).[3] The analysis of mortality focuses on the time span 2000-2010.

Consistent with the literature, I find a strong association between college education and mortality for both men and women. In an OLS specification, individuals with at least some college have a 10-year mortality rate 20% lower than those with a high school degree or less.

I implement a simple selection test proposed by Dan A. Black, Joonhwi Joo, Robert LaLonde, Jeffrey Smith and Evan J. Taylor (2016). There is strong evidence of positive selection on longevity

[2] NUMIDENT is an acronym for the Social Security Administration's "Numerical Identification System."

[3] I would like to acknowledge several colleagues with whom I am collaborating on the place-of-birth matching algorithm used in constructing data files: Martha Bailey, Bryan Stuart, and Reed Walker.

into college education for women. That is, women with a higher unobserved inclination towards education tend to have lower mortality whether they receive college education or not. For women, selection explains all of the differences in observed mortality rates between high and low education groups. For men, I only find evidence of positive selection on the *untreated* outcome. I do not find selection on the treated outcome, which suggests a “reverse Roy" pattern of selection for men.

In the MTE analysis, I find a heterogeneous treatment effects of college on mortality. For women, I estimate no treatment effect on old-age mortality. For men, MTE results show “essential heterogeneity" in treatment effects; individuals with unobserved characteristics that make them more likely to attend college have smaller treatment effects on mortality. I estimate an average treatment effect on mortality of zero for individuals who had at least some college, but find that college attendance would have reduced mortality by about 15% among individuals who did not attend college. Taken at face value results suggest that for the cohorts of men I study, a policy that increased college attendance at the margin would have produced a substantial decline in mortality among affected men.

Studies such as Pedro Carneiro, James J. Heckman and Edward J. Vytlacil (2011) and Moffitt (2008) suggest that the impact of college education on wages is concentrated among individuals whose unobservable characteristics make them most likely to attend college. Both papers find that individuals sort into college based on gains. These MTEs of education on income, in combination with my results, are consistent with David Cutler, Angus Deaton and Adriana Lleras-Muney (2006) contention that the positive impact of education on longevity is likely not through any effect on adult income. I find that the individuals with little gains in terms of income receive large benefits in mortality.

My work contributes to an emerging literature that uses modern econometric methods to evaluate the effect of education on health and mortality. Much of the previous work has used changes in mandatory schooling laws, which exploit variation in the lower parts of the schooling distribution. The results of these studies have been mixed. Lleras-Muney (2005) finds large negative effects of increased education on mortality in the U.S. for whites born 1901–1925. Mazumder (2008) and Dan A. Black, Yu-Chieh Hsu and Lowell J. Taylor (2015) find no evidence of causal effects using somewhat different data and/or approaches. Valerie Albouy and Laurent Lequien (2009) and Damon Clark and Heather Royer (2013) estimate the effect of education in France and the UK, respectively, and find no evidence of an effect. Using Swedish data, Costas Meghir, Marten Palme and Emilia Simeonova (2012) find that negative effects on mortality induced by changes in compulsory schooling are concentrated among men from low SES backgrounds.

Also related to my study is recent empirical work from Kasey Buckles, Andreas Hagemann, Ofer Malamud, Melinda Morrill and Abigail Wozniak (2016), who use variation in draft avoidance behavior during the Vietnam War and find that college education reduces mortality. In contrast to my work, Buckles et al. (2016) focus on mortality among those aged 28–65, and study men only (women were not subject to the military draft). Also, they did not estimate MTEs.[4]

The paper proceeds as follows: In the next section, I discuss the literature on the relationship between education and health. In section 3, I explain the definition and estimation of marginal treatment effects. In section 4, I develop a model of education and mortality which fits the MTE framework. The next section describes the data. In sections 6 and 7, I show results from the selection test and main results from the MTE analysis. The following section gives discussion of the results and possible mechanisms. The final section concludes.

1.2 RELATIONSHIP BETWEEN EDUCATION AND HEALTH

The positive correlation between education and health is well documented. Many factors contribute.

First, there are many reasons why even in the absence of a causal effect of education on health, we may see an association between the two. Adverse health circumstances experienced in infancy and early childhood can harm prospects for learning and reduce human capital accumulation,[5] and the same health deficits that impair learning early in life likely have an adverse impact on adult health. For instance, poor childhood nutrition may affect adult health (Fogel, 1997, 2004), as might adverse health events *in utero* (Barker, 1998). Further, expectations about increased longevity could induce individuals to acquire higher levels of human capital (Ram and Schultz, 1979).

In addition, models of human capital investment—pioneered by Becker (1967), Mincer (1974), and Griliches (1977), and further developed by Card (2001), James J. Heckman, Lance Lochner and Petra E. Todd (2006) and James J. Heckman, John Eric Humphries and Gregory Veramendi (2016), among others—lead us to understand that along unobservable dimensions, well-educated individuals differ from those with less education. There are two distinct forms of systematic selection into higher education at work. People may be *sorting on gains* (e.g., if those who have the most to gain are most likely to acquire schooling) and there may be *selection bias* unrelated to gains (e.g., if there is an individual fixed effect in the earnings function that is correlated with completed schooling). Even if education has no direct impact on health and longevity, we would observe an education-mortality gradient if components of income-related selection are correlated with health-related outcomes.

Finally, there are many theories as to why education may have a causal impact on health and

[4]Buckles and co-authors uses data at a cohort by birth-state level, and have two endogenous regressors (college and veteran status), making MTE estimation impossible.

[5]See Janet Currie (2009, 2011) and Douglas Almond and Janet Currie (2011) for comprehensive discussions of the evidence on this point.

mortality. The theory of human capital accumulation has been adapted to investments in health by Grossman (1972, 2000), Mark R. Rosenzweig and T. Paul Schultz (1982), David M. Cutler and Adriana Lleras-Muney (2010) and Lochner (2011), among others.[6] Grossman (1972) argues that education can directly impact mortality and health by increasing the marginal productivity of health inputs, which he calls "productive efficiency." David M. Cutler and Adriana Lleras-Muney (2010) suggest a number of channels through which education can improve health outcomes, and note that "better educated people are less likely to smoke, less likely to be obese, less likely to be heavy drinkers, more likely to drive safely and live in a safe house, and more likely to use preventative care."[7] Many studies provide additional evidence on this topic, including Gabriella Conti, James J. Heckman and Sergio Urzua (2010), who identify heterogeneous beneficial causal effects of education on health behaviors, including smoking and exercise. Walque (2007) and Grimard and Parent (2007) find evidence that college education reduces smoking rates among men.

Education generally causes increases in earnings (Card, 1999). If higher income allows individuals to afford better health care and nutrition, then education will have a positive effect on longevity. Cutler and Lleras-Muney (2010) suggest that about 20% of the education-health gradient can be explained by changes in economic resources. Caution is warranted when it comes to interpretation, however, because of the possibility of reverse causality; Smith (1999) argues that poor adult health causes decreases in income.

Clearly, the links between health and education are multi-faceted. One point of view, articulated by David Cutler, Angus Deaton and Adriana Lleras-Muney (2006), suggests a central role for education: "It seems clear that much of the link between income and health is a result of the latter causing the former, rather than the reverse. There is most likely a direct positive effect of education on health." In an essay on the injustice of health inequality, Deaton (2013) provides a set of arguments for why it is important that we develop a conceptually clear empirically-grounded understanding of these relationships. For instance, if there is little impact of adult income on health, a transfer policy that increases income among the poor may have little impact of health inequalities.

[6] Lochner (2011) offers a theoretical synthesis and overview of the literature.

[7] David M. Cutler and Adriana Lleras-Muney (2010) are careful to note that these associations cannot be interpreted as causal in their framework. They do provide an informed judgment that "education seems to influence cognitive ability, and cognitive ability in turn leads to healthier behaviors. As best we can tell, the impact of cognitive ability is not so much what one knows but how one processes information. Everyone 'knows' that smoking is bad and seat belts are useful but the better educated may understand it better."

1.3 Marginal Treatment Effects

1.3.1 Model

I first introduce the standard Marginal Treatment Effects framework used in Heckman and Vytlacil (1999). The outcome variable is Y, which I define as a 10-year survival dummy. D is a binary indicator for treatment, which I define as 1 if the individual has at least some college education, and 0 if the individual is a high school graduate or less. I discuss the reasoning behind this definition in the next section. I have observable characteristics X and instruments Z. Let Y_1 denote the outcome when an individual is treated and Y_0 denote the outcome when an individual is not treated. Crucially, when an individual is treated we only observe Y_1, and not the missing counter-factual Y_0, and vice-versa. Then, $Y = DY_1 + (1-D)Y_0$. Suppose that the conditional mean functions for Y are as follows:

$$Y_{1i} = \mu_1(X_i) + U_{1i} \tag{1.1}$$

$$Y_{0i} = \mu_0(X_i) + U_{0i}, \tag{1.2}$$

where the vector X_i in empirical specification below include observable characteristics—birth state fixed effects, birth-county economic variables, cohort effects—that may affect survival in either treatment condition.

The indicator variable D_i reflects individual choice. The choice model is given by

$$D_i = \mathbb{1}(f(X_i, Z_i) > V_i); \tag{1.3}$$

the college attendance decision is influenced by both X_i and the instrument Z_i.

Let F_V be the CDF of V_i, $F_V(V_i) = U_{Di}$ and $F_V(f(X_i, Z_i)) = P(X_i, Z_i)$. We can rewrite (1.3) as

$$D_i = \mathbb{1}(P(X_i, Z_i) > U_{Di}), \tag{1.4}$$

where $P(X, Z) = E[D|X, Z]$ is the propensity score and U_D is distributed uniform over $(0, 1)$.[8] For Z to be a valid instrument, we need conditional independence: $(U_0, U_1, U_D)Z|X$. Vytlacil (2002) shows the equivalence of the latent index model shown above and the standard IV assumption of monotonicity (Joshua D. Angrist, Guido W. Imbens and Donald B. Rubin, 1996).

In this set-up, it is surely possible that there will be a relationship between errors in the outcome equations (U_0, U_1) and the error in the choice equation, U_D. The term U_D represents all the unobservable determinants in the college attendance decision. These could include parents' wealth

[8] The use of the uniform distribution is a convenient normalization.

and educational attainment, latent “ability" or motivation, the individual’s personal discount factor and childhood health. Going forward it is helpful to take note that in this formulation agents with low U_D are *more* inclined to attend college. Thus low U_D is likely associated with high parental SES, high ability, stronger childhood health, etc.

Figure 1.1 outlines three possible patterns of correlation between the error terms:

First, suppose that the treatment effect is homogeneous and individuals from advantaged backgrounds are generally healthier, i.e., have high values of both U_{0i} and U_{1i}, and are more likely to attend college. Thus, individuals with lower U_{Di} will tend to have $U_{1i} = U_{0i} > 0$. This pattern, shown in Panel A, is the familiar “ability bias” (Griliches, 1977). In this instance, if we estimate the impact of college on survival in an OLS regression, we overstate the treatment effect.

Second, Panel B shows a case in which there is heterogeneity in gains from college attendance (in terms of increased longevity) and those individuals who have the most to gain attend college. In other words, there is selection on gains as in Roy (1951). Here individuals with high values of $(U_{1i} - U_{0i})$ disproportionately have U_{Di} close to 0. Now it is possible that OLS understates the treatment effect of college among those most likely to attend college—those with U_D close to 0—while overestimating the potential value of college on those who do not attend.

Finally, Panel C illustrates an interesting alternative. Perhaps children from advantaged backgrounds build strong health capital and develop good health-related habits that persist into adulthood regardless of college education, *and* those same children are disproportionately likely to attend college (i.e., U_{0i} and U_{Di} are negatively correlated). Individuals that benefit the most from college are lower SES and have high U_D. This has been called a “reverse Roy" selection pattern. In the example illustrated, OLS overstates the impact of college on those who attend college, but does not necessarily overstate the potential impact of college on those who are not currently attending. The MTE approach is designed to help us distinguish among possible patterns.

We work with the conditional expectation functions of the error terms,

$$c_1(x, p) = E[U_1 | X = x, U_D = p] \tag{1.5}$$

$$c_0(x, p) = E[U_0 | X = x, U_D = p]. \tag{1.6}$$

Then the marginal treatment effect is defined as

$$\begin{aligned} \Delta^{MTE}(x, p) &= E[Y_1 - Y_0 | X = x, U_D = p] && (1.7) \\ &= \mu_1(x) - \mu_0(x) + c_1(x, p) - c_0(x, p). && (1.8) \end{aligned}$$

The MTE is the expected difference in potential outcomes (i.e., the average treatment effect) for individuals who have observables $X = x$ and are indifferent between getting treatment and not when $P(X, Z) = p$.

1.3.2 Estimation through Local IV and Separate Approach

Define $K(x, p) \equiv E[Y|X = x, P(X, Z) = p]$. As shown by Heckman and Vytlacil (1999), the marginal treatment effect can be estimated by taking the derivative of the expectation of the outcome with respect to P, as follows:

$$\Delta^{MTE}(x, u_D) = \Delta^{LIV}(x, u_D) = \left.\frac{\partial K(x, p)}{\partial p}\right|_{p=u_D}. \tag{1.9}$$

This is the method of local instrumental variables.

Heckman and Vytlacil (2007) examine a second approach, which is based on estimating $E[Y|X, P, D = 1]$ and $E[Y|X, P, D = 0]$ separately rather than estimating $E[Y|X, P]$. In parametric estimation, the parametric assumptions for this approach are somewhat stronger. For example, if estimating an MTE which is linear in parameters, in the separate approach both $E[Y_1|X, U_D]$ and $E[Y_0|X, U_D]$ are assumed to be linear in X and U_D, whereas in Local IV only the difference, the MTE, is assumed to be linear. This can lead to large increases in precision of the estimates.[9]

Define K_1 and K_0 as such:

$$K_1(x, p) \equiv E[Y|X = x, P(X, Z) = p, D = 1] = \frac{\int_0^p c_1(x, u_D) du_D}{p} + \mu_1(x) \tag{1.10}$$

$$K_0(x, p) \equiv E[Y|X = x, P(X, Z) = p, D = 0] = \frac{\int_p^1 c_0(x, u_D) du_D}{1 - p} + \mu_0(x). \tag{1.11}$$

Differentiating with respect to p and solving for c_1 and c_0 we get

$$c_1(x, p) = p\frac{\partial K_1(x, p)}{\partial p} + K_1(x, p) - \mu_1(x) \tag{1.12}$$

$$c_0(x, p) = -(1 - p)\frac{\partial K_0(x, p)}{\partial p} + K_0(x, p) - \mu_0(x). \tag{1.13}$$

Plugging the above into (1.8), we get[10]

$$\Delta^{MTE}(x, p) = p\frac{\partial K_1(x, p)}{\partial p} + K_1(x, p) + (1 - p)\frac{\partial K_0(x, p)}{\partial p} - K_0(x, p). \tag{1.14}$$

[9] In unpublished work in progress, I discuss the relative merits of each estimation strategy and show solutions when the stronger parametric assumptions of the separate approach are incorrect (Taylor, 2016).

[10] Notice that both $\mu_j(X)$ cancel out. In general, these are not identifiable separately from K_j. See Heckman and Vytlacil (2007).

To generalize these two methods, define

$$\tilde{K}(x,p,d) \equiv E[Y|X=x, P(X,Z)=p, D=d] = dK_1(x,p) + (1-d)K_0(x,p). \tag{1.15}$$

Then

$$\begin{aligned}\Delta^{MTE}(x,u_D) &= \tilde{K}(x,u_D,1) - \tilde{K}(x,u_D,0) + \left.\frac{\partial \tilde{K}(x,p,0)}{\partial p}\right|_{p=u_D} \\ &\quad + u_D\left(\left.\frac{\partial \tilde{K}(x,p,1)}{\partial p}\right|_{p=u_D} - \left.\frac{\partial \tilde{K}(x,p,0)}{\partial p}\right|_{p=u_D} \right).\end{aligned} \tag{1.16}$$

To estimate the MTE, we estimate the function $\tilde{K}(x,p,d)$, and substitute it into the equation above.

1.4 Model of Education and Mortality

I introduce and develop a simple model of educational attainment and mortality. The model builds on the canonical model of Becker (1967). Individuals maximize expected utility $U(y,S)$, which is a function of earnings (y) and years of schooling (S). For my purposes, y can include the value of health and longevity, as in Lochner (2011). Alternatively, it may be that individuals are unaware of benefits from education in terms of health improvements or increased longevity, or in any event that they do not take them into consideration when making human capital investments. For simplicity, let utility take a separable form, $U_i(y_i,S_i) = f(y_i) - \phi_i(S_i)$, and then let $y_i = g_i(S_i)$, as in Card (1994). $\phi_i(S_i)$ represents the total costs of schooling, including monetary and non-monetary costs. Finally, let $m_i(S_i)$ represent the individuals expected mortality rate as a function of their schooling.

In my research design, the key assumption is that the instrument, college proximity, alters the cost curve, $\phi_i(\cdot)$, but does not have an impact on the mortality function $m_i(\cdot)$ nor the function that generates returns from schooling $g_i(\cdot)$.

Note that for the instrument to be valid, it can only shift individuals *across* the treated or untreated states and cannot shift people to different educational attainments that are *within* either the treated or untreated state. This follows from the MTE assumption that $(U_1,U_0)Z|X$.[11] If we rewrite the schooling cost curve, $\phi_i(S_i) = \phi_i'(S_i) + \lambda_i \mathbb{1}(S_i > 12)$, then the instrument can only affect λ_i, but not $\phi_i'(\cdot)$. The assumption is that the distance to the nearest college will only affect the fixed cost of first attending college, while having little impact on the marginal cost of each

[11] Alternatively, it could be the case that the mortality function $m_i(S_i)$ is flat except for a discontinuity at $S_i = 12$. Given the results from the health literature which show constant gains in health throughout the education distribution, see Cutler and Lleras-Muney (2010), this makes the latter improbable.

year of school thereafter. Similarly, the instrument must also not affect the marginal cost curve for $S < 12$.

This is plausible if most of the cost reduction coming from proximity is a one-time "transition cost" (e.g., migration cost), which includes all monetary and non-monetary costs. I will return to concerns about biases that are introduced if college proximity also affects marginal costs beyond the first year of college.

This model guides the choice of the definition of the treated state as "some college" and the untreated state as "high school or less." If the cutoff for treatment was defined as at least a college degree and some college was part of the untreated state, then the instrument could not affect the initial cost of going to college and only the margin between attending and graduating. This latter set of assumptions is less credible.

For each individual, there is an optimal level of schooling that is less than or equal to 12-years (high school or less), and an optimal level of schooling that is greater than 12 years (at least some college). These can be written as $S_{0i} = \arg\max_{S \leq 12} U_i(y, S)$ and $S_{1i} = \arg\max_{S > 12} U_i(y, S)$. Then $Y_{0i} = m_i(S_{0i})$ and $Y_{1i} = m_i(S_{1i})$. The individual attends college if $U_i(y_i, S_{1i}) > U_i(y_i, S_{0i})$. Figure 1.2 shows an example of the cost curve, $\phi(S)$ and benefit curve $f(g(S))$. The instrument can make parallel shifts of the cost curve for $S > 12$, which do not affect the optimal conditional choices S_0 and S_1. Thus, the treatment effect does not simply encompass the marginal gain between 12 and 13 years of school. The treatment effect evaluates the gain from college attendance, relative to non-attendance, for an individual who makes the preferred schooling decision within each treatment state.

1.5 Data

1.5.1 Census and NUMIDENT

As discussed in the introduction, many papers measuring the effect of education on mortality in the U.S., such as Lleras-Muney (2005), Black, Hsu and Taylor (2015), Buckles et al. (2016), have exploited variation at a state-by-cohort level, and measured state-by-cohort mortality rates using census and vital statistics data. While these strategies can be used to estimate an IV, they are unsuitable for the evaluation of essential heterogeneity in treatment effects. To estimate MTE, we require a large sample of individual level data which contains both information on educational attainment and mortality, and we need an instrument which can be measured at the individual level at an early age.

The primary data in the analysis comes from two sources, the 2000 U.S. Census and the NUMIDENT from the Social Security Administration. The two sources have been linked using the Protected Identification Key (PIK). Every individual in the NUMIDENT file is given a unique PIK, which is then matched to the 2000 Census based on personal information.

The 2000 U.S. Census contains information on completed years of education, race, gender and age. The NUMIDENT has basic demographic information, as well as the individual's date of birth and date of death if they are deceased. The NUMIDENT also contains a 12-character place of birth field and a 2-character state of birth field. This is used to match the individual to a physical location. The algorithm used to perform this match is an augmented version of the algorithm my co-authors and I developed in Dan A. Black, Seth G. Sanders, Evan J. Taylor and Lowell J. Taylor (2015). We take place names from the U.S. Geological Service's Geographic Names Information System (GNIS). The GNIS is a list of all place names in the United States, including both current and historic. From this we match individuals to a county of birth.[12] A relatively small percentage of individuals cannot be matched to place of birth, either because the place of birth field is missing, because we cannot determine their place of birth from the 12-character string, or because do not have PIKs and cannot be linked to the NUMIDENT. These people are not included in the analysis.

My sample contains almost 5 million white individuals born in the continental United States from 1911–1940, who are in the 2000 U.S. long-form Census and can be linked to a birthplace through NUMIDENT.[13] I do not include individuals that have imputed education, whose age in the Census was five or more years different than their age in the NUMIDENT, and whose race or gender was not the same in the Census and the NUMIDENT.[14]

Of the 2.7 million women in the data set, 34% attended at least some college. Men in the data attended college at higher rates; 43% of the 2.2 million men had at least some college. For the dependent variable in my analysis, I use a 10-year survival indicator from 2000–2010. The indicator is equal to zero if the individual is deceased in 2010 or earlier, and one otherwise. In the sample, the average 10-year mortality rate was 32% for women and 39% for men.

1.5.2 College Data

The college data come from the Integrated Postsecondary Education Data System (IPEDS), published by the National Center for Education Statistics (NCES). The IPEDS data extend back to the year 1980, so the data contain all 2-year and 4-year colleges that were open in the United States in 1980, but not colleges that closed prior to 1980.[15] IPEDS contains information on the location as well as the year that the college opened.

[12]For individuals who were born in a county that has changed borders, the algorithm will match to the current county of the individual's birthplace.

[13]A separate analysis for blacks showed the instruments were weak, even when restricting the distance measure to historically black college and universities. More work is needed to estimate the effect of college education on blacks.

[14]When there is a discrepancy between birth year in the Census and NUMIDENT, I use the NUMIDENT birth year.

[15]This may affect the power of the instrument, but not the exclusion restriction as long as locations that had a closed college are no different than locations that never had a college along unobservable characteristics that affect mortality.

1.5.3 Instrument

Using these data, I construct the instrument—proximity from the individual's county of birth to the nearest college, including both two and four-year colleges. From the NUMIDENT data it is possible for most individuals to construct a city of birth rather than county of birth. However, for all individuals, I use distance from the individual's birth county, even if it is possible to determine the city of birth. The county of birth measure is less likely to violate the exclusion restriction, given that all economic covariates are at a county level. This approach also allows for geographic sorting within a county. To approximate the average distance faced by a resident of a county, I calculate the distance from the centroid of the county to the nearest college. The measure includes all colleges that are open when the individual turns 17.

There are several potential threats to exogeneity of the instrument. First, parents who have a strong preference towards sending their children to higher education may choose to live in counties nearer to colleges. If the children of these parents are also of higher "ability" or are different in some other unobserved way that is correlated with later life mortality, then the exclusion restriction would be violated. In this situation, all estimated treatment effects would be biased upwards. If these parents instead choose to live in places that are observably wealthier without thought to how close the nearest college is, then the instrument remains valid.

Secondly, places that have colleges may be healthier for reasons beyond the effect of increased education for the individuals that are born there. If, controlling for other wealth measures, counties near colleges are healthier—perhaps because they have better access to medical care or because they have better primary and secondary public education—then the exclusion restriction would be violated. Again, in this case, estimated treatment effects would be biased upwards.

A third possible violation is that the presence of college can shift individuals not only across the no college versus some college threshold, but to different educational attainment categories within these groups. For example, consider an individual who was born far away from a college and dropped out of high school. It would be a violation of the exclusion restriction for that individual to get a high school degree in the state of the world where she is born close to the college. As shown in the model, the presence of the college can only affect the marginal cost of moving from the 12th year of school to the 13th year of school. In this case, the instrument will only shift individuals *across* the some college threshold. Note that this is a stricter set of assumptions than needed to use the instrument on a continuous measure of education. This assumption is necessary because of the use of the binary measure of education, which is needed to estimate a marginal treatment effect.

I control for several observed characteristics of the individual's birth county, to eliminate the potential effects of colleges being located in wealthier areas. These are an array of county-level economic characteristics from near the birth years and present day. From the Bureau of Economic Analysis, I use the average per capita income from 1969-2010 to control for the modern day wealth of a county. These data only go back to 1969, and the Census data on prosperity measures vary by decade. I include the median housing value in 1930, the percentage of households which owned a radio in 1940, and the median household income, by decile, in 1950.[16] While these old measures are likely somewhat noisy measures of wealth, they are generally strongly correlated with each other, as shown in Table 1.1.

1.6 Selection Test

Dan A. Black, Joonhwi Joo, Robert LaLonde, Jeffrey Smith and Evan J. Taylor (2016) provide a simple selection test with a binary dependent variable and an instrument. The test is two separate OLS regressions for treated and untreated individuals—regressing the outcome Y on both X and Z:

$$Y_{1i} = \alpha_1 X_i + \beta_1 Z_i + \epsilon_{1i}, \quad \text{for } D_i = 1, \quad \text{and} \tag{1.17}$$

$$Y_{0i} = \alpha_0 X_i + \beta_0 Z_i + \epsilon_{0i}, \quad \text{for } D_i = 0. \tag{1.18}$$

$\beta_1 \neq 0$ implies selection on unobserved variables on the treated outcome. That is, the unobserved variables that enter the choice equation are correlated with the unobserved variables that enter the treated outcome equation. Similarly, $\beta_0 \neq 0$ implies selection on the untreated outcome.

If the instrument is valid, in the absence of selection $E[Y_1|X,Z]$ and $E[Y_0|X,Z]$ are determined by X only. With positive selection on Y_1, we expect that as we decrease the distance to college we increase the marginal propensity score and move further into the U_D distribution. This in turn decreases the expected value of the outcome, as in Figure 1.3. Thus, in this case $\beta_1 > 0$. Similar intuition shows that with positive selection on the untreated outcome, Y_0, we have $\beta_0 > 0$.

These tests also provide a sanity check on the validity of the exclusion restriction. Given the evidence discussed in the prior literature, our prior on the nature of selection into education suggest that it should be weakly positive on both the treated and untreated outcomes. Therefore, $\beta_1 < 0$ and $\beta_0 < 0$ imply that either selection into education is negative or that the exclusion restriction is violated such that counties further from colleges have lower survival rates. This would constitute evidence that the exclusion restriction is violated, as negative selection into college on health goes strongly against our priors.

1.6.1 Results for the Selection Test

Estimates for the selection tests for women and men appear in Table 1.2. The results for women show evidence of positive selection on both the treated and untreated outcomes. For men, there is

[16] Where data is missing, I impute zeros and include an indicator variable for missing data in all specifications.

strong evidence of positive selection on the untreated outcome, but no evidence of selection on the treated outcome. Any bias coming from the instrument, in which individuals born in closer proximity to colleges are more likely to live longer, would bias these coefficients downward, making it more difficult to find positive selection. These tests provide credible evidence of positive selection into education in terms of mortality.

Results from the selection tests foreshadow the MTE analysis. Consider results for men. Among those who attend college, expected longevity Y_1 (conditional on X) is approximately the same for those who are born close to a college and those born far from a college ($\beta_1 \approx 0$). Among who do not attend, expected longevity Y_0 is lower for men born close to a college than those born far away ($\beta_0 > 0$). These suggest a pattern of selection as in Panel C of Figure 1.1—“reverse Roy" sorting. The MTE curve for survival is upward-sloping in U_D. For women, the similarity in the coefficients from the selection regressions may imply a flatter MTE curve, and that selection likely explains a large portion of the observed differences in mortality between the treated and untreated.

1.7 Results

1.7.1 OLS Regression

To explore basic relationships, I begin with OLS regressions of college education on 10-year survival, using variants of the following specification:

$$Y_i = \alpha + \beta X_{c(i)} + \lambda_{b(i)} + \zeta_{s(i)} + \gamma D_i + \epsilon_i, \tag{1.19}$$

where $Y_i = 1$ if an individual survives past year 2010, and $Y_i = 0$ if the individual is deceased in year 2010 or earlier. $X_{c(i)}$ are birth county covariates, $\lambda_{b(i)}$ are cohort fixed effects, $\zeta_{s(i)}$ are birth state fixed effects, and $D_i = 1$ if the individual has at least some college.

Results are presented in Tables 1.3 and 1.4. Women with at least some college have a 10-year survival rate 6 percentage point higher than women with high school or less. This is roughly a 20% reduction in the 10-year mortality rate. The results for men are similar. Men who attended college have a 21% lower 10-year mortality rate than men who did not.[17]

Birth-county covariates and even birth-county fixed effects do not dramatically change the result after controlling for state level fixed effects. The coefficients on college for both men and

[17]For a comparison, Lleras-Muney (2005) found in OLS that each additional year of schooling reduces 10-year mortality rate by 0.036, using 1901-1925 birth cohorts in the U.S. with mortality measured from 1960-1980. The difference in education between the high and low groups was four years, which suggest roughly 0.02 reduction in 10-year mortality per extra year of schooling. Lleras-Muney (2005) use state-cohort level data rather than individual level data, which makes a direct comparison difficult. However, it is consolatory that the estimates are not substantially different.

women are slightly larger when including birth-county fixed effects than when just including birth-state fixed effects. This indicates that, conditional on education and birth state, individuals have a *higher* survival rate if they were born in a county with *lower* average college attendance. This is consistent with a story of positive selection. With positive selection, as the college attendance rate rises within a county, it digs further into the “ability" distribution, reducing the average survival rates of both treated and untreated groups. Without selection, this would indicate that counties with higher education rates, conditional on the birth state, were less healthy places. This seems implausible.

The estimates with birth-state fixed effects and birth-county fixed effects are fairly similar, which is somewhat reassuring for the validity of the instrument. If estimates from a birth-county fixed effects model were substantially lower that with birth-state fixed effects, it would indicate that individuals born in counties with high educational attainment had much better health outcomes, even conditional on their own education. This would be damaging evidence for the validity of the instrument.

1.7.2 IV Regression

For comparison to OLS and MTE results, I present instrumental variable regressions. For men and women, I estimate first stages using linear, quadratic and cubic polynomials in the distance to nearest college. First stage results are presented in Tables 1.5 and 1.6. F-stats from the different first stage models are between 22 and 37, well above the rule-of-thumb cutoffs for weak instruments (James H. Stock, Jonathan W. Wright and Motohiro Yogo, 2002). For women, an increase in distance of 100 miles reduces college attendance by three percentage points, which amounts to an effect of approximately 9%. The instrument has a larger effect on men. A 100 mile increase in distance reduces college rates by five percentage points, approximately 12%.

IV results are presented in Table 1.7. In none of the specifications for men or women are the point estimates on college education significantly different than zero. If we made the (likely incorrect) assumption of a homogeneous treatment effect, these results would not provide clear evidence on the magnitude or sign of the effect. If treatment effects are heterogeneous, then IV estimates are a weighted average of marginal treatment effects.[18]

1.7.3 MTE

I start with a simplifying assumption common in the estimation of MTEs that $(U_0, U_1, U_D)(X, Z)$, which implies that the MTE is additively separable in U_D and X.[19] I allow the treatment effect to

[18] Heckman and Vytlacil (2005) provide formulas for weights. They note that weights are not always guaranteed to be non-negative.

[19] This assumption has been used in Carneiro, Heckman and Vytlacil (2011) and Moffitt (2008) among others.

vary on unobserved and observed variables, but restrict to no interactions between the observed and unobserved variables in the treatment effect. Here X includes birth-state and cohort fixed effects, as well as county level economic covariates.

To estimate the propensity score, I use a logit model. The model includes interactions between a cubic polynomial in the instrument and the birth state. I also include interactions between all covariates and decade of birth, to allow for the effect of all variables to vary over time. This specification includes a large number of parameters so we may be worried about weak instruments. However, results are robust to the first-stage model specification. Less flexible models raise concerns about misspecification of the first stage, which would bias our estimates of the MTE.

Densities of the propensity score estimates from this model are shown for women and men in Figures 1.4 and 1.5, respectively. Display of the histogram has been trimmed at 5% on both tails, per U.S. Census Bureau guidelines. The trimmed observations are still included in all analysis.

As discussed above, MTE can be estimated in two different ways: local instrumental variables and the separate approach, which estimates the second stage separately for treated and untreated individuals. When estimating a linear MTE, the parametric assumptions behind the separate approach are slightly stronger than for local IV.[20] In the separate approach, both $E[Y_1|X, U_D]$ and $E[Y_0|X, U_D]$ are assumed to be linear in X and U_D, whereas in local IV only the difference, the MTE, is assumed to be linear. In this case, the slightly stronger assumption of the separate approach leads to huge efficiency gains, whether allowing for heterogeneity in the treatment in X or not. The separate approach also allows estimation of both conditional expectation curves, $E[Y_1|X, U_D]$ and $E[Y_0|X, U_D]$, whereas local IV only estimates the difference. Thus, the separate approach can inform us about the nature of heterogeneity and selection.

I estimate the marginal treatment effect model as follows:

$$\Delta^{MTE}(X, U_D) = \alpha + \beta U_D + \gamma X. \tag{1.20}$$

Results for the slope coefficient on U_D when using the separate estimation approach appear in Table 1.8. For women and men, the point estimates are positive. This suggests that individuals with unobserved characteristics that make them more likely to attend college had lower treatment effects on survival (recall the low U_D is associated with higher unobserved proclivity towards college). For men, results using the preferred first stage specification, flexible logit, are significant at 10% level. Results are similar when using a less flexible logit or linear probability model in the first-stage. In these specifications I include a cubic in the instrument, but do not include interactions of the instrument with birth-state. These specifications still include full interactions with decade of birth in the first-stage. All subsequent analyses use only the preferred flexible logit first-stage

[20]Brinch, Mogstad and Wiswall (forthcoming, 2016) show how these assumptions can be used to estimate a MTE in cases where local IV cannot, such as with a binary instrument.

specification.

Table 1.9 shows estimates using local IV and the separate approach with and without allowing for heterogeneity in the treatment effect in X. Column (1) shows a constant effect model. Columns (2) and (3) show results using local IV. Standard errors are five to nine times larger using local IV than the separate approach. The slightly stronger parametric assumptions under the separate approach induce huge efficiency gains. The standard errors in Column (3), which does not restrict $\gamma = 0$, are too large for meaningful interpretation of the parameters.

Columns (3) and (5) allow the treatment effect to vary on observed characteristics of the individuals birth county and state.[21] As in Moffitt (2008), allowing for heterogeneity on observed and unobserved variables results in much larger standard errors. For women, the sign of β flips when making the restriction that $\gamma = 0$. As it seems improbable that the treatment effect could vary on unobserved variables but not observed variables, estimates in columns (2) and (4) are given for completeness only. However, for men, the restricted and unrestricted models are almost identical under the separate approach.

1.7.4 Average Treatment Effects for Treated and Untreated

The marginal treatment effect is only identified for U_D over the support of the propensity score, $P(X, Z)$. However, using linear extrapolation we can calculate average treatment effects on the treated and untreated. The perils of such extrapolation are well known; results should be interpreted with caution. I calculate these effects because it adds substantial clarity to the patterns of selection and treatment heterogeneity found in the MTE analysis, and to give estimates of the average treatment effects implied by the results.

Results are presented in Table 1.10. The *treated* group is comprised of individuals with at least some college, and the *untreated* have a high school degree or less. Birth cohorts have been balanced across treated and untreated groups, so as to eliminate any effect rising educational attainment over time.

Results for women show average treatment effects are close to zero for both treated and untreated groups. Differences in survival rates between the treated and untreated groups shows the strong patterns of positive selection into college education. My findings are consistent with an interpretation that selection explains essentially all of the observed difference in mortality rates between college-educated women and women with no college. This can be seen by comparing the differences in conditional outcomes between the groups (0.074 when untreated and 0.050 when treated) to the OLS estimate (0.061).

For men, we see evidence of positive selection only on the *untreated* outcome, mirroring the results found in the Black et al. selection test. I find evidence of a beneficial treatment effect for

[21] Note that caution should be taken in interpretation of α in these models.

the individuals who did not attend college. The point estimate for the average treatment on the untreated is close to OLS result for men, suggesting a rather large effect.

Table 1.11 shows the same analysis, but with all observed covariates balanced across the untreated and treated individuals. Results are largely unchanged, showing that the selection and heterogeneity patterns found are almost completely driven by unobserved rather than observed characteristics. How this analysis would change if we could observe more individual-level characteristics which are observed in other data sets, such as parental education or some measure of ability, is an open question.

1.8 Mechanisms and Discussion

In summary, my study of marginal treatment effects imply a large positive effect for men who did not attend college, and no effect for those who did. I find no evidence of a causal impact of college education on the mortality of women. In this section, I explore implications of these findings and possible mechanisms consistent with the empirical results.

1.8.1 Income

While I cannot directly separate the effects of education and income, the MTE analysis provides suggestive evidence on the causal impact of income on mortality.

Using data from the NLSY, Pedro Carneiro, James J. Heckman and Edward J. Vytlacil (2011) find the marginal treatment effect of college education on income is positive and decreasing in U_D. That is, individuals with unobservables that make them more likely to attend college are the ones who benefit the most in terms of income. With parametric models, Carneiro, Heckman and Vytlacil estimate that the average treatment effect on the treated of a year of college is a 14% increase in earnings, and the average treatment on the untreated is zero. This is consistent with the findings of Robert Moffitt (2008), who estimates the MTE of higher education on income using data from the UK.[22] In combination, these results and my findings imply that men who experience the largest income gains from college have the lowest benefits in terms of reduced mortality. The individuals who would have received near-zero gains in income from attending college had large longevity benefits. This is consistent with arguments of Cutler, Deaton and Lleras-Muney (2006) and Deaton (2013), who believe that the impact of income on mortality is small.

[22]For my analysis I do not have some of the variables observed in other studies, e.g., parental education. Thus, there is not a direct one-to-one correlation between the U_D in my work and the error term in those studies. However, they should be strongly correlated given that treatment status cannot change for any given observed variables. These papers also study later cohorts than the ones in my analysis.

1.8.2 Behavioral and Environmental

As noted earlier, there are a number of papers that discuss the potential pathways by which education might reduce mortality, and there is a literature that provides evidence about the causal impact of education on health behaviors and outcomes, such as smoking, drinking, and BMI. However, we know little about heterogeneity in the effect of education on these behaviors. It is possible that this heterogeneity is driving the results found here. For example, if individuals from high socio-economic backgrounds (which are associated with low U_D) are already predisposed not to smoke,[23] then it is plausible that the treatment effect of education on smoking is smaller for individuals with lower U_D. Further research is needed to know if the behavioral effects of education on health-related decisions and activities are heterogeneous across observed and unobserved variables.

For the cohorts I study, it is possible that education has an effect on the types of jobs individuals worked, which in turn had health impacts. There may have been important gender differences in any such relationships as well. Table 1.12 shows counts by gender and education from the 1960 Census for three different occupational categories that have may have been especially high risk in terms of health.[24] While these are a select group of occupations, the differences by gender and education are striking. Men from these cohorts without college education were much more likely to work in potentially dangerous factory or mining jobs. In all probability, men from advantaged socio-economic backgrounds, associated with low U_D, were less likely to work in the dangerous jobs regardless of their education. Men from poorer backgrounds, on the other hand, may have only been able to avoid working the most dangerous jobs if they acquired a college education. Women were less likely to be working in low-skill mining or smelting jobs irrespective of their education. Again, further research is needed to understand the causal effects of education on work-related health conditions.

1.8.3 Migration

There has been little attention to the role of lifetime migration in the relationship between education and mortality. Janna Johnson and Evan J. Taylor (2016) and Dan A. Black, Seth G. Sanders, Evan J. Taylor and Lowell J. Taylor (2015) show that migration had a negative causal effect on older-age mortality for individuals born in the early 20th century U.S. These papers also show that well educated individuals were more likely to migrate.[25]

[23]Soteriades and DiFranza (2003) found that parental education and income were strongly negatively correlated with adolescent smoking rates.

[24]The differences in college attainment rates between this table and the 2000 Census arise from the differences in cohort distribution caused by mortality prior to 2000.

[25]In ongoing work, Daniel Aaronson, Bhash Mazumder, Seth G. Sanders and Evan J. Taylor (2016) find that the mortality benefits of Rosenwald Schools for African-Americans in the South were positive but counter-balanced by

With my data, I can provide some additional evidence about how education might impact migration. Using the same method to estimate MTEs of higher education on survival, I examine MTEs of college education on lifetime migration. I define lifetime mobility as living outside one's birth state as of 2000.

OLS results, shown in Table 1.13, indicate that higher educated people have higher lifetime migration rates. Conditional on education, individuals that migrate have lower survival rates. The estimated association between survival and migration is small relative to the association with college education, but it is non-trivial.

I estimate MTEs of college education on migration. Estimates of β are -0.531 (s.e. 0.096) for men and -0.409 (s.e. 0.094) for women. I estimate average effects of 0.355 (s.e. 0.079) for men and 0.262 (s.e. 0.090) for women. College education induces higher lifetime mobility,[26] and individuals with unobservables that make them most likely to get college education have larger average treatment effects of education on migration. Individuals with high U_D have only a small treatment effect of education on migration.

These results are consistent with the possibility that some of the heterogeneity in the impact of education on mortality could be the consequence of the auxiliary effect of education on migration. Individuals with low U_D have the largest effect of education on migration, which could mute any positive direct effect of education on mortality. However, this reasoning is somewhat speculative, because we cannot directly estimate the effect of migration on mortality in this framework. Johnson and Taylor (2016) suggests that the effect of migration on mortality is not homogeneous across the population. Furthermore, MTEs on migration do not differ much between men and women, so they do not explain gender differences in the effect of education on mortality.

1.8.4 Exclusion Restriction and Bias

As discussed above, most of the potential violations of the exclusion restriction point to an upward bias in the treatment effects. Thus results might best be interpreted as upper bounds on the true effects. However, all the average treatment effects estimated are near zero except for the treatment effect on men with high school education or less. If all estimates exhibit meaningful upward bias, this implies that for most groups the true effect of education is to *increase* mortality. This strongly flies against our priors on models of education and health. It also seems unlikely that the instrument would cause positive bias only in the estimated effect on low-educated men, while leading to unbiased estimated effects for the other groups.

the negative effect of education on migration and migration on health.

[26] This result is consistent with Malamud and Wozniak (2012).

1.9 CONCLUSION

In this paper I estimate the marginal treatment effects of college education on mortality. OLS estimates show a strong association between college attendance and reduced mortality. Positive selection into education, for both men and women, accounts for part of that association. I find no evidence of a causal effect of college education on mortality for women; all of the observed differences in mortality rates between education groups can be explained by selection. For men, education reduces mortality, but only for individuals whose unobserved characteristics that make them relatively less likely to attend college. Thus there is evidence of essential heterogeneity in the treatment of college on mortality; men with the largest potential gains were those who did not attend college.

My findings, in conjunction with prior research on the impact of college education on income, suggest that income is not the channel through which education reduces mortality. Longevity benefits are highest for the individuals that have the smallest income gains. As Deaton (2013) argues, it may be that "education directly promotes health . . . because the knowledge and life lessons learned in school and college enable people to take better care of themselves and to take good advantage of the health care system when they need it." For example, a recent study by Case and Deaton (2015) shows that mortality rates for whites aged 45–54 has been increasing over the last 15 years, in part due to increase in opioid use. They find this mortality increase has been confined to individuals with high school education or less.

Cutler, Deaton and Lleras-Muney (2006) hypothesize that the "greater speed of introduction of new health-relevant knowledge and technology will tend to raise the health gradient . . ." If so, my results may well generalize to current cohorts of young people. My study of MTEs indicates that policies that increase college access at the margin will have health benefits for the men induced into higher education. For women, my results are less optimistic in terms of predicting improvements in women's longevity, but college attendance may nonetheless have important payoffs in terms of the health of the next generation.[27]

On the other hand, generalization on the basis of historical precedent is necessarily speculative. If these causal effects are driven primarily by the role of education in shifting men out of dangerous factory and mining jobs, college may now have a smaller effect on health outcomes than it did for previous generations, as these jobs become less common (and also likely safer). Changes in health care and medicine in the next 40 and 50 years could have dramatic and unpredictable impacts on any possible old-age mortality effects for current youth. Any efforts to predict how these changes will affect the mortality gains from education are fraught. Further research on the mechanisms driving the results found in this paper is crucial for evaluating the external validity for

[27] Janet Currie and Enrico Moretti (2003) show that maternal education at the college level increases infant health.

contemporary generations.

Table 1.1: Correlations between County Level Covariates

	Median Housing Value, 1930	Radio Percent, 1940	Median Family Income, 1950
Median Housing Value, 1930	1		
Percent of HH with Radio, 1940	0.585	1	
Median Family Income, 1950	0.575	0.832	1

Correlation coefficients between county level covariates, not weighted by population. Median Family Income is measured in deciles, with ten being the highest and one being the lowest

Table 1.2: Black et al. [2016] Selection Test

	Women		Men	
	Treated (College)	Untreated (HS or Less)	Treated (College)	Untreated (HS or Less)
College Proximity (100s Miles)	0.00582	0.00681	-0.00278	0.00906
	(0.00359)	(0.00376)	(0.00357)	(0.00381)
Two-sided P-Value	0.105	0.070	0.436	0.018
Observations	923,000	1,786,000	954,000	1,266,000

Estimated OLS regression coefficients on the instrument. The dependent variable is ten-year survival. Regressions are estimated separately for those with at least some college (“College") and those with high school education or less. County level economic covariates, birth state and cohort fixed effects are included.

Table 1.3: OLS Results on 10-year Survival, Women

	(1)	(2)	(3)	(4)
College Attendance	0.0625	0.0607	0.0613	0.0609
	(0.0007)	(0.0007)	(0.0007)	(0.0007)
Avg Per Capita Income (1969–2010)				0.0002
				(0.0002)
Median Housing Value (1930)				0.0006
				(0.0004)
Percent of HHs with Radio (1940)				0.0030
				(0.0081)
Cohort FE	x	x	x	x
Birth State FE		x		x
Birth County FE			x	
Observations (rounded)	2,709,000	2,709,000	2,709,000	2,709,000
R-squared	0.186	0.188	0.189	0.188
Clusters	3103	3103	3103	3103

Standard errors are clustered at the birth-county level. Mortality is measured from 2000–2010. "College Attendance" indicates at least some college. Fixed effects (FE) in deciles of 1950 Median Family Income also included in (4). Mean of the dependent variable is 0.69.

Table 1.4: OLS Results on 10-year Survival, Men

	(1)	(2)	(3)	(4)
College Attendance	0.0855	0.0833	0.0841	0.0833
	(0.0011)	(0.0009)	(0.0009)	(0.0009)
Avg Per Capita Income (1969–2010)				0.0002
				(0.0002)
Median Housing Value (1930)				0.0006
				(0.0005)
Percent of HHs with Radio (1940)				-0.0185
				(0.0081)
Cohort FE	x	x	x	x
Birth State FE		x		x
Birth County FE			x	
Observations (rounded)	2,220,000	2,220,000	2,220,000	2,220,000
R-squared	0.180	0.185	0.187	0.185
Clusters	3105	3105	3105	3105

Standard errors are clustered at the birth-county level. Mortality is measured from 2000–2010. "College Attendance" indicates at least some college. Fixed effects (FE) in deciles of 1950 Median Family Income also included in (4). Mean of the dependent variable is 0.62.

Table 1.5: Linear First Stage on College, Women

	(1)	(2)	(3)
College Proximity (100s Miles)	-0.0296	-0.0947	-0.1682
	(0.0059)	(0.0130)	(0.0225)
Proximity Squared		0.0814	0.2800
		(0.0130)	(0.0460)
Proximity Cubed			-0.1190
			(0.0244)
Avg Per Capita Income (1969-2010)	-0.0005	-0.0004	-0.0003
	(0.0005)	(0.0005)	(0.0004)
Median Housing Value (1930)	0.0107	0.0100	0.0095
	(0.0011)	(0.0011)	(0.0011)
Percent of HHs with Radio (1940)	0.1454	0.1427	0.1451
	(0.0206)	(0.0204)	(0.0203)
Observations (rounded)	2,709,000	2,709,000	2,709,000
Partial F-Statistic	25.2	26.7	21.7
Clusters	3103	3103	3103

All regression include birth-state and cohort fixed effects, and deciles of birth-county 1950 Median Family Income. The dependent variable is an indicator equal to one for at least some college. College proximity measures the distance from the birth-county to the nearest college. Standard errors are clustered at the birth-county level.

Table 1.6: Linear First Stage on College, Men

	(1)	(2)	(3)
College Proximity (100s Miles)	-0.0506	-0.1390	-0.2438
	(0.0083)	(0.0179)	(0.0325)
Proximity Squared		0.1110	0.3940
		(0.0172)	(0.0650)
Proximity Cubed			-0.1700
			(0.0329)
Avg Per Capita Income, (1969-2010)	-0.0033	-0.0032	-0.0031
	(0.0010)	(0.0010)	(0.0009)
Median Housing Value (1930)	0.0229	0.0220	0.0213
	(0.0020)	(0.0019)	(0.0019)
Percent of HHs with Radio (1940)	0.2557	0.2522	0.2552
	(0.0296)	(0.0293)	(0.0292)
Observations (rounded)	2,220,000	2,220,000	2,220,000
Partial F-Statistic	36.9	30.7	22.4
Clusters	3105	3105	3105

All regression include birth-state and cohort fixed effects, and deciles of birth-county 1950 Median Family Income. The dependent variable is an indicator equal to one for at least some college. College proximity measures the distance from the birth-county to the nearest college. Standard errors are clustered at the birth-county level.

Table 1.7: IV Results on 10-year Survival

	Women			Men		
	(1)	(2)	(3)	(4)	(5)	(6)
Deg(Z), 1st Stage:	Linear	Quadratic	Cubic	Linear	Quadratic	Cubic
College Attendance	-0.1561	-0.0509	-0.0558	-0.0289	0.0321	0.0537
	(0.1059)	(0.0670)	(0.0614)	(0.0567)	(0.0423)	(0.0404)
Avg Per Capita	0.0000	0.0001	0.0001	-0.0002	0.0000	0.0001
Inc (1970-2010)	(0.0002)	(0.0002)	(0.0002)	(0.0004)	(0.0003)	(0.0003)
Median Housing	0.0030	0.0019	0.0020	0.0033	0.0019	0.0014
Value (1930)	(0.0013)	(0.0009)	(0.0008)	(0.0015)	(0.0011)	(0.0011)
Pct of HHs with	0.0377	0.0209	0.0217	0.0130	-0.0042	-0.0102
Radio (1940)	(0.0194)	(0.0136)	(0.0129)	(0.0179)	(0.0142)	(0.0138)
Obs. (rounded)	2,709,000	2,709,000	2,709,000	2,220,000	2,220,000	2,220,000
Clusters	3103	3103	3103	3105	3105	3105

All regressions include birth-state and cohort fixed effects, and deciles of birth-county 1950 Median Family Income. Mortality is measured from 2000–2010. College Attendance indicates at least some college. Standard errors are clustered at the birth-county level.

Table 1.8: Estimated Slope of the MTE of College Attendance on Survival

Panel A. Women

First Stage:	(1) Flexible Logit	(2) Linear Probability	(3) Logit
$\hat{\beta}$	0.052	0.059	0.082
	(0.040)	(0.046)	(0.046)

Panel B. Men

First Stage:	(1) Flexible Logit	(2) Linear Probability	(3) Logit
$\hat{\beta}$	0.086	0.074	0.079
	(0.050)	(0.062)	(0.061)

All regressions include birth-state and cohort fixed effects, and deciles of birth-county 1950 Median Family Income. Mortality measured from 2000–2010. College Attendance indicates at least some college. Standard errors are clustered at the birth-county level.

Table 1.9: Parameter Estimates of $\Delta^{MTE}(X, U_D) = \alpha + \beta U_D + \gamma X$, with and without Heterogeneity in Observables X

Panel A. Women

	(1) Constant Effect	(2) Local IV	(3) Local IV	(4) Separate Approach	(5) Separate Approach
$\hat{\alpha}$	-0.003	0.086	-0.094	0.014	-0.042
	(0.024)	(0.048)	(0.118)	(0.025)	(0.030)
$\hat{\beta}$		-0.241	0.597	-0.049	0.052
		(0.102)	(0.396)	(0.014)	(0.040)
$\gamma = 0$	yes	yes	no	yes	no

Panel A. Men

	(1) Constant Effect	(2) Local IV	(3) Local IV	(4) Separate Approach	(5) Separate Approach
$\hat{\alpha}$	0.028	0.004	-0.067	-0.012	0.005
	(0.018)	(0.038)	(0.106)	(0.019)	(0.027)
$\hat{\beta}$		0.061	0.291	0.085	0.086
		(0.080)	(0.285)	(0.014)	(0.050)
$\gamma = 0$	yes	yes	no	yes	no

Mortality measured from 2000–2010. Standard errors are clustered at the birth-county level.

Table 1.10: Average Treatment Effects for the Treated and Untreated

Panel A. Women

	Untreated	Treated	Difference
Average Y_0	0.665	0.739	0.074
	(0.003)	(0.031)	(0.030)
Average Y_1	0.679	0.728	0.050
	(0.028)	(0.003)	(0.029)
Average Treatment, $Y_1 - Y_0$	0.014	-0.010	-0.024
	(0.028)	(0.030)	(0.035)
Observations	1,786,000	923,000	

Panel B. Men

	Untreated	Treated	Difference
Average Y_0	0.586	0.666	0.080
	(0.003)	(0.025)	(0.024)
Average Y_1	0.648	0.668	0.020
	(0.023)	(0.003)	(0.023)
Average Treatment, $Y_1 - Y_0$	0.062	0.002	-0.060
	(0.023)	(0.025)	(0.032)
Observations	1,266,000	954,000	

Birth cohorts are weighted for balance across treatment and untreated. Calculations are conducted using extrapolation of the Marginal Treatment Effect. Standard errors clustered at the county level.

Table 1.11: Average Treatment Effects for the Treated and Untreated with Balanced Observables

Panel A. Women

	Untreated	Treated	Difference
Average Y_0	0.667	0.737	0.070
	(0.003)	(0.031)	(0.029)
Average Y_1	0.680	0.727	0.047
	(0.028)	(0.003)	(0.028)
Average Treatment, $Y_1 - Y_0$	0.013	-0.010	-0.023
	(0.028)	(0.030)	(0.034)
Observations	1,786,000	923,000	

Panel B. Men

	Untreated	Treated	Difference
Average Y_0	0.589	0.663	0.075
	(0.003)	(0.024)	(0.023)
Average Y_1	0.650	0.666	0.016
	(0.022)	(0.003)	(0.022)
Average Treatment, $Y_1 - Y_0$	0.061	0.003	-0.058
	(0.023)	(0.025)	(0.030)
Observations	1,266,000	954,000	

All observed covariates are balanced across the treated and untreated. Calculations are based on extrapolation of the Marginal Treatment Effect. Standard errors clustered at the county level.

Table 1.12: Gender and Education in Selected Dangerous Occupations, 1960

	Men		Women	
Occupation	College	HS or Less	College	HS or Less
Furnacemen, smeltermen, and pourers	75	1,375	1	41
Mine operatives and laborers, n.e.c.	595	10,481	5	48
Motormen, mine, factory, logging camp, etc.	6	415	0	1
Total observations	328,869	915,512	254,468	1,029,656

Author's calculations using the 1960 IPUMS, Ruggles et al. (2015). Counts are for whites born in the U.S., 1911–1940. "College" indicates at least some college.

Table 1.13: OLS Regressions of Location and Survival Rates

Panel A. Women

	Dependent Variable		
	Live Outside Birth State	10-Year Survival	10-Year Survival
College Attendance	0.1464		0.0623
	(0.0037)		(0.0007)
Live Outside Birth State		-0.0008	-0.0095
		(0.0009)	(0.0008)
Observations (rounded)	2,709,000	2,709,000	2,709,000
Clusters	3103	3103	3103

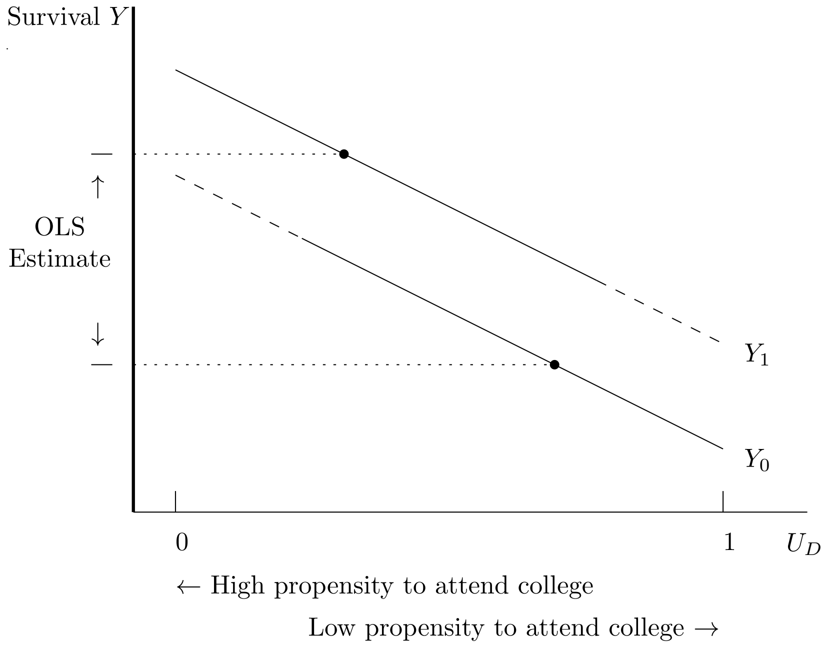
Panel B. Men

	Dependent Variable		
	Live Outside Birth State	10-Year Survival	10-Year Survival
College Attendance	0.1912		0.0844
	(0.0038)		(0.0009)
Live Outside Birth State		0.0106	-0.0058
		(0.0008)	(0.0007)
Observations (rounded)	2,220,000	2,220,000	2,220,000
Clusters	3105	3105	3105

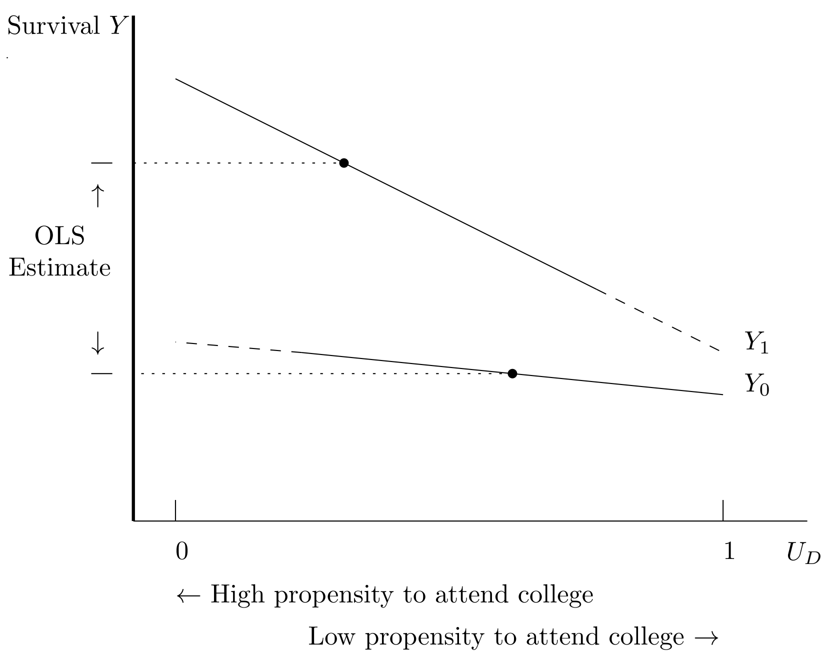
"Live Outside Birth State" is an indicator equal to 1 if an individual lives in a different state than the state of birth as of 2000. Mortality measured from 2000–2010. All regression include birth-state and cohort fixed effects. Standard errors clustered at birth-county level.

Figure 1.1: Patterns of Selection into Treatment (College Attendance)

A. Ability Bias



B. Selection on Gains from Treatment



C. Treatment Effect Highest among those Selecting Out of Treatment

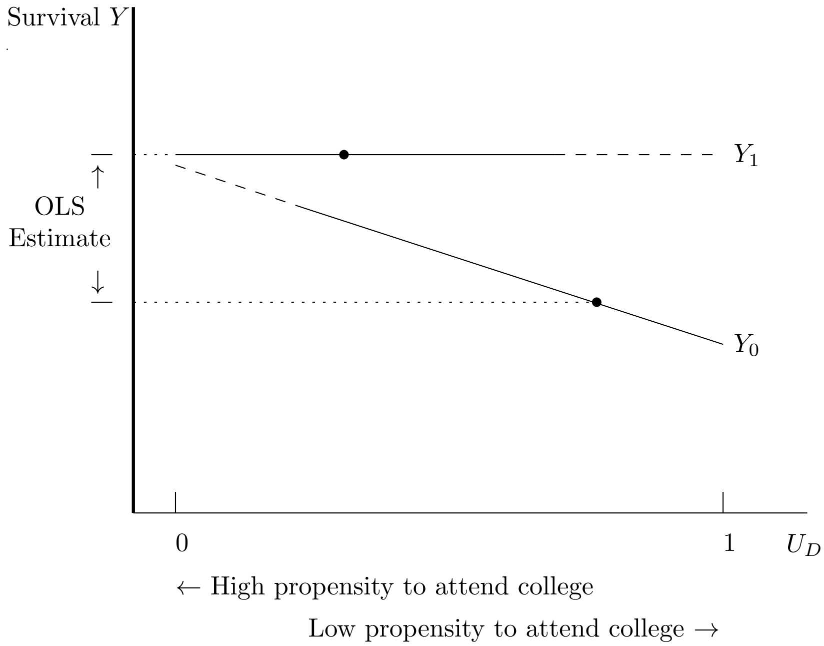
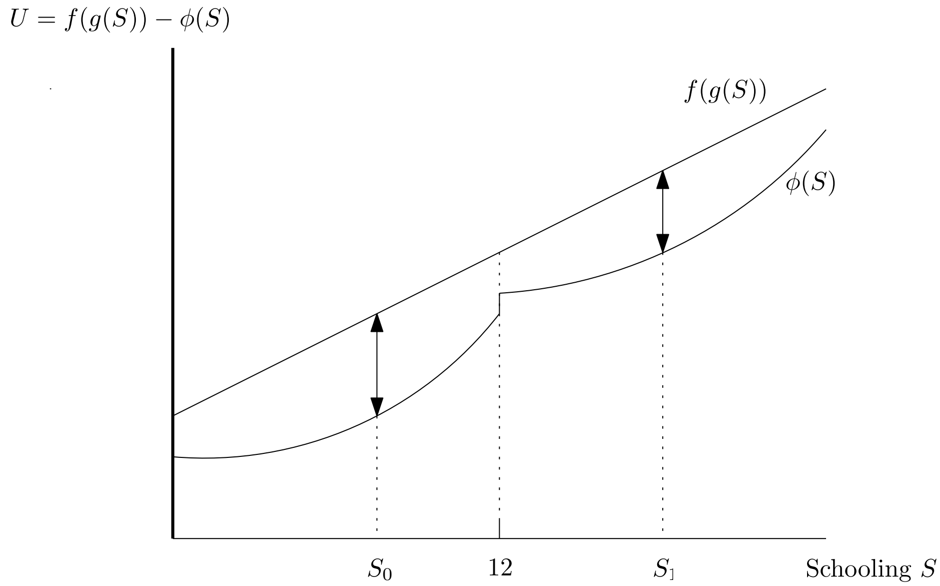
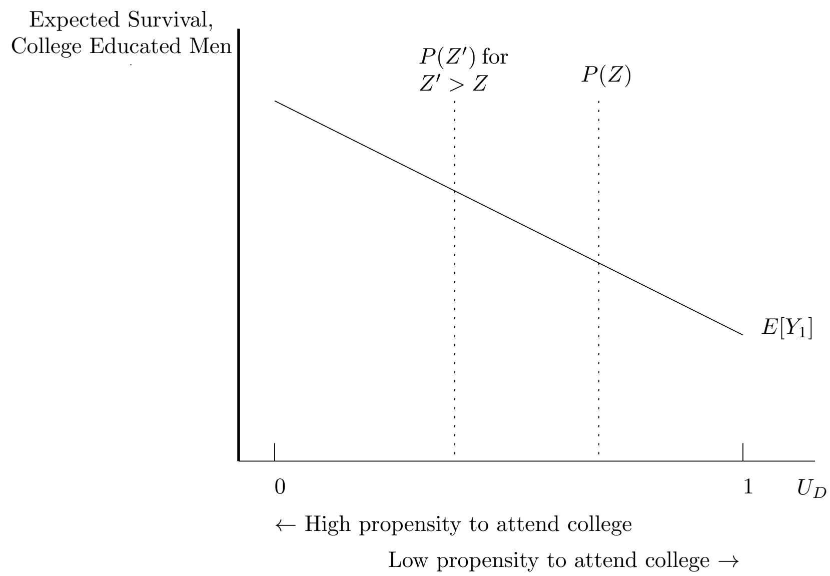


Figure 1.2: An Example of Benefit and Cost Curves for Schooling Choice



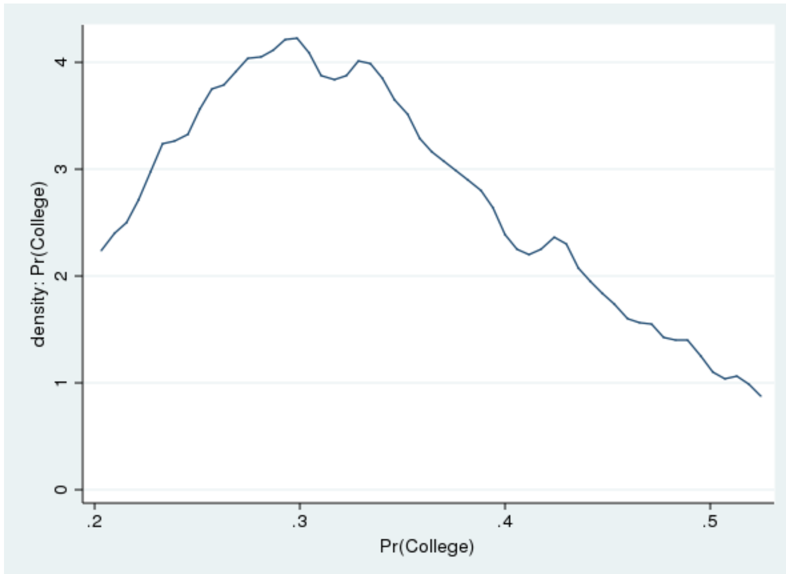
The instrument can only induce parallel shifts of the cost curve, $\phi(S)$ for $S > 12$, and have no effect on the benefit curve, $f(g(S))$. S_0 shows the optimal choice of schooling for $S \leq 12$ and S_1 shows the optimal choice of schooling for $S > 12$.

Figure 1.3: An Example of Positive Selection on the Treated Outcome



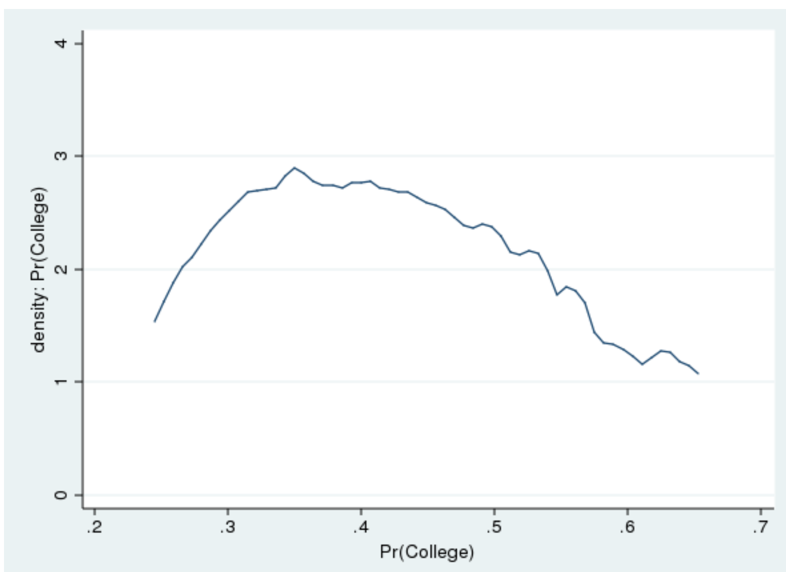
When P = P(Z), all individuals to left of the right most vertical line become treated. When P = P(Z') the indifference line shifts left, which increases the average outcome for treated individuals.

Figure 1.4: Histogram of Propensity Scores for College, White Women born 1911-1940



Density is trimmed at 5 pct on both sides per Census guidelines

Figure 1.5: Histogram of Propensity Scores for College, White Men born 1911-1940



Density is trimmed at 5 pct on both sides per Census guidelines

Chapter 2 Social Interactions and Location Decisions: Evidence from U.S. Mass Migration

2.1 Introduction

A large and growing literature finds that social interactions influence many economic outcomes, including crime, education, and employment (for recent reviews, see Blume et al., 2011; Epple and Romano, 2011; Munshi, 2011; Topa, 2011). While research has long-recognized the effect of location decisions on individual and aggregate economic outcomes, there is little evidence on the importance of social interactions in location decisions, and even less evidence on the types of individuals or economic conditions for which social interactions are most important. Evidence on the role of social interactions in location decisions would inform theoretical models of migration, the equilibrium of local labor markets, and the impacts of policies that affect migration incentives.

This paper provides new evidence on the magnitude and nature of social interactions in location decisions. We focus on the mass migrations in the mid-twentieth century of African Americans from the South and whites from the Great Plains. The millions of moves in these episodes yield particularly valuable settings for studying the long-run effects of social interactions on location decisions. We use confidential administrative data that measure town of birth and county of residence at old age for most of the U.S. population born from 1916-1936. Detailed geographic information allows us to distinguish birth town-level social interactions from other determinants of location decisions, such as expected wages or moving costs. For example, we observe that 51 percent of African-American migrants born from 1916-1936 in Pigeon Creek, Alabama moved to Niagara County, New York, while less than six percent of black migrants from nearby towns moved to the same county.

We develop a new, intuitive method of characterizing social interactions in location decisions. The social interactions (SI) index allows us to estimate the strength of social interactions for each receiving and sending location, which we then relate to locations' economic characteristics. We show that existing methods may mischaracterize the strength of social interactions in our setting. In particular, the widely used approach of Bayer, Ross and Topa (2008) could estimate strong social interactions for popular destinations even if social interactions are relatively weak, and as a result could misstate the overall strength of social interactions.[1] Our method does not suffer from this problem. Under straightforward and partly testable assumptions, the SI index identifies the effect of social interactions and maps directly to social interaction models.

We find very strong social interactions among Southern black migrants and smaller interactions among whites from the Great Plains. Our estimates imply that if we observed one randomly chosen African American move from a birth town to some destination county, then on average 1.9 additional black migrants from that birth town would make the same move. For white migrants from the Great Plains, the average is only 0.4, and results for Southern whites are similarly small. Interpreted through the social interactions model of Glaeser, Sacerdote and Scheinkman (1996), our estimates imply that 49 percent of African-American migrants chose their long-run destination because of social interactions, while 16 percent of Great Plains whites were similarly influenced.

To understand the nature of social interactions in location decisions, we examine whether economic characteristics of receiving and sending locations are associated with stronger social interactions. Social interactions among African Americans were stronger in destinations with a higher share of 1910 employment in manufacturing, a particularly attractive sector for black workers in our sample. This evidence highlights an important role for job referrals in determining location decisions, and suggests that job referrals were more valuable in destinations with better employment opportunities. We also find that social interactions were weaker in destinations that were further away and less connected by railroads, pointing to the importance of access to information and low mobility costs. Social interactions were stronger in destinations with fewer African Americans in 1900, suggesting that networks helped migrants find opportunities in new places. In addition, social interactions were stronger in sending counties with higher literacy rates in 1920, suggesting that education facilitated social interactions.

Several pieces of evidence support the validity of our empirical strategy. Our research design asks whether individuals born in the same town were more likely to live in the same destination in old age than individuals born in nearby towns. This design implies that SI index estimates should not change when controlling for observed birth town covariates, because geographic proximity controls for the relevant determinants of location decisions. Reassuringly, our estimates are essentially unchanged when adding several covariates. We also estimate strong social interactions in certain locations, like Rock County, Wisconsin, for which rich qualitative work supports our findings (Bell, 1933; Rubin, 1960; Wilkerson, 2010).

This paper makes three contributions. First, we develop a new method of characterizing the magnitude and nature of social interactions. Our approach integrates previous work by Glaeser, Sacerdote and Scheinkman (1996) and Bayer, Ross and Topa (2008), has desirable theoretical and statistical properties, and can be used to study social interactions in a variety of other settings. Sec-

[1]This potential problem also applies to studies of social interactions in employment and other outcomes.

ond, we provide new evidence on the importance of social interactions in location decisions and the types of individuals and economic conditions for which social interactions are most important. Previous work shows that individuals tend to migrate to the same areas, often broadly defined, as other individuals from the same town or country, but does not isolate the role of social interactions (Bartel, 1989; Bauer, Epstein and Gang, 2005; Beine, Docquier and Ozden, 2011; Giuletti, Wahba and Zenou, 2014; Spitzer, 2014).[2] Third, our results inform landmark migration episodes that have drawn interest from economists for almost a century (Scroggs, 1917; Smith and Welch, 1989; Carrington, Detragiache and Vishwanath, 1996; Collins, 1997; Boustan, 2009, 2011; Hornbeck, 2012; Hornbeck and Naidu, 2014; Johnson and Taylor, 2016; Black et al., 2015; Collins and Wanamaker, 2015). Our results complement the small number of possibly unrepresentative historical accounts suggesting that social interactions were important in these migration episodes (Rubin, 1960; Gottlieb, 1987; Gregory, 1989).

Our paper also complements recent work by Chay and Munshi (2015). They find that, above a threshold, migrants born in counties with higher plantation crop intensity tend to move to fewer locations, as measured by a Herfindahl-Hirschman Index, and show that this non-linear relationship accords with a network formation model with fixed costs of participation. We find some evidence that social interactions were stronger in denser sending communities, consistent with the results in Chay and Munshi (2015). We differ in our empirical methodology, study of white migrants from the Great Plains and South, and examination of how social interactions vary across destinations.

2.2 Historical Background on Mass Migration Episodes

The Great Migration saw nearly six million African Americans leave the South from 1910 to 1970 (Census, 1979). Although migration was concentrated in certain destinations, like Chicago, Detroit, and New York, other cities also experienced dramatic changes. For example, Chicago's black population share increased from two to 32 percent from 1910-1970, while Racine, Wisconsin experienced an increase from 0.3 to 10.5 percent (Gibson and Jung, 2005). Migration out of the South increased from 1910-1930, slowed during the Great Depression, and then resumed forcefully from 1940 to the 1970's. Panel A of Figure 2.2 shows that the vast majority of African American migrants born from 1916-1936, who comprise our analysis sample, moved out of the South between 1940 and 1960. Most of these migrants moved between age 15 and 35 (Panel A of Appendix Figure A.1).

Several factors contributed to the exodus of African Americans from the South. World War I, which simultaneously increased labor demand among Northern manufacturers and decreased labor supply from European immigrants, helped spark the Great Migration, although many un-

[2] One exception is Chen, Jin and Yue (2010), who study the impact of peer migration on temporary location decisions in China, but lack detailed geographic information on where individuals move.

derlying causes existed long before the war (Scroggs, 1917; Scott, 1920; Gottlieb, 1987; Marks, 1989; Jackson, 1991; Collins, 1997; Gregory, 2005). Underlying causes included a less developed Southern economy, the decline in agricultural labor demand due to the boll weevil's destruction of crops (Scott, 1920; Marks, 1989, 1991; Lange, Olmstead and Rhode, 2009), widespread labor market discrimination (Marks, 1991), and racial violence and unequal treatment under Jim Crow laws (Tolnay and Beck, 1991).

Migrants tended to follow paths established by railroad lines: Mississippi-born migrants predominantly moved to Illinois and other Midwestern states, and South Carolina-born migrants predominantly moved to New York and Pennsylvania (Scott, 1920; Carrington, Detragiache and Vishwanath, 1996; Collins, 1997; Boustan, 2011; Black et al., 2015). Labor agents, offering paid transportation, employment, and housing, directed some of the earliest migrants, but their role diminished sharply after the 1920's, and most individuals paid for the relatively expensive train fares themselves (Gottlieb, 1987; Grossman, 1989).[3] African-American newspapers from the largest destinations circulated throughout the South, providing information on life in the North (Gottlieb, 1987; Grossman, 1989).[4] Blacks attempting to leave the South sometimes faced violence (Scott, 1920; Henri, 1975).

A small number of historical accounts suggest a role for social interactions in location decisions. Social networks, consisting primarily of family, friends, and church members, provided valuable job references or shelter (Rubin, 1960; Gottlieb, 1987). For example, Rubin (1960) finds that migrants from Houston, Mississippi had close friends or family at two-thirds of all initial destinations.[5] These accounts motivate our focus on birth town-level social interactions.

The experience of John McCord captures many important features of early black migrants' location decision.[6] Born in Pontotoc, Mississippi, nineteen-year-old McCord traveled in search of higher wages in 1912 to Savannah, Illinois, where a fellow Pontotoc-native connected him with a job. McCord moved to Beloit, Wisconsin in 1914 after hearing of employment opportunities and quickly began working as a janitor at the manufacturer Fairbanks Morse and Company. After two years in Beloit, McCord spoke to his manager about returning home for a vacation. The manager asked McCord to recruit workers during the trip. McCord returned with 18 unmarried men, all of whom were soon hired. Thus began a persistent flow of African Americans from Pontotoc to Beloit: among individuals born from 1916-1936, 14 percent of migrants from Pontotoc lived in Beloit's county at old age (see Table 2.2, discussed below).

[3]In 1918, train fare from New Orleans to Chicago cost $22 per person, when Southern farmers' daily wages typically were less than $1 and wages at Southern factories were less than $2.50 (Henri, 1975).

[4]The *Chicago Defender*, perhaps the most prominent African-American newspaper of the time, was read in 1,542 Southern towns and cities in 1919 (Grossman, 1989).

[5]Rubin (1960) studied individuals from Houston, Mississippi because so many migrants from Houston moved to Beloit, Wisconsin, so this is clearly not a representative sample.

[6]The following paragraph draws on Bell (1933). See also Knowles (2010).

Migration out of the Great Plains has received less academic attention than the Great Migration, but nonetheless represents a landmark reshuffling of the U.S. population. Considerable out-migration from the Great Plains started around 1930 (Johnson and Rathge, 2006). Among whites born in the Great Plains from 1916-1936, the most rapid out-migration occurred from 1940-1960, as seen in Panel B of Figure 2.2. Most of these migrants left the Great Plains by age 35 (Panel B of Appendix Figure A.1). Explanations for the out-migration include the decline in agricultural prices due to the Great Depression, a drop in agricultural productivity due to drought, and the mechanization of agriculture (Gregory, 1989; Curtis White, 2008; Hurt, 2011; Hornbeck, 2012). Some historical work points to an important role for social interactions in location decisions (Jamieson, 1942; Gregory, 1989).[7]

The mass migrations out of the South and Great Plains share several traits. Both episodes featured millions of people making long-distance moves in search of better economic and social opportunities. Furthermore, both episodes saw a similar share of the population undertake long-distance moves. Figure 2.3 shows that 97 percent of blacks born in the South and 90 percent of whites born in the Great Plains lived in their birth region in 1910, and out-migration reduced this share to 75 percent for both groups by 1970. Both African American and white migrants experienced discrimination in many destinations, although African Americans faced more severe discrimination and had less wealth (Gregory, 2005).

2.3 Estimating Social Interactions in Location Decisions

2.3.1 Data on Location Decisions

We use confidential administrative data to measure location decisions made during two historical mass migration episodes. In particular, we use the Duke University SSA/Medicare data, which covers over 70 million individuals who received Medicare Part B from 1976-2001. The data contain race, sex, date of birth, date of death (if deceased), and the ZIP code of residence at old age (death or 2001, whichever is earlier). In addition, the data include a 12-character string with self-reported birth town information, which is matched to places, as described in Black et al. (2015). We use the data to measure long-run migration flows from birth town to destination county for individuals born from 1916-1936.[8] This sample lies at the center of both mass migration episodes and likely contains very few parent-child pairs. To improve the reliability of our estimates, we restrict the sample to birth towns with at least ten migrants and, separately for each birth state, combine all destination counties with less than ten migrants.

[7] Jamieson (1942) finds that almost half of migrants to Marysville, California had friends or family living there.

[8] Our sample begins with the 1916 cohort because coverage rates are low for prior years (Black et al., 2015) and ends with 1936 because that is the last cohort available in the data.

Panels A and B of Figure 2.4 display the states we include in the South and Great Plains. For migration out of the South, we study individuals born in Alabama, Georgia, Florida, Louisiana, Mississippi, North Carolina, and South Carolina. We define a migrant as someone who moved out of the 11 former Confederate states.[9] For migration out of the Great Plains, we study individuals born in Kansas, Oklahoma, Nebraska, North Dakota, and South Dakota. We define a migrant as someone who moved out of the Great Plains and a border region, shaded in light grey in Panel B.[10] We make these choices to focus on the long-distance moves that characterize both migration episodes.

Our data capture long-run location decisions, as we only observe an individual's location at birth and old age. We cannot identify return migration: if an individual moved from Mississippi to Wisconsin, then returned to Mississippi at age 60, we do not count that person as a migrant. We also do not observe individuals who die before age 65 or do not enroll in Medicare. We discuss the implications of these measurement issues below.

2.3.2 ECONOMETRIC MODEL: THE SOCIAL INTERACTIONS INDEX

We first introduce some notation and discuss the basic idea underlying our approach to estimating social interactions.[11] Let $D_{i,j,k} = 1$ if migrant i moves from birth town j to destination county k and $D_{i,j,k} = 0$ if migrant i moves elsewhere. The probability of a migrant born in town j choosing destination k is $P_{j,k} \equiv \mathbb{E}[D_{i,j,k}]$. This probability reflects individuals' preferences, resources, and the expected return to migration, but does not depend on other individuals' realized location decisions. The number of people who move from birth town j to destination k is $N_{j,k} \equiv \sum_{i \in j} D_{i,j,k}$, and the number of migrants from birth town j is $N_j \equiv \sum_k N_{j,k}$.

Positive social interactions increase the variance of individuals' decisions (Glaeser, Sacerdote and Scheinkman, 1996; Bayer, Ross and Topa, 2008; Graham, 2008). To see this, imagine that we observed multiple realizations of $N_{j,k}$ from a fixed data generating process. The across-realization variance of location decisions for a single birth town-destination county pair would be

$$\begin{aligned} \mathbb{V}[N_{j,k}] &= \sum_{i \in j} \mathbb{V}[D_{i,j,k}] + \sum_{i \neq i' \in j} \mathbb{C}[D_{i,j,k}, D_{i',j,k}] \\ &= N_j P_{j,k}(1 - P_{j,k}) + N_j(N_j - 1)C_{j,k}, \end{aligned} \tag{2.1}$$

where $C_{j,k} \equiv \sum_{i \neq i' \in j} \mathbb{C}[D_{i,j,k}, D_{i',j,k}]/(N_j(N_j - 1))$ is the average covariance of location decisions

[9] These include the seven states already listed, plus Arkansas, Tennessee, Texas, and Virginia.

[10] This border region includes Arkansas, Colorado, Iowa, Minnesota, Missouri, Montana, New Mexico, Texas, and Wyoming.

[11] See Brock and Durlauf (2001) and Blume et al. (2011) for comprehensive discussions of various approaches to estimating social interactions.

for two migrants from the same town. Positive social interactions ($C_{j,k} > 0$) clearly increase the variance of location decisions. If, counterfactually, we observed multiple realizations of $N_{j,k}$, we could directly estimate $\mathbb{V}[N_{j,k}]$ and $P_{j,k}$, which leads to an estimate of $C_{j,k}$ from equation (2.1). Because we observe a single set of location decisions for each (j, k) pair, we use an econometric model to estimate social interactions.

A natural starting point for our econometric model is the widely used approach of Bayer, Ross and Topa (2008), who propose an empirical strategy that uses detailed geographic data to identify social interactions. Extending their model to our setting yields

$$D_{i,j(i),k} D_{i',j(i'),k} = \alpha_{g,k} + \sum_{j \in g} \beta_{j,k} 1[j(i) = j(i') = j] + \epsilon_{i,i',k}, \tag{2.2}$$

where $j(i)$ is the birth town of migrant i, and both i and i' live in birth town group g. As described below, we define birth town groups in two ways: counties and square grids independent of county borders. The fixed effect $\alpha_{g,k}$ equals the average propensity of migrants from birth town group g to co-locate in destination k, and $\beta_{j,k}$ equals the additional propensity of individuals from the same birth town j to co-locate in k.[12] Equation (2.2) allows location decision determinants to vary arbitrarily at the birth town group-destination level through $\alpha_{g,k}$ (e.g., because of differences in migration costs due to railroad lines or highways).

To better understand the reduced-form model in equation (2.2), we show how to map the parameters of the extended Bayer, Ross and Topa (2008) model, $(\alpha_{g,k}, \beta_{j,k})$, into classic parameters governing social interactions, $(P_{j,k}, C_{j,k})$. Doing so requires two assumptions. The most important assumption is that $P_{j,k}$ is constant across birth towns in the same group:

Assumption 1. $P_{j,k} = P_{j',k}$ for different birth towns in the same birth town group, $j \neq j' \in g$.

Assumption 1 formalizes the idea that there are no ex-ante differences across nearby birth towns in the value of moving to destination k. For example, this assumes away the possibility that migrants from Pigeon Creek, Alabama had preferences or human capital particularly suited for Niagara Falls, New York relative to migrants from a nearby town, such as Oaky Streak, which is six miles away. This assumption attributes large differences in realized moving propensities across nearby towns to social interactions. Assumption 1 covers the probability of choosing a destination, conditional on migrating; we make no assumptions regarding out-migration probabilities.

Assumption 1 is plausible in our setting. Preferences for destination features (e.g., wages or climate) likely did not vary sharply across nearby birth towns, and individuals had little information about most destinations outside of what was relayed through social networks. Furthermore, African

[12] Bayer, Ross and Topa (2008) study the propensity of workers that live in the same census block to work in the same census block, beyond the propensity of workers living in the same block group (a larger geographic area) to work in the same block. In their initial specification, $\alpha_{g,k}$ does not vary by k, and $\beta_{j,k}$ does not vary by j or k. In other specifications, they allow the slope coefficient to depend on observed characteristics of the pair (i, i').

Americans tended to work in different industries in the North and South, suggesting a negligible role for human capital specific to a destination county that differed across nearby towns. The fixed effect $\alpha_{g,k}$ accounts for broader variation in human capital, such as the fact that some Great Plains migrants chose specific destinations in California to pick cotton (Gregory, 1989). Conditional on migrating, the cost of moving to a given destination likely did not vary sharply across nearby towns.[13]

Importantly, Assumption 1 yields a testable prediction. The assumption relies on geographic proximity to control for the relevant determinants of location decisions, which implies that using birth town-level covariates to explain moving probabilities should not affect estimates of $P_{j,k}$ or our social interactions estimates. As discussed in detail below, we test this prediction and find evidence consistent with Assumption 1.

The second assumption is that social interactions occur only among individuals from the same birth town:

Assumption 2. $\mathbb{C}[D_{i,j,k}, D_{i',j',k}] = 0$ for individuals from different birth towns, $j \neq j'$.

Assumption 2 allows us to map the parameters of the extended Bayer, Ross and Topa (2008) model, $(\alpha_{g,k}, \beta_{j,k})$, into the key parameters governing social interactions, $(P_{j,k}, C_{j,k})$. Positive social interactions across nearby towns, which violate Assumption 2, would lead us to underestimate the strength of town-level social interactions, $\beta_{j,k}$.

Under Assumptions 1 and 2, the slope coefficient in equation (2.2) equals the covariance of location decisions from birth town j to destination k: $\beta_{j,k} = C_{j,k}$.[14] In addition, the fixed effect in equation (2.2) equals the squared moving probability: $\alpha_{g,k} = P_{g,k}^2$, where $P_{g,k}$ is the probability of moving from birth town group g to destination k. This analysis demonstrates that the Bayer, Ross and Topa (2008) model uses the covariance of decisions to measure social interactions.

In certain settings, the Bayer, Ross and Topa (2008) model could mischaracterize the strength of social interactions. To see this, let $\mu_{j,k} \equiv \mathbb{E}[D_{i,j,k}|D_{i',j,k} = 1]$ be the probability that a migrant moves from birth town j to destination k, conditional on a randomly chosen migrant from birth town j making the same move. Slight manipulation of the definition of the covariance of location

[13] Assumption 1 is not violated if the cost of moving to all destinations varied sharply across birth towns (e.g., because of proximity to a railroad), as we focus on where people move, conditional on migrating.

[14] Proof:

$$
\begin{aligned}
\beta_{j,k} &= \mathbb{E}[D_{i,j(i),k} D_{i',j(i'),k} | j(i) = j(i') = j] - \mathbb{E}[D_{i,j(i),k} D_{i',j(i'),k} | j(i) \neq j(i')] \\
&= \mathbb{E}[D_{i,j(i),k} D_{i',j(i'),k} | j(i) = j(i') = j] - (\mathbb{E}[D_{i,j,k}])^2 \\
&= \mathbb{C}[D_{i,j,k}, D_{i',j,k}] = C_{j,k}
\end{aligned}
$$

The first line follows directly from equation (2.2). The second line follows from Assumptions 1 and 2. The third line follows from the definition of covariance.

decisions yields

$$C_{j,k} = P_{g,k} \left(\mu_{j,k} - P_{g,k} \right). \tag{2.3}$$

Equation (2.3) shows that variation in $C_{j,k}$ arises from two sources: the probability of moving to a destination, $P_{g,k}$, and the "marginal social interaction effect," $\mu_{j,k} - P_{g,k}$. For example, $C_{j,k}$ could be large for a popular destination like Chicago because $P_{g,k}$ is large, even if $\mu_{j,k} - P_{g,k}$ is small. For less popular destinations, $\mu_{j,k} - P_{g,k}$ could be large, but $C_{j,k}$ will be small if $P_{g,k}$ is sufficiently small. Because $P_{g,k}$ varies tremendously across destinations in our setting, the covariance of location decisions, $C_{j,k}$, or any aggregation of $C_{j,k}$ is not an attractive measure of social interactions.[15]

To characterize the strength of social interactions for receiving and sending locations, we propose an intuitive social interactions (SI) index that equals the expected increase in the number of people from birth town j that move to destination county k when an arbitrarily chosen person i is observed to make the same move,

$$\Delta_{j,k} \equiv \mathbb{E}[N_{-i,j,k} | D_{i,j,k} = 1] - \mathbb{E}[N_{-i,j,k} | D_{i,j,k} = 0], \tag{2.4}$$

where $N_{-i,j,k}$ is the number of people who move from j to k, excluding person i. A positive value of $\Delta_{j,k}$ indicates positive social interactions in moving from j to k, while $\Delta_{j,k} = 0$ indicates no social interactions.

The SI index, $\Delta_{j,k}$, possesses several attractive properties as a method of measuring social interactions. The SI index permits meaningful comparisons of social interactions across heterogeneous receiving and sending locations. In addition, the SI index is consistent with and can be mapped directly to multiple simple structural models. The weak structural assumptions embedded in the SI index are valuable because of the considerable uncertainty about the true model. For example, suppose that all migrants in town j form coalitions of size s, all members of a coalition move to the same destination, and all coalitions move independently of each other. In this case, the SI index for each destination k depends only on the structural parameter s ($\Delta_{j,k} = s - 1$), while the covariance of location decisions depends on additional parameters that complicate comparisons across receiving and sending locations ($C_{j,k} = (s-1)P_{g,k}(1 - P_{g,k})/(N_j - 1)$). As another example, we connect the SI index to the model of Glaeser, Sacerdote and Scheinkman (1996) in Section 2.4.5. Another attractive property of the SI index that we demonstrate below is that it can be estimated non-parametrically with increasingly available data. The SI index could be used to study social interactions for many outcomes besides location choices, such as where individuals

[15] This issue likely arises in other applications, as there is considerable variation in the probability of working at specific locations or establishments.

work.

In Appendix A.1, we show that the SI index can be written as

$$\Delta_{j,k} = \frac{(\mu_{j,k} - P_{g,k})(N_j - 1)}{1 - P_{g,k}} = \frac{C_{j,k}(N_j - 1)}{P_{g,k} - P_{g,k}^2}. \tag{2.5}$$

Several features of equation (B21) are noteworthy. First, the SI index depends on the classic parameters governing social interactions, $(P_{g,k}, C_{j,k})$. Second, the SI index increases in the marginal social interaction effect, $\mu_{j,k} - P_{g,k}$. If migrants move independently of each other, then $\mu_{j,k} - P_{g,k} = \Delta_{j,k} = 0$. Third, the SI index scales down $C_{j,k}$ for more popular destinations, as $P_{g,k} << 0.5$ is the relevant range in our setting. Finally, the SI index does not necessarily increase in the number of migrants from birth town j, N_j, as the marginal social interaction effect might decrease in N_j.[16]

2.3.3 Estimating the Social Interactions Index

As suggested by equation (B21), estimating the SI index is straightforward. We first define birth town groups, and then non-parametrically estimate the underlying parameters $P_{g,k}$, $P_{g,k}^2$, and $C_{j,k}$.

We define birth town groups in two ways. Our preferred approach balances the inclusion of very close towns, for which Assumption 1 likely holds, with the inclusion of towns that are further away and lead to a more precise estimate of $P_{g,k}$. We divide each birth state into a grid of squares with sides x^* miles long and choose x^* separately for each state using cross validation.[17] Given x^*, the location of the grid is determined by a single latitude-longitude reference point.[18] Results are very similar across four different reference points, so we average estimates across them.

An alternative definition of a birth town group is a county. If the value of choosing a destination varied sharply across county borders in the sending region, then this definition would be appropriate. However, differences across counties, such as local government policies and elected

[16] In addition, $-1 \leq \Delta_{j,k} \leq N_j - 1$. At the upper bound, all migrants from j move to the same location, while at the lower bound, migrants displace each other one-for-one.

[17] That is,

$$x^* = \arg\min_x \sum_j \sum_k \left(N_{j,k}/N_j - \hat{P}_{g(x),-j,k} \right)^2,$$

where $\hat{P}_{g(x),-j,k} = \sum_{j' \neq j \in g(x)} N_{j',k} / \sum_{j' \neq j \in g(x)} N_{j'}$ is the average moving propensity from the birth town group of size x, excluding moves from town j. If there is only one town within a group g, then we define $\hat{P}_{g(x),-j,k}$ to be the statewide moving propensity. We search over even integers for convenience.

[18] In a related but substantively different setting, Billings and Johnson (2012) use cross validation in estimating the degree of industrial specialization. Duranton and Overman (2005) and Billings and Johnson (2012) estimate specialization parameters that do not require the aggregation of decisions at a spatial level. In contrast, we aggregate decisions at the receiving and sending county level to examine whether observed economic characteristics are related with social interactions.

officials, do not necessarily imply that counties are better birth town groups, as what matters is whether these differences affect the probability of choosing a destination, conditional on migrating. An advantage of cross-validation is that it facilitates comparisons across birth states, which differ widely in average county size. We emphasize results based on cross validation in the main text and include results based on counties as birth town groups in the appendix.[19]

We estimate the probability of moving from birth town group g to destination county k as the total number of people who move from g to k divided by the total number of migrants in g,

$$\widehat{P_{g,k}} = \frac{\sum_{j \in g} N_{j,k}}{\sum_{j \in g} N_j}. \tag{2.6}$$

We estimate the squared moving probability using the closed-form solution implied by equation (2.2),[20]

$$\widehat{P^2_{g,k}} = \frac{\sum_{j \in g} \sum_{j' \neq j \in g} N_{j,k} N_{j',k}}{\sum_{j \in g} \sum_{j' \neq j \in g} N_j N_{j'}}, \tag{2.7}$$

and the covariance of location decisions using the closed-form solution implied by equation (2.2),

$$\widehat{C_{j,k}} = \frac{N_{j,k}(N_{j,k} - 1)}{N_j(N_j - 1)} - \widehat{P^2_{g,k}}. \tag{2.8}$$

The final component of the SI index is the number of migrants from birth town j, N_j.

Given $(\widehat{P_{g,k}}, \widehat{P^2_{g,k}}, \widehat{C_{j,k}}, N_j)$, we can estimate the SI index, $\Delta_{j,k}$, using equation (B21). However, each estimate $\widehat{\Delta_{j,k}}$ depends primarily on a single birth town observation. To conduct inference, increase the reliability of our estimates, and decrease the number of parameters reported, we aggregate SI index estimates across all birth towns in each state for each destination county,

$$\widehat{\Delta_k} = \sum_j \left(\frac{\widehat{P_{g(j),k}} - \widehat{P^2_{g(j),k}}}{\sum_{j'} \widehat{P_{g(j'),k}} - \widehat{P^2_{g(j'),k}}} \right) \widehat{\Delta_{j,k}}, \tag{2.9}$$

where $g(j)$ is the group of town j. The destination level SI index estimate, $\widehat{\Delta_k}$, is robust to small estimates of $P_{g,k}$, which can blow up estimates of $\Delta_{j,k}$. The weighting scheme used in equation

[19] Appendix Figures A.2 and A.3 describe the number of birth towns per group when groups are defined using cross validation for Southern black and Great Plains white migrants. The median number of towns per group is 15 for African Americans and 39 for whites from the Great Plains. Appendix Figures A.4 and A.5 describe the number of towns per county. All groups used in estimation have at least two towns in them, because we cannot estimate $C_{j,k}$ or $P^2_{j,k}$ without multiple towns in the same group.

[20] Equation (2.7) yields an unbiased estimate of $P^2_{j,k}$ under Assumptions 1 and 2. In contrast, simply squaring $\widehat{P_{g,k}}$ would result in a biased estimate.

(2.9) arises naturally from assuming that $\Delta_{j,k}$ does not vary across birth towns within a state.[21] The destination-level SI index estimate, $\widehat{\Delta_k}$, allows us to identify the destinations for which social interactions were particularly important and the economic characteristics associated with stronger social interactions.

We also construct birth county-level SI index estimates by aggregating across destinations and towns within birth county c,

$$\widehat{\Delta_c} = \sum_k \sum_{j \in c} \left(\frac{\widehat{P_{g(j),k}} - \widehat{P^2_{g(j),k}}}{\sum_{k'} \sum_{j' \in c} \widehat{P_{g(j'),k'}} - \widehat{P^2_{g(j'),k'}}} \right) \widehat{\Delta_{j,k}}. \tag{2.10}$$

Birth county-level SI index estimates have similar conceptual and statistical properties as destination-level SI index estimates.

To facilitate exposition, we have described estimation of the SI index in terms of four distinct components, $(\widehat{P_{g,k}}, \widehat{P^2_{g,k}}, \widehat{C_{j,k}}, N_j)$. However, the SI index estimates depend only on observed population flows, and equation (2.9) forms the basis of an exactly identified generalized method of moments (GMM) estimator. To estimate the variance of $\widehat{\Delta_k}$, we treat the birth town group as the unit of observation and use a standard GMM variance estimator. This is akin to calculating heteroskedastic robust standard errors clustered at the birth town group level.[22] Appendix A.2 contains details.

2.3.4 An Extension to Assess the Validity of Our Empirical Strategy

The key threat to our empirical strategy is that the ex-ante value of moving to some destination differs across nearby birth towns in the same group. If, contrary to this threat, Assumption 1 were true, then geographic proximity would adequately control for the relevant determinants of location decisions, and using birth town-level covariates to explain moving probabilities would not affect SI index estimates.

To assess this threat, we allow moving probabilities to depend on town level covariates,

$$P_{j,k} = \rho_{g,k} + X_j \pi_k, \tag{2.11}$$

where $\rho_{g,k}$ is a birth town group-destination fixed effect, and X_j is a vector of town-level covariates whose effect on the moving probability can differ across destinations. X_j contains an

[21]When assuming $\Delta_{j,k} = \Delta_k \forall j$, the derivation in Appendix A.1 yields $\Delta_k = \left(\sum_j C_{j,k}(N_j - 1)\right) / \left(\sum_j P_{g(j),k}(1 - P_{g(j),k})\right)$, which leads directly to the estimator in equation (2.9).

[22]Treating birth town groups as the units of observation has no impact on the point estimate, $\widehat{\Delta_k}$. We estimate clustered standard errors because the estimates $\widehat{P_{g,k}}$ and $\widehat{P^2_{g,k}}$ are common to all birth towns within g.

indicator for being along a railroad, an indicator for having above-median black population share, and four indicators corresponding to population quintiles.[23] These covariates, available from the Duke SSA/Medicare data and the railroad information used in Black et al. (2015), capture potentially relevant determinants of location decisions. For example, migrants born in larger towns might have had more human capital or information, and these resources might have made certain destinations more attractive, so that our SI index estimates might reflect the role of birth town population size instead of social interactions; if this were the case, then our SI index estimates would be attenuated when controlling for birth town population.

With this extension, we construct an alternative SI index estimate using an alternative moving probability estimate, $\widetilde{P_{j,k}}$, equal to the fitted value from the OLS regression

$$\frac{N_{j,k}}{N_j} = \rho_{g,k} + X_j \pi_k + e_{j,k}. \tag{2.12}$$

We also use fitted values from a separate OLS regression, implied by equation (2.11), to form an alternative squared moving probability estimate, $\widetilde{P^2_{j,k}}$.[24] We estimate all equations separately for each birth state.[25] Similarity between the baseline and alternative SI index estimates would provide support for our empirical strategy.[26]

2.4 Results: Social Interactions in Location Decisions

2.4.1 Social Interactions Index Estimates

Table 2.1 provides an overview of the long-run population flows that we use to estimate social interactions. Our data contain 1.3 million African Americans born in the South from 1916-1936, 1.9 million whites born in the Great Plains, and 2.6 million whites born in the South. In old age, 42 percent of blacks born in the South and 35 percent of whites born in the Great Plains lived

[23]We construct percentiles for black population share and population separately for each birth state.

[24]We estimate $\widetilde{P^2_{j,k}}$ using fitted values from the OLS regression

$$\frac{N_{j,k}}{N_j}\frac{N_{j',k}}{N_{j'}} = \rho_{g(j),k}\rho_{g(j'),k} + X_j \pi_k \rho_{g(j'),k} + X_{j'}\pi_k \rho_{g(j),k} + (X_j \pi_k)(X_{j'}\pi_k) + e'_{j,j',k}$$

for different birth towns, $j \neq j'$.

[25]When estimating the variance of our SI index estimates under this extension, we ignore the variance that arises because $\widetilde{P_{j,k}}$ and $\widetilde{P^2_{j,k}}$ rely on OLS estimates. Accounting for this variance would make our estimates with and without covariates appear even more similar when performing statistical tests.

[26]An alternative approach to assessing the validity of Assumption 1 is testing whether the parameter vector $\pi_k = 0$ in equation (2.12). We prefer to test the difference in SI index estimates because this approach allows us to assess the statistical and substantive significance of any differences.

outside their birth region, while only nine percent of whites born in the South lived elsewhere.[27] We focus on Southern-born blacks and Great Plains-born whites in the main text, and leave results for Southern-born whites for the appendix. Appendix Table A.1 shows that, on average, there were 142 migrants per birth town for African Americans from the South, and 181 migrants per birth town for whites from the Great Plains.

We begin with some examples that illustrate how we identify social interactions. Table 2.2 shows the birth town to destination county migration flows that would be most unlikely in the absence of social interactions. Panel A shows that 10-50 percent of African-American migrants from each of these birth towns lived in the same destination county in old age, while 0.1-1.6 percent of migrants from each birth state lived in the same county. The observed moving propensities are 49-65 standard deviations larger than what would be expected if migrants moved independently of each other according to the statewide moving propensities. The estimated moving probabilities, $\widehat{P_{g,k}}$, exceed the statewide moving propensities, suggesting a meaningful role for local conditions in determining location decisions. Most importantly, the observed moving propensities are much larger than the estimated moving probabilities, consistent with positive covariance and SI index estimates. The results in Panel B for Great Plains whites are similar.

To summarize the importance of social interactions for all location decisions in our data, Table 2.3 reports averages of destination-level SI index estimates. Our data contain 516,712 black migrants from the South and 644,523 white migrants from the Great Plains.[28] For African Americans, unweighted averages of the destination-level SI index, $\widehat{\Delta_k}$, across all destination counties vary from 0.46 (Louisiana) to 0.90 (Mississippi). Averages weighted by the number of migrants in each destination vary from 0.81 (Florida) to 2.62 (South Carolina) and are larger because we estimate stronger social interactions in destinations that received more migrants. We prefer the weighted average as a summary measure because it better reflects the experience of a randomly chosen migrant and depends less on our decision to combine destination counties with fewer than 10 migrants. Across all states, the migrant-weighted average of destination-level SI index estimates is 1.94; this means that when we observe one randomly chosen African American move from a birth town to a destination county, then on average 1.94 additional black migrants from that birth town would make the same move. Panel B contains results for white moves out of the Great Plains. The weighted average of destination-level SI index estimates for whites is 0.38, only one-fifth the size of the black average.[29] These results indicate that African American migrants relied

[27] Census data show that return migration was quite low among Southern-born blacks and much higher among Southern-born whites (Gregory, 2005).

[28] The number of migrants in Table 2.3 differs slightly from the implied number of migrants in Table 2.1 because we exclude individuals from birth towns with fewer than 10 migrants when we estimate the SI index.

[29] Appendix Table A.2 displays the lengths of the square grid chosen by cross validation. Appendix Table A.3 shows that results are similar when we define birth town groups using counties. For Southern blacks, the linear (rank) correlation between the destination-level SI index estimates using cross validation and counties is 0.858 (0.904). For

more heavily on social networks in making their long-run location decisions. Historical context suggests that one explanation for this finding is that African Americans used social networks to overcome their lack of resources or the discrimination they faced in many destinations.

We provide a more complete picture of social interactions in Figure 2.5, which plots the distribution of destination-level SI index estimates.[30] Social interactions were particularly strong for some destinations and relatively weak for most destinations. As described below, our empirical approach allows us to examine whether destinations' economic characteristics can explain this considerable heterogeneity. Across the board, SI index estimates for African Americans are larger than those for whites.

To examine social interactions more closely, Figure 2.6 plots the spatial distribution of destination-level SI index estimates for Mississippi-born blacks. We estimate strong social interactions for several destinations: 23 counties have a SI index estimate greater than 3 and 58 counties have a SI index estimate between 1 and 3. These counties lie in the Midwest and, to a lesser degree, the Northeast. The figure also shows that African Americans moved to a relatively small number of destination counties, consistent with limited opportunities, information, or interest in moving to many places in the U.S.[31] We estimate strong social interactions ($\widehat{\Delta_k} > 3$) in Rock County, Wisconsin, consistent with historical accounts suggesting strong social interactions for Mississippi-born blacks in Beloit, which is located in Rock County (Bell, 1933; Rubin, 1960; Wilkerson, 2010). Figure 2.7 maps the destination-level SI index estimates for whites from North Dakota. We find little evidence of strong social interactions, although one exception is San Joaquin county ($\hat{\Delta}_k > 3$), an area described memorably in *The Grapes of Wrath* (Steinbeck, 1939).[32] Unlike black migrants, whites moved to a large number of destinations throughout the U.S. The difference between the number of destinations chosen by Mississippi blacks and North Dakota whites is striking, especially because there were almost 30,000 more migrants from Mississippi. Appendix Figures A.9 and A.10, for Southern Carolina-born blacks and Kansas-born whites, show similar patterns.

To assess the validity of our empirical strategy, we examine whether SI index estimates change when using birth town-level covariates to explain moving probabilities. Assumption 1 implies that geographic proximity adequately controls for the relevant determinants of location decisions, and so additional covariates should have no impact. Table 2.4 reports weighted averages of destination-level SI index estimates with and without covariates. When we examine birth states individually,

whites from the Great Plains, the linear (rank) correlation is 0.965 (0.891). Appendix Table A.4 shows that average SI index estimates for whites from the South are somewhat smaller than for whites from the Great Plains.

[30] Appendix Figure A.6 displays the associated t-statistic distributions, and Appendix Figures A.7 and A.8 display analogous results for whites from the South. A destination county can appear multiple times in these figures because we estimate destination-level SI indices separately for each birth state.

[31] In Figure 2.6, the counties in white received less than 10 migrants.

[32] In *The Grapes of Wrath*, the Joad family travels from Oklahoma to the San Joaquin Valley. Gregory (1989) notes that the (fictional) Joads were poorer than many migrants from the Great Plains.

there are no substantively or statistically significant differences between the two sets of estimates. When pooling all Southern states together, the estimates are very similar in magnitude (1.94 and 1.92) and statistically indistinguishable ($p = 0.76$). When pooling all Great Plains states together, the estimates again are very similar in magnitude (0.38 and 0.36), but are statistically distinguishable ($p = 0.02$). In addition, the destination-level SI index estimates with and without covariates are highly correlated: the linear (rank) correlation is 0.914 (0.992) for blacks from the South and 0.939 (0.988) for whites from the Great Plains. On net, this evidence indicates that geographic proximity adequately controls for the relevant determinants of location decisions and supports the validity of our empirical strategy.

Table 2.5 shows that our results are not driven by migration from the largest birth towns or migration to the largest destinations and, relatedly, that there is limited heterogeneity in SI index estimates on these dimensions. Birth town size could be correlated with unobserved determinants of social interactions and location decisions, such as the level of social and human capital or information about destinations. However, it is not clear beforehand whether social interactions will vary with the size of receiving or sending locations. For reference, column 1 of Table 2.5 reports weighted averages of destination-level SI index estimates when including all birth towns and destinations. In column 2, we exclude birth towns with at least 20,000 residents in 1920 when estimating each destination-level SI index.[33] Column 3 excludes destination counties that intersect with the ten largest non-Southern consolidated metropolitan statistical areas (CMSAs) as of 1950, in addition to counties that received less than 10 migrants.[34] We exclude both large birth towns and large destinations in column 4. The average SI index estimates are similar across all four specifications for both Southern blacks and Great Plains whites.[35]

A widely noted feature of the Great Migration is the tendency of migrants to move along vertical pathways established by railroads, which reduced the cost of moving to destinations on the same line and increased the flow of information. Social interactions might have benefitted from reduced migration costs and increased information, or social interactions might have drawn migrants to destinations that they would not consider otherwise. Table 2.6 displays weighted averages of destination-level SI index estimates for different regions, demonstrating that social interactions among African Americans clearly follow vertical migration patterns. The largest SI index esti-

[33]The excluded birth towns are Birmingham, Mobile, and Montgomery, Alabama; Jacksonville, Miami, Pensacola, and Tampa, Florida; Atlanta, Augusta, Columbus, Macon, and Savannah, Georgia; Baton Rouge, New Orleans, and Shreveport, Louisiana; Jackson and Meridian, Mississippi; Asheville, Charlotte, Durham, Raleigh, Wilmington, and Winston-Salem, North Carolina; Charleston, Greenville, and Spartanburg, South Carolina; Hutchinson, Kansas City, Topeka, and Wichita, Kansas; Lincoln and Omaha, Nebraska; Fargo, North Dakota; Muskogee, Oklahoma City, and Tulsa, Oklahoma; and Sioux Falls, South Dakota

[34]The ten CMSAs are New York, Chicago, Los Angeles, Philadelphia, Boston, Detroit, Washington, D.C., San Francisco, Pittsburgh, and St. Louis. The first nine of these are also the largest non-Great Plains (and border region) CMSAs.

[35]Appendix Table A.5 reports similar results for Southern-born whites.

mates in the Northeast come from the Carolinas, while the largest estimates in the Midwest are among migrants from Mississippi and Alabama, and the largest estimates in the West come from Louisiana.[36] Panel B displays weighted averages by region for Great Plains whites.[37] Social interactions among Great Plains whites were much stronger in the Midwest and West, where moving costs were lower, than the Northeast or South. These patterns suggest that lower migration costs and greater information facilitated social interactions.

To further understand the nature of social interactions, we examine whether the location decisions of African American migrants influenced white migrants from the same Southern birth town, and vice versa. While blacks and whites could have shared information about opportunities in the North, the high segregation in the Jim Crow South makes cross-race social interactions unlikely. Appendix A.3 describes how we estimate cross-race social interactions. Appendix Table A.7 displays little evidence of cross-race interactions, indicating that social interactions operated within racial groups. In addition, there is little correlation between destination-level SI index estimates for blacks and whites from the South: the linear (rank) correlation is 0.076 (0.149). This implies that our SI index estimates do not simply reflect unobserved characteristics of certain Southern towns.

2.4.2 Addressing Measurement Error due to Incomplete Migration Data

SI 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 (B21) implies that SI index estimates will be attenuated because we undercount the number of migrants from each town, N_j.

More specifically, suppose that we are interested in the effect of social interactions on location decisions at age 40. Denote the number of migrants that survive to age 40 by N_j^{40}, and assume for simplicity that this equals the observed number of migrants divided by a scaling factor, $N_j^{40} = N_j/\alpha$. To approximate the coverage rate α, we divide the number of individuals in the

[36]The Northeast region includes Connecticut, Delaware, Washington, D.C., Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont, and West Virginia. The Midwest region includes Illinois, Indiana, Iowa, Kansas, Kentucky, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, Oklahoma, South Dakota, and Wisconsin. The West region includes Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming. The South region includes Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, and Virginia. These regions vary from Census-defined regions because we define the South to be the former Confederate states.

[37]Appendix Table A.6 reports region-specific results for Southern-born whites.

Duke/SSA Medicare data by the number of individuals in decennial census data.[38] Across birth states, the average coverage rate is 52.3% for African Americans from the South and 69.3% for whites from the Great Plains (see Appendix Table A.8), which implies that $N_j^{40} \approx 1.91N_j$ for Southern blacks and $N_j^{40} \approx 1.44N_j$ for Great Plains whites. As an approximate measurement error correction, SI index estimates should be multiplied by a factor of 1.91 for Southern blacks and 1.44 for Great Plains whites. Appendix Table A.9 presents results that reflect state-specific coverage rate adjustments. The weighted average of destination level SI index estimates is 3.71 for Southern blacks and 0.55 for Great Plains whites. Adjusting for incomplete data under a missing at random assumption increases the magnitude of SI index estimates and increases the gap between black and white social interaction estimates.

Appendix A.4 describes the consequences of measurement error when we relax the missing at random assumption. We derive a lower bound on the SI index and show that estimates of this lower bound still reveal sizable social interactions.

2.4.3 THE ROLE OF FAMILY MIGRATION

The SI index might capture the influence of family members from the same birth town on migrants' location decisions. While family migration does not threaten our empirical strategy, it would be interesting to know the extent to which social interactions occur within the family. Unfortunately, we do not observe family relationships and have limited ability to study this issue directly. We can examine whether our results stem entirely from the migration of heterosexual couples. If this were true, there would be no social interactions among men only or women only. Appendix Table A.9 shows that SI index estimates are similar in magnitude among men and women, implying that our results do not simply reflect the migration of couples.[39] Our sample likely contains very few sets of parents and children, since we only include individuals born from 1916-1936.

A related question is whether differences in family structure explain differences in social interactions between black and white migrants. As a first step, we use the 1940 Census to measure the average within-household family size for individuals born from 1916-1936. African Americans from the South had families that were 17 percent larger than whites from the Great Plains (6.16 vs. 5.25). This difference is too small to explain our finding that average SI index estimates are 410 percent larger among blacks than whites.[40] To construct an upper bound on extended family

[38]We use the 1960 Census to construct coverage rates for individuals born from 1916-1925 and the 1970 Census for individuals born from 1926-1935.

[39]The similarity between men and women is not surprising given the relative sex balance among migrants in this period (Gregory, 2005).

[40]The weighted average of SI index estimates in Table 2.3 is 1.938 for blacks and 0.380 for whites, and (1.938-0.380)/0.380 = 4.1. When adjusting for incomplete migration data under the missing at random assumption (Appendix Table A.9), social interactions among African Americans are 582 percent larger than among Great Plains whites.

size, we use the 100 percent sample of the 1940 Census to count the average number of individuals in a county born from 1916-1936 with the same last name (Minnesota Population Center and Ancestry.com, 2013). We find that Southern black family networks likely were no more than 270 percent larger than those for Great Plains whites (54.5 versus 14.7). This upper bound is sizable, but still less than the 410 percent difference in social interaction strength. In sum, differences in family structure might explain some, but not all, of the differences in social interactions between black and white migrants.

2.4.4 Social Interactions and Economic Characteristics of Receiving and Sending Locations

To better understand why social interactions affected location decisions, we relate SI index estimates to economic characteristics of receiving and sending locations. We focus on African American migrants because social interactions were more important for this group.

We first consider the economic characteristics of receiving locations. Employment opportunities were among the most important characteristics of a destination, and relatively high wages and demand for workers made manufacturing jobs particularly attractive. In the presence of imperfect information, networks might have directed their members to destinations with more manufacturing employment.[41] This is the story of John McCord, told in Section 2.2. Because individuals living in the South almost certainly had more information about employment opportunities in the largest destinations, the imperfect information channel suggests a stronger relationship between social interactions and manufacturing employment intensity in smaller destinations. In contrast, if information about employment opportunities was widely known, then social interactions might not be stronger in destinations with more manufacturing. Pecuniary moving costs, which were largely determined by railroads and physical distance, represented another key characteristic of destinations. Lower moving costs could have fostered social interactions by facilitating the transmission of information. On the other hand, migrants might have been willing to pay high moving costs only if they received information or benefits from a network. Ultimately, these relationships must be determined empirically.

To explore these hypotheses, we regress destination-level SI index estimates on county-level covariates. Column 1 of Table 2.7 shows that social interactions were significantly larger in destinations with a higher 1910 manufacturing employment share: a one standard deviation increase in the 1910 manufacturing employment share is associated with an increase in the SI index of 0.22 people.[42] Column 2 shows that the positive relationship between manufacturing employment and

[41] There is a large literature on social networks and employment opportunities (recent examples include Topa, 2001; Munshi, 2003; Ioannides and Loury, 2004; Bayer, Ross and Topa, 2008; Hellerstein, McInerney and Neumark, 2011; Beaman, 2012; Burks et al., 2015; Schmutte, 2015; Heath, 2016).

[42] Appendix Table A.10 contains summary statistics. Appendix Figure A.11 plots the bivariate relationship between

social interactions was almost four times larger in smaller destinations.[43] We also find that social interactions were significantly stronger in destinations that could be reached by rail directly or with one stop from the birth state and destinations that were closer to the birth state. We also find that social interactions were stronger in destinations with a smaller black population share in 1900, suggesting that networks helped migrants find opportunities in new places. One possible concern is that these results do not reflect characteristics of destination counties, but instead characteristics of birth states linked to destinations via vertical migration patterns. Column 3 indicates that this concern is unimportant, as adding birth state fixed effects has very little impact.[44]

We next consider the relationship between social interactions and economic characteristics of sending counties. Social networks could have been particularly valuable in locating jobs or housing for migrants from poorer communities who had fewer resources to engage in costly search. Alternatively, resources that facilitated migration might have been a prerequisite for social interactions to influence location decisions. Another potentially important characteristic is population density among African-Americans, which could have encouraged stronger social networks because of more frequent interactions (Chay and Munshi, 2015). We also consider black literacy rates and exposure to Rosenwald schools, which improved educational attainment among Southern blacks in this period (Aaronson and Mazumder, 2011). The relationship between education and social interactions is theoretically ambiguous, as education could promote social ties in the South while also increasing the return to choosing a non-network destination. In addition, we examine whether social interactions were stronger in counties with greater access to railroads, which could have facilitated the transmission of information through both network and non-network channels.

Table 2.8 displays results from regressing birth county-level SI index estimates on birth county characteristics. We find a positive but insignificant relationship between the strength of social interactions and the 1920 farm ownership rate among African Americans, which we use to proxy for assets. Social interactions were significantly stronger in counties with higher density and literacy rates in 1920.[45] Results are similar when we include birth state fixed effects.[46] The estimates in column 2 imply that a one standard deviation increase in log black density is associated with a 1.08 person increase in the SI index, and a one standard deviation increase in the black literacy rate is

SI index estimates and 1910 manufacturing employment share, showing the considerable variation in the manufacturing employment share across destinations.

[43]Small destination counties are those that do not intersect with the ten largest non-South CMSAs in 1950 (New York, Chicago, Los Angeles, Philadelphia, Boston, Detroit, Washington, D.C., San Francisco, Pittsburgh, and St. Louis).

[44]Results are qualitatively similar using counties to define birth town groups (Appendix Table A.11). Results for Great Plains whites and Southern whites are in Appendix Tables A.12 and A.13.

[45]Using a different empirical strategy, Chay and Munshi (2015) also find a positive relationship between social interactions and black population density.

[46]We include birth state fixed effects to mitigate the possibility that our results are driven by destination factors, such as labor demand, that are linked to certain areas of the South through vertical migration patterns.

associated with a 0.48 person increase.[47] We find little evidence that social interactions varied with railroad exposure, although the standard errors are fairly large.

2.4.5 CONNECTING THE SOCIAL INTERACTIONS INDEX TO A SIMPLE STRUCTURAL MODEL

Next, we connect the SI index to the simple structural model of social interactions from Glaeser, Sacerdote and Scheinkman (1996). The additional assumptions in their model allow us to estimate the share of migrants that chose their long-run location because of social interactions, a parameter that complements our SI index in intuitively describing the size of social interactions. This connection also demonstrates that our SI index integrates the model of Glaeser, Sacerdote and Scheinkman (1996) and the general identification strategy of Bayer, Ross and Topa (2008).

Migrants, indexed on a circle by $i \in \{1, \ldots, N_j\}$, are either a "fixed agent" or a "complier." Fixed agents choose their location independently of other migrants, while a complier i chooses the same destination as the neighbor, $i-1$. The probability that a migrant is a complier equals χ, assumed for simplicity to be constant across birth towns and destinations for a given birth state. The covariance of location decisions for migrants i and $i+n$ is $\mathbb{C}[D_{i,j,k}, D_{i+n,j,k}] = P_{g,k}(1-P_{g,k})\chi^n$. Hence, the average covariance of location decisions implied by the model is

$$C_{j,k}(\chi; P_{g,k}, N_j) \equiv \frac{\sum_{i \in j} \sum_{i' \neq i \in j} \mathbb{C}[D_{i,j,k}, D_{i',j,k}]}{N_j(N_j-1)} \tag{2.13}$$

$$= \frac{2P_{g,k}(1-P_{g,k}) \sum_{a=1}^{N_j-1} (N_j - a)\chi^a}{N_j(N_j-1)}. \tag{2.14}$$

In the absence of social interactions, there are no compliers, and the covariance of location decisions equals zero.[48]

Substituting the expression for $C_{j,k}$ in equation (2.14) into the expression for the SI index in equation (B21) yields

$$\Delta_{j,k} = 2 \sum_{a=1}^{N_j-1} (1 - a/N_j)\chi^a. \tag{2.15}$$

[47] Appendix Table A.14 contains summary statistics for birth county characteristics.

[48] Glaeser, Sacerdote and Scheinkman (1996) measure social interactions using the normalized variance of outcomes, which in our model is

$$\mathbb{V}\left[\sum_{i=1}^{N_j} \frac{D_{i,j,k} - P_{g,k}}{N_j}\right] = \frac{P_{g,k}(1-P_{g,k})}{N_j} + \left(\frac{N_j - 1}{N_j}\right) C_{j,k}(\chi; P_{g,k}, N_j).$$

With a sufficiently large number of migrants, we obtain $\Delta_{j,k} = 2\chi/(1-\chi)$. Because the destination-level SI index, Δ_k, is just a weighted average of $\Delta_{j,k}$, and the average destination-level SI index, denoted Δ, is just a weighted average of Δ_k, we can estimate the probability that an individual is a complier as

$$\hat{\chi} = \frac{\hat{\Delta}}{2 + \hat{\Delta}}. \tag{2.16}$$

As seen in Table 2.9, we estimate that between 29 (Florida) and 57 percent (South Carolina) of black migrants chose their long-run location because of social interactions. There is considerable variation across destination regions.[49] For example, of Mississippi-born migrants, 32 percent of Northeast-bound, 57 percent of Midwest-bound, and 34 percent of West-bound migrants chose their location because of social interactions. Among whites from the Great Plains, between 11 (Kansas) and 19 percent (North Dakota) of migrants chose their destination because of social interactions. Although estimates of χ depend on stronger assumptions than are needed to estimate the SI index, they help illustrate the considerable impact of social interactions on location decisions for Southern blacks and the smaller impact among whites.[50]

Explicit connections to structural models also allow us to refine the interpretation of the SI index. One parameter of interest, which we denote $\theta_{j,k}$, is the number of additional people induced to move from birth town j to destination k by moving one migrant along this path. The relationship between $\Delta_{j,k}$ and $\theta_{j,k}$ depends on the underlying structural model. In the coalition model, where all migrants in birth town j form coalitions of size s, all members of a coalition move to the same destination, and all coalitions move independently of each other, $\theta_{j,k}^{\text{COA}} = \Delta_{j,k} = s - 1$. In the model of Glaeser, Sacerdote and Scheinkman (1996), $\theta_{j,k}^{\text{GSS}} = 0.5\Delta_{j,k}$.[51] The difference between $\Delta_{j,k}$ and $\theta_{j,k}$ stems from the weak structural assumptions needed to estimate the SI index. The weakness of these assumptions, and the ability to map the SI index directly to several structural

[49] Assuming that χ is constant across destinations implies that it should not vary across different regions. Nonetheless, we find the rescaled regional estimates to be informative. Appendix A.5 contains a richer model that allows the probability of complying to vary with birth town and destination.

[50] Estimates of χ would be larger if we used estimates of the SI index that accounted for measurement error due to incomplete migration data.

[51] In the Glaeser, Sacerdote and Scheinkman (1996) model, migrant i has the following effect on migrant $i+n$,

$$\mathbb{E}[D_{i+n,j,k} | D_{i,j,k} = 1, D_{1,j,k}, \ldots, D_{i-1,j,k}] - \mathbb{E}[D_{i+n,j,k} | D_{i,j,k} = 0, D_{1,j,k}, \ldots, D_{i-1,j,k}] = \chi^n,$$

which implies that

$$\theta_{j,k}^{\text{GSS}}(i) = \mathbb{E}[N_{-i,j,k} | D_{i,j,k} = 1, D_{1,j,k}, \ldots, D_{i-1,j,k}] - \mathbb{E}[N_{-i,j,k} | D_{i,j,k} = 0, D_{1,j,k}, \ldots, D_{i-1,j,k}] = \sum_{a=1}^{N_j - i} \chi^a.$$

As $N_j \to \infty$, $\theta_{j,k}^{\text{GSS}}(i) \to \chi/(1-\chi) = 0.5\Delta_{j,k}$.

models, are valuable features of our approach.

2.5 Conclusion

This paper provides new evidence on the magnitude and nature of social interactions in location decisions. We use confidential administrative data to study over one million long-run location decisions made during two landmark migration episodes by African Americans born in the U.S. South and whites born in the Great Plains. We formulate a novel social interactions (SI) index that characterizes the strength of social interactions for each receiving and sending location. The SI index allows us to estimate the overall magnitude of social interactions and the degree to which social interactions were associated with economic characteristics of receiving and sending locations. The SI index can be used for other outcomes and settings to provide a deeper understanding of social interactions in economic decisions.

We find very strong social interactions among Southern black migrants and smaller interactions among whites. Estimates of our social interactions (SI) index imply that if we observed one randomly chosen African American move from a birth town to some destination county, then on average 1.9 additional black migrants from that birth town would make the same move. For white migrants from the Great Plains, the average is only 0.4, and results for Southern whites are similarly small. Interpreted through the social interactions model of Glaeser, Sacerdote and Scheinkman (1996), our estimates imply that 49 percent of African-American migrants chose their long-run destination because of social interactions, while 16 percent of Great Plains whites were similarly influenced. One interpretation of our results is that African Americans relied on social networks more heavily to overcome the more intense discrimination they faced in labor and housing markets. In addition, our results suggest that social interactions were particularly important in providing African American migrants with information about attractive employment opportunities, and that social interactions played a larger role in less costly moves. Our results also suggest that educational attainment in the South facilitated social interactions.

These results shed new light on migration decisions. Social interactions play a major role in our setting, especially for migrants with fewer opportunities and resources. Our results suggest that social interactions help migrants mitigate the substantial information frictions that characterize long-distance location decisions. Social interactions likely play an important role in contemporaneous rural-to-urban migrations in the developing world, which resemble the historical migration episodes we study on several dimensions. Our results also suggest that long-run location decisions will more effectively shift population to areas with a high marginal product of labor if there are pioneer migrants who can facilitate these costly moves. Policies that seek to direct migration to certain areas should account for the role of social interactions.

Our results also have implications for economic outcomes besides migration. Birth town social networks continued to operate after location decisions had been made, and the Great Migration generated considerable variation in the strength of social networks across destinations. In other work, we use this variation to study the relationship between crime and social connectedness in U.S. cities (Stuart and Taylor, 2017*b*).

Table 2.1: Location at Old Age, 1916-1936 Cohorts

		Percent Living in Location		
		Outside Birth	In Birth Region	
	People	Region	Birth State	Other State
Birth State	(1)	(2)	(3)	(4)
Panel A: Southern Blacks				
Alabama	209,128	47.2%	39.5%	13.3%
Florida	79,237	26.1%	67.1%	6.8%
Georgia	218,357	36.3%	44.2%	19.5%
Louisiana	179,445	32.4%	52.7%	14.9%
Mississippi	218,759	56.1%	28.9%	15.0%
North Carolina	200,999	40.2%	49.7%	10.1%
South Carolina	163,650	43.4%	41.9%	14.7%
Total	1,269,575	41.8%	44.0%	14.1%
Panel B: Great Plains Whites				
Kansas	462,490	30.4%	43.3%	26.3%
Nebraska	374,265	36.0%	42.0%	22.0%
North Dakota	210,199	44.1%	31.8%	24.1%
Oklahoma	635,621	31.8%	41.6%	26.6%
South Dakota	196,266	40.4%	35.4%	24.2%
Total	1,878,841	34.6%	40.3%	25.1%
Panel C: Southern Whites				
Alabama	469,698	9.8%	62.1%	28.1%
Florida	231,071	12.7%	68.5%	18.8%
Georgia	454,286	7.4%	65.5%	27.1%
Louisiana	384,601	8.7%	71.1%	20.2%
Mississippi	275,147	11.0%	57.0%	32.0%
North Carolina	588,674	8.5%	71.6%	19.8%
South Carolina	238,697	6.6%	70.6%	22.8%
Total	2,642,174	9.0%	66.9%	24.0%

Notes: Column 1 contains the number of people from the 1916-1936 birth cohorts observed in the Duke SSA/Medicare data. Columns 2-4 display the share of individuals living in each location at old age (2001 or date of death, if earlier). Figure 2.4 displays birth regions.
Source: Duke SSA/Medicare data

Table 2.2: Extreme Examples of Correlated Location Decisions, Southern Blacks and Great Plains Whites

Birth Town (1)	Largest City in Destination County (2)	Total Birth Town Migrants (3)	Town-Destination Flow (4)	Destination Share of Birth Town Migrants (5)	Destination Share of Birth State Migrants (6)	SD under Independent Binomial Moves (7)	Moving Probability Estimate (8)	Social Interaction Index Estimate (9)
Panel A: Southern Blacks								
Pigeon Creek, AL	Niagara Falls, NY	85	43	50.6%	0.5%	64.5	4.5%	8.5
Marion, AL	Fort Wayne, IN	1311	200	15.3%	0.7%	63.7	3.8%	8.8
Greeleyville, SC	Troy, NY	215	34	15.8%	0.1%	62.2	1.7%	15.2
Athens, AL	Rockford, IL	649	64	9.9%	0.2%	61.0	2.0%	5.6
Pontotoc, MS	Janesville, WI	456	62	13.6%	0.2%	59.4	3.3%	6.5
New Albany, MS	Racine, WI	599	97	16.2%	0.4%	58.7	4.9%	11.4
West, MS	Freeport, IL	336	35	10.4%	0.1%	56.9	0.8%	6.2
Gatesville, NC	New Haven, CT	176	88	50.0%	1.6%	51.8	8.1%	7.1
Statham, GA	Hamilton, OH	75	22	29.3%	0.3%	50.0	3.0%	4.4
Cochran, GA	Paterson, NJ	259	62	23.9%	0.6%	49.4	4.1%	6.3
Panel B: Great Plains Whites								
Krebs, OK	Akron, OH	210	32	15.2%	0.1%	82.6	0.3%	7.4
Haven, KS	Elkhart, IN	144	22	15.3%	0.1%	51.1	0.4%	6.9
McIntosh, SD	Rupert, ID	299	20	6.7%	0.1%	50.9	0.6%	4.8
Hull, ND	Bellingham, WA	55	24	43.6%	0.5%	44.6	1.5%	4.3
Lindsay, NE	Moline, IL	226	29	12.8%	0.2%	41.5	0.4%	5.2
Corsica, SD	Holland, MI	253	26	10.3%	0.2%	39.6	0.4%	6.3
Corsica, SD	Grand Rapids, MI	253	34	13.4%	0.3%	37.2	0.7%	6.0
Montezuma, KS	Merced, CA	144	21	14.6%	0.3%	32.7	0.9%	2.7
Hillsboro, KS	Fresno, CA	407	65	16.0%	0.9%	32.0	1.2%	2.2
Henderson, NE	Fresno, CA	146	32	21.9%	0.7%	31.1	0.8%	2.2

Notes: Each panel contains the most extreme examples of correlated location decisions, as determined by column 7. Column 7 equals the difference, in standard deviations, of the actual moving propensity (column 5) relative to the prediction with independent moves following a binomial distribution governed by the statewide moving propensity (column 6). Column 8 equals the estimated probability of moving from town j to county k using observed location decisions from nearby towns, where the birth town group is defined by cross validation. Column 9 equals the destination-level SI index estimate for the relevant birth state. When choosing these examples, we restrict attention to town-destination pairs with at least 20 migrants.

Source: Duke SSA/Medicare data

Table 2.3: Average Social Interactions Index Estimates, by Birth State

Birth State	Number of Migrants (1)	Unweighted Average (2)	Weighted Average (3)
Panel A: Black Moves out of South			
Alabama	96,269	0.770	1.888
		(0.049)	(0.195)
Florida	19,158	0.536	0.813
		(0.052)	(0.117)
Georgia	77,038	0.735	1.657
		(0.048)	(0.177)
Louisiana	55,974	0.462	1.723
		(0.039)	(0.478)
Mississippi	120,454	0.901	2.303
		(0.050)	(0.313)
North Carolina	78,420	0.566	1.539
		(0.039)	(0.130)
South Carolina	69,399	0.874	2.618
		(0.054)	(0.301)
All States	516,712	0.736	1.938
		(0.020)	(0.110)
Panel B: White Moves out of Great Plains			
Kansas	139,374	0.128	0.255
		(0.007)	(0.024)
Nebraska	134,011	0.141	0.361
		(0.008)	(0.082)
North Dakota	92,205	0.174	0.464
		(0.012)	(0.036)
Oklahoma	200,392	0.112	0.453
		(0.008)	(0.036)
South Dakota	78,541	0.163	0.350
		(0.009)	(0.026)
All States	644,523	0.137	0.380
		(0.004)	(0.022)

Notes: Column 2 is an unweighted average of destination-level SI 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 are in parentheses.
Source: Duke SSA/Medicare data

Table 2.4: Average Social Interactions Index Estimates, With and Without Birth Town Covariates

	Include Covariates		p-value of
	No	Yes	difference
Birth State	(1)	(2)	(3)
Panel A: Black Moves out of South			
Alabama	1.888	1.852	0.763
	(0.195)	(0.189)	
Florida	0.813	0.742	0.401
	(0.117)	(0.119)	
Georgia	1.657	1.689	0.658
	(0.177)	(0.175)	
Louisiana	1.723	1.651	0.862
	(0.478)	(0.474)	
Mississippi	2.303	2.295	0.967
	(0.313)	(0.306)	
North Carolina	1.539	1.482	0.149
	(0.130)	(0.127)	
South Carolina	2.618	2.636	0.827
	(0.301)	(0.304)	
All States	1.938	1.917	0.764
	(0.110)	(0.108)	
Panel B: White Moves out of Great Plains			
Kansas	0.255	0.233	0.112
	(0.024)	(0.024)	
Nebraska	0.361	0.349	0.504
	(0.082)	(0.082)	
North Dakota	0.464	0.445	0.456
	(0.036)	(0.035)	
Oklahoma	0.453	0.439	0.241
	(0.036)	(0.036)	
South Dakota	0.350	0.331	0.145
	(0.026)	(0.026)	
All States	0.380	0.363	0.021
	(0.022)	(0.022)	

Notes: All columns contain weighted averages of destination-level SI index estimates, $\hat{\Delta}_k$, where the weights are the number of people who move from each state to destination k. Column 2 controls for birth town-level covariates as described in the text. Column 3 reports the p-value from testing the null hypothesis that the two columns are equal. Birth town groups are defined by cross validation. Standard errors are in parentheses.
Source: Duke SSA/Medicare data

Table 2.5: Average Social Interactions Index Estimates, by Size of Birth Town and Destination

Exclude Largest Birth Towns:	No	Yes	No	Yes
Exclude Largest Destinations:	No	No	Yes	Yes
Birth State	(1)	(2)	(3)	(4)
Panel A: Black Moves out of South				
Alabama	1.888	1.784	2.056	2.189
	(0.195)	(0.149)	(0.285)	(0.268)
Florida	0.813	0.607	1.323	1.231
	(0.117)	(0.061)	(0.229)	(0.215)
Georgia	1.657	1.458	1.696	1.772
	(0.177)	(0.092)	(0.170)	(0.133)
Louisiana	1.723	1.106	0.971	0.960
	(0.478)	(0.095)	(0.182)	(0.176)
Mississippi	2.303	2.299	2.085	2.032
	(0.313)	(0.304)	(0.210)	(0.205)
North Carolina	1.539	1.451	0.743	0.687
	(0.130)	(0.126)	(0.064)	(0.059)
South Carolina	2.618	2.556	1.784	1.742
	(0.301)	(0.283)	(0.241)	(0.234)
All States	1.938	1.791	1.755	1.783
	(0.110)	(0.089)	(0.108)	(0.102)
Panel B: White Moves out of Great Plains				
Kansas	0.255	0.220	0.243	0.228
	(0.024)	(0.019)	(0.021)	(0.019)
Nebraska	0.361	0.253	0.265	0.253
	(0.082)	(0.014)	(0.019)	(0.017)
North Dakota	0.464	0.464	0.527	0.531
	(0.036)	(0.036)	(0.046)	(0.046)
Oklahoma	0.453	0.395	0.450	0.427
	(0.036)	(0.029)	(0.040)	(0.038)
South Dakota	0.350	0.339	0.387	0.381
	(0.026)	(0.026)	(0.034)	(0.033)
All States	0.380	0.331	0.374	0.361
	(0.022)	(0.012)	(0.016)	(0.016)

Notes: All columns contain weighted averages of destination-level SI 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: Duke SSA/Medicare data

Table 2.6: Average Social Interactions Index Estimates, by Destination Region

	Destination Region			
	Northeast	Midwest	West	South
	(1)	(2)	(3)	(4)
Panel A: Black Moves out of South				
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: White Moves out of Great Plains				
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 SI index estimates, $\hat{\Delta}_k$, where the weights are the number of people who move from each state to destination k. See footnote 36 for region definitions. We do not estimate social interactions for blacks who move to the South. Birth town groups are defined by cross validation. Standard errors are in parentheses.
Source: Duke SSA/Medicare data

Table 2.7: Social Interactions Index Estimates and Destination County Characteristics, Black Moves out of South

Dependent variable: Destination-level SI index estimate	(1)	(2)	(3)
Manufacturing employment share, 1910	1.561***	0.582	0.551
	(0.401)	(0.406)	(0.420)
Manufacturing employment share by small destination indicator		1.739***	1.825***
		(0.581)	(0.575)
Small destination indicator		0.291**	0.327***
		(0.118)	(0.117)
Direct railroad connection from birth state	0.315***	0.334***	0.347***
	(0.111)	(0.111)	(0.127)
One-stop railroad connection from birth state	0.224***	0.214***	0.180**
	(0.077)	(0.075)	(0.078)
Log distance from birth state	-0.364***	-0.339***	-0.319***
	(0.060)	(0.061)	(0.059)
Log population, 1900	0.099***	0.117***	0.125***
	(0.037)	(0.035)	(0.037)
Percent African-American, 1900	-2.026***	-1.931***	-1.846***
	(0.331)	(0.315)	(0.309)
Birth state fixed effects			x
R2	0.093	0.103	0.115
N (birth state-destination county pairs)	1,469	1,469	1,469
Destination counties	371	371	371

Notes: The sample contains only counties that received at least 10 migrants. Birth town groups are defined by cross validation. Standard errors, clustered by destination county, are in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Sources: Duke SSA/Medicare data, Haines and ICPSR (2010) data, and Black et al. (2015) data

Table 2.8: Social Interactions Index Estimates and Birth County Characteristics, Black Moves out of South

Dependent variable: Birth county-level SI index estimate	(1)	(2)
African-American farm ownership rate, 1920	1.854	2.123
	(1.353)	(1.390)
Log African-American density, 1920	1.099*	1.027*
	(0.562)	(0.565)
Rosenwald school exposure	-0.981	-1.202*
	(0.656)	(0.687)
African-American literacy rate, 1920	3.680**	5.128**
	(1.574)	(2.094)
Railroad exposure	-0.309	-0.268
	(0.423)	(0.442)
Percent African-American, 1920	0.606	0.880
	(1.684)	(1.589)
Birth state fixed effects		x
R2	0.090	0.097
N (birth counties)	549	549

Notes: The dependent variable is the birth county level social interaction estimate. Railroad exposure is the share of migrants in a county that lived along a railroad. Rosenwald exposure is the average Rosenwald coverage experienced over ages 7-13. Heteroskedastic robust standard errors are in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

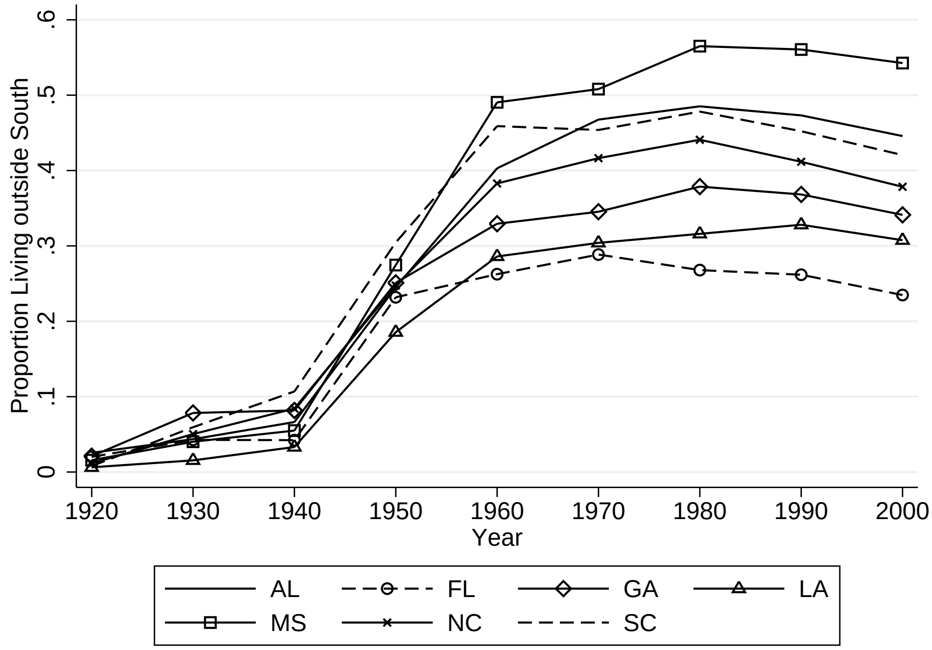
Sources: Duke SSA/Medicare data, Haines and ICPSR (2010) data, Aaronson and Mazumder (2011) data, and Black et al. (2015) data

Table 2.9: Estimated Share of Migrants That Chose Their Destination Because of Social Interactions

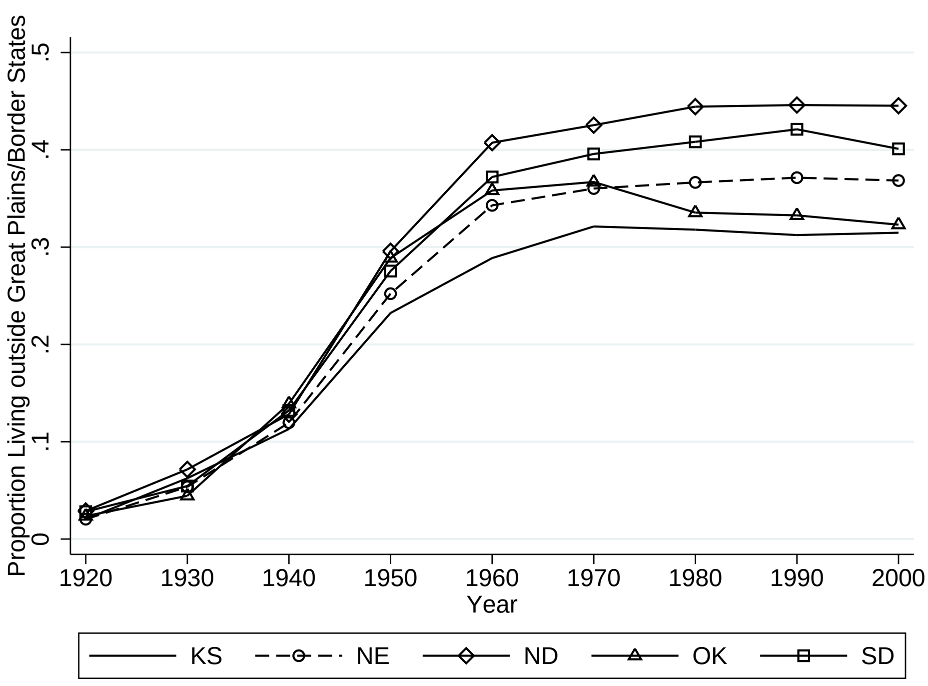
	Destination Region				
	All	Northeast	Midwest	West	South
Birth State	(1)	(2)	(3)	(4)	(5)
Panel A: Black Moves out of South					
Alabama	0.486	0.382	0.541	0.289	-
	(0.026)	(0.031)	(0.031)	(0.069)	-
Florida	0.289	0.328	0.284	0.117	-
	(0.030)	(0.039)	(0.043)	(0.042)	-
Georgia	0.453	0.436	0.508	0.170	-
	(0.026)	(0.039)	(0.038)	(0.070)	-
Louisiana	0.463	0.123	0.363	0.520	-
	(0.069)	(0.039)	(0.042)	(0.084)	-
Mississippi	0.535	0.316	0.571	0.341	-
	(0.034)	(0.025)	(0.036)	(0.028)	-
North Carolina	0.435	0.456	0.312	0.085	-
	(0.021)	(0.022)	(0.042)	(0.017)	-
South Carolina	0.567	0.592	0.379	0.095	-
	(0.028)	(0.029)	(0.032)	(0.023)	-
All States	0.492	0.482	0.530	0.412	-
	(0.014)	(0.016)	(0.022)	(0.060)	-
Panel B: White Moves out of Great Plains					
Kansas	0.113	0.038	0.184	0.123	0.025
	(0.009)	(0.009)	(0.032)	(0.012)	(0.003)
Nebraska	0.153	0.039	0.180	0.174	0.031
	(0.029)	(0.007)	(0.032)	(0.037)	(0.004)
North Dakota	0.188	0.051	0.168	0.208	0.023
	(0.012)	(0.012)	(0.020)	(0.015)	(0.004)
Oklahoma	0.185	0.025	0.163	0.213	0.036
	(0.012)	(0.003)	(0.032)	(0.015)	(0.003)
South Dakota	0.149	0.030	0.195	0.160	0.028
	(0.010)	(0.006)	(0.022)	(0.012)	(0.005)
All States	0.160	0.035	0.178	0.181	0.030
	(0.008)	(0.003)	(0.013)	(0.010)	(0.002)

Notes: Table contains estimates and standard errors of $\chi = \Delta/(2+\Delta)$, the share of migrants that chose their destination because of social interactions, based on weighted average estimates from column 3 of Table 2.3 and columns 1-4 of Table 2.6. Standard errors, estimated using the Delta method, are in parentheses.
Source: Duke SSA/Medicare data

Figure 2.1: Proportion Living Outside Birth Region, 1916-1936 Cohorts, by Birth State and Year



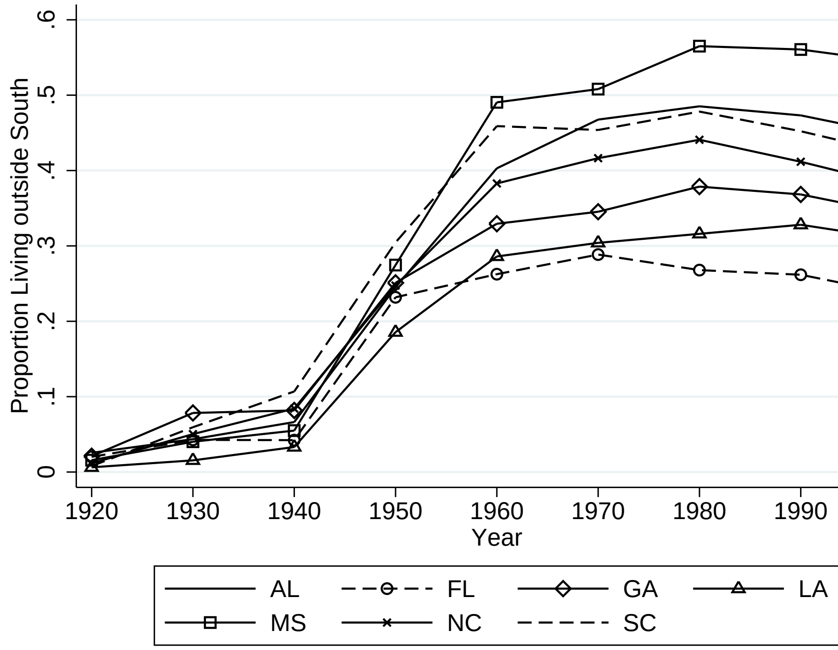
(a) Southern Blacks



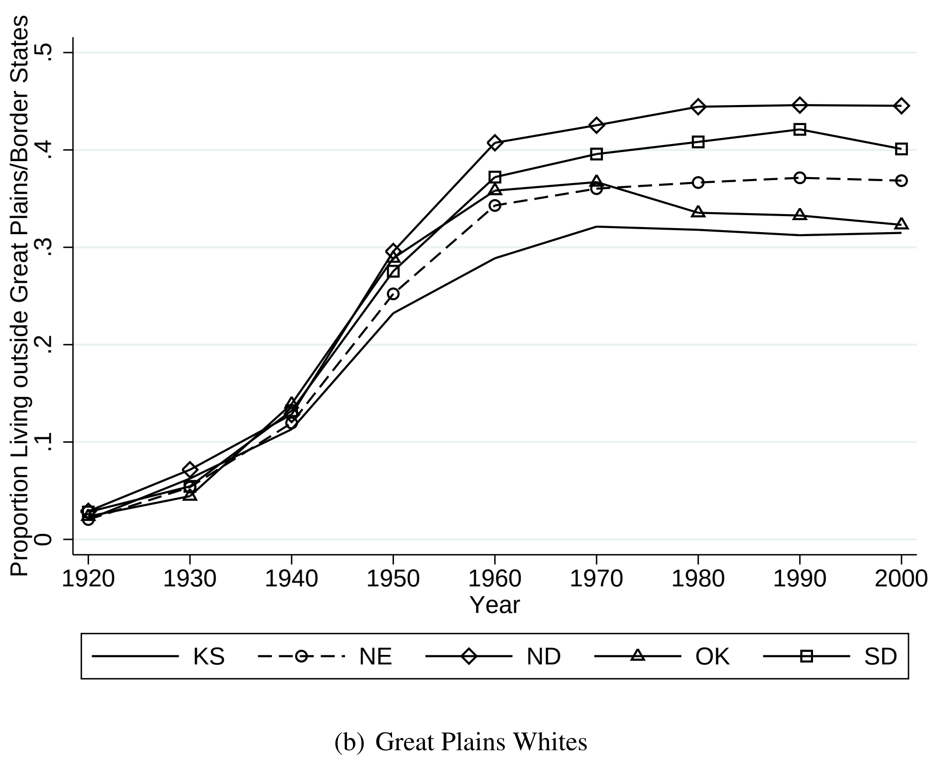
(b) Great Plains Whites

Notes: Figure 2.4 displays birth regions.
Source: Ruggles et al. (2010) data

Figure 2.2: Proportion Living Outside Birth Region, 1916-1936 Cohorts, by Birth State and Year



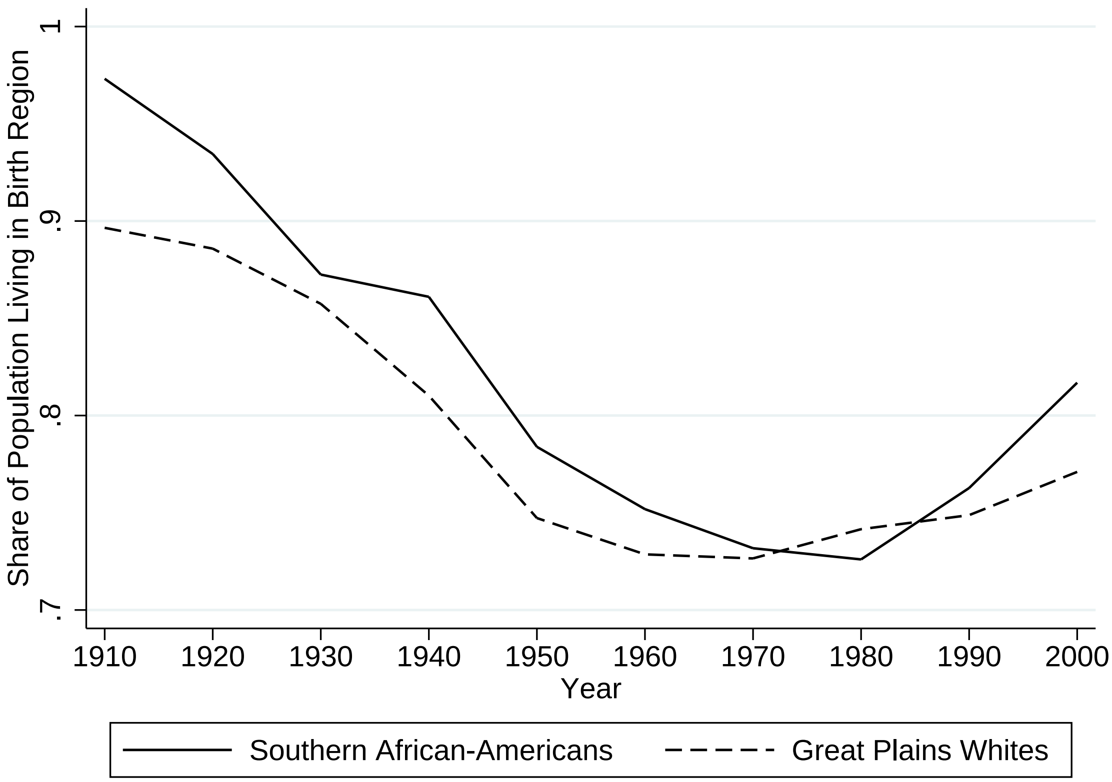
(a) Southern Blacks



(b) Great Plains Whites

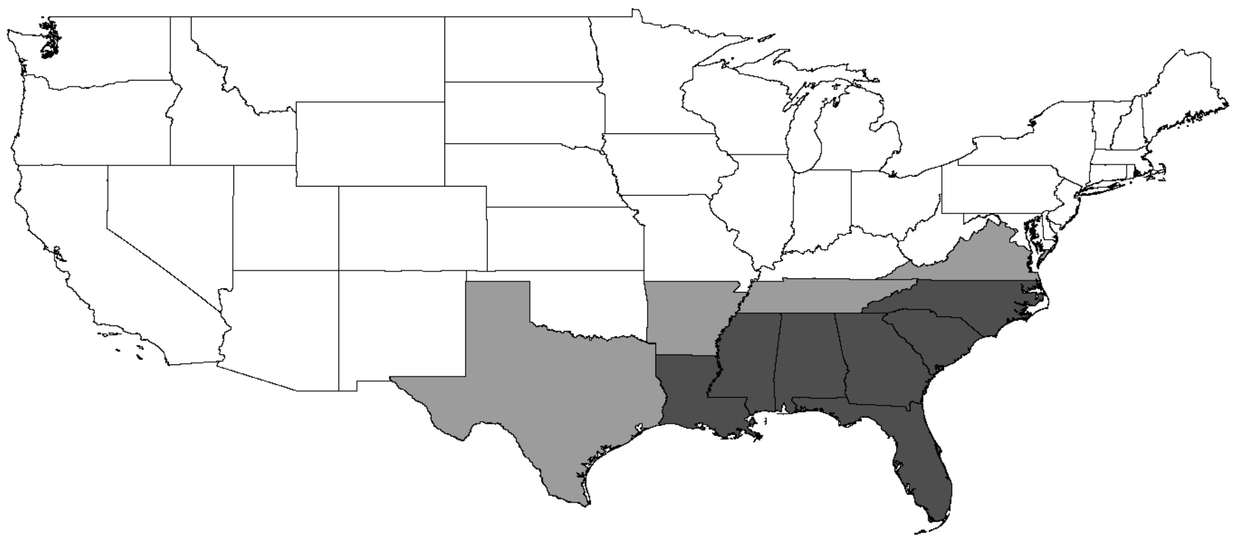
Notes: Figure 2.4 displays birth regions.
Source: Ruggles et al. (2010) data

Figure 2.3: Trajectory of Migrations out of South and Great Plains

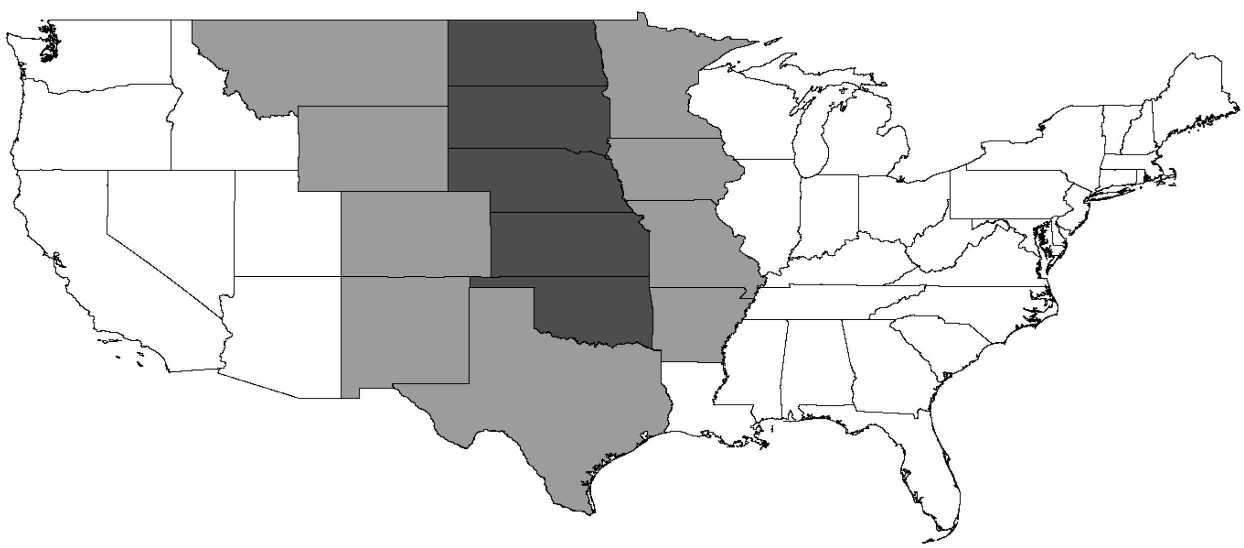


Notes: The solid line shows the proportion of blacks from the seven Southern birth states we analyze (dark grey states in Figure **??**) living in the South (light and dark grey states) at the time of Census enumeration. The dashed line shows the proportion of whites from the Great Plains states living in the Great Plains or Border States. We do not impose age or cohort restrictions for this graph.
Source: Ruggles et al. (2010) data

Figure 2.4: Geographic Coverage



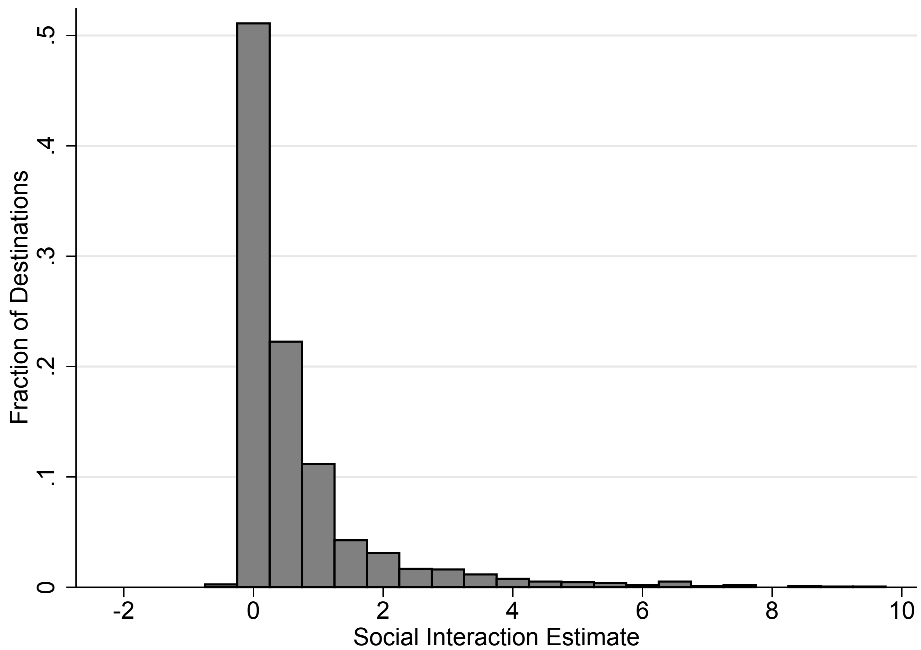
(a) South



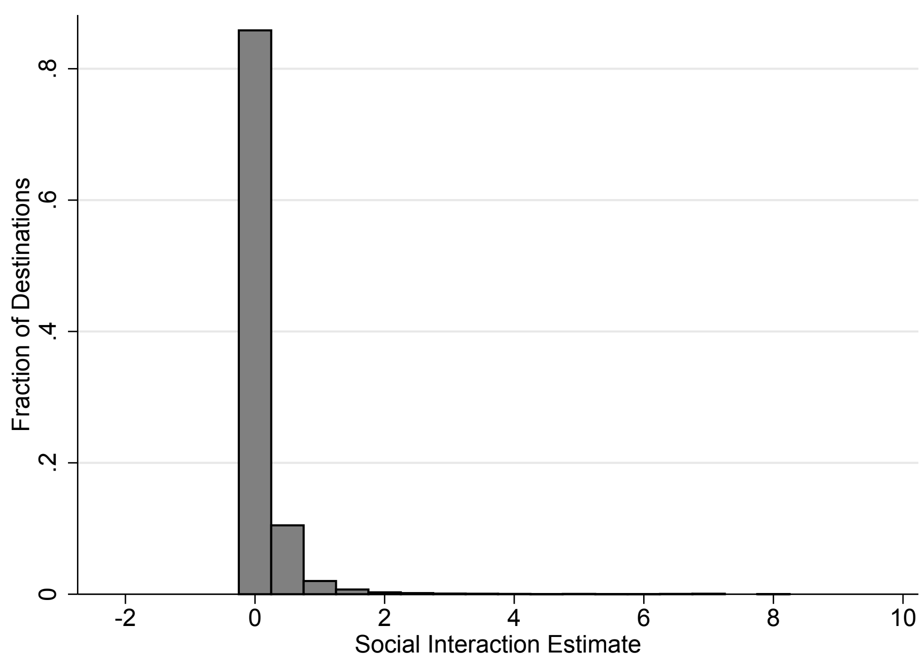
(b) Great Plains

Notes: For the South, our sample includes migrants born in the seven states in dark grey (Alabama, Georgia, Florida, Louisiana, Mississippi, North Carolina, South Carolina). A migrant is someone who at old age lives outside of the former Confederate states, which are the dark and light grey states. For the Great Plains, our sample includes migrants born in the five states in dark grey (Kansas, Nebraska, North Dakota, Oklahoma, South Dakota). A migrant is someone who at old age lives outside of the Great Plains states and the surrounding border area.

Figure 2.5: Distribution of Destination-Level Social Interactions Index Estimates



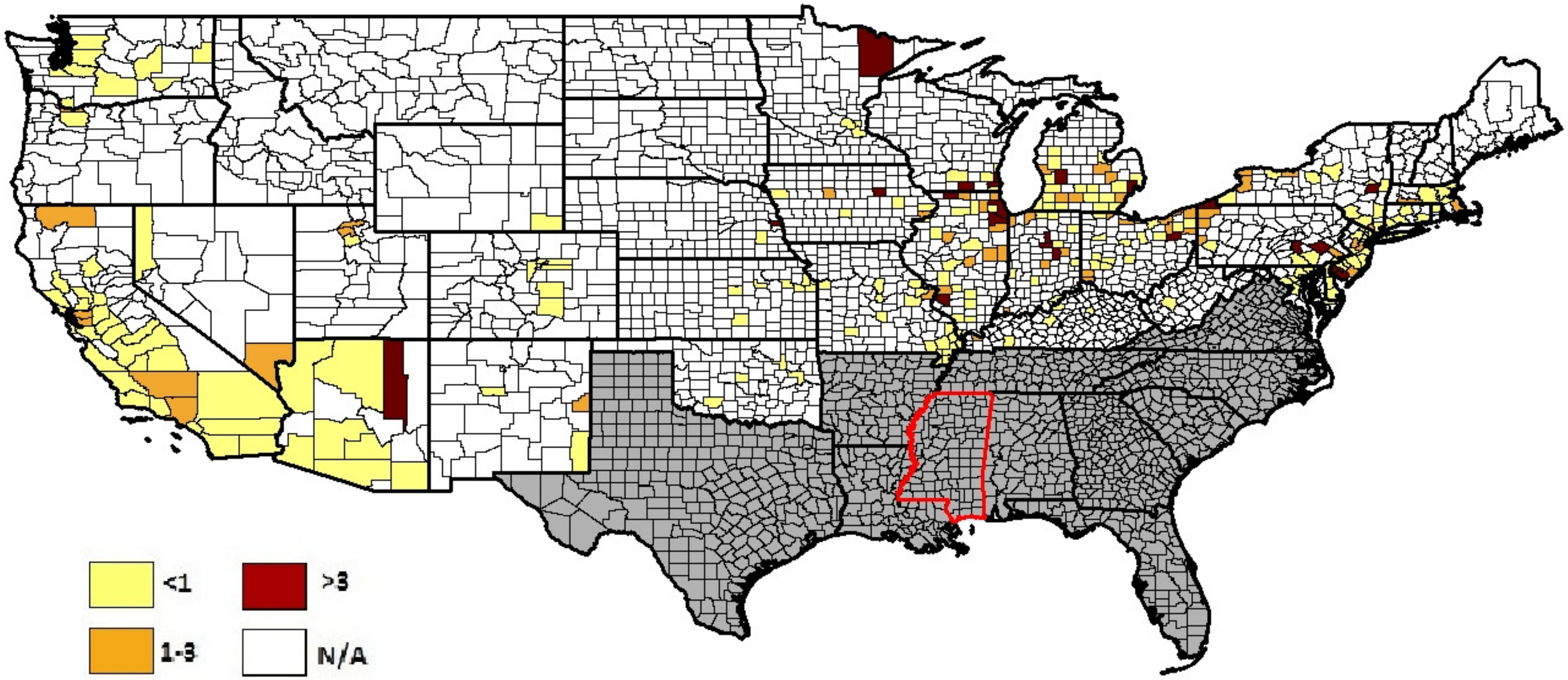
(a) Black Moves out of South



(b) White Moves out of Great Plains

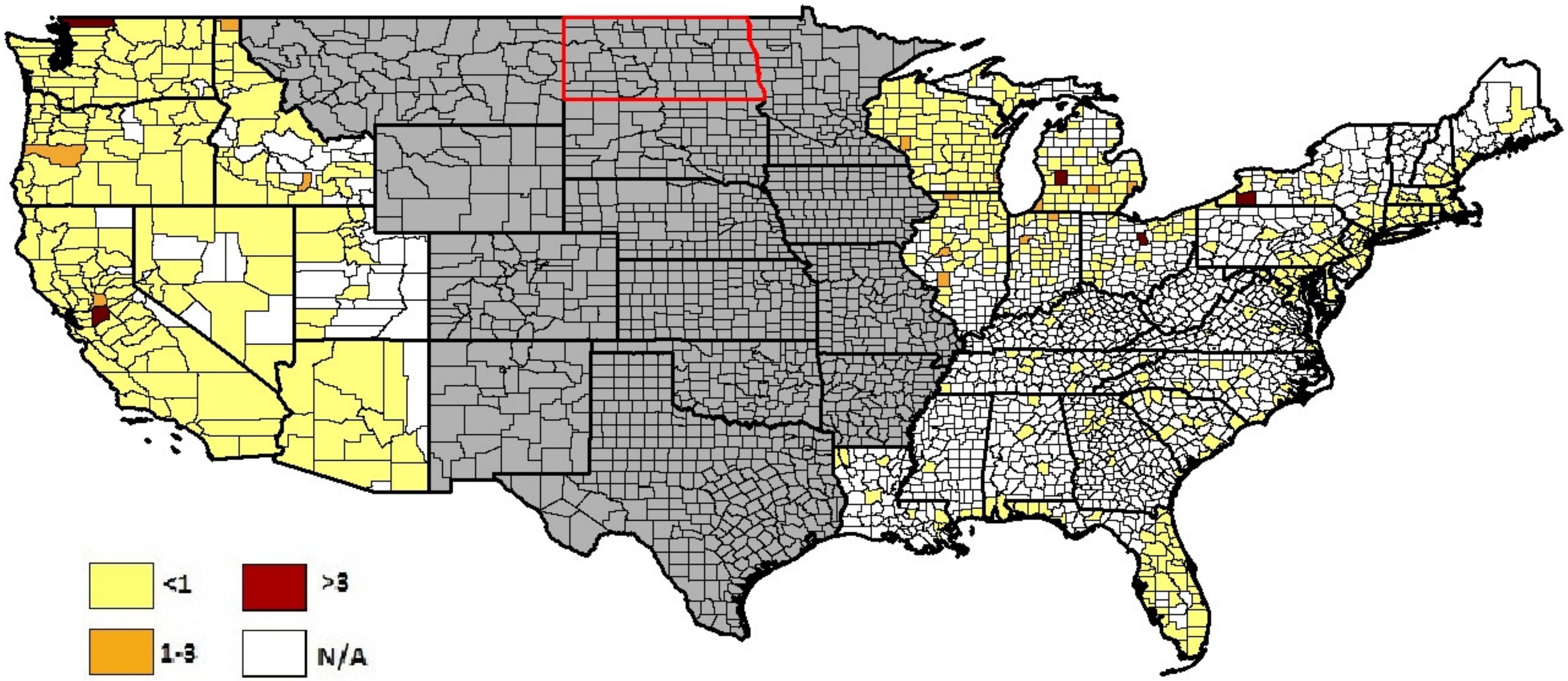
Notes: Bin width is 1/2. Birth town groups are defined by cross validation. Panel (a) omits the estimate $\hat{\Delta}_k = 11.4$ from Mississippi to Racine County, WI, $\hat{\Delta}_k = 15.2$ from South Carolina to Rensselaer County, NY, and $\hat{\Delta}_k = 18.1$ from Florida to St. Joseph County, IN.
Source: Duke SSA/Medicare data

Figure 2.6: Spatial Distribution of Destination-Level Social Interactions Index Estimates, Mississippi-born Blacks



Notes: Figure displays destination-level SI index estimates, $\hat{\Delta}_k$, across U.S. counties for Mississippi-born black migrants. The South is shaded in grey, with Mississippi outlined in red. Destinations to which less than 10 migrants moved are in white. Among all African-American estimates, $\hat{\Delta}_k = 3$ corresponds to the 95th percentile, and $\hat{\Delta}_k = 1$ corresponds to the 81st percentile.
Source: Duke SSA/Medicare data

Figure 2.7: Spatial Distribution of Destination-Level Social Interactions Index Estimates, North Dakota-born Whites



Notes: See note to Figure 2.6. Among all Great Plains white estimates, $\hat{\Delta}_k = 3$ is greater than the 99th percentile, and $\hat{\Delta}_k = 1$ corresponds to the 98th percentile.
Source: Duke SSA/Medicare data

Chapter 3 The Effect of Social Connectedness on Crime: Evidence from the Great Migration

3.1 Introduction

For almost 200 years, the enormous variance of crime rates across space has intrigued social scientists and policy makers (Guerry, 1833; Quetelet, 1835; Weisburd, Bruinsma and Bernasco, 2009). Standard covariates explain a modest amount of cross-city variation in crime, which suggests a potential role for social influences. One possible explanation is peer effects, whereby an individual is more likely to commit crime if his peers commit crime (e.g., Case and Katz, 1991; Glaeser, Sacerdote and Scheinkman, 1996; Damm and Dustmann, 2014). A non-rival explanation is that cities differ in the degree of social connectedness, or the strength of relationships between individuals. Despite vast academic and public interest in the related concept of social capital, concerns about reverse causality and omitted variables seriously limit existing evidence on the effect of social connectedness on crime.

This paper uses a new source of variation in social connectedness to estimate its effect on crime. Social interactions in the migration of millions of African Americans out of the U.S. South from 1915-1970 generated plausibly exogenous variation across destinations in the concentration of migrants that came from the same birth town. For example, consider Beloit, Wisconsin and Middletown, Ohio, two cities similar along many dimensions, including the total number of Southern black migrants that moved there. Around 18 percent of Beloit's black migrants came from Pontotoc, Mississippi, while less than five percent of Middletown's migrants came from any single town. Historical accounts trace the sizable migration from Pontotoc to Beloit to a single influential migrant getting a job in 1914 at a manufacturer in search of workers. Furthermore, qualitative evidence suggests that Southern birth town networks translated into strong community ties in the North. Guided by a simple economic model, we proxy for social connectedness using a Herfindahl-Hirschman Index of birth town to destination city population flows for individuals born from 1916-1936 who we observe in the Duke SSA/Medicare dataset.

Economic theory does not make an unambiguous prediction about whether social connectedness will increase or decrease crime. Social connectedness could increase crime by reinforcing unproductive norms or providing trust that facilitates criminal activity, as with the Ku Klux Klan, Mafia, or gangs (Fukuyama, 2000; Putnam, 2000). Alternatively, social connectedness could decrease crime by increasing the probability that criminals are identified and punished (Becker, 1968) or by facilitating the development of cognitive and non-cognitive skills during childhood (Heckman, Stixrud and Urzua, 2006).

We estimate regressions that relate cross-city differences in crime from 1960-2009 to cross-city differences in social connectedness. We control for the number of Southern black migrants that live in each city to adjust for differences in the overall attractiveness of cities to black migrants, and we control for a rich set of demographic and economic variables, plus state-by-year fixed effects, that might influence crime. We measure city-level crime data using FBI Uniform Crime Reports, which are widely available starting in 1960.

We find that social connectedness leads to sizable reductions in crime rates. At the mean, a one standard deviation increase in social connectedness leads to a precisely estimated 14.1 percent decrease in murder, the best measured crime in FBI data. Our estimates imply that replacing Middletown's social connectedness with that of Beloit would decrease murders by 25.4 percent, robberies by 35.2 percent, and motor vehicle thefts by 22.9 percent. By comparison, the estimates in Chalfin and McCrary (2015) imply that a similar decrease in murders would require a 38 percent increase in the number of police officers. The elasticity of crime with respect to social connectedness ranges from -0.05 to -0.25 across the seven commonly studied index crimes of murder, rape, robbery, assault, burglary, larceny, and motor vehicle theft, and is statistically distinguishable from zero for every crime besides larceny. As predicted by our economic model, the effect of social connectedness on city-level crime rates is stronger in cities with a higher African American population share. Social connectedness reduces crimes that are more and less likely to have witnesses, which suggests that an increased probability of detection is not the only operative mechanism.

The substantial reductions in crime due to social connectedness are not permanent. We estimate significant negative effects of social connectedness in each decade from 1960-1999, and much smaller and insignificant effects from 2000-2009. The attenuated effects from 2000-2009 appear to reflect a decline in the effective strength of social connectedness, as Southern black migrants aged and eventually died. From 1980-1989, social connectedness reduces murders attributed to African American adults and especially African American youth, who belong to the generation of the migrants' children and grandchildren. Social connectedness also reduces murders attributed to non-blacks, consistent with an important role of peer effects.

Several pieces of evidence support the validity of our empirical strategy. Historical accounts point to the importance of migrants who were well connected in their birth town and who worked for an employer in search of labor in establishing concentrated migration flows from Southern birth towns to Northern cities (Scott, 1920; Bell, 1933; Gottlieb, 1987; Grossman, 1989). Many of the initial location decisions were made in the 1910's, over 40 years before we estimate effects on crime. Consistent with the dominant role of idiosyncratic factors, social connectedness is not correlated with crime rates from 1911-1916 or in a consistent manner with economic or demographic covariates from 1960-2000.[1] One potential threat to our empirical strategy is that migrants from the same birth town tended to move to cities with low unobserved determinants of crime and these unobserved determinants of crime persisted over time. We provide evidence that this threat is unimportant by showing that the estimated effect of social connectedness on crime after 1965 is very similar when we control for the 1960-1964 crime rate. We also show that our results are robust to controlling for the share of migrants in each destination that moved there because of social interactions, a variable we obtain by estimating a novel structural model of social interactions in location decisions. Consequently, our estimates likely reflect the effect of social connectedness per se, as opposed to unobserved characteristics of certain migrants.

This paper contributes most directly to the literature studying how characteristics of social networks affect crime. Arguably the best available evidence comes from Sampson, Raudenbush and Earls (1997), who examine the neighborhood-level relationship in Chicago between crime and proxies for collective efficacy, defined as "social cohesion among neighbors combined with their willingness to intervene on behalf of the common good" (p. 918). Despite extremely rich data, their proxies could be correlated with unobserved determinants of crime.[2] We contribute by providing a new source of plausibly exogenous variation in social connectedness and new evidence. We also use a simple economic model to highlight the important interaction between social connectedness and peer effects.

We also contribute to the literature in economics studying the impact of social capital and trust on various outcomes, including growth and development (Knack and Keefer, 1997; Miguel, Gertler and Levine, 2005), government efficiency and public good provision (La Porta et al., 1997; Alesina, Baqir and Easterly, 1999, 2000), financial development (Guiso, Sapienza and Zingales, 2004), and the repayment of microfinance loans (Karlan, 2005, 2007; Cassar, Crowley and Wydick, 2007; Feigenberg, Field and Pande, 2013). We differ from most of this work by focusing on social connectedness, as opposed to social capital or trust, and by using plausibly exogenous cross-city variation in social connectedness.[3] Several papers also examine the determinants of social

[1]The one exception is that social connectedness is positively correlated with the share of a destination's work force employed in manufacturing, a relatively attractive sector for African American migrants (Stuart and Taylor, 2017*a*). We control for a city's manufacturing employment share in our regressions.

[2]Sampson, Raudenbush and Earls (1997) acknowledge that "causal effects were not proven" (p. 923) in their study.

[3]Social connectedness is a broader concept than social capital, trust, or collective efficacy. For example, social connectedness might reduce crime by increasing the probability that criminals are identified, and this behavior typically is not included in definitions of social capital, trust, or collective efficacy. At the same time, our measure might capture social capital that was transported from South to North.

capital and trust (Alesina and Ferrara, 2000; Glaeser et al., 2000; Glaeser, Laibson and Sacerdote, 2002; Karlan et al., 2009; Sapienza, Toldra-Simats and Zingales, 2013). Our results point to the importance of social interactions in location decisions in generating social connectedness.

More broadly, there is enormous interest in the causes and consequences of criminal activity and incarceration in U.S. cities, especially for African Americans (Freeman, 1999; Neal and Rick, 2014; Evans, Garthwaite and Moore, 2016), and this paper demonstrates the importance of social connectedness among African Americans in reducing crime. We also add to the literature on the consequences of the Great Migration for migrants and cities (e.g., Scroggs, 1917; Smith and Welch, 1989; Carrington, Detragiache and Vishwanath, 1996; Collins, 1997; Boustan, 2009, 2011; Hornbeck and Naidu, 2014; Black et al., 2015). This paper draws on Stuart and Taylor (2017*a*), which examines the importance of social interactions in location decisions for African American migrants in more detail.

3.2 Historical Background on the Great Migration

The Great Migration saw nearly six million African Americans leave the South from 1910 to 1970 (Census, 1979).[4] Although migration was concentrated in certain destinations, like Chicago, Detroit, and New York, other cities also experienced dramatic changes. For example, Chicago's black population share increased from two to 32 percent from 1910-1970, while Racine, Wisconsin experienced an increase from 0.3 to 10.5 percent (Gibson and Jung, 2005). Migration out of the South increased from 1910-1930, slowed during the Great Depression, and then resumed forcefully from 1940 to the 1970's.

Several factors contributed to the exodus of African Americans from the South. World War I, which simultaneously increased labor demand among Northern manufacturers and decreased labor supply from European immigrants, helped spark the Great Migration, although many underlying causes existed long before the war (Scroggs, 1917; Scott, 1920; Gottlieb, 1987; Marks, 1989; Jackson, 1991; Collins, 1997; Gregory, 2005). Underlying causes included a less developed Southern economy, the decline in agricultural labor demand due to the boll weevil's destruction of crops (Scott, 1920; Marks, 1989, 1991; Lange, Olmstead and Rhode, 2009), widespread labor market discrimination (Marks, 1991), and racial violence and unequal treatment under Jim Crow laws (Tolnay and Beck, 1991).

Migrants tended to follow paths established by railroad lines: Mississippi-born migrants predominantly moved to Illinois and other Midwestern states, and South Carolina-born migrants predominantly moved to New York and Pennsylvania (Scott, 1920; Carrington, Detragiache and Vishwanath, 1996; Collins, 1997; Boustan, 2011; Black et al., 2015). Labor agents, offering paid

[4] Parts of this section come from Stuart and Taylor (2017*a*).

transportation, employment, and housing, directed some of the earliest migrants, but their role diminished sharply after the 1920's, and most individuals paid for the relatively expensive train fares themselves (Gottlieb, 1987; Grossman, 1989).[5] African-American newspapers from the largest destinations circulated throughout the South, providing information on life in the North (Gottlieb, 1987; Grossman, 1989).[6] Blacks attempting to leave the South sometimes faced violence (Scott, 1920; Henri, 1975).

Historical accounts and recent quantitative work indicate that social interactions strongly affected location decisions during the Great Migration. Initial migrants, most of whom moved in the 1910's, chose their destination primarily in response to economic opportunity. Migrants who worked for an employer in search of labor and were well connected in their birth town linked friends, family, and acquaintances to jobs and shelter in the North, sometimes leading to persistent migration flows from birth town to destination city (Rubin, 1960; Gottlieb, 1987; Stuart and Taylor, 2017*a*). Stuart and Taylor (2017*a*) show that birth town-level social interactions strongly influenced the location decisions of African American migrants from the South. These social interactions mirror vertical migration patterns established by railroad lines and were stronger in destinations with more manufacturing employment, a particularly attractive sector for black workers during this time.

The experience of John McCord captures many important features of early black migrants' location decision.[7] Born in Pontotoc, Mississippi, nineteen-year-old McCord traveled in search of higher wages in 1912 to Savannah, Illinois, where a fellow Pontotoc-native connected him with a job. McCord moved to Beloit, Wisconsin in 1914 after hearing of employment opportunities and quickly began working as a janitor at the manufacturer Fairbanks Morse and Company. After two years in Beloit, McCord spoke to his manager about returning home for a vacation. The manager asked McCord to recruit workers during the trip. McCord returned with 18 unmarried men, all of whom were soon hired. Thus began a persistent flow of African Americans from Pontotoc to Beloit: among individuals born from 1916-1936, 14 percent of migrants from Pontotoc lived in Beloit's county at old age (Stuart and Taylor, 2017*a*).[8]

Qualitative evidence documents the importance of social ties among African Americans from the same birth town for life in the North. For example, roughly 1,000 of Erie, Pennsylvania's 11,600 African American residents once lived in Laurel, Alabama, and almost half had family connections to Laurel, leading an Erie resident to say, "I'm surrounded by so many Laurelites here,

[5] In 1918, train fare from New Orleans to Chicago cost $22 per person, when Southern farmers' daily wages typically were less than $1 and wages at Southern factories were less than $2.50 (Henri, 1975).

[6] The *Chicago Defender*, perhaps the most prominent African-American newspaper of the time, was read in 1,542 Southern towns and cities in 1919 (Grossman, 1989).

[7] The following paragraph draws on Bell (1933). See also Knowles (2010).

[8] This is 68 times larger than the percent of migrants from Mississippi that lived in Beloit's county at old age.

it's like a second home" (Associated Press, 1983). Nearly forty percent of the migrants in Decatur, Illinois came from Brownsville, Tennessee, and Brownsville high school reunions took place in Decatur from the 1980's to 2000's (Laury, 1986; Smith, 2006).[9] As described by a Brownsville native, "Decatur's a little Brownsville, really" (Laury, 1986).

3.3 A Simple Model of Crime and Social Connectedness

This section describes a simple model of crime and social connectedness. Social connectedness, or the strength of relationships between individuals, could reduce crime through multiple channels, including by increasing the probability that criminals are identified and punished or by facilitating the development of human capital during childhood. We use the model to derive an empirical measure of social connectedness, and we show how the effect of social connectedness on crime depends on peer effects.

3.3.1 Individual Crime Rates

We focus on a single city and characterize individuals by their age and social ties. For simplicity, we consider a static model in which each younger individual makes a single decision about whether to commit crime, while older individuals do not commit crime. Each individual belongs to one of three groups: blacks with ties to the South ($\tau_i = s$), blacks without ties to the South ($\tau_i = n$), and all others ($\tau_i = w$). Older individuals have a tie to the South if they were born there. Younger individuals have a tie to the South if at least one of their parents, who are older individuals, was born in the South. We index younger individuals by i and older individuals by o.

For a younger individual who is black with ties to the South, we model the probability of committing crime as

$$\mathbb{E}[C_i|\tau_i = s, j_i = j] = \alpha^s + \beta^s\,\mathbb{E}[C_{-i}] + \sum_o \gamma^s_{i,o,j}, \tag{3.1}$$

where $C_i = 1$ if person i commits crime and $C_i = 0$ otherwise, and j_i denotes the birth town of i's parents. Equation (3.1) is a linear approximation to the optimal crime rule from a utility-maximizing model in which the relative payoff of committing crime depends on three factors. First, α^s, which is common to all individuals of type s, captures all non-social determinants of crime (e.g., due to police or employment opportunities). Second, an individual's decision to commit crime depends on the expected crime rate among his peers, $\mathbb{E}[C_{-i}]$. Finally, the effect of social connectedness is $\sum_o \gamma^s_{i,o,j}$, where $\gamma^s_{i,o,j}$ is the influence of older individual o on younger individual

[9]The 40 percent figure comes from the Duke SSA/Medicare dataset, described below.

i. This reduced-form representation captures several possible channels through which social connectedness might affect crime. For example, older individuals might reduce crime among younger individuals by increasing the probability a criminal is identified and punished (Becker, 1968) or by increasing younger individuals' stock of cognitive and non-cognitive skills, which boost earnings in the non-crime labor market (Heckman, Stixrud and Urzua, 2006). Alternatively, social connectedness could increase crime by reinforcing unproductive norms or providing trust that facilitates criminal activity, as with the Ku Klux Klan, Mafia, or gangs (Fukuyama, 2000; Putnam, 2000). Ethnographic work describing African American families and kinship networks suggests crime-reducing effects of social connectedness (Stack, 1970).

Motivated by the qualitative evidence described in Section 3.2, we model social connectedness as a function of whether the parents of individual i share a birth town with individual o. In particular, $\gamma^s_{i,o,j} = \gamma^s_H$ if the individuals share a birth town connection, $j_i = j_o$, and $\gamma^s_{i,o,j} = \gamma^s_L$ otherwise. We assume that younger blacks with ties to the South are only influenced by older blacks with ties to the South, so that $\gamma^s_{i,o,j} = 0$ if $\tau_i \neq \tau_o$. Given these assumptions, the effect of social connectedness on person i is a weighted average of the high connectedness effect (γ^s_H) and the low connectedness effect (γ^s_L),

$$\sum_o \gamma^s_{i,o,j} = \frac{N^s_{j,0}}{N^s_0}\gamma^s_H + \left(1 - \frac{N^s_{j,0}}{N^s_0}\right)\gamma^s_L, \tag{3.2}$$

where $N^s_{j,0}$ is the number of older individuals of type s from birth town j, and $N^s_0 = \sum_j N^s_{j,0}$ is the total number of older individuals in the city. Because social interactions depend on birth town connections, the older generation's migration decisions lead to differences in expected crime rates for younger individuals with ties to different birth towns.

The Herfindahl-Hirschman Index emerges as a natural way to measure social connectedness in this model. In particular, the probability that a randomly chosen African American with ties to the South commits crime is

$$\mathbb{E}[C_i|\tau_i = s] = \alpha^s + \beta^s\,\mathbb{E}[C_{-i}] + \gamma^s_L + (\gamma^s_H - \gamma^s_L)\mathrm{HHI}^s, \tag{3.3}$$

where $\mathrm{HHI}^s \equiv \sum_j (N^s_{j,0}/N^s_0)^2$ is the Herfindahl-Hirschman Index of birth town to destination city population flows for African Americans with ties to the South.[10] The direct effect of social connectedness on the type s crime rate is $\gamma^s_H - \gamma^s_L$. One reasonable case is $\gamma^s_H < \gamma^s_L < 0$, so that older individuals discourage younger individuals from committing crime, and the effect is stronger among individuals who share a birth town connection. Expressions analogous to equation (3.3)

[10] In deriving equation (3.3), we assume that each Southern birth town accounts for the same share of individuals in the younger and older generations, so that $N^s_{j,0}/N^s_0 = N^s_{j,1}/N^s_1 \forall j$, where $N^s_{j,1}$ is the number of younger individuals of type s with a connection to birth town j, and $N^s_1 = \sum_j N^s_{j,1}$ is the total number of younger individuals.

exist for African American youth without ties to the South ($\tau_i = n$) and non-black youth ($\tau_i = w$).

3.3.2 City-Level Crime Rates

We next consider the equilibrium of this model, in which peer effects can accentuate or attenuate the effect of social connectedness on crime. We use HHI to measure social connectedness and allow peer effects to differ by the type of peer, leading to the following equilibrium,

$$\bar{C}^s = F^s(\alpha^s, \text{HHI}^s, \bar{C}^s, \bar{C}^n, \bar{C}^w) \tag{3.4}$$

$$\bar{C}^n = F^n(\alpha^n, \text{HHI}^n, \bar{C}^s, \bar{C}^n, \bar{C}^w) \tag{3.5}$$

$$\bar{C}^w = F^w(\alpha^w, \text{HHI}^w, \bar{C}^s, \bar{C}^n, \bar{C}^w), \tag{3.6}$$

where $\bar{C}^\tau$ is the crime rate among younger individuals of type τ, and F^τ characterizes the equilibrium crime rate responses. The equilibrium crime rate vector $(\bar{C}^s, \bar{C}^n, \bar{C}^w)$ is a fixed point of equations (3.4)-(3.6).

We are interested in the effect of social connectedness among African Americans with ties to the South, HHI^s, on equilibrium crime rates. Equations (3.4)-(3.6) imply that

$$\frac{d\bar{C}^s}{d\text{HHI}^s} = \frac{\partial F^s}{\partial \text{HHI}^s}\left(\frac{(1-J_{22})(1-J_{33}) - J_{23}J_{32}}{\det(I-J)}\right) \equiv \frac{\partial F^s}{\partial \text{HHI}^s} m^s \tag{3.7}$$

$$\frac{d\bar{C}^n}{d\text{HHI}^s} = \frac{\partial F^s}{\partial \text{HHI}^s}\left(\frac{J_{23}J_{31} + J_{21}(1-J_{33})}{\det(I-J)}\right) \equiv \frac{\partial F^s}{\partial \text{HHI}^s} m^n \tag{3.8}$$

$$\frac{d\bar{C}^w}{d\text{HHI}^s} = \frac{\partial F^s}{\partial \text{HHI}^s}\left(\frac{J_{21}J_{32} + J_{31}(1-J_{22})}{\det(I-J)}\right) \equiv \frac{\partial F^s}{\partial \text{HHI}^s} m^w, \tag{3.9}$$

where I is the 3×3 identity matrix and J, a sub-matrix of the Jacobian of equations (3.4)-(3.6), captures the role of peer effects.[11] Equations (3.7)-(3.9) depend on the direct effect of HHI^s on crime among blacks with ties to the South, $\partial F^s/\partial\text{HHI}^s$, times a peer effect multiplier, given by m^s, m^n, and m^w. We assume the equilibrium is stable, which essentially means that peer effects are not too large.[12] For example, if $J_{11} \equiv \partial F^s/\partial \bar{C}^s \geq 1$, and there are no cross-group peer effects,

[11] In particular,

$$J \equiv \begin{bmatrix} \partial F^s/\partial \bar{C}^s & \partial F^s/\partial \bar{C}^n & \partial F^s/\partial \bar{C}^w \\ \partial F^n/\partial \bar{C}^s & \partial F^n/\partial \bar{C}^n & \partial F^n/\partial \bar{C}^w \\ \partial F^w/\partial \bar{C}^s & \partial F^w/\partial \bar{C}^n & \partial F^w/\partial \bar{C}^w \end{bmatrix},$$

and J_{ab} is the (a, b) element of J. m^s is the $(1, 1)$ element of $(I - J)^{-1}$, m^n is the $(2, 1)$ element, and m^w is the $(3, 1)$ element.

[12] The technical assumption underlying stability is that the spectral radius of J is less than one. This condition is analogous to the requirement in linear-in-means models that the slope coefficient on the endogenous peer effect is less than one in absolute value (e.g., Manski, 1993).

then a small increase in the crime rate among individuals of type s leads to an equilibrium where all individuals of type s commit crime. In contrast, a small change in any group's crime rate does not lead to a corner solution in a stable equilibrium.

Our first result is that if social connectedness reduces the crime rate of African Americans with ties to the South, then social connectedness reduces the crime rate of all groups, as long as the equilibrium is stable and peer effects (i.e., elements of J) are non-negative. $d\bar{C}^s/d\text{HHI}^s \leq 0, d\bar{C}^n/d\text{HHI}^s \leq 0$, and $d\bar{C}^w/d\text{HHI}^s \leq 0$ if $\partial F^s/\partial \text{HHI}^s < 0$, the equilibrium is stable, and peer effects are non-negative.

In a stable equilibrium with non-negative peer effects, the crime-reducing effect of social connectedness among Southern blacks is not counteracted by higher crime rates among other groups. Hence, equilibrium crime rates of all groups weakly decrease in Southern African American HHI. With negative cross-group peer effects, the reduction in crime rates among Southern blacks could lead to higher crime by other groups. Proposition 3.3.2 is not surprising, and we provide a proof in Appendix B.1.

Because of data limitations, most of our empirical analysis examines the city-level crime rate, $\bar{C}$, which is a weighted average of the three group-specific crime rates,

$$\bar{C} = P^b[P^{s|b}\bar{C}^s + (1 - P^{s|b})\bar{C}^n] + (1 - P^b)\bar{C}^w, \tag{3.10}$$

where P^b is the black population share and $P^{s|b}$ is the share of the black population with ties to the South. Proposition 3.3.2 provides sufficient, but not necessary, conditions to ensure that Southern black HHI decreases the city-level crime rate, $\bar{C}$, when the direct effect is negative. There exist situations in which cross-group peer effects are negative, but an increase in HHI^s still decreases in the city-level crime rate.

Our second result is that the effect of Southern black social connectedness on the city-level crime rate decreases (or, increases in magnitude) with the black population share for certain peer effect parametrizations. $d\bar{C}/d\text{HHI}^s$ decreases with P^b if $\partial F^s/\partial \text{HHI}^s < 0$, the equilibrium is stable, and cross-group peer effects are non-negative and sufficiently small.

We assume that the effect of HHI^s on each group's crime rate does not depend on the black population share, yielding[13]

$$\frac{d^2\bar{C}}{d\text{HHI}^s dP^b} = P^{s|b}\frac{d\bar{C}^s}{d\text{HHI}^s} + (1 - P^{s|b})\frac{d\bar{C}^n}{d\text{HHI}^s} - \frac{d\bar{C}^w}{d\text{HHI}^s}. \tag{3.11}$$

[13]It is not clear whether we would expect, say, $d\bar{C}^s/d\text{HHI}^s$ to be more or less negative in cities with higher P^b. The effect could decrease in magnitude if the higher black population share diluted existing community ties, or the effect could increase in magnitude if the higher black population share reinforced community ties. The former case makes Proposition 3.3.2 less likely to hold, while the latter case makes it more likely.

Two jointly sufficient conditions for Proposition 3.3.2 are (a): $d\bar{C}^s/d\text{HHI}^s < d\bar{C}^w/d\text{HHI}^s$ and (b): $d\bar{C}^n/d\text{HHI}^s \leq d\bar{C}^w/d\text{HHI}^s$. If Southern black social connectedness leads to greater crime reductions among both groups of African Americans, relative to non-blacks, then the total effect will be larger in magnitude in cities with a higher black population share. In this case, Proposition 3.3.2 occurs mechanically. The nature of peer effects determines whether conditions (a) and (b) are satisfied, and we provide precise conditions in Appendix B.1.

As a simple example, suppose there are no cross-group peer effects between blacks and non-blacks ($J_{13} = J_{23} = J_{31} = J_{32} = 0$). In this case, an increase in HHI^s does not affect the crime rate among non-blacks, so condition (a) holds. Condition (b) requires that an increase in HHI^s must not increase crime among blacks without ties to the South, which will be true if peer effects between the two groups of African Americans are non-negative. As shown in Appendix B.1, the formal conditions in this example are a stable equilibrium and $J_{21} \geq 0$.

In sum, we expect that higher social connectedness among African Americans with ties to the South will reduce the city-level crime rate (Proposition 3.3.2). We also expect that the effect will be stronger in cities with a higher black population share (Proposition 3.3.2). Furthermore, the effect of social connectedness among African Americans with ties to the South on the city-level crime rate depends critically on the nature of a peer effects, an issue we examine more fully in Section 3.6 after presenting our baseline results.

3.4 Data and Empirical Strategy

3.4.1 Data on Crime, Social Connectedness, and Control Variables

To estimate the effect of social connectedness on crime, we use three different data sets. We measure annual city-level crime counts using FBI Uniform Crime Report (UCR) data for 1960-2009, available from ICPSR. UCR data contain voluntary monthly reports on the number offenses reported to police, which we aggregate to the city-year level.[14] We focus on the seven commonly studied index crimes: murder and non-negligent manslaughter ("murder"), forcible rape ("rape"), robbery, assault, burglary, larceny, and motor vehicle theft. Murder is the best measured crime, and robbery and motor vehicle theft are also relatively well-measured (Blumstein, 2000; Tibbetts, 2012). Because missing observations are indistinguishable from true zeros, we drop any city-year in which any of the three property crimes (burglary, larceny, and motor vehicle theft) equal zero. We also use annual population estimates from the Census Bureau in the UCR data.

The Duke SSA/Medicare dataset provides the birth town-to-destination city population flows

[14] We use Federal Information Processing System (FIPS) place definitions of cities. We follow Chalfin and McCrary (2015) in decreasing the number of murders for year 2001 in New York City by 2,753, the number of victims of the September 11 terrorist attack.

that underlie our measure of social connectedness. The data contain sex, race, date of birth, date of death (if deceased), and the ZIP code of residence at old age (death or 2001, whichever is earlier) for over 70 million individuals who received Medicare Part B from 1976-2001. In addition, the data include a 12-character string with self-reported birth town information, which is matched to places, as described in Black et al. (2015). We focus on individuals born from 1916-1936 in the former Confederate states, which we refer to as the South.[15] We restrict our main analysis sample to cities that received at least 25 Southern-born African American migrants in the Duke dataset to improve the reliability of our estimates.

Census city data books provide numerous city-level covariates for 1960, 1970, 1980, 1990, and 2000. These data are only available for cities with at least 25,000 residents in 1960, 1980, and 1990, and we apply the same restriction for 1970 and 2000. We limit our sample to cities in the Northeast, Midwest, and West Census regions to focus on the cross-region moves that characterize the Great Migration. Our main analysis sample excludes the 14 cities with 1980 population greater than 500,000, as we found considerable measurement error in murder counts for these cities.[16] Appendix Tables B.1 and B.2 provide summary statistics.

3.4.2 Estimating the Effect of Social Connectedness on Crime

Our main estimating equation is

$$Y_{k,t} = \exp[\ln(\text{HHI}_k)\delta + \ln(N_k)\theta + X'_{k,t}\beta] + \epsilon_{k,t}, \tag{3.12}$$

where $Y_{k,t}$ is the number of crimes in city k in year t. The key variable of interest is our proxy for social connectedness among African Americans with ties to the South, $\text{HHI}_k = \sum_j (N_{j,k}/N_k)^2$, where $N_{j,k}$ is the number of migrants from birth town j that live in destination city k, and $N_k \equiv \sum_j N_{j,k}$ is the total number of migrants. A Herfindahl-Hirschman Index is a natural way to measure social connectedness, as shown in Section 3.3, and approximately equals the probability that two randomly chosen migrants living in city k share a birth town.[17] $X_{k,t}$ is a vector of covariates, including log population and other variables described below, and $\epsilon_{k,t}$ captures unobserved

[15] Coverage rates decline considerably for earlier and later cohorts (Black et al., 2015; Stuart and Taylor, 2017*a*).

[16] In particular, we constructed annual murder counts using the FBI UCR data, which are not broken down by age, race, or sex, and the FBI Age-Sex-Race (ASR) data, which are. Both data sets should yield the same number of murders in a city, but substantial discrepancies exist in the largest cities (see Appendix Figure B.1). We do not know why the murder counts differ between these data sets.

[17] The probability that two randomly chosen migrants living in city k share a birth town is

$$\mathbb{P}[j_i = j_{i'}] = \sum_j \mathbb{P}[j_i = j_{i'} | j_{i'} = j]\, \mathbb{P}[j_i = j] = \sum_j \frac{N_{j,k} - 1}{N_k - 1} \frac{N_{j,k}}{N_k} \approx \text{HHI}_k.$$

determinants of crime.[18] We use an exponential function in equation (3.12) because there are no murders for many city-year observations (Appendix Table B.1). We cluster standard errors at the city level to allow for arbitrary autocorrelation in the unobserved determinants of crime.[19]

The key parameter of interest is δ, which we interpret as the elasticity of the crime *rate* with respect to HHI_k, our proxy of social connectedness, because we control for log population. If social connectedness reduces the city-level crime rate, as predicted by Proposition 3.3.2, then $\delta < 0$.

We estimate δ using cross-city variation in social connectedness, conditional on the total number of migrants and other covariates. To identify δ, we make the following conditional independence assumption,

$$\epsilon_{k,t}\text{HHI}_k|(N_k, X_{k,t}). \tag{3.13}$$

Condition (3.13) states that, conditional on the number of migrants living in city k and the vector of control variables, social connectedness is independent of unobserved determinants of crime from 1960-2009. This condition allows the total number of migrants, N_k, to depend arbitrarily on unobserved determinants of crime, $\epsilon_{k,t}$.[20]

We include several control variables in $X_{k,t}$ that bolster the credibility of condition (3.13). State-by-year fixed effects flexibly account for determinants of crime that vary over time at the state-level, due to changes in economic conditions, police enforcement, government spending, and other factors. Demographic covariates include log population, percent black, percent female, percent age 5-17, percent age 18-64, percent age 65 and older, percent at least 25 years old with a high school degree, percent at least 25 years old with a college degree, and log city area. Economic covariates include log median family income, unemployment rate, labor force participation rate, and manufacturing employment share.[21] We observe log population in every year and, with a few exceptions, we observe the remaining demographic and economic covariates every ten years from 1960-2000.[22] In explaining crime in year t, we only use covariates corresponding to the decade in which t lies. We allow coefficients for all covariates besides log population to vary across decades to account for possible changes in the importance of economic and demographic covariates.

[18]Because equation (3.12) includes $\ln(\text{HHI}_k)$, $\ln(N_k)$, and log population, our estimate of δ would be identical if we used city population as the denominator of HHI_k.

[19]Equation (3.12) emerges from a Poisson model, but consistent estimation of (δ, θ, β) does not require any restriction on the conditional variance of the error term (e.g., Wooldridge, 2002).

[20]Condition (3.13) does not guarantee identification of the other parameters in equation (3.12) besides δ. For example, identification of θ requires exogenous variation in the total number of migrants in each city. Boustan (2011) provides one possible strategy for such an approach, but we do not pursue that here.

[21]Stuart and Taylor (2017*a*) find that the manufacturing employment share predicts the strength of social interactions in location decisions among Southern black migrants, which leads to higher social connectedness.

[22]The exceptions are percent female (not observed in 1960), percent at least 25 years old with a high school degree and a college degree (not observed in 2000), log median family income (not observed in 2000), and manufacturing share (not observed in 2000). For decades in which a covariate is not available, we use the adjacent decade.

Several pieces of evidence support the validity of condition (3.13). First, variation in social connectedness stems from location decisions made 50 years before we estimate effects on crime. As described in Section 3.2, initial migrants in the 1910's chose their destination in response to economic opportunity, and idiosyncratic factors, like a migrant's ability to persuade friends and family to join them, strongly influenced whether other migrants followed.[23]

Table 3.1 shows that social connectedness is not correlated with homicide rates from 1911-1914. In particular, we regress $\ln(\text{HHI}_k)$ on $\ln(N_k)$ and log homicide rates from 1911-1914, which we observe from historical mortality statistics published for cities with at least 100,000 residents in 1920 (Census, 1922). We find no significant relationship between social connectedness and early century crime rates. This conclusion holds when we use inverse probability weights to make this sample of cities more comparable to our main analysis sample on observed characteristics.[24] These results partially dismiss the possibility that social connectedness is correlated with extremely persistent unobserved determinants of crime, which would threaten our empirical strategy.

If anything, limitations in the data used to construct HHI_k could lead us to understate any negative effect of social connectedness on crime. We construct HHI_k using migrants' location at old age, measured at some point from 1976-2001. As a result, migration after 1960, when we first measure crime, could influence HHI_k and the estimated effect on crime, δ. If migrants with a higher concentration of friends and family nearby were less likely to out-migrate in response to higher crime shocks, $\epsilon_{k,t}$, then HHI_k would be larger in cities with greater unobserved determinants of crime. This would bias our estimate of δ upwards, making it more difficult to conclude that social connectedness reduces crime. Reassuringly, Table 3.2 reveals very low migration rates during this period among African Americans who were born from 1916-1936 in the South and living in the North. Around 90 percent of individuals stayed in the same county for the five-year periods from 1955-1960, 1965-1970, 1975-1980, 1985-1990, and 1995-2000.[25] This table suggests that our inability to construct HHI_k using migrants' location before 1960 is relatively unimportant.

Table 3.3 provides additional indirect evidence in support of condition (3.13) by showing that social connectedness is not systematically correlated with most demographic or economic covari-

[23]For example, Scott (1920) writes, "The tendency was to continue along the first definite path. Each member of the vanguard controlled a small group of friends at home, if only the members of his immediate family. Letters sent back, representing that section of the North and giving directions concerning the route best known, easily influenced the next groups to join their friends rather than explore new fields. In fact, it is evident throughout the movement that the most congested points in the North when the migration reached its height, were those favorite cities to which the first group had gone" (p. 69).

[24]We do not adjust the standard errors in columns 3-4 for the use of inverse probability weights. As a result, the p-values for these columns are likely too small, which further reinforces our finding of no significant relationship. Appendix Table B.3 compares the observed characteristics of cities for which we do and do not observe 1911-1914 mortality rates.

[25]Available data do not allow us to examine whether out-migration rates vary with the concentration of friends and family living nearby, which is the type of behavior that would affect HHI_k.

ates. The lack of systematic correlations with observed variables suggests that social connectedness is not correlated with unobserved determinants of crime, $\epsilon_{k,t}$. We regress log HHI on various covariates for the 228 cities observed in every decade from 1960 to 2000. To facilitate comparisons, we normalize all variables, separately for each decade, to have mean zero and standard deviation one. Only the log number of migrants and the manufacturing employment share are consistently correlated with log HHI. The negative correlation between log HHI and the log number of migrants arises because a large number of migrants necessarily came from many sending towns, due to the small size of Southern towns relative to Northern cities. The positive correlation between log HHI and the manufacturing employment share arises because social interactions in location decisions guided migrants to destinations with ample manufacturing employment, which was especially attractive to African American workers (Stuart and Taylor, 2017*a*). The bottom panel reports p-values from tests that demographic or economic covariates (besides the manufacturing employment share) are unrelated to log HHI. We fail to reject this null hypothesis at standard significance levels from 1960-1980, providing support for condition (3.13). There is a significant relationship between social connectedness and covariates in 1990 and 2000, but this does not necessarily provide evidence against condition (3.13) because social connectedness might have affected these later outcomes.[26] Appendix Table B.4 shows results when adding a number of covariates measured among African-Americans.

Figure 3.1 further describes the cross-city variation in social connectedness by plotting log HHI and the log number of Southern black migrants. Our regressions identify the effect of social connectedness on crime with variation in HHI conditional on the number of migrants in a city (and other covariates), which is variation in the vertical dimension of Figure 3.1. Except for cities with at least 500,000 residents in 1980, there is considerable variation in log HHI conditional on the log number of migrants. Figure 3.2 shows that social connectedness stems largely from the location decisions of a single sending town. Sixty-seven percent of the variation in log HHI is explained by the leading term of log HHI, which equals the log squared share of migrants from the top sending town. This finding reinforces the importance of idiosyncratic features of migrants and birth towns in generating variation in social connectedness.[27]

[26]The significant relationship between social connectedness and demographic covariates in 1990 and 2000 is driven by a negative relationship between social connectedness and the percent of the population age 0-4. Social connectedness could lower birth rates by increasing the opportunity cost of having children (by increasing human capital). The significant relationship between social connectedness and economic covariates in 1990 is driven by a negative relationship between social connectedness and log median income. Social connectedness and log median income are not significantly correlated in other decades.

[27]Appendix Table B.5 displays the relationship between log HHI and estimates of social capital, based mainly on 1990 county-level data, from Rupasingha, Goetz and Freshwater (2006). Raw correlations between log HHI and various measures of social capital are positive, but small and indistinguishable from zero. After controlling for the log number of migrants and state fixed effects, these correlations shrink even further. The social capital estimates of Rupasingha, Goetz and Freshwater (2006) depend on the density of membership organizations, voter turnout for

3.5 The Effect of Social Connectedness on Crime

3.5.1 Effects on City-Level Crime Rates

Motivated by the model in Section 3.3, we estimate the effect of social connectedness on city-level crime rates (Proposition 3.3.2) and whether this effect is stronger in cities with a higher African American population share (Proposition 3.3.2).

Table 3.4 shows that social connectedness leads to sizable and statistically significant reductions in murder, rape, robbery, assault, burglary, and motor vehicle theft. The table reports estimates of equation (3.12) for an unbalanced panel of 471 cities.[28] As seen in column 1, our estimated elasticity of the murder rate with respect to HHI is -0.181 (0.034). The estimates for robbery and motor vehicle theft, two other well-measured crimes in the FBI data, are -0.251 (0.035) and -0.163 (0.041). These results are consistent with Proposition 3.3.2.

Because social connectedness reduces crimes that are more and less likely to have witnesses, an increased probability of detection likely is not the only operative mechanism. Burglary and motor vehicle theft are less likely to have witnesses than rape, robbery, or assault, yet our estimates are roughly comparable for all of these crimes.[29] As a result, the effect of social connectedness on crime probably stems in part from other mechanisms, such as an improvement in cognitive or non-cognitive skills.

Simple examples help illustrate the sizable effects of social connectedness on crime. First, consider Middletown, Ohio and Beloit, Wisconsin. These cities are similar in their total number of Southern black migrants, 1980 population, and 1980 black population share, but Beloit's HHI is over four times as large as in Middletown (0.057 versus 0.014).[30] The estimates in Table 3.4 imply that replacing Middletown's HHI with that of Beloit would decrease murders by 25.4 percent, robberies by 35.2 percent, and motor vehicle thefts by 22.9 percent. By comparison, the estimates in Chalfin and McCrary (2015) imply that a similar decrease in murders would require a 38 percent increase in the number of police officers.[31] The effect of social connectedness is even larger in other examples. HHI in Decatur, Illinois is almost twenty times larger than that of Albany, NY (0.118

presidential elections, response rates for the decennial Census, and the number of non-profit organizations. The weak correlation between log HHI and the county-level social capital estimates is not particularly surprising, given the different time periods involved and, more importantly, the fact that these social capital estimates do not isolate social ties among African Americans.

[28]Appendix Table B.6 displays results for all covariates in the regressions.

[29]Unlike larceny or motor vehicle theft, a robbery features the use of force or threat of force. Consequently, robberies are witnessed by at least one individual (the victim).

[30]For Middletown and Beloit, the number of Southern black migrants is 376 and 407; the 1980 population is 35,207 and 43,719; and the 1980 percent black is 11.3 and 12.0.

[31]Chalfin and McCrary (2015) estimate an elasticity of murder with respect to police of -0.67, almost four times the size of our estimated elasticity of murder with respect to social connectedness.

versus 0.006).[32] Replacing Albany's HHI with that of Decatur would decrease murders by 53.9 percent, robberies by 74.8 percent, and motor vehicle thefts by 48.6 percent. While these effects are sizable, they are reasonable in light of the tremendous variation in crime rates across cities (Appendix Table B.2).

Table 3.5 demonstrates that our results are robust to various sets of control variables. We focus on the effect of social connectedness on murder, given its importance for welfare and high measurement quality, and we restrict the sample to the 228 cities observed in every decade. Our baseline specification in column 1 yields an estimate of δ of -0.244 (0.041). Estimates are very similar when excluding demographic or economic covariates (columns 2-3) and somewhat attenuated when excluding both sets of covariates or replacing state-year fixed effects with region-year fixed effects (columns 4-5). The estimate is even larger in magnitude when not controlling for the log number of migrants and is very similar when using ten indicator variables to control flexibly for the number of migrants (columns 6-7).[33] Controlling for log HHI and the log number of Southern white migrants and foreign immigrants has little impact on the estimate (column 8).[34] Results are similar when we control for the share of migrants that chose their destination because of social interactions (column 9); this variable controls for unobserved characteristics of migrants that could confound our results, as detailed below.

Table 3.6 provides some evidence that the effect of social connectedness on crime is stronger in cities with a higher African American population share. We estimate equation (3.12) separately for each tercile of cities' 1960 African American population share. Across increasing levels of the black population share, the estimated effect of HHI on murder is -0.017 (0.124), -0.085 (0.052), and -0.213 (0.051). A similar pattern exists for other crimes, including robbery and motor vehicle theft. Point estimates for the highest percent black tercile are negative and statistically significant across all crimes, while point estimates for the lowest percent black tercile are indistinguishable from zero for six out of seven crimes.[35] Moving from the 25th to 75th percentile of HHI (0.008 to 0.028) has essentially no effect on the murder rate in cities in the bottom tercile of black population share. For the middle tercile, increasing HHI across the interquartile range leads to 0.6 fewer murders per 100,000 residents, relative to a base of 5.4 murders per 100,000; the effect is 3.4 fewer murders per 100,000 residents at the highest percent black tercile, relative to a base of 12.8

[32]For Decatur and Albany, the number of Southern black migrants is 760 and 874; the 1980 population is 94,081 and 101,727; and the 1980 percent black is 14.6 and 15.9.

[33]For identification purposes, we strongly prefer to control for the log number of migrants. We estimate the regression in column 6 to demonstrate that the strong relationship between log HHI and the log number of migrants does not account for the negative coefficient on log HHI.

[34]We use country of birth to construct HHI for immigrants.

[35]However, standard errors for estimates in the lowest percent black tercile are quite large, and we cannot reject equality of coefficients in the low and high terciles for murder ($t = -1.46$) or robbery ($t = -1.42$), but can for motor vehicle theft ($t = -2.35$).

murders per 100,000. The results in Table 3.6 are consistent with Proposition 3.3.2 of the model, which predicts a stronger effect of social connectedness on city-level crime rates in cities with a higher black population share because a higher share of individuals in these cities have social ties to African Americans from the South.

3.5.2 EFFECTS OVER TIME

Table 3.7 shows that the effect of social connectedness on crime is generally smaller in magnitude from 2000-2009 relative to 1960-1999. We estimate equation (3.12) separately for each decade.[36] Focusing on the best measured crimes of murder, robbery, and motor vehicle theft, we see significant negative effects of social connectedness in each decade from 1960-1999, and much smaller and insignificant effects from 2000-2009.

One possible explanation for the attenuated effects from 2000-2009 is a decline in the effective strength of social connectedness over time. Reductions in crime in 1960 were likely driven by individuals who were born around 1940 to mothers born around 1915.[37] More generally, the individuals most affected by social connectedness were likely the children and grandchildren of post-war migrants and the grandchildren or great-grandchildren of the earliest group of migrants. As a result, the crime-reducing effect of social connectedness might have declined as the original migrants died. A second possible explanation is that individuals committing crime in the 2000's, when crime rates were relatively low (see Figure 3.3), were inframarginal and not affected by social connectedness.

The attenuated effects from 2000-2009 appear to reflect a decline in the effective strength of social connectedness, as opposed to an interaction between the level of crime and the effect of social connectedness. Figure 3.5 shows that fewer black children had ties to the South from 2000-2009 compared to previous decades. We characterize individuals age 14-17 who are living in the North, Midwest, or West regions as having a tie to the South if they or an adult in their household were born in the South. The share of black children with ties to the South declines from 67 percent in 1980 to 33 percent in 2000 and 20 percent in 2010. We also examine whether the effect of social connectedness from 2000-2009 differs across cities with higher and lower predicted crime rates. In particular, we estimate equation (3.12) using data from 1995-1999 and use the coefficients from this regression to predict cities' crime rates from 2000-2009 based on their economic and demographic covariates.[38] There is little evidence of a negative effect of social connectedness

[36] To ensure that our results are not driven by changes in the sample over time, we limit the sample in Table 3.7 to cities that appear in at least five years of every decade.

[37] The highest offending rate for murder is between ages 18-24 (Fox, 2000).

[38] We include $\ln(\text{HHI}_k)$ and $\ln(N_k)$ in the 1995-1999 regression, but replace these variables with their mean when constructing predicted crime rates. We also use state-specific linear trends in place of state-by-year fixed effects for these regressions.

from 2000-2009 even for the cities with higher predicted crime rates (Appendix Table B.7).

Figure 3.4 plots the evolution of crime rates from 1960-2009 for two hypothetical cities with HHI at the 75th and 25th percentiles and average values of other covariates. Crime rates rose much more slowly from 1960-1990 in cities with higher social connectedness. Crime rates for cities with high and low social connectedness converged after 1990. Adding up the effect of social connectedness on crime rates from 1960-2009 implies that the city with HHI at the 75th percentile had 139 fewer murders and 10,822 fewer motor vehicle thefts per 100,000 residents over this period.

3.5.3 Effects by Age and Race of Offender over Time

Table 3.8 shows that social connectedness leads to particularly large reductions in murders committed by black youth. From 1980-1989, the elasticity of murders committed by black youth with respect to social connectedness is -0.761 (0.175), almost four times the size of the elasticity of murders committed by non-black youth.[39] The effect of social connectedness on murders committed by black youth declines over time, consistent with the decline in social ties seen in Figure 3.5. The effect of social connectedness on murders committed by black adults declines more slowly over time, consistent with social connectedness having persistent effects on cohorts. Peer effects provide a natural explanation for the reduction in crime among non-blacks, as described in our model.

3.5.4 Threats to Empirical Strategy and Additional Robustness Checks

A key potential threat to our empirical strategy is that cities with higher social connectedness had lower unobserved determinants of crime, $\epsilon_{k,t}$. For example, if migrants from the same birth town moved to cities with low unobserved determinants of crime, and these unobserved characteristics persisted over time, then our estimate of δ could be biased downwards. We have already presented indirect evidence against this threat by showing that log HHI is not correlated with homicide rates from 1911-1916 (Table 3.1) or most demographic and economic covariates from 1960-2009 (Table 3.3).

To provide more direct evidence against this threat, we estimate the effect of social connectedness on crime for each five-year interval from 1965-2009 while controlling for the log average crime rate from 1960-1964.[40] Figure 3.6 shows that the effect of social connectedness on murder is nearly identical when controlling for the 1960-1964 crime rate. These results directly rule out the possibility that our estimates are driven by a persistent correlation between HHI and unobserved

[39] FBI data provide the age, race, and sex of offenders for crimes resulting in arrest starting in 1980.

[40] Controlling for the average log crime rate is unattractive because many cities report zero murders in a given year.

determinants of crime from 1960-forward.[41]

Another possible concern is that HHI reflects unobserved characteristics of migrants who chose the same destination as other individuals from their birth town. Census data show that Southern black migrants living in a state or metropolitan area with a higher share of migrants from their birth state have less education and income (Appendix Table B.8). As a result, migrants who followed their birth town network likely had less education and earnings capacity than other migrants. This negative selection in terms of education and earnings could generate a positive correlation between HHI_k and $\epsilon_{k,t}$, making it more difficult for us to estimate a negative effect of social connectedness on crime. At the same time, migrants who followed their birth town network might have displayed greater cooperation or other "pro-social" behaviors. To address this possibility, we estimate a structural model of social interactions in location decisions. As described in Appendix B.2, the model allows us to estimate the share of migrants in each destination that moved there because of social interactions. When used as a covariate in equation (3.12), this variable proxies for unobserved characteristics of migrants that chose to follow other migrants from their birth town. Column 9 of Table 3.5 shows that the estimated effect of social connectedness on murder barely changes when we control for the share of migrants that chose their destination because of social interactions.[42] Consequently, our results appear to reflect the effect of social connectedness per se, as opposed to unobserved characteristics of certain migrants.

Appendix Table B.9 shows that our results are robust to including the 14 largest cities that are excluded from the main analysis, estimating negative binomial models, dropping outliers of the dependent variable, and measuring HHI using birth county to destination county population flows.[43]

3.6 UNDERSTANDING THE ROLE OF PEER EFFECTS

We now use the model in Section 3.3 to examine the role of peer effects in mediating the relationship between social connectedness and city-level crime rates. The model connects the total effect of HHI on city-level crime, δ, to the effect of HHI on crime for blacks with ties to the South and peer effects. In particular, equations (3.7)-(3.10) imply that the elasticity of the city-level crime rate with respect to Southern black HHI, δ, can be written

$$\delta = \varepsilon^s r^s \left[P^b (P^{s|b} m^s + (1 - P^{s|b}) m^n) + (1 - P^b) m^w \right], \tag{3.14}$$

[41] The similarity of the results in Figure 3.6 is not driven by a weak relationship between the log average crime rate from 1960-1964 and crime rates from 1965-forward.

[42] Results are nearly identical when we use quadratic, cubic, or quartics in this variable.

[43] We prefer equation (3.12) over the negative binomial specification because it requires fewer assumptions to generate consistent estimates of δ (e.g., Wooldridge, 2002).

where $\delta \equiv (d\bar{C}/d\text{HHI}^s)(\text{HHI}^s/\bar{C})$ is the parameter of interest in our regressions, $\varepsilon^s \equiv (\partial F^s/\partial \text{HHI}^s)$ (HHI^s/F^s) captures the direct effect of HHI on the crime rate of blacks with ties to the South, $r^s \equiv \bar{C}^s/\bar{C}$ is the ratio of the crime rate among blacks with ties to the South to the overall crime rate, P^b is the black population share, $P^{s|b}$ is the share of blacks with ties to the South, and m^s, m^n, and m^w are peer effect multipliers defined in equations (3.7)-(3.10).

We use equation (3.14) to examine which direct effect (ε^s) and peer effect (m^s, m^n, m^w) parametrizations are consistent with our central estimate of δ for murder. We set the black population share $P^b = 0.13$ and the share of the black population with ties to the South $P^{s|b} = 0.67$.[44] We do not observe the crime rate among blacks with ties to the South. In the FBI data, half of the murders resulting in arrest are attributed to African Americans. If crime rates are equal among blacks with and without ties to the South, then $r^s = 3.8$.[45]

We make several simplifying assumptions about peer effects. First, we assume that own-group peer effects are equal across all three groups.[46] Second, we assume that cross-group peer effects between non-blacks and both groups of African Americans are equal. Third, we assume that cross-group peer effects are symmetric in terms of elasticities.[47] The first assumption implies that $J_{11} = J_{22} = J_{33}$, and the second implies that $J_{12} = J_{21}$, $J_{13} = J_{23}$, and $J_{31} = J_{32}$. Letting E_{ab} denote the elasticity form of J_{ab}, these three assumptions imply that $E_{11} = E_{22} = E_{33}$, $E_{12} = E_{21}$, and $E_{13} = E_{23} = E_{31} = E_{32}$.

We draw on previous empirical work to guide our parametrization of peer effects. As detailed in Appendix B.3, the literature suggests on-diagonal values of J (own-group peer effects) between 0 and 0.5 and off-diagonal values of J (cross-group peer effects) near zero (Case and Katz, 1991; Glaeser, Sacerdote and Scheinkman, 1996; Ludwig and Kling, 2007; Damm and Dustmann, 2014).[48] We consider on-diagonal values of J of 0, 0.25, and 0.5. We allow for sizable peer effects between African Americans with and without ties to the South, and we parametrize the cross-race effects so that elasticities equal 0 or 0.1. Given values of $(r^s, P^b, P^{s|b}, m^s, m^n, m^w)$ and our estimate of δ, equation (3.14) yields a unique value for ε^s. Equations (3.7)-(3.9) then allow us to to solve for the effect of a change in Southern black HHI on crime rates for each group.[49]

[44]The black population share in our sample is 0.13 in 1980. As seen in Figure 3.5, the share of African American youth living in the North with ties to the South is 0.67.

[45]If crime rates are equal among blacks with and without ties to the South, then $\bar{C}^s = \bar{C}^b$, where $\bar{C}^b \equiv C^b/N^b$ is the crime rate among all blacks. As a result, $r^s = (C^b/N^b)/(C/N) = (C^b/C)/(N^b/N) = 0.5/0.13$, where C and N are the total number of crimes and individuals. To the extent that blacks with ties to the South commit less crime than blacks without ties to the South, we will overstate r^s and understate the direct effect, ε^s.

[46]We are aware of no evidence suggesting that own-group peer effects differ for black versus non-black youth.

[47]Given the differences in crime rates between blacks and non-blacks, we believe that assuming symmetric cross-group elasticities is more appropriate than assuming symmetric cross-group linear effects (J).

[48]Estimates from previous work are valuable, but are not necessarily comparable to each other or our setting, as they rely on different contexts, identification strategies, data sources, and crime definitions.

[49]In particular: $(d\bar{C}^s/d\text{HHI}^s)(\text{HHI}^s/\bar{C}^s) = \varepsilon^s m^s$, $(d\bar{C}^n/d\text{HHI}^s)(\text{HHI}^s/\bar{C}^n) = \varepsilon^s m^n(\bar{C}^s/\bar{C}^n)$, and $(d\bar{C}^w/d\text{HHI}^s)(\text{HHI}^s/\bar{C}^w) = \varepsilon^s m^w(\bar{C}^s/\bar{C}^w)$. Our assumption that crime rates are equal among blacks with and

Table 3.9 maps the estimated effect of social connectedness on the city-level murder rate, $\hat{\delta}$, to the effect on murder rates of various groups under different peer effect parametrizations.[50] We consider a one standard deviation increase in HHI, equal to 0.78, which decreases the total murder rate by 14.1 percent according to the estimate in Table 3.4. This implies a decrease in the murder rate of blacks with ties to the South between 42.2 percent, when there are no cross-group peer effects (column 1), and 21.2 percent, when peer effects operate across all groups (column 7). The murder rate of blacks without ties to the South decreases by 0-24.2 percent, while the murder rate of non-blacks decreases by 0-8.0 percent. Depending on the parametrization, up to 82 percent of the effect on blacks with ties to the South is driven by peer effects. The existing evidence on peer effects suggests placing the most emphasis on columns 3 and 4, which imply that a one standard deviation increase in HHI reduces the murder rate of African Americans with ties to the South by 37.3 and 30.1 percent and reduces the murder rate of African Americans without ties to the South by 9.9 and 8.7 percent.[51] In columns 3 and 4, peer effects account for 30.2 and 32.6 percent of the effect on blacks with ties to the South. Peer effects clearly could play an important role in amplifying the effect of social connectedness on crime.

3.7 Conclusion

This paper estimates the effect of social connectedness on crime across U.S. cities from 1960-2009. We use a new source of variation in social connectedness stemming from social interactions in the migration of millions of African Americans out of the South. A one standard deviation increase in social connectedness leads to a precisely estimated 14 percent decrease in murder. We find that social connectedness also leads to sizable reductions in rapes, robberies, assaults, burglaries, and motor vehicle thefts. As predicted by our economic model, social connectedness leads to greater reductions in the city-level crime rate in cities with a higher African American population share. Social connectedness reduces crimes that are more and less likely to have witnesses, which suggests that an increased probability of detection is not the only mechanism through which social connectedness reduces crime.

Our results highlight the importance of birth town level social ties in reducing violent and property crimes in U.S. cities. In principle, similar social ties among immigrants could reduce crime and generate other desirable outcomes. While the benefits of these social ties must be weighed against any possible offsetting effects (e.g., on assimilation), the characteristics of social networks

without ties to the South implies that $\bar{C}^s/\bar{C}^n = 1$. The same assumption, combined with the fact that half of murders are attributed to blacks in the UCR data, implies that $\bar{C}^s/\bar{C}^w = (1 - P^b)/P^b = 6.69$.

[50] Under all peer effect parametrizations in Table 3.9, the equilibrium is stable, and Propositions 3.3.2 and 3.3.2 are true.

[51] The results in Table 3.8, which show significant effects of social connectedness on non-black crime, suggest sizable peer effects between non-blacks and blacks.

could prove valuable in achieving difficult economic and social milestones.

In future work, we plan to use our new source of variation in social connectedness to study its effects on a variety of other economic outcomes, such as schooling, employment, marriage, and fertility. Evidence on these effects is of independent interest and would improve our understanding of the negative effects on crime documented in this paper.

Table 3.1: The Relationship between Social Connectedness and 1911-1916 Homicide Rates

	Dependent variable: Log HHI, Southern black migrants			
	(1)	(2)	(3)	(4)
Log mean homicide rate, 1911-1916	0.010	0.073	0.050	-0.012
	(0.147)	(0.101)	(0.216)	(0.088)
p-value	[0.948]	[0.476]	[0.817]	[0.896]
		(0.055)		(0.043)
Log number, Southern black migrants		x		x
Inverse probability weighted			x	x
R2	0.00	0.43	0.00	0.67
N (cities)	46	46	46	46

Notes: The sample contains cities in the North, Midwest, and West Census regions with at least 100,000 residents in 1920. We exclude homicide rates based on less than five deaths in constructing the mean homicide rate from 1911-1916. In columns 3-4, we use inverse probability weights (IPWs) because the sample of cities for which we observe homicide rates from 1911-1916 differs on various characteristics from our main analysis sample. We construct IPWs using fitted values from a logit model, where the dependent variable is an indicator for a city having homicide rate data for at least one year from 1911-1916, and the explanatory variables are log population, percent black, percent female, percent with a high school degree or more, percent with a college degree or more, log land area, log median family income, unemployment rate, labor force participation rate, and manufacturing employment share, all measured in 1980. Unlike our main analysis sample, we do not restrict the sample to cities with less than 500,000 residents in 1980. Heteroskedastic-robust standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$
Sources: Census (1922, p. 64-65) , Duke SSA/Medicare data, Census city data book

Table 3.2: Five-Year Migration Rates, Southern Black Migrants Living Outside of the South

	1955-1960 (1)	1965-1970 (2)	1975-1980 (3)	1985-1990 (4)	1995-2000 (5)
Percent living in same state	93.1	95.5	96.2	96.0	95.9
Same county	86.4	90.4	93.8	77.2	93.8
Same house	33.0	54.0	72.8	77.2	79.1
Different house	53.4	36.4	21.0	-	14.7
Different county	-	4.3	2.4	-	2.1
Unknown	6.7	0.8	-	18.8	-
Percent living in different state	6.9	4.5	3.8	4.0	4.1
Not in South	4.0	2.8	1.4	1.2	1.0
In South	2.9	1.6	2.4	2.9	3.1

Notes: Sample restricted to African Americans who were born in the South from 1916-1936 and were living in the North, Midwest, or West regions five years prior to the census year. For 2000, column 3 equals the percent living in the same PUMA.
Sources: Census IPUMS, 1960-2000

Table 3.3: The Relationship between Social Connectedness and City Covariates, 1960-2000

	Dependent variable: Log HHI, Southern black migrants					
Year covariates are measured:	-	1960	1970	1980	1990	2000
	(1)	(2)	(3)	(4)	(5)	(6)
Log number, Southern	-0.839***	-0.834***	-0.834***	-0.813***	-0.727***	-0.737***
black migrants	(0.040)	(0.066)	(0.072)	(0.078)	(0.082)	(0.072)
Log population		0.013	-0.009	-0.020	-0.065	0.006
		(0.062)	(0.067)	(0.075)	(0.085)	(0.083)
Percent black		0.011	-0.013	-0.005	-0.059	-0.063
		(0.053)	(0.060)	(0.075)	(0.067)	(0.058)
Percent female		0.017	-0.036	-0.004	-0.011	-0.013
		(0.047)	(0.058)	(0.076)	(0.077)	(0.055)
Percent age 5-17		-0.131	0.089	0.161	0.557**	0.324
		(0.151)	(0.204)	(0.242)	(0.248)	(0.292)
Percent age 18-64		-0.117	0.044	0.164	0.586**	0.499
		(0.122)	(0.211)	(0.250)	(0.260)	(0.319)
Percent age 65+		-0.029	0.109	0.236	0.521***	0.393*
		(0.094)	(0.146)	(0.198)	(0.187)	(0.200)
Percent with high school degree		-0.052	-0.065	-0.178*	-0.037	-0.046
		(0.115)	(0.117)	(0.096)	(0.076)	(0.079)
Percent with college degree		0.149**	0.101	0.076	0.118*	0.047
		(0.073)	(0.064)	(0.051)	(0.064)	(0.063)
Log area, square miles		-0.028	0.021	0.022	0.031	-0.021
		(0.049)	(0.060)	(0.065)	(0.073)	(0.078)
Log median family income		-0.032	-0.028	-0.002	-0.238***	-0.070
		(0.085)	(0.084)	(0.089)	(0.089)	(0.065)
Unemployment rate		0.115*	0.147*	0.027	0.001	0.057
		(0.060)	(0.079)	(0.070)	(0.079)	(0.060)
Labor force participation rate		0.024	0.085	0.017	0.106	-0.047
		(0.025)	(0.052)	(0.091)	(0.100)	(0.051)
Manufacturing employment		0.225***	0.166***	0.142**	0.162***	0.190***
share		(0.058)	(0.061)	(0.055)	(0.048)	(0.045)
State fixed effects	x	x	x	x	x	x
Adjusted R2	0.742	0.769	0.763	0.756	0.762	0.769
N (cities)	228	228	228	228	228	228
p-value: Wald test that parameters equal zero						
Demographic covariates		0.239	0.631	0.280	0.022	0.001
Economic covariates		0.121	0.104	0.983	0.012	0.066

Notes: Sample restricted to cities with less than 500,000 residents in 1980. We normalize all variables, separately for each regression, to have mean zero and standard deviation one. For the Wald tests, demographic covariates include log population, percent black, percent female, percent age 5-17, percent age 18-64, percent age 65+, percent with high school degree, percent with college degree, and log area. Economic covariates include log median family income, unemployment rate, and labor force participation rate (but not manufacturing employment share). Heteroskedastic-robust standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Sources: Duke SSA/Medicare data, Census city data book

Table 3.4: The Effect of Social Connectedness on Crime, 1960-2009

	Dependent variable: Number of offenses reported to police						
	Murder (1)	Rape (2)	Robbery (3)	Assault (4)	Burglary (5)	Larceny (6)	Motor Vehicle Theft (7)
Log HHI, Southern black migrants	-0.181***	-0.083**	-0.251***	-0.142***	-0.095***	-0.049	-0.163***
	(0.034)	(0.035)	(0.035)	(0.042)	(0.022)	(0.030)	(0.041)
Log number, Southern black migrants	x	x	x	x	x	x	x
Demographic covariates	x	x	x	x	x	x	x
Economic covariates	x	x	x	x	x	x	x
State-year fixed effects	x	x	x	x	x	x	x
Pseudo R2	0.773	0.838	0.931	0.913	0.938	0.926	0.906
N (city-years)	18,854	17,690	18,854	18,854	18,854	18,854	18,854
Cities	471	471	471	471	471	471	471

Notes: Table displays estimates of equation (3.12). Sample restricted to cities with less than 500,000 residents in 1980. Demographic covariates include log population, percent black, percent age 5-17, percent age 18-54, percent 65+, percent female, percent with high school degree, percent with college degree, and log area. Economic covariates include log median family income, unemployment rate, labor force participation rate, and manufacturing employment share. Standard errors, clustered at the city level, are in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Sources: FBI UCR, Duke SSA/Medicare data, Census city data book

Table 3.5: The Effect of Social Connectedness on Murder, 1960-2009, Robustness

	Dependent variable: Number of murders reported to police								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log HHI, Southern black migrants	-0.244***	-0.269***	-0.228***	-0.163**	-0.157***	-0.342***	-0.222***	-0.234***	-0.278***
	(0.041)	(0.044)	(0.046)	(0.073)	(0.054)	(0.042)	(0.045)	(0.044)	(0.053)
Log number, Southern black migrants	x	x	x	x	x			x	x
Demographic covariates	x		x		x	x	x	x	x
Economic covariates	x	x			x	x	x	x	x
State-year fixed effects	x	x	x	x		x	x	x	x
Region-year fixed effects					x				
Indicators for number of Southern black migrants							x		
Log HHI, Southern white migrants								x	
Log number, Southern white migrants								x	
Log HHI, immigrants								x	
Log number, immigrants								x	
Share of Southern black migrants influenced by social interactions									x
Pseudo R2	0.805	0.796	0.801	0.764	0.787	0.803	0.805	0.805	0.805
N (city-years)	11,284	11,284	11,284	11,284	11,284	11,284	11,284	11,284	11,284
Cities	228	228	228	228	228	228	228	228	228

Notes: Table displays estimates of equation (3.12). Sample restricted to cities with less than 500,000 residents in 1980 that also are observed in every decade from 1960-2000. Demographic covariates include log population, percent black, percent age 5-17, 18-64, and 65+, percent female, percent of population at least 25 years old with a high school degree, percent of population at least 25 years old with a college degree, and log of area in square miles. Economic covariates include log median family income, unemployment rate, labor force participation rate, and manufacturing employment share. Indicators for the number of Southern black migrants correspond to deciles. Column 9 includes an estimate of the share of migrants that chose their destination because of social interactions. We estimate this variable using a structural model of social interactions in location decisions, as described in the text. Standard errors, clustered at the city level, are in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Sources: FBI UCR, Duke SSA/Medicare data, Census city data book

Table 3.6: The Effect of Social Connectedness on Crime, 1960-2009, by Percent Black Tercile

	Dependent variable: Number of offenses reported to police						
	Murder	Rape	Robbery	Assault	Burglary	Larceny	Motor Vehicle Theft
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coefficient on Log HHI, Southern Black Migrants by Percent Black Tercile							
Low	-0.017	-0.118	-0.062	-0.184	-0.067	-0.154*	0.072
	(0.124)	(0.157)	(0.136)	(0.120)	(0.083)	(0.092)	(0.150)
Medium	-0.085	0.053	-0.091	-0.051	-0.043	-0.006	-0.056
	(0.052)	(0.067)	(0.072)	(0.067)	(0.043)	(0.047)	(0.071)
High	-0.213***	-0.195***	-0.264***	-0.280***	-0.117***	-0.147**	-0.304***
	(0.051)	(0.066)	(0.040)	(0.073)	(0.032)	(0.057)	(0.056)

Notes: Table displays estimates of equation (3.12). Sample restricted to cities with less than 500,000 residents in 1980. Regressions include the same covariates used in Table 3.4. Percent black is measured in 1960, and the tercile cutoffs are 0.022 and 0.075. Standard errors, clustered at the city level, are in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$
Sources: FBI UCR, Duke SSA/Medicare data, Census city data book

Table 3.7: The Effect of Social Connectedness on Crime, 1960-2009, by Decade

	Dependent variable: Number of offenses reported to police						
	Murder	Rape	Robbery	Assault	Burglary	Larceny	Motor Vehicle Theft
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coefficient on Log HHI, Southern Black Migrants by Decade							
1960-69	-0.121**	-0.313***	-0.368***	-0.265***	-0.145***	-0.087	-0.198**
	(0.062)	(0.112)	(0.082)	(0.098)	(0.054)	(0.064)	(0.078)
1970-79	-0.273***	-0.220***	-0.327***	-0.179**	-0.133***	-0.033	-0.219***
	(0.055)	(0.046)	(0.057)	(0.082)	(0.031)	(0.045)	(0.067)
1980-89	-0.313***	-0.181***	-0.374***	-0.099	-0.174***	-0.089	-0.307***
	(0.050)	(0.057)	(0.059)	(0.075)	(0.033)	(0.059)	(0.074)
1990-99	-0.285***	-0.068	-0.300***	-0.150***	-0.116***	-0.064	-0.277***
	(0.080)	(0.064)	(0.058)	(0.054)	(0.040)	(0.046)	(0.076)
2000-09	-0.059	0.127**	-0.089	-0.129**	-0.039	-0.033	-0.038
	(0.062)	(0.061)	(0.058)	(0.059)	(0.043)	(0.041)	(0.067)

Notes: Table displays estimates of equation (3.12). Sample contains 240 cities that have less than 500,000 residents in 1980 and appear in at least five years of every decade from 1960-2009. Regressions include the same covariates used in Table 3.4. Standard errors, clustered at the city level, are in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$
Sources: FBI UCR, Duke SSA/Medicare data, Census city data book

Table 3.8: The Effect of Social Connectedness on Murder, 1980-2009, by Age-Race Group and Decade

	Dependent variable: Number of murders resulting in arrest for age-race group				
	All	Black Youth	Black Adults	Non-Black Youth	Non-Black Adults
	(1)	(2)	(3)	(4)	(5)
Coefficient on Log HHI, Southern Black Migrants by Decade					
1980-89	-0.210***	-0.761***	-0.355***	-0.200	-0.162
	(0.069)	(0.175)	(0.078)	(0.203)	(0.089)
1990-99	-0.224***	-0.305***	-0.247**	-0.458***	-0.278***
	(0.084)	(0.118)	(0.098)	(0.176)	(0.101)
2000-09	-0.148	-0.195	-0.086	-0.297	-0.227*
	(0.102)	(0.200)	(0.121)	(0.271)	(0.120)

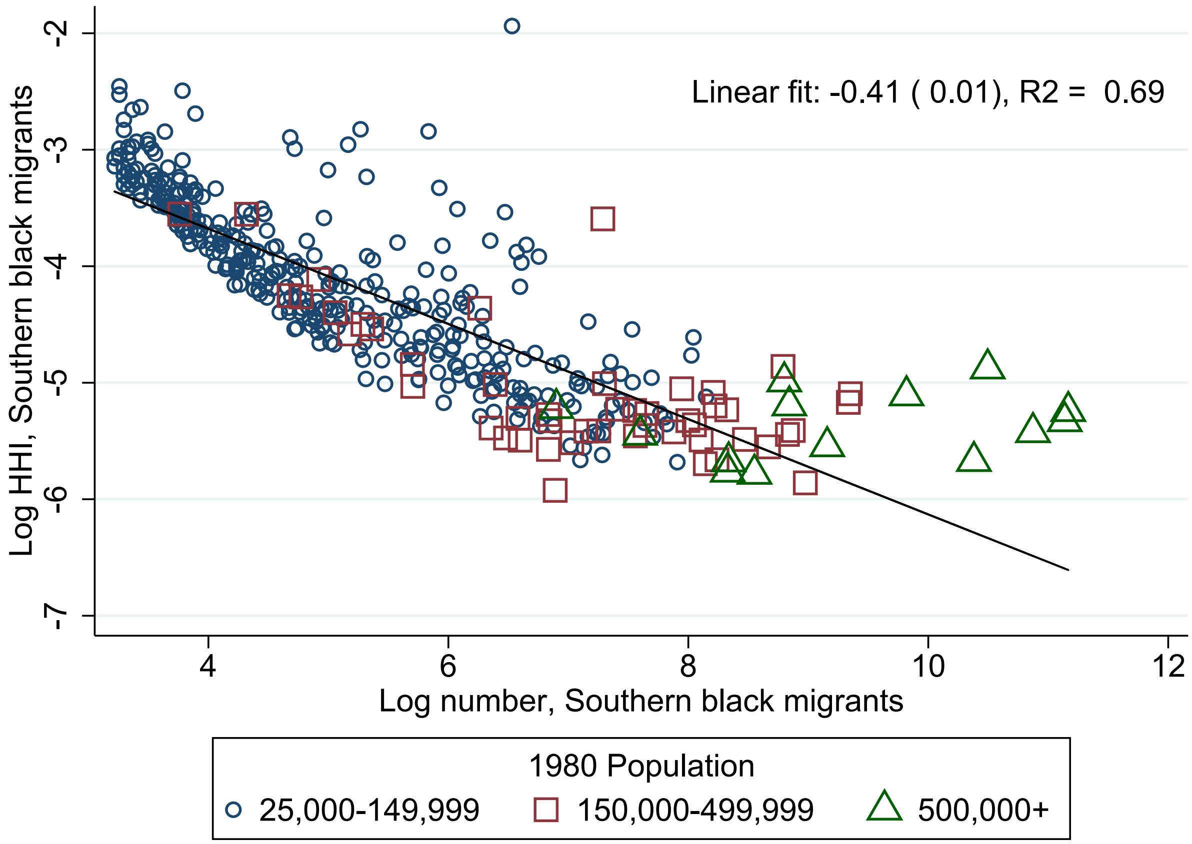
Notes: Table displays estimates of equation (3.12). Sample contains 298 cities that have less than 500,000 residents in 1980 and appear in at least five years of every decade from 1980-2009. Regressions include the same covariates used in Table 3.4. Standard errors, clustered at the city level, are in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$
Sources: FBI UCR, Duke SSA/Medicare data, Census city data book

Table 3.9: The Role of Peer Effects in the Effect of Social Connectedness on Crime

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Peer effect parametrization							
$J_{11} = J_{22} = J_{33}$ (own-group)	0	0.25	0.25	0.25	0.5	0.5	0.5
$J_{12} = J_{21}$ (cross-group, black)	0	0	0.2	0.2	0	0.4	0.4
$J_{13} = J_{23}$ (cross-race, non-black on black)	0	0	0	0.67	0	0	0.67
$J_{31} = J_{32}$ (cross-race, black on non-black)	0	0	0	0.015	0	0	0.015
Implied peer effect elasticities							
$E_{11} = E_{22} = E_{33}$ (own-group)	0	0.25	0.25	0.25	0.5	0.5	0.5
$E_{12} = E_{21}$ (cross-group, black)	0	0	0.2	0.2	0	0.4	0.4
$E_{13} = E_{23}$ (cross-race, non-black on black)	0	0	0	0.1	0	0	0.1
$E_{31} = E_{32}$ (cross-race, black on non-black)	0	0	0	0.1	0	0	0.1
Implied peer effect multipliers							
m^s (blacks with ties to South)	1	1.33	1.44	1.48	2	5.56	8.92
m^n (blacks without ties to South)	0	0	0.38	0.43	0	4.44	7.81
m^w (non-black)	0	0	0	0.04	0	0	0.50
Percent change in murder rate due to one standard deviation increase in HHI, Southern Black Migrants							
City-level murder rate	-14.1	-14.1	-14.1	-14.1	-14.1	-14.1	-14.1
Murder rate among non-blacks	0	0	0	-5.2	0	0	-8.0
Murder rate among blacks	-28.3	-28.3	-28.3	-23.1	-28.3	-28.3	-20.3
Among blacks without ties to South	0	0	-9.9	-8.7	0	-24.2	-18.5
Among blacks with ties to South	-42.2	-42.2	-37.3	-30.1	-42.2	-30.3	-21.2
Direct effect of HHI	-42.2	-31.6	-26.0	-20.3	-21.1	-5.4	-2.4
Peer effect	0	-10.5	-11.3	-9.8	-21.1	-24.8	-18.8

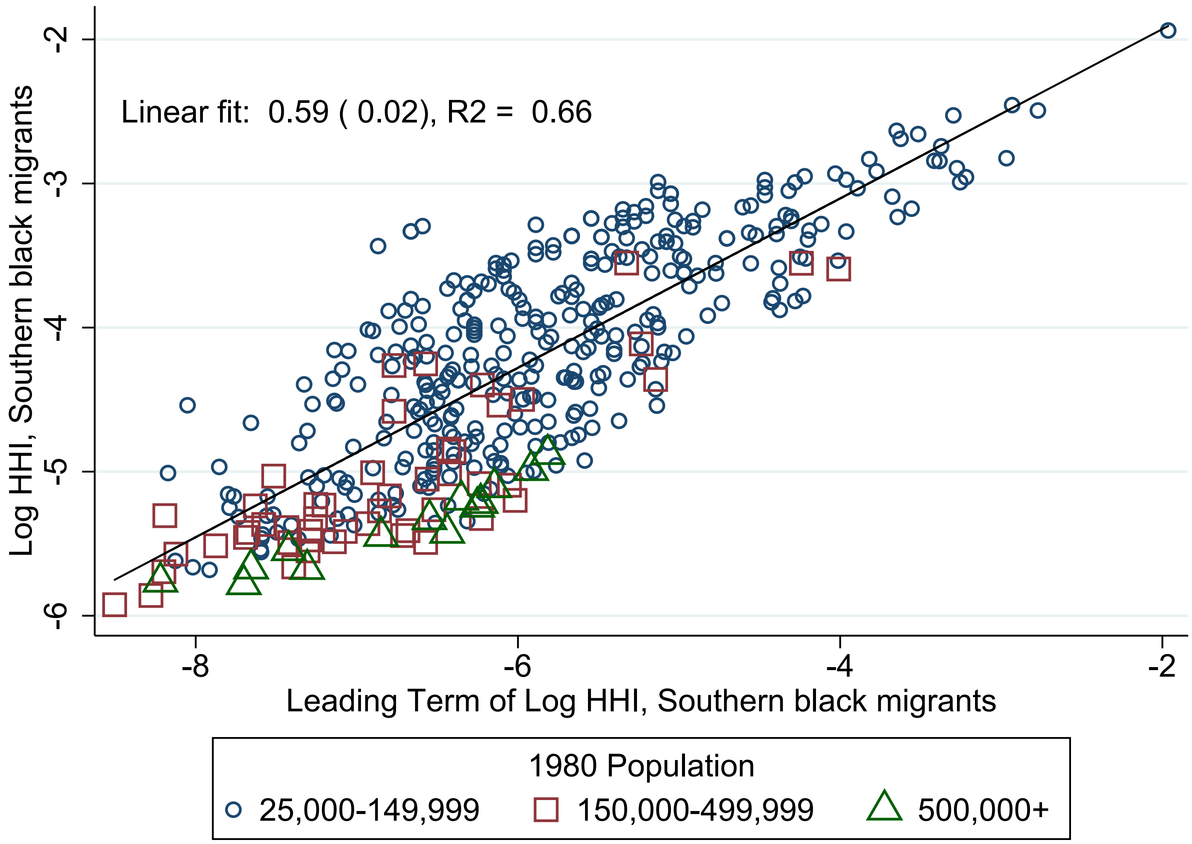
Notes: The top half of Table 3.9 describes the peer effect parametrizations that we consider. The bottom half decomposes the effect of a one standard deviation increase in social connectedness into changes in murder rates among different groups. See text for details.
Sources: FBI UCR, Duke SSA/Medicare data, Census city data book

Figure 3.1: The Relationship between Social Connectedness and the Number of Southern Black Migrants



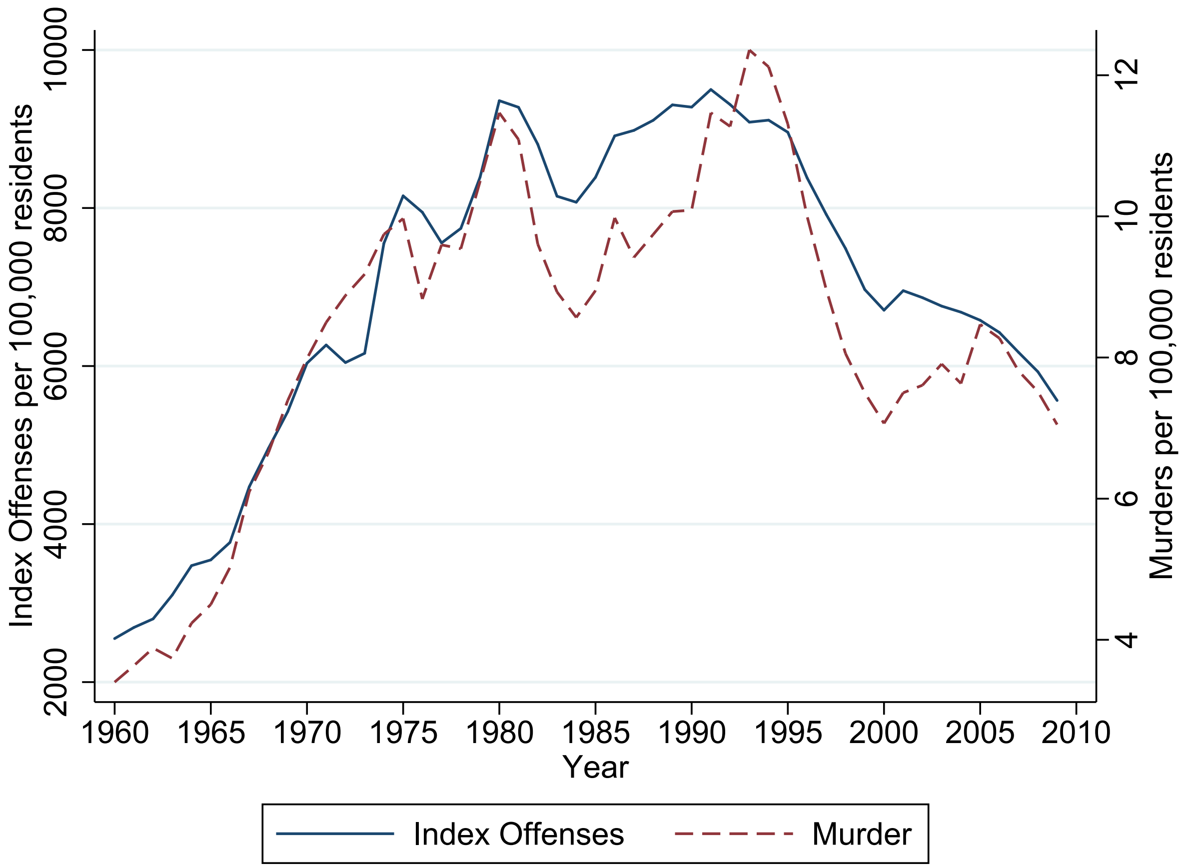
Notes: Figure contains 418 cities. Our main analysis sample excludes the 14 cities with at least 500,000 residents in 1980.
Source: Duke SSA/Medicare data

Figure 3.2: The Top Sending Town Accounts for Most of the Variation in Social Connectedness



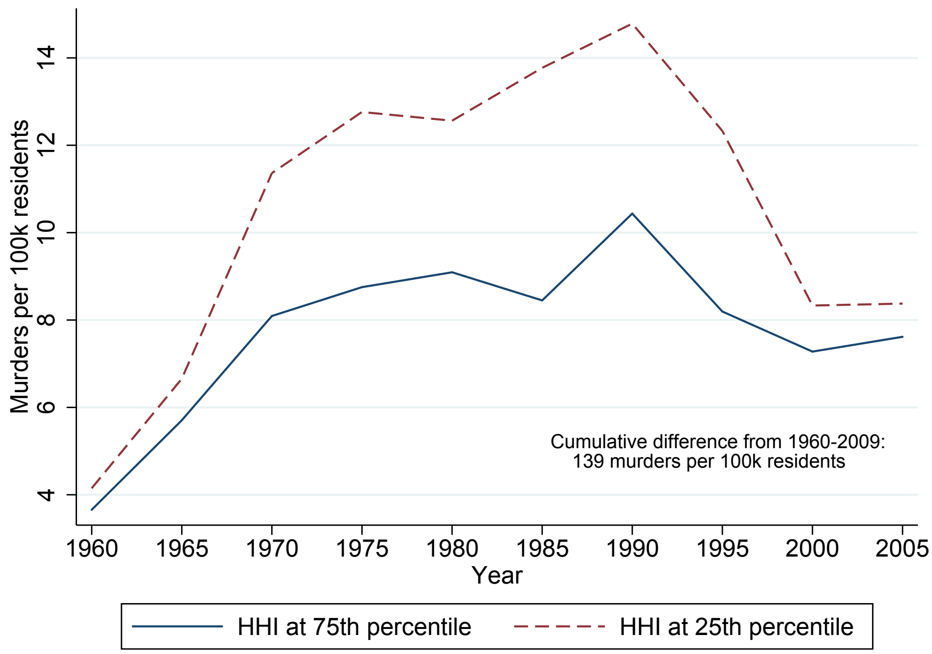
Notes: The leading term of HHI equals the log squared percent of migrants from the top sending town. Figure contains 418 cities. Our main analysis sample excludes the 14 cities with at least 500,000 residents in 1980.
Source: Duke SSA/Medicare data

Figure 3.3: The Evolution of Crime Rates Over Time

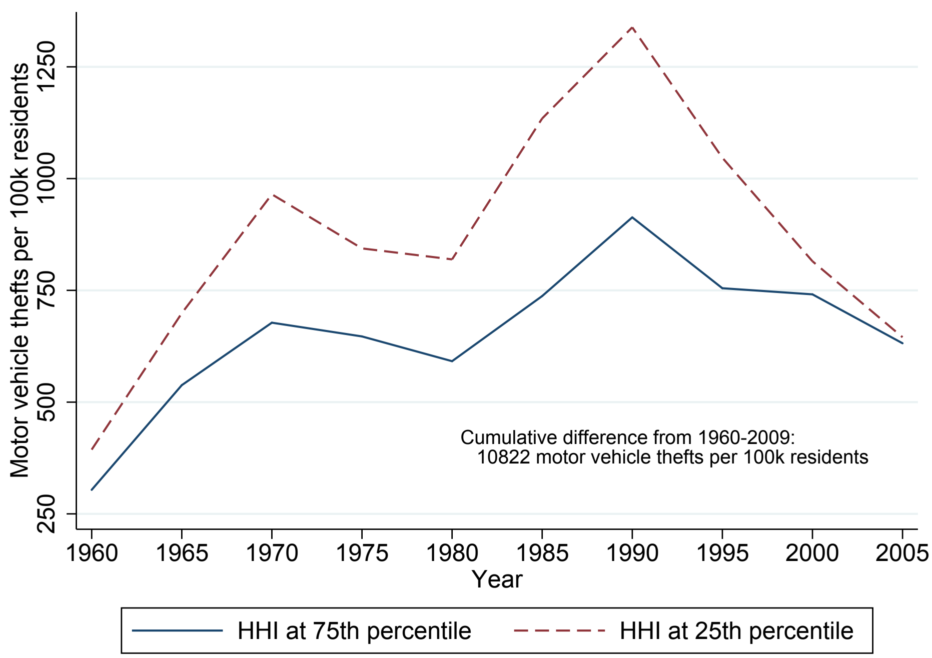


Notes: Index offenses include murder, rape, robbery, aggravated assault, burglary, larceny theft, and motor vehicle theft. Sample restricted to cities in our main analysis sample with less than 500,000 residents in 1980.
Source: FBI UCR

Figure 3.4: Social Connectedness and the Evolution of Crime Rates Over Time



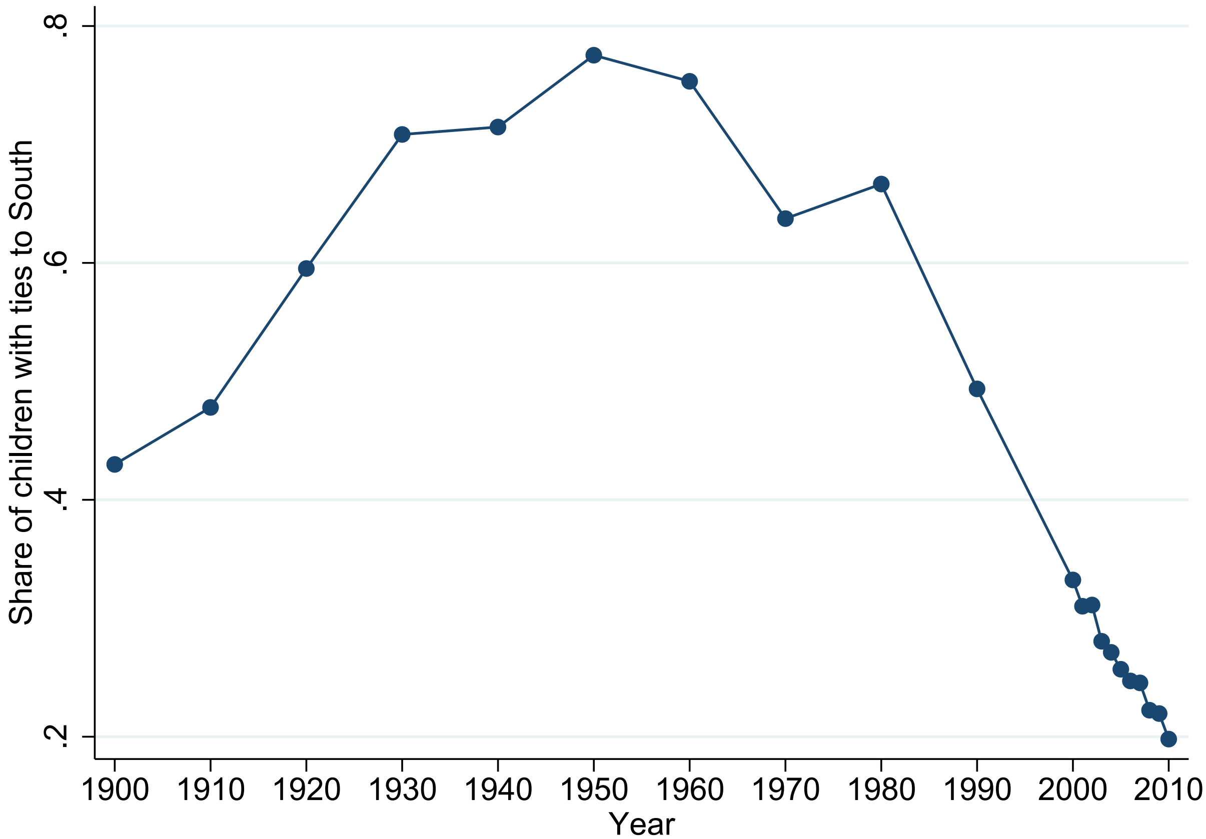
(a) Murder



(b) Motor Vehicle Theft

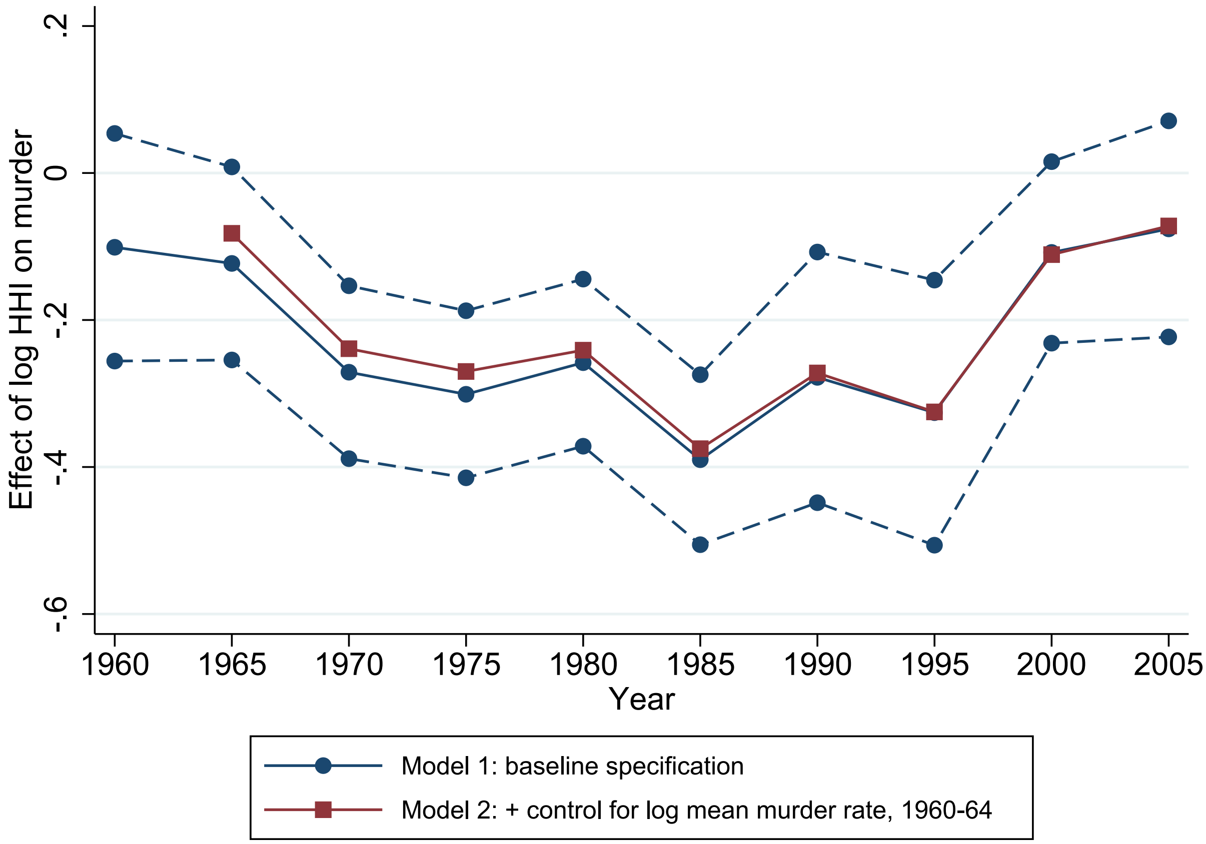
Notes: For each five year period from 1960-2009, we estimate equation (3.12) and take the level of covariates associated with the average crime rate. We then plot the murder rate associated with the 75th and 25th percentiles of HHI. Sources: FBI UCR, Duke SSA/Medicare data, Census city data book

Figure 3.5: The Share of African American Children Living in the North with Ties to the South



Notes: Figure plots the share of individuals age 14-17 who are living in the North, Midwest, or West regions who were born in the South or live in the same household as an adult born in the South.
Sources: IPUMS Decennial Census (1900-2000) and American Community Survey (2001-2010)

Figure 3.6: The Effect of Social Connectedness on Murder, Robustness to Controlling for 1960-1964 Murder Rate



Notes: Figure shows point estimates and 95-percent confidence intervals from estimating equation (3.12) separately for year 1960-64, 1965-69, and so on. Model 1 includes the same covariates used in Table 3.4, and model 2 additionally controls for the log mean murder rate from 1960-64.
Sources: FBI UCR, Duke SSA/Medicare data, Census city data book

APPENDICES

A Appendix for Chapter 2

A.1 Derivation of Social Interactions Index

Appendix A.1 derives the expression for the social interactions (SI) index in equation (B21).

First, recall the definition of the SI 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

$$\Delta_{j,k} = (N_j - 1)\left(\mathbb{E}[D_{i',j,k}|D_{i,j,k} = 1] - \mathbb{E}[D_{i',j,k}|D_{i,j,k} = 0]\right),\ i' \neq i. \tag{A1}$$

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

$$P_{j,k} = \mathbb{E}[D_{i',j,k}|D_{i,j,k} = 1]P_{j,k} + \mathbb{E}[D_{i',j,k}|D_{i,j,k} = 0](1 - P_{j,k}). \tag{A2}$$

Using the definition $\mu_{j,k} \equiv \mathbb{E}[D_{i',j,k}|D_{i,j,k} = 1]$ and rearranging equation (A2) yields

$$\mathbb{E}[D_{i',j,k}|D_{i,j,k} = 0] = \frac{P_{j,k}(1 - \mu_{j,k})}{1 - P_{j,k}}. \tag{A3}$$

Hence, we have

$$\begin{aligned}
\mathbb{E}[D_{i',j,k}|D_{i,j,k} = 1] - \mathbb{E}[D_{i',j,k}|D_{i,j,k} = 0] &= \mu_{j,k} - \frac{P_{j,k}(1 - \mu_{j,k})}{1 - P_{j,k}} && \text{(A4)}\\
&= \frac{\mu_{j,k} - P_{j,k}}{1 - P_{j,k}}. && \text{(A5)}
\end{aligned}$$

Substituting equation (A5) into equation (A1) yields

$$\Delta_{j,k} = (N_j - 1)\left(\frac{\mu_{j,k} - P_{j,k}}{1 - P_{j,k}}\right). \tag{A6}$$

Applying the law of iterated expectations to the first term of the covariance of location decisions,

$C_{j,k}$, yields

$$C_{j,k} \equiv \mathbb{E}[D_{i',j,k} D_{i,j,k}] - \mathbb{E}[D_{i',j,k}]\, \mathbb{E}[D_{i,j,k}] \tag{A7}$$

$$= \mathbb{E}[D_{i',j,k} | D_{i,j,k} = 1] P_{j,k} - (P_{j,k})^2 \tag{A8}$$

Using the definition of $\mu_{j,k}$ and rearranging yields $\mu_{j,k} - P_{j,k} = C_{j,k}/P_{j,k}$. Substituting this expression into (A6), and noting that Assumption 1 implies that $P_{j,k} = P_{g,k}$, yields equation (B21).

A.2 Method of Moments Formulation

A.2.1 Basic Model

As described in the text, we can derive the destination-level SI 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 SI index, but the latter approach allows us to use the method of moments to estimate standard errors.

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

$$\Delta_k = \Delta_{j,k} = \frac{C_{j,k}(N_j - 1)}{P_{j,k} - P_{j,k}^2}. \tag{A9}$$

It is useful to rewrite this as

$$\Delta_k \left(P_{j,k} - P_{j,k}^2\right) - C_{j,k}(N_j - 1) = 0. \tag{A10}$$

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

$$\Delta_k Y_{g,k} - X_{g,k} = 0, \tag{A11}$$

where

$$X_{g,k} \equiv \sum_{j \in g} C_{j,k}(N_j - 1) \tag{A12}$$

$$Y_{g,k} \equiv \sum_{j \in g} P_{j,k} - P_{j,k}^2. \tag{A13}$$

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,

$$\widehat{X_{g,k}} = X_{g,k} + u_{g,k}^X \tag{A14}$$
$$\widehat{Y_{g,k}} = Y_{g,k} + u_{g,k}^Y. \tag{A15}$$

This allows us to rewrite equation (A11),

$$\Delta_k \widehat{Y_{g,k}} - \widehat{X_{g,k}} + (\Delta_k u_{g,k}^Y - u_{g,k}^X) = 0. \tag{A16}$$

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

$$\mathbb{E}\left[\Delta_k \widehat{Y_{g,k}} - \widehat{X_{g,k}}\right] = 0. \tag{A17}$$

Equation (A17) is the basis of our method of moments estimator. The sample analog is

$$\frac{1}{G}\sum_g \left(\widehat{\Delta_k}\widehat{Y_{g,k}} - \widehat{X_{g,k}}\right) = 0, \tag{A18}$$

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

$$\widehat{\Delta_k} = \frac{\sum_j \widehat{C_{j,k}}(N_j - 1)}{\sum_{j'} \widehat{P_{j',k}} - \widehat{P_{j',k}^2}}. \tag{A19}$$

Equation (A19) is identical to equation (2.9).

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

$$\mathbb{E}\begin{bmatrix} \Delta_1 \widehat{Y_{g,1}} - \widehat{X_{g,1}} \\ \vdots \\ \Delta_K \widehat{Y_{g,K}} - \widehat{X_{g,K}} \end{bmatrix} \equiv \mathbb{E}\left[\boldsymbol{f}(\boldsymbol{w}_g, \boldsymbol{\Delta})\right] = \boldsymbol{0}, \tag{A20}$$

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

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

is

$$\sqrt{G}(\hat{\mathbf{\Delta}} - \mathbf{\Delta}) \xrightarrow{d} \mathcal{N}\left[\mathbf{0}, \hat{\boldsymbol{F}}^{-1}\hat{\boldsymbol{S}}(\hat{\boldsymbol{F}}')^{-1}\right], \tag{A21}$$

where

$$\hat{\boldsymbol{F}} = \frac{1}{G}\sum_g \left.\frac{\partial \boldsymbol{f}_g}{\partial \mathbf{\Delta}'}\right|_{\hat{\mathbf{\Delta}}} \tag{A22}$$

$$= \frac{1}{G}\sum_g \begin{bmatrix} \widehat{Y_{g,1}} & 0 & 0 & \cdots & 0 \\ 0 & \widehat{Y_{g,2}} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & \widehat{Y_{g,K}} \end{bmatrix} \tag{A23}$$

and

$$\hat{\boldsymbol{S}} = \frac{1}{G}\sum_g \boldsymbol{f}(\boldsymbol{W}_g, \hat{\mathbf{\Delta}})\boldsymbol{f}(\boldsymbol{W}_g, \hat{\mathbf{\Delta}})'. \tag{A24}$$

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

A.2.2 Comparing Estimates from Two Models

The method of moments 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 $Y_{g,k}$:

$$\widehat{X^1_{g,k}} = X_{g,k} + u^X_{g,k} \tag{A25}$$

$$\widehat{Y^1_{g,k}} = Y_{g,k} + u^Y_{g,k} \tag{A26}$$

$$\widehat{X^2_{g,k}} = X_{g,k} + v^X_{g,k} \tag{A27}$$

$$\widehat{Y^2_{g,k}} = Y_{g,k} + v^Y_{g,k}. \tag{A28}$$

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

$$\frac{1}{G}\sum_g \left(\begin{array}{c} \widehat{\Delta^1_k}\widehat{Y^1_{g,k}} - \widehat{X^1_{g,k}} \\ \widehat{\Delta^2_k}\widehat{Y^2_{g,k}} - \widehat{X^2_{g,k}} \end{array} \right) \tag{A29}$$

This simply stacks the two estimates of the destination-level SI 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 SI 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

$$\hat{t} = \frac{\widehat{\Delta^1} - \widehat{\Delta^2}}{\left(\widehat{\mathbb{V}}[\widehat{\Delta^1} - \widehat{\Delta^2}]\right)^{1/2}}. \tag{A30}$$

Given destination-level SI 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 SI 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 social interaction 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 $\widehat{\mathbb{V}}[\widehat{\Delta^1} - \widehat{\Delta^2}]$ as the appropriately weighted sum of

$$\widehat{\mathbb{V}}[\widehat{\Delta^1_k} - \widehat{\Delta^2_k}] = \widehat{\mathbb{V}}[\widehat{\Delta^1_k}] + \widehat{\mathbb{V}}[\widehat{\Delta^2_k}] - 2\widehat{\mathbb{C}}[\widehat{\Delta^1_k}, \widehat{\Delta^2_k}] \tag{A31}$$

which we obtain from the method of moments variance estimate.

A.3 Estimating Cross-Group Social Interactions

When estimating cross-group social interactions, 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,

$$\Delta^{b|w}_{j,k} \equiv \mathbb{E}[N^b_{j,k} | D^w_{i,j,k} = 1] - \mathbb{E}[N^b_{j,k} | D^w_{i,j,k} = 0]. \tag{A32}$$

The steps described in Appendix A.1 yield

$$\Delta^{b|w}_{j,k} = \frac{C^{b,w}_{j,k} N^b_j}{P^w_{j,k}(1 - P^w_{j,k})}, \tag{A33}$$

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

$$D^{b}_{i,j(i),k}D^{w}_{i',j(i'),k} = \alpha_{g,k} + \sum_{j\in g}\beta^{b,w}_{j,k}1[j(i) = j(i') = j] + \epsilon_{i,i',k}. \tag{A34}$$

This model is analogous to equation (2.2) in the text and yields the following covariance estimator,

$$\widehat{C^{b,w}_{j,k}} = \frac{N^{b}_{j,k}N^{w}_{j,k}}{N^{b}_{j}N^{w}_{j}} - \frac{\sum_{j\in g}\sum_{j'\neq j\in g}N^{b}_{j,k}N^{w}_{j',k}}{\sum_{j\in g}\sum_{j'\neq j\in g}N^{b}_{j}N^{w}_{j'}}. \tag{A35}$$

We estimate the destination-level SI index as

$$\hat{\Delta}^{b|w}_{k} = \sum_{j}\left(\frac{\widehat{P^{w}_{j,k}} - \widehat{(P^{w}_{j,k})^2}}{\sum_{j'}\widehat{P^{w}_{j',k}} - \widehat{(P^{w}_{j',k})^2}}\right)\hat{\Delta}^{b|w}_{j,k}. \tag{A36}$$

We only estimate social interactions 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.

A.4 Additional Detail on Measurement Error due to Incomplete Migration Data

Appendix A.4 discusses the implications of measurement error due to incomplete migration data without making a missing at random (MAR) assumption. We derive a lower bound on the social interactions (SI) index and show that estimates of this lower bound still reveal sizable social interactions.

As described in the text, the SI 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, we assume 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 $\Delta^{*}_{j,k}$, $C^{*}_{j,k}$, and N^{*}_{j} be the true values of the SI index, covariance of location decisions, and number of migrants. The true parameters are connected through the equation

$$\Delta^{*}_{j,k} = \frac{C^{*}_{j,k}(N^{*}_{j} - 1)}{P_{g,k} - P^{2}_{g,k}}. \tag{A37}$$

As in the text, we let α denote the coverage rate, defined by the relationship between the observed number of migrants, N_j, and the true number of migrants,

$$N_j = \alpha N_j^*. \tag{A38}$$

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

$$C_{j,k}^* = \alpha^2 C_{j,k} + 2\alpha(1-\alpha) C_{j,k}^{\text{in, out}} + (1-\alpha)^2 C_{j,k}^{\text{out, out}}, \tag{A39}$$

where $C_{j,k}$ is the covariance of location decisions between migrants who are in our data, $C_{j,k}^{\text{in, out}}$ is the average covariance of location decisions between a migrant who is in our data and a migrant who is not, and $C_{j,k}^{\text{out, out}}$ 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 ($C_{j,k}^{\text{in, out}}$ and $C_{j,k}^{\text{out, out}}$) could differ from the covariance of location decisions between migrants who are in our data ($C_{j,k}$). As a result, the SI 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 SI index under this more general measurement error model. However, we can construct a lower bound for the strength of social interactions. In particular, we make the extreme assumptions that there are no social interactions between migrants in and out of our data, so that $C_{j,k}^{\text{in, out}} = 0$, and that there are no social interactions between migrants out of our data, so that $C_{j,k}^{\text{out, out}} = 0$. In this case, equations (A37), (A38), and (A39) imply that

$$\Delta_{j,k}^* \geq \alpha \Delta_{j,k}, \tag{A40}$$

so that we can estimate a lower bound on the true SI index by multiplying the estimated SI index by the coverage rate.[52] The average coverage rate is 52.3% for African American migrants from the South and 69.3% for white migrants from the Great Plains. Combined with the average destination-level SI index estimates from Table 2.3, we estimate a lower bound for the SI index of 1.014 for African Americans and 0.263 for whites. These lower bounds, which depend on extremely

[52]Proof: If $C_{j,k}^{\text{in, out}} = C_{j,k}^{\text{out, out}} = 0$, equations (A37), (A38), and (A39) imply

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

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 (B21) in the text. One could also construct upper bounds, but these are not particularly informative.

conservative assumptions about the migration behavior of individuals not in our data, still reveal sizable social interactions, especially among African Americans.

A.5 A Richer Model of Local Social Interactions

This section extends the local social interactions model in Section 2.4.5. In particular, we allow the probability that a migrant follows his neighbor to vary with birth town and destination.

Migrants from birth town j are indexed on a line by $i \in \{1, \ldots, N_j\}$, where N_j is the total number of migrants from town j. For migrant i, destination k belongs to one of three preference groups: high (H_i), medium (M_i), or low (L_i). The high preference group contains a single destination. In the absence of social interactions, the destination in H_i is most preferred, and destinations in M_i are preferred over those in L_i.[53] A migrant never moves to a destination in L_i. A migrant chooses a destination in M_i if and only if his neighbor, $i-1$, chooses the same destination. A migrant chooses a destination in H_i if his neighbor chooses the same destination or his neighbor selects a destination in L_i.[54]

Migrants from the same birth town can differ in their preferences over destinations. The probability that destination k is in the high preference group for a migrant from town j is $h_{j,k} \equiv \mathbb{P}[k \in H_i | i \in j]$, and the probability that destination k is in the medium preference group is $m_{j,k} \equiv \mathbb{P}[k \in M_i | i \in j]$.

The probability that migrant i moves to destination k given that his neighbor moves there is

$$\rho_{j,k} \equiv \mathbb{P}[D_{i,j,k} = 1 | D_{i-1,j,k} = 1] = \mathbb{P}[k \in H_i] + \mathbb{P}[k \in M_i] \tag{A41}$$

$$= h_{j,k} + m_{j,k}, \tag{A42}$$

where $D_{i,j,k}$ equals one if migrant i moves from j to k and zero otherwise.

The probability that destination k is in the medium preference group, conditional on not being in the high preference group, is $\nu_{j,k} \equiv \mathbb{P}[k \in M_i | k \notin H_i, i \in j]$. The conditional probability definition for $\nu_{j,k}$ implies that $m_{j,k} = \nu_{j,k}(1 - h_{j,k})$. We use $\nu_{j,k}$ to derive a simple sequential estimation approach.

[53] The assumption that H_i is a non-empty singleton ensures that migrant i has a well-defined location decision in the absence of social interactions. We could allow H_i to contain many destinations and specify a decision rule among the elements of H_i. This extension would complicate the model without adding any new insights.

[54] This model shares a similar structure as Glaeser, Sacerdote and Scheinkman (1996) in that some agents imitate their neighbors. However, we differ from Glaeser, Sacerdote and Scheinkman (1996) in that we model the interdependence between various destinations (i.e., this is a multinomial choice problem) and allow for more than two types of agents.

In equilibrium, the probability that a randomly chosen migrant i moves from j to k is

$$
\begin{aligned}
P_{j,k} \equiv \mathbb{P}[D_{i,j,k} = 1] &= \mathbb{P}[D_{i-1,j,k} = 1, k \in H_i] + \mathbb{P}[D_{i-1,j,k} = 1, k \in M_i] \\
&\quad + \sum_{k' \neq k} \mathbb{P}[D_{i-1,j,k'} = 1, k \in H_i, k' \in L_i] && \text{(A43)} \\
&= P_{j,k} h_{j,k} + P_{j,k} \nu_{j,k}(1 - h_{j,k}) + \sum_{k' \neq k} P_{j,k'} h_{j,k}(1 - \nu_{j,k'}) && \text{(A44)} \\
&= P_{j,k} \nu_{j,k} + \left(\sum_{k'=1}^{K} P_{j,k'}(1 - \nu_{j,k'}) \right) h_{j,k}. && \text{(A45)}
\end{aligned}
$$

The first term on the right hand side of equation (B16) is the probability that a migrant's neighbor moves to k, and k is in the migrant's high preference group; in this case, social interaction reinforces the migrant's desire to move to k. The second term is the probability that a migrant follows his neighbor to k because of social interactions. The third term is the probability that a migrant resists the pull of social interactions because town k is in the migrant's high preference group and the neighbor's chosen destination is in the migrant's low preference group.

The average covariance of location decisions implied by the richer model is[55]

$$
C_{j,k} = \frac{2P_{j,k}(1 - P_{j,k}) \sum_{s=1}^{N_j - 1} (N_j - s) \left(\frac{\rho_{j,k} - P_{j,k}}{1 - P_{j,k}} \right)^s}{N_j (N_j - 1)}. \tag{A46}
$$

Substituting equation (B22) into equation (B21) and simplifying yields[56]

$$
\Delta_{j,k} = \frac{2(\rho_{j,k} - P_{j,k})}{1 - \rho_{j,k}}, \tag{A47}
$$

which can be rearranged to show that

$$
\rho_{j,k} = \frac{2P_{j,k} + \Delta_{j,k}}{2 + \Delta_{j,k}}. \tag{A48}
$$

We could use equation (B24) to estimate $\rho_{j,k}$ with our estimates of $P_{j,k}$ and $\Delta_{j,k}$.

Equations (B15) and (B18), plus the fact that $m_{j,k} = \nu_{j,k}(1 - h_{j,k})$, imply that

$$
\rho_{j,k} = \nu_{j,k} + \frac{P_{j,k}(1 - \nu_{j,k})^2}{\sum_{k'=1}^{K} P_{j,k'}(1 - \nu_{j,k'})}. \tag{A49}
$$

[55]This follows from the fact that the covariance of location decisions for individuals i and $i + n$ is $\mathbb{C}[D_{i,j,k}, D_{i+n,j,k}] = P_{j,k}(1 - P_{j,k}) \left(\frac{\rho_{j,k} - P_{j,k}}{1 - P_{j,k}} \right)^n$.

[56]Equation (B23) results from taking the limit as $N_j \to \infty$, and so relies on N_j being sufficiently large.

We could use equation (B25) to estimate $\nu_j \equiv (\nu_{j,1}, \ldots, \nu_{j,K})$ using our estimates of $(P_{j,1}, \ldots, P_{j,K}, \rho_{j,1}, \ldots, \rho_{j,K})$. In addition, we could use equation (B18) to estimate $h_{j,k}$ with our estimates of $\rho_{j,k}$ and $\nu_{j,k}$. Finally, we could estimate $m_{j,k}$ using the fact that $m_{j,k} = \rho_{j,k} - h_{j,k}$.

Table A.1: Number of Birth Towns and Migrants, by Birth State

Birth State	Birth Towns (1)	Migrants (2)	Migrants Per Town (3)
Panel A: Black Moves out of South			
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: White Moves out of Great Plains			
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: Duke SSA/Medicare data

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

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

Notes: Column 1 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: Duke SSA/Medicare data

Table A.3: Average Destination Level Social Interactions Index Estimates, Birth Town Groups Defined by Cross Validation and Counties

	Cross Validation		Counties	
Type of Average:	Unweighted	Weighted	Unweighted	Weighted
Birth State	(1)	(2)	(3)	(4)
Panel A: Black Moves out of South				
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: White Moves out of Great Plains				
Kansas	0.128	0.255	0.106	0.194
	(0.007)	(0.024)	(0.008)	(0.028)
North Dakota	0.174	0.464	0.156	0.385
	(0.012)	(0.036)	(0.010)	(0.029)
Nebraska	0.141	0.361	0.121	0.399
	(0.008)	(0.082)	(0.009)	(0.117)
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 SI 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: Duke SSA/Medicare data

Table A.4: Average Social Interactions Index Estimates, White Moves out of South

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 SI 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: Duke SSA/Medicare data

Table A.5: Average Social Interactions Index Estimates, By Size of Birth Town and Destination, White Moves out of South

Exclude Largest Birth Towns:	No	Yes	No	Yes
Exclude Largest Destinations:	No	No	Yes	Yes
Birth State	(1)	(2)	(3)	(4)
Alabama	0.516	0.458	0.531	0.481
	(0.052)	(0.045)	(0.071)	(0.062)
Florida	0.072	0.074	0.134	0.030
	(0.100)	(0.012)	(0.082)	(0.009)
Georgia	0.117	0.101	0.119	0.088
	(0.021)	(0.012)	(0.019)	(0.013)
Louisiana	0.269	0.207	0.198	0.143
	(0.071)	(0.022)	(0.035)	(0.017)
Mississippi	0.186	0.185	0.135	0.134
	(0.021)	(0.022)	(0.013)	(0.013)
North Carolina	0.412	0.395	0.337	0.319
	(0.040)	(0.037)	(0.040)	(0.034)
South Carolina	0.094	0.090	0.058	0.055
	(0.029)	(0.023)	(0.013)	(0.012)
All States	0.280	0.254	0.262	0.223
	(0.021)	(0.013)	(0.021)	(0.015)

Notes: All columns contain weighted averages of destination-level SI 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: Duke SSA/Medicare data

Table A.6: Average Social Interactions Index Estimates, by Destination Region, White Moves out of South

	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: All columns contain weighted averages of destination-level SI index estimates, $\hat{\Delta}_k$, where the weights are the number of people who move from each state to destination k. See footnote 36 for region definitions. We do not estimate social interactions for Southern-born whites who move to the South. Birth town groups are defined by cross validation. Standard errors are in parentheses. Source: Duke SSA/Medicare data

Table A.7: Average Cross-Race Social Interactions Index Estimates, Southern White and Black Migrants

Birth State	All Counties (1)	Excluding Largest CMSAs (2)
Panel A: Blacks 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: Whites 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.7 contains weighted averages of cross-race destination-level SI index estimates. Birth town groups are defined by cross validation. Standard errors in parentheses.
Source: Duke SSA/Medicare data

Table A.8: 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 Blacks							
Alabama	70.2%	78.6%	55.2%	55.0%	55.4%	47.7%	62.8%
Florida	62.3%	83.3%	51.9%	53.2%	50.9%	45.8%	57.4%
Georgia	67.2%	72.8%	48.9%	47.5%	50.1%	43.2%	55.5%
Louisiana	67.9%	84.4%	57.3%	57.4%	57.2%	51.3%	63.2%
Mississippi	77.3%	74.6%	57.7%	57.7%	57.6%	50.4%	65.2%
North Carolina	68.5%	72.4%	49.6%	46.7%	51.9%	42.9%	56.5%
South Carolina	75.3%	61.6%	46.4%	43.6%	48.8%	39.3%	55.3%
All States	70.4%	74.2%	52.3%	51.2%	53.2%	45.5%	59.5%
Panel B: Great Plains Whites							
Kansas	75.9%	92.3%	70.1%	68.9%	71.3%	64.8%	76.0%
Nebraska	75.2%	93.2%	70.0%	69.8%	70.3%	65.6%	74.8%
North Dakota	76.1%	89.6%	68.1%	64.6%	71.7%	64.6%	71.8%
Oklahoma	75.8%	89.8%	68.1%	67.2%	69.0%	62.8%	73.2%
South Dakota	78.3%	91.0%	71.3%	70.5%	72.1%	64.3%	79.6%
All States	76.0%	91.2%	69.3%	68.1%	70.4%	64.2%	74.7%

Notes: Column 1 reports the number of individuals in the Duke SSA/Medicare dataset divided by the number of individuals in the 1960/1970 Census. Column 2 reports the share of individuals in the Duke SSA/Medicare dataset for whom birth town and destination county is identified. Columns 3-7 reports the number of individuals in the Duke SSA/Medicare dataset for whom birth town and destination county is identified divided by the number of individuals in the 1960/1970 Census. In all columns, we use the 1960 Census for individuals born from 1916-1925 and the 1970 Census for individuals born from 1926-1936. The sample includes individuals living inside and outside their birth region.

Source: Duke SSA/Medicare data and Ruggles et al. (2010) data

Table A.9: Average Social Interactions 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: Black Moves out of South					
Alabama	3.420	1.542	1.891	1.739	1.874
	(0.353)	(0.160)	(0.204)	(0.197)	(0.185)
Florida	1.567	0.725	0.832	0.650	0.980
	(0.226)	(0.116)	(0.168)	(0.145)	(0.154)
Georgia	3.389	1.317	2.069	2.072	1.566
	(0.362)	(0.153)	(0.246)	(0.281)	(0.144)
Louisiana	3.007	1.533	1.218	1.280	2.015
	(0.834)	(0.408)	(0.478)	(0.296)	(0.689)
Mississippi	3.990	1.759	2.273	1.769	2.353
	(0.542)	(0.244)	(0.331)	(0.267)	(0.323)
North Carolina	3.104	1.414	1.729	1.742	1.561
	(0.263)	(0.137)	(0.150)	(0.164)	(0.128)
South Carolina	5.643	2.543	3.141	3.223	2.630
	(0.648)	(0.262)	(0.433)	(0.423)	(0.276)
All States	3.713	1.648	2.064	1.965	1.972
	(0.197)	(0.088)	(0.123)	(0.113)	(0.118)
Panel B: White Moves out of Great Plains					
Kansas	0.364	0.185	0.197	0.248	0.185
	(0.034)	(0.020)	(0.018)	(0.025)	(0.015)
Nebraska	0.515	0.221	0.290	0.333	0.268
	(0.117)	(0.063)	(0.056)	(0.071)	(0.053)
North Dakota	0.681	0.317	0.361	0.445	0.324
	(0.054)	(0.027)	(0.034)	(0.037)	(0.024)
Oklahoma	0.665	0.320	0.345	0.361	0.382
	(0.053)	(0.029)	(0.028)	(0.031)	(0.031)
South Dakota	0.491	0.220	0.274	0.325	0.236
	(0.037)	(0.020)	(0.023)	(0.027)	(0.018)
All States	0.552	0.258	0.297	0.338	0.294
	(0.031)	(0.017)	(0.016)	(0.019)	(0.016)

Notes: Table A.9 reports weighted averages of destination-level SI index estimates, adjusted for incomplete migration data using the coverage rates in Appendix Table A.8. Birth town groups are defined by cross validation. Standard errors in parentheses.
Source: Duke SSA/Medicare data

Table A.10: Summary Statistics, Destination County Characteristics

Variable	Mean	S.D.
Panel A: Black Moves out of South (N=1469)		
SI index estimate, $\widehat{\Delta_k}$	0.732	1.373
Manufacturing employment share, 1910	0.240	0.140
Direct railroad connection from birth state	0.093	0.291
One-stop railroad connection from birth state	0.557	0.497
Log distance from birth state	6.684	0.517
Log population, 1900	11.004	1.105
Percent African-American, 1900	0.045	0.082
Panel B: White Moves out of Great Plains (N=3822)		
SI index estimate, $\widehat{\Delta_k}$	0.140	0.441
Manufacturing employment share, 1910	0.169	0.134
Direct railroad connection from birth state	0.112	0.315
One-stop railroad connection from birth state	0.504	0.500
Log distance from birth state	6.788	0.355
Log population, 1900	10.122	1.080
Percent African-American, 1900	0.121	0.197
Panel C: White Moves Out of South (N=3153)		
SI index estimate, $\widehat{\Delta_k}$	0.131	0.566
Manufacturing employment share, 1910	0.195	0.141
Direct railroad connection from birth state	0.084	0.278
One-stop railroad connection from birth state	0.492	0.500
Log distance from birth state	6.766	0.593
Log population, 1900	10.418	1.143
Percent African-American, 1900	0.038	0.077

Notes: The unit of observation is a birth state-destination county pair. Sample includes destination counties that existed from 1900-2000 and for which we estimate a SI index.

Sources: Duke SSA/Medicare data, Haines and ICPSR (2010) data

Table A.11: Social Interactions Index Estimates and Destination County Characteristics, Black Moves out of South, Birth Town Groups Defined by Counties

Dependent variable: Destination-level SI index estimate	(1)	(2)	(3)
Manufacturing employment share, 1910	1.472**	0.381	0.362
	(0.604)	(0.372)	(0.386)
Manufacturing employment share by small destination indicator		1.932***	1.965***
		(0.727)	(0.683)
Small destination indicator		0.331**	0.349***
		(0.133)	(0.123)
Direct railroad connection from birth state	0.348***	0.370***	0.391***
	(0.114)	(0.117)	(0.142)
One-stop railroad connection from birth state	0.222**	0.210**	0.189**
	(0.092)	(0.087)	(0.093)
Log distance from birth state	-0.246***	-0.220***	-0.230***
	(0.070)	(0.077)	(0.062)
Log population, 1900	0.081*	0.101**	0.102**
	(0.045)	(0.041)	(0.046)
Percent African-American, 1900	-1.530***	-1.434***	-1.443***
	(0.291)	(0.318)	(0.288)
Birth state fixed effects			x
R2	0.055	0.064	0.073
N (birth state-destination county pairs)	1,469	1,469	1,469
Destination counties	371	371	371

Notes: The sample contains only counties that received at least 10 migrants. Birth town groups are defined by counties. Standard errors, clustered by destination county, are in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Sources: Duke SSA/Medicare data, Haines and ICPSR (2010) data, and **?** data

Table A.12: Social Interactions Index Estimates and Destination County Characteristics, White Moves out of Great Plains

Dependent variable: Destination-level SI index estimate	(1)	(2)	(3)
Manufacturing employment share, 1910	-0.008	-0.261***	-0.257***
	(0.080)	(0.088)	(0.088)
Manufacturing employment share by small destination indicator		0.319***	0.316**
		(0.123)	(0.123)
Small destination indicator		0.009	0.008
		(0.035)	(0.034)
Direct railroad connection from birth state	0.213***	0.212***	0.203***
	(0.043)	(0.043)	(0.045)
One-stop railroad connection from birth state	0.089***	0.084***	0.084***
	(0.018)	(0.018)	(0.017)
Log distance from birth state	0.062*	0.074**	0.065*
	(0.035)	(0.037)	(0.038)
Log population, 1900	0.008	0.016**	0.016*
	(0.008)	(0.008)	(0.008)
Percent African-American, 1900	-0.228***	-0.240***	-0.237***
	(0.032)	(0.034)	(0.033)
Birth state fixed effects			x
R2	0.031	0.034	0.035
N (birth state-destination county pairs)	3,822	3,822	3,822
Destination counties	1148	1148	1148

Notes: The sample contains only counties that received at least 10 migrants. Birth town groups are defined by cross validation. Standard errors, clustered by destination county, are in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Sources: Duke SSA/Medicare data, Haines and ICPSR (2010) data, and **?** data

Table A.13: Social Interactions Index Estimates and Destination County Characteristics, White Moves out of South

Dependent variable: Destination-level SI index estimate	(1)	(2)	(3)
Manufacturing employment share, 1910	0.382**	-0.081	-0.074
	(0.161)	(0.141)	(0.141)
Manufacturing employment share by small destination indicator		0.627***	0.638***
		(0.219)	(0.222)
Small destination indicator		0.167***	0.171***
		(0.052)	(0.053)
Direct railroad connection from birth state	0.068	0.072*	0.082**
	(0.041)	(0.041)	(0.040)
One-stop railroad connection from birth state	0.058**	0.051**	0.057***
	(0.022)	(0.021)	(0.021)
Log distance from birth state	-0.049***	-0.055***	-0.022
	(0.019)	(0.019)	(0.019)
Log population, 1900	-0.016	-0.017	-0.010
	(0.013)	(0.012)	(0.011)
Percent African-American, 1900	-0.260***	-0.346***	-0.262***
	(0.097)	(0.096)	(0.092)
Birth state fixed effects			x
R2	0.013	0.017	0.028
N (birth state-destination county pairs)	3,153	3,153	3,153
Destination counties	728	728	728

Notes: The sample contains only counties that received at least 10 migrants. Birth town groups are defined by cross validation. Standard errors, clustered by destination county, are in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Sources: Duke SSA/Medicare data, Haines and ICPSR (2010) data, and **?** data

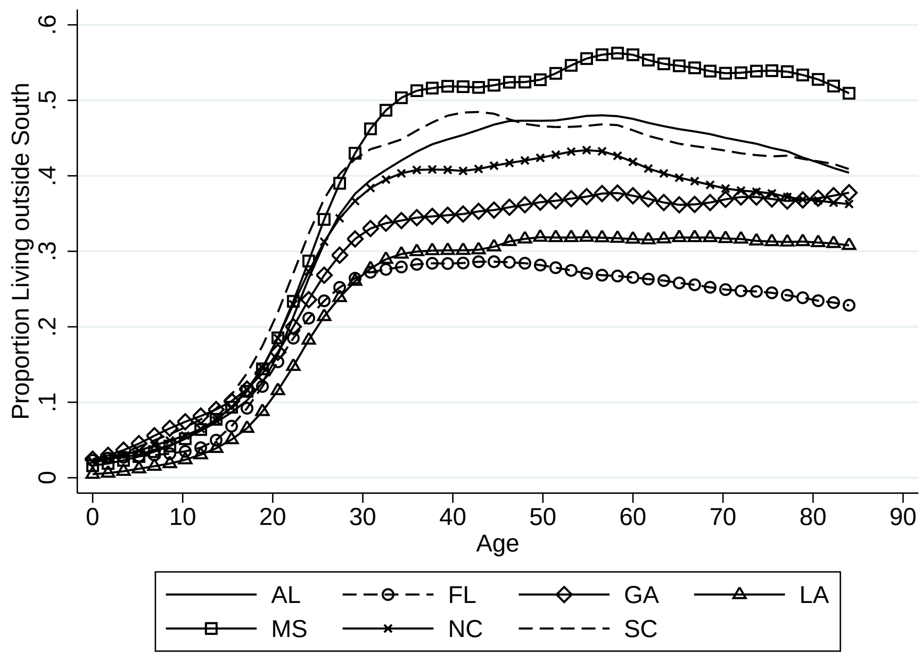
Table A.14: Summary Statistics, Birth County Characteristics, Black Moves out of South

Variable	Mean	S.D.
SI index estimate, $\widehat{\Delta_c}$	1.721	3.544
African-American farm ownership rate, 1920	0.318	0.246
Log African-American density, 1920	2.534	1.055
Rosenwald school exposure	0.204	0.217
African-American literacy rate, 1920	0.705	0.093
Railroad exposure	0.542	0.405
Percent African-American, 1920	0.408	0.209

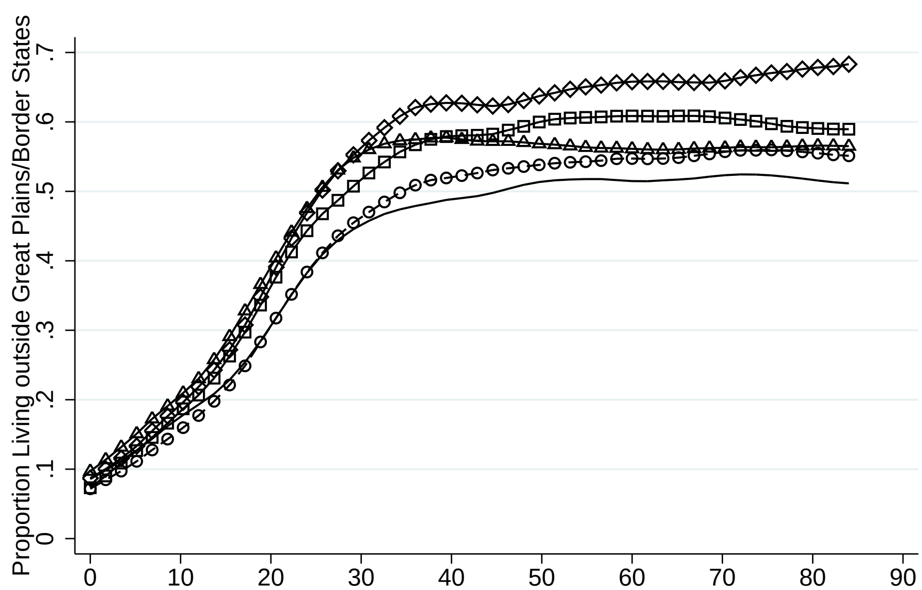
Notes: Sample includes Southern counties containing at least one town with at least 10 black migrants in the Duke data (N=549). Railroad exposure is the share of migrants in a county that lived along a railroad. Rosenwald school exposure is the average Rosenwald coverage experienced over ages 7-13.

Sources: Duke SSA/Medicare data, Haines and ICPSR (2010) data, Aaronson and Mazumder (2011) data, and **?** data

Figure A.1: Proportion Living Outside Birth Region, 1916-1936 Cohorts, by Birth State and Age



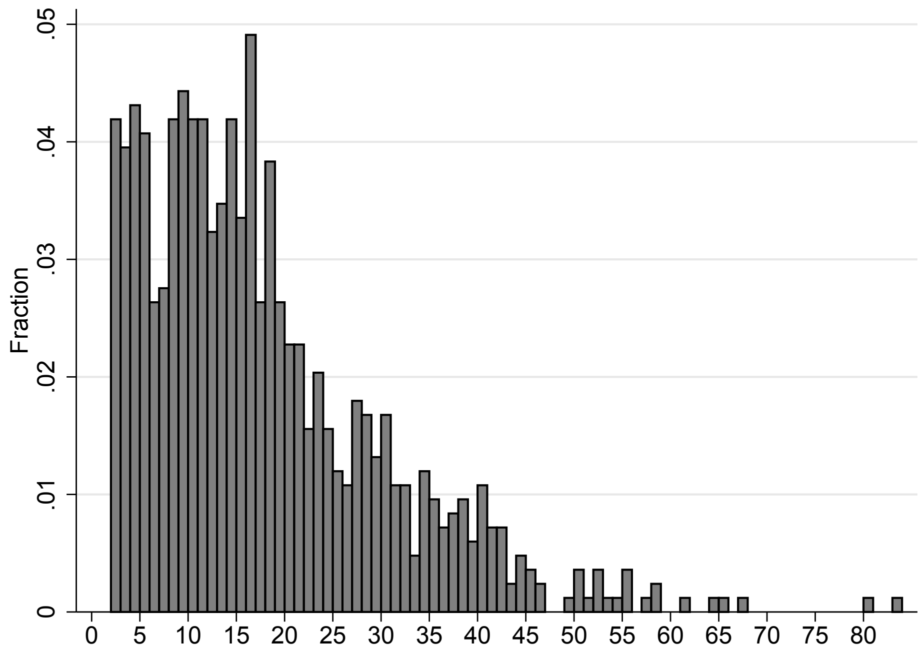
(a) Southern Blacks



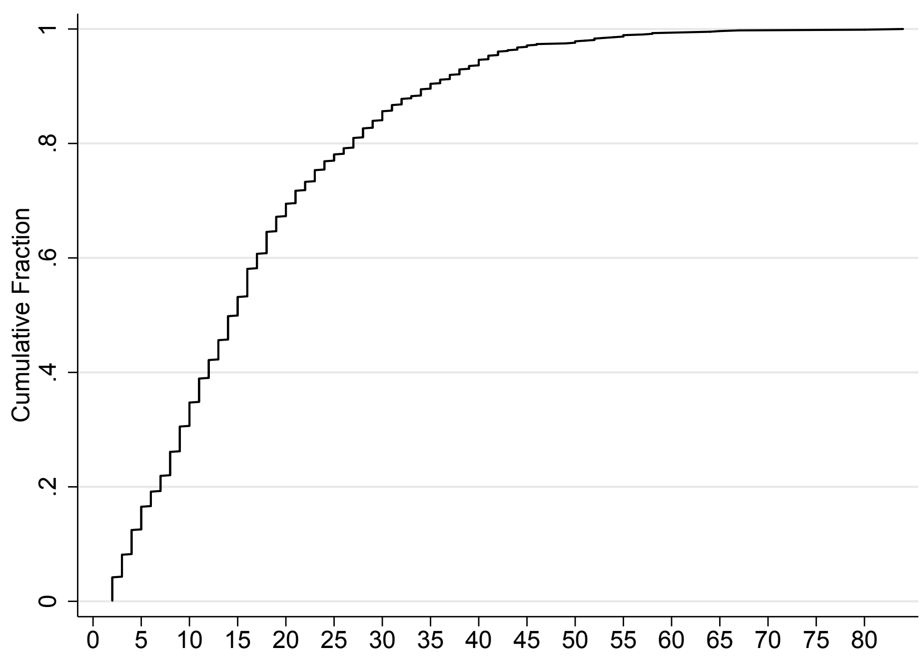
(b) Great Plains Whites

Notes: Figure A.1 displays the locally mean-smoothed relationships. Figure 2.4 displays birth regions.
Source: Ruggles et al. (2010) data

Figure A.2: Number of Towns per Birth Town Group, Cross Validation, Black Moves out of South



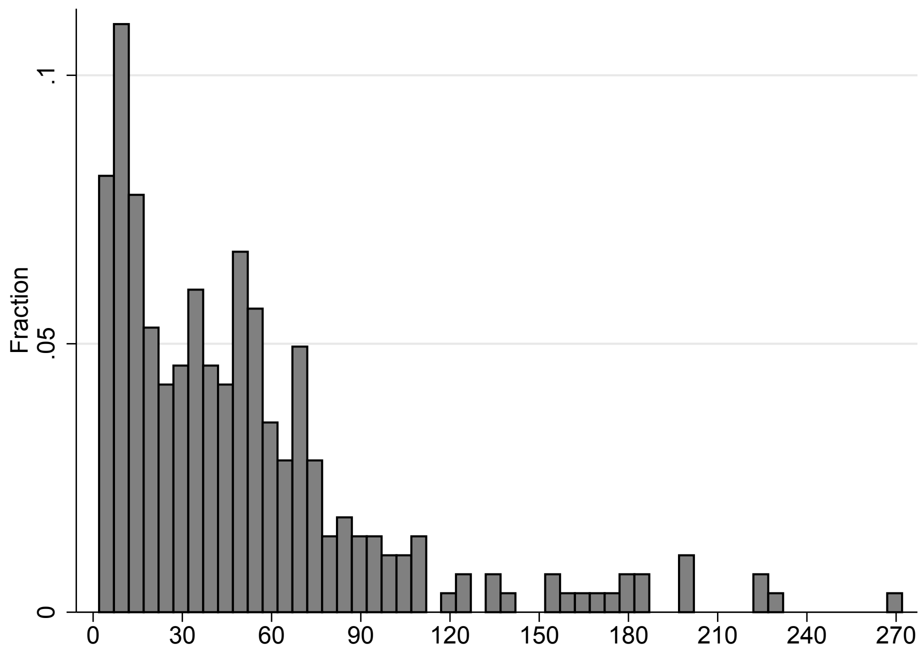
(a) Histogram



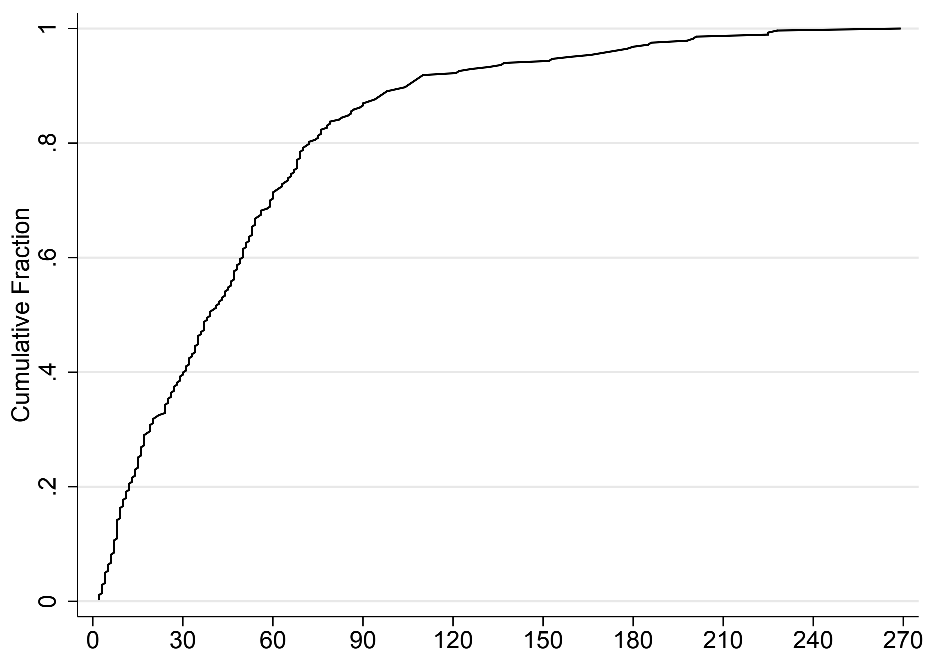
(b) Cumulative Distribution

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

Figure A.3: Number of Towns per Birth Town Group, Cross Validation, White Moves out of Great Plains



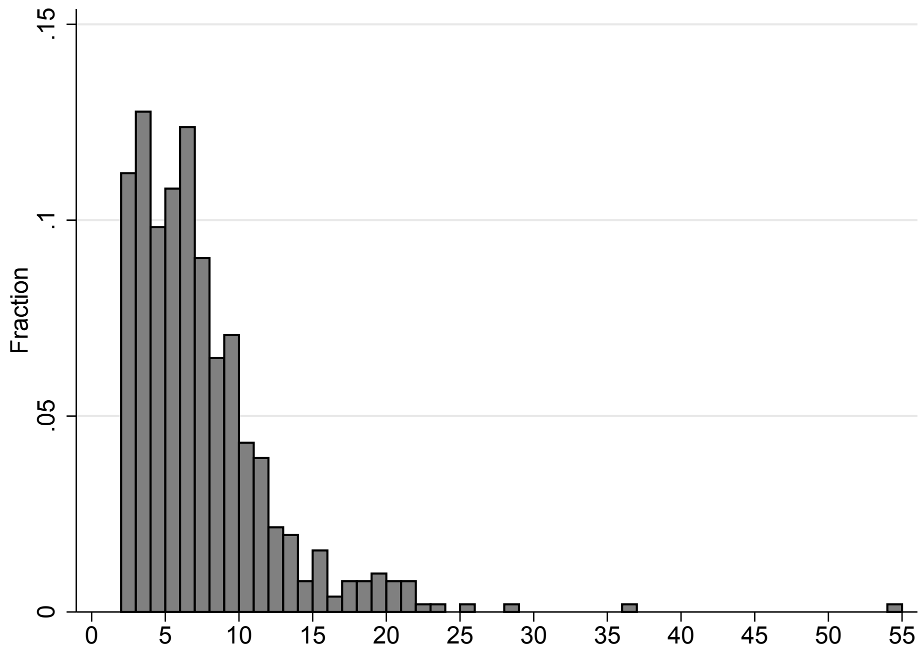
(a) Histogram



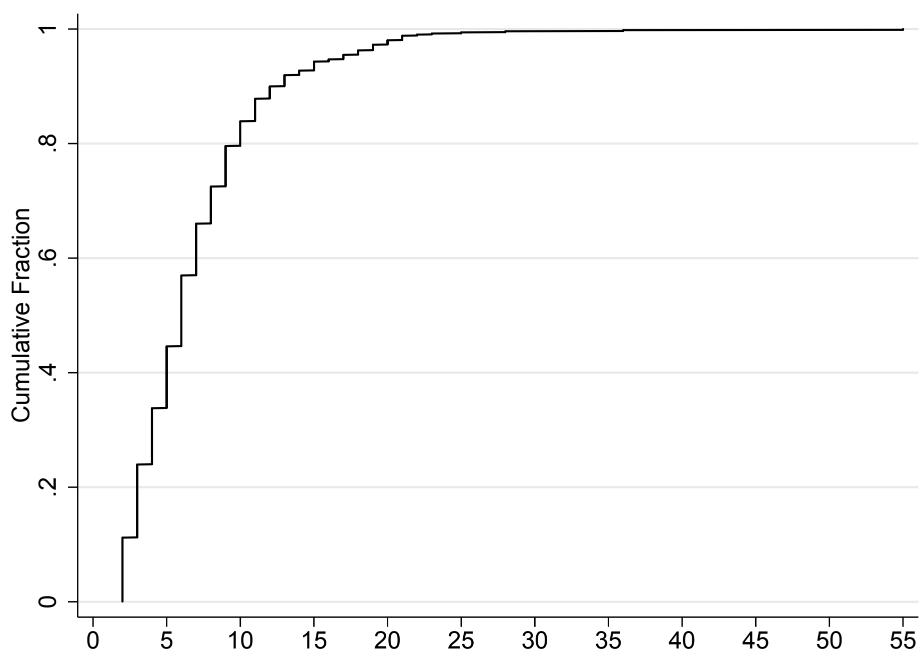
(b) Cumulative Distribution

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

Figure A.4: Number of Towns per County, Black Moves out of South



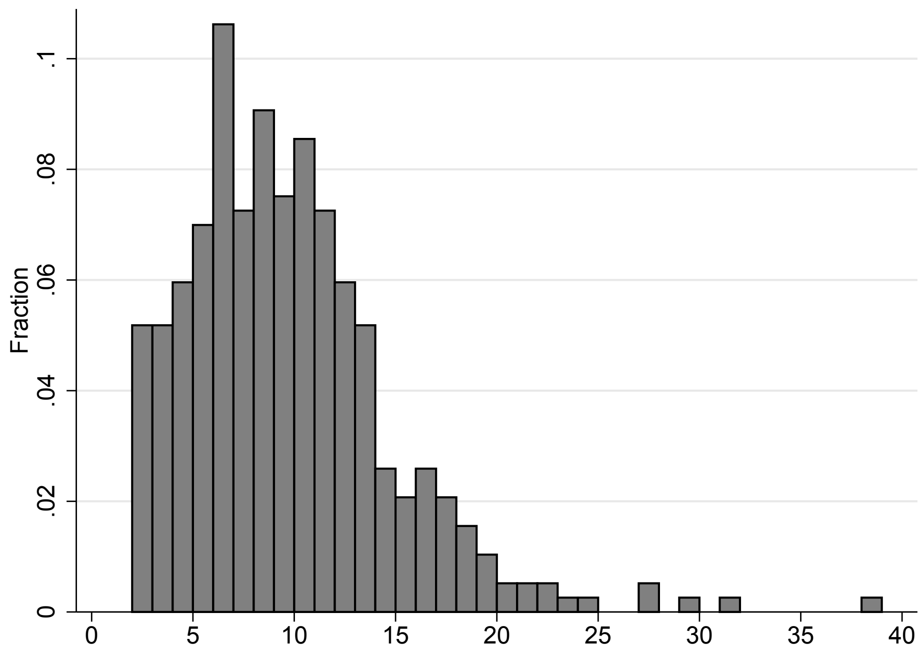
(a) Histogram



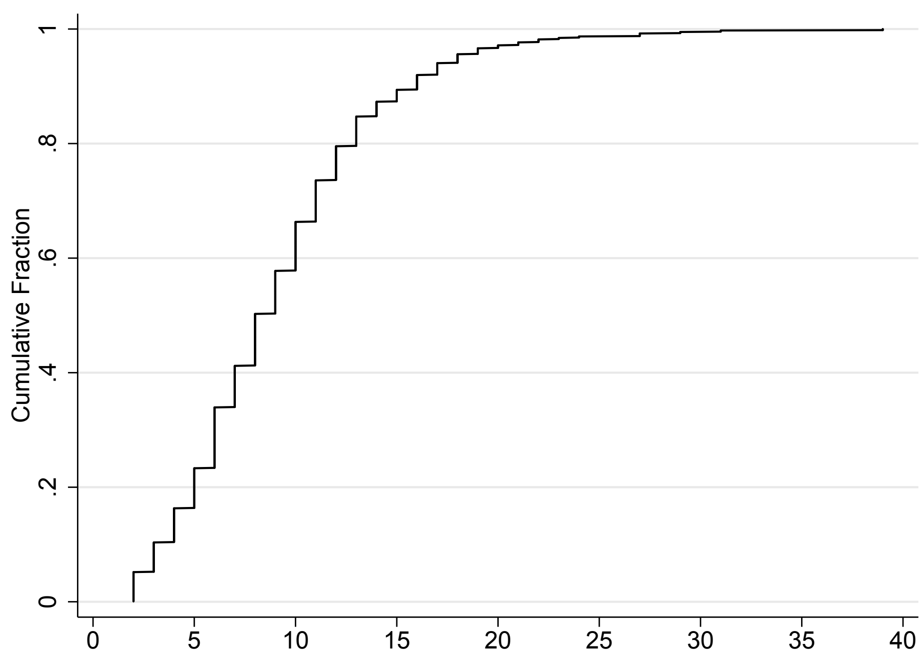
(b) Cumulative Distribution

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

Figure A.5: Number of Towns per County, White Moves out of Great Plains



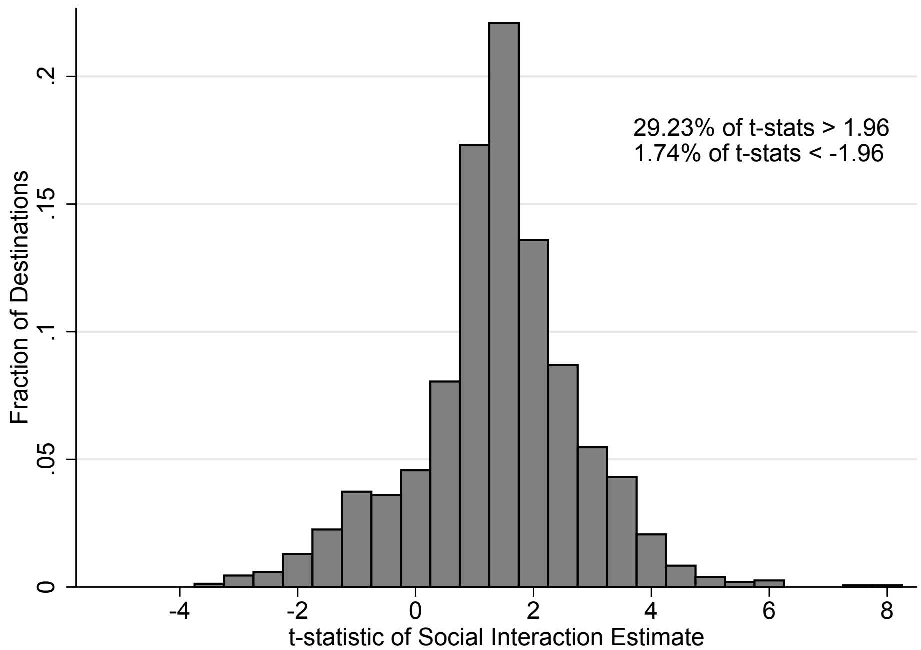
(a) Histogram



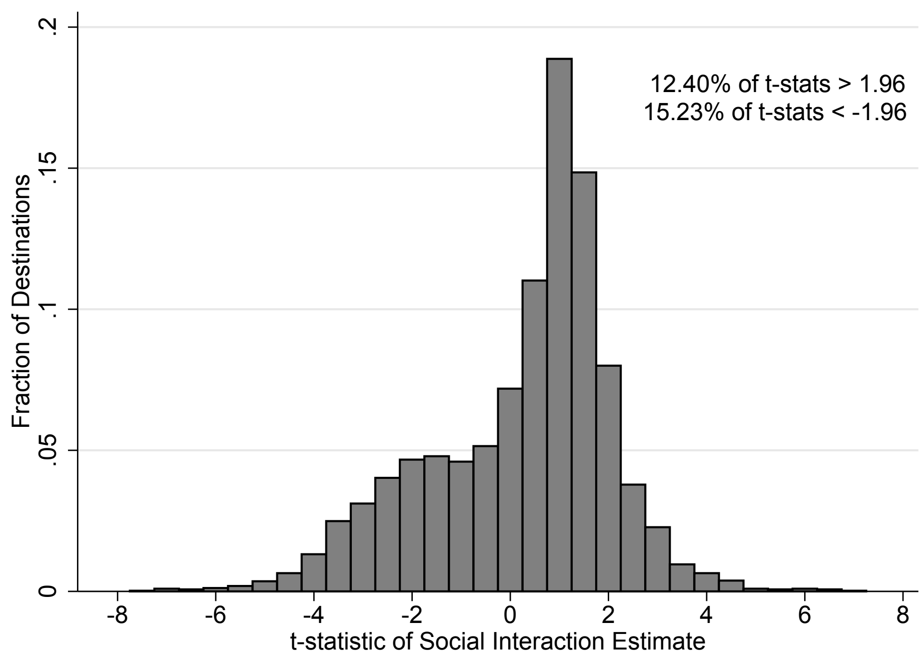
(b) Cumulative Distribution

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

Figure A.6: Distribution of Destination-Level Social Interactions Index t-statistics



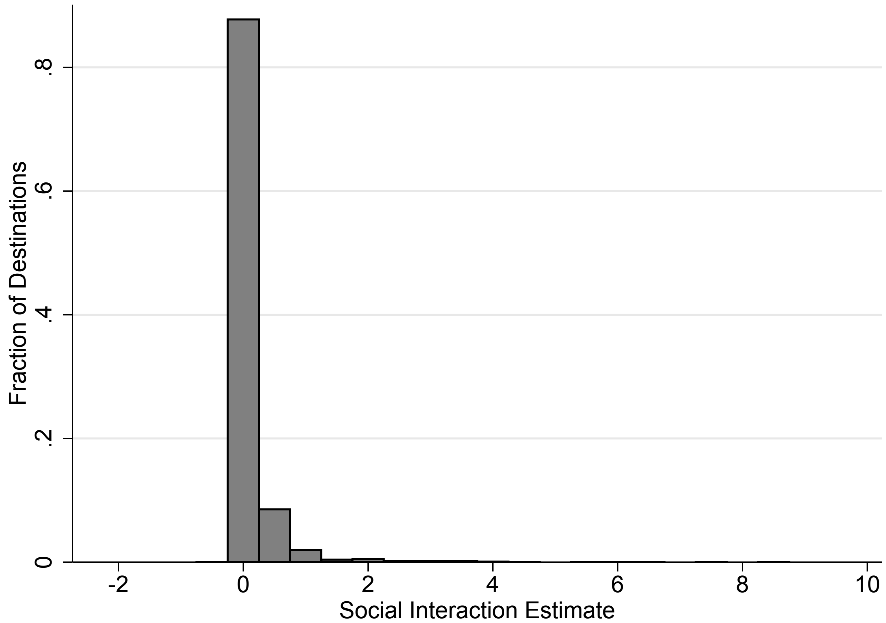
(a) Black Moves out of South



(b) White Moves out of Great Plains

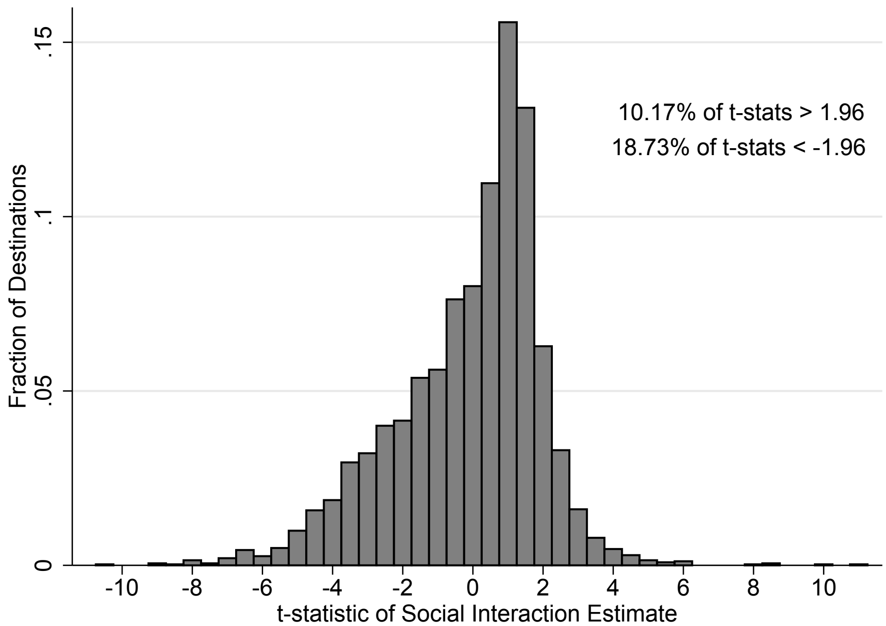
Notes: Bin width is 1/2. Birth town groups are defined by cross validation. Panel (a) omits the t-statistic of 13.7 from South Carolina to Hancock, WV.
Source: Duke SSA/Medicare data

Figure A.7: Distribution of Destination-Level Social Interactions Index Estimates, White Moves out of South



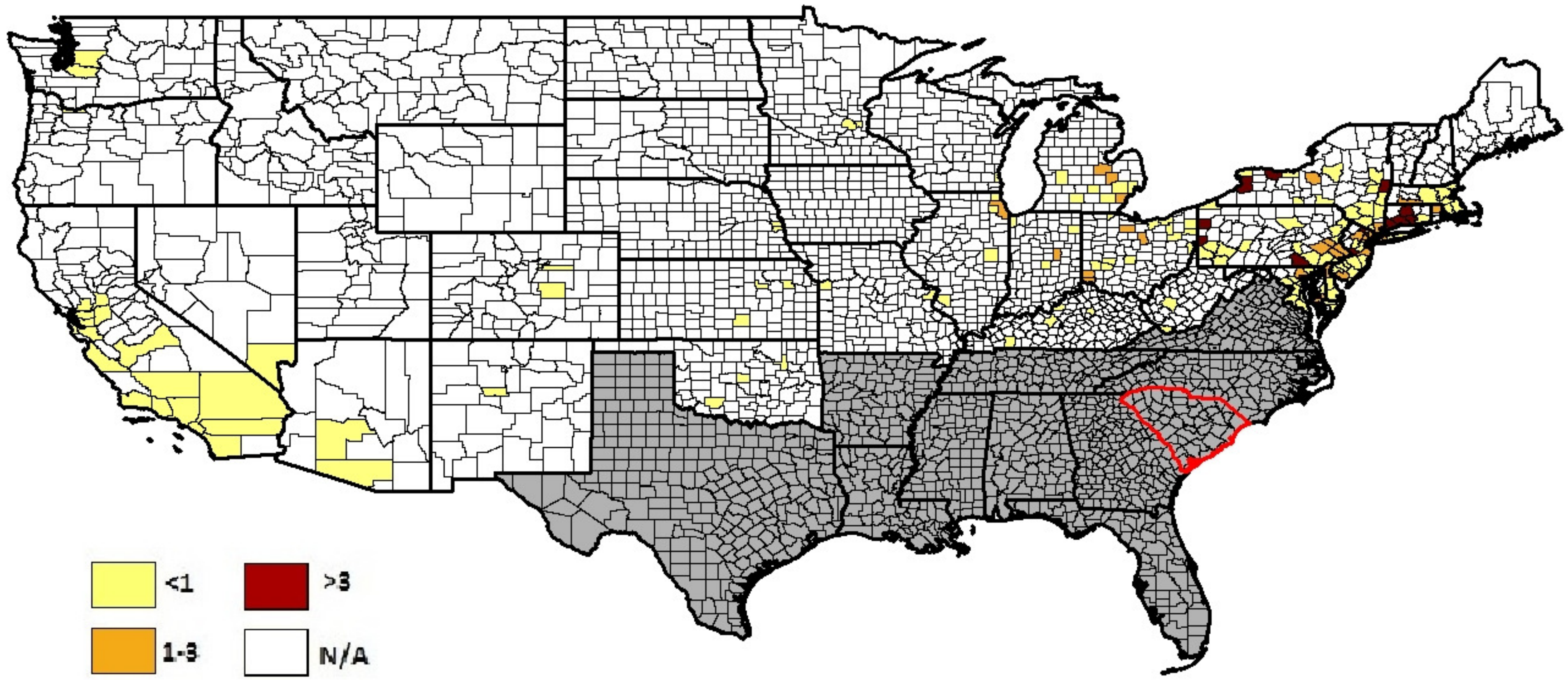
Notes: Bin width is 1/2. Figure omits estimate of $\hat{\Delta}_k = 19.3$ from Alabama to St. Joseph County, IN.
Source: Duke SSA/Medicare data

Figure A.8: Distribution of Destination-Level Social Interactions Index t-statistics, White Moves out of South



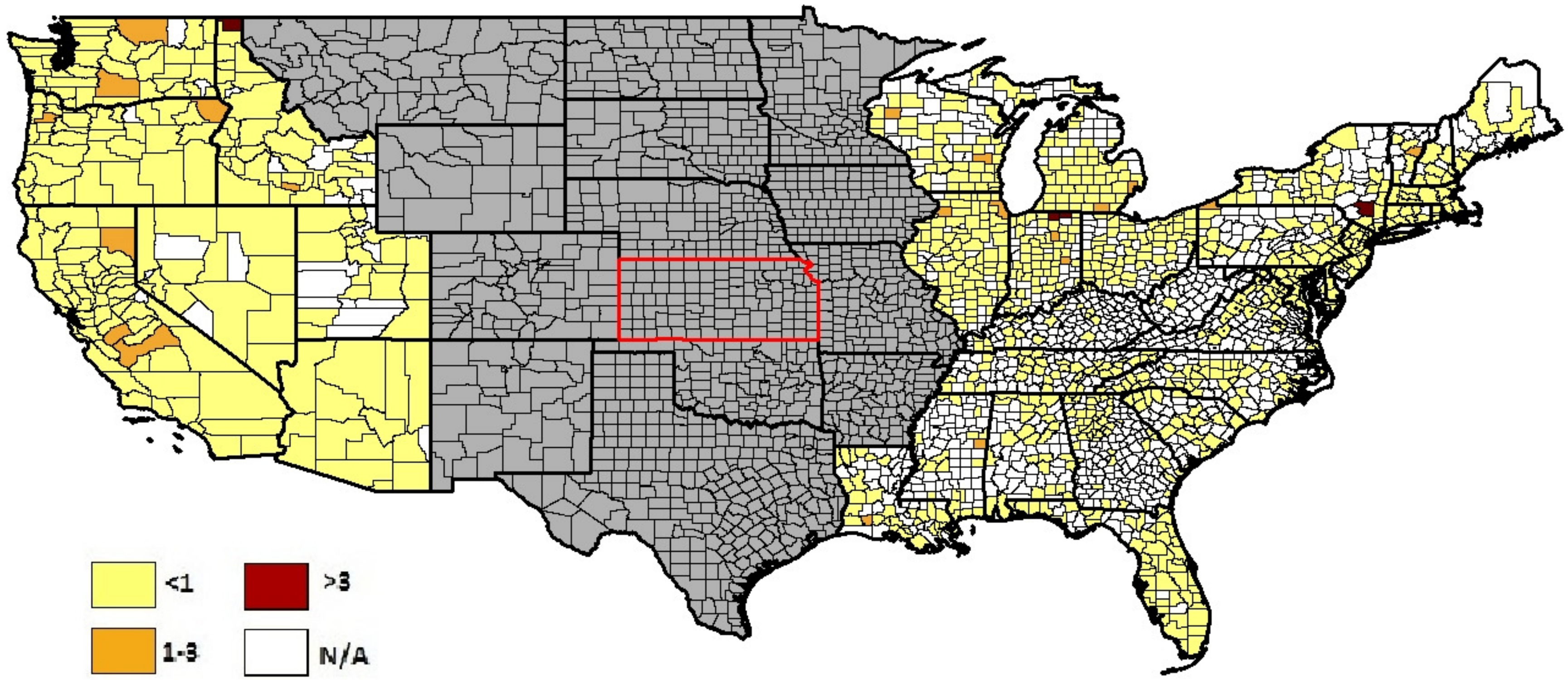
Note: Bin width is 1/2.
Source: Duke SSA/Medicare data

Figure A.9: Spatial Distribution of Destination-Level Social Interactions Index Estimates, South Carolina-born Blacks



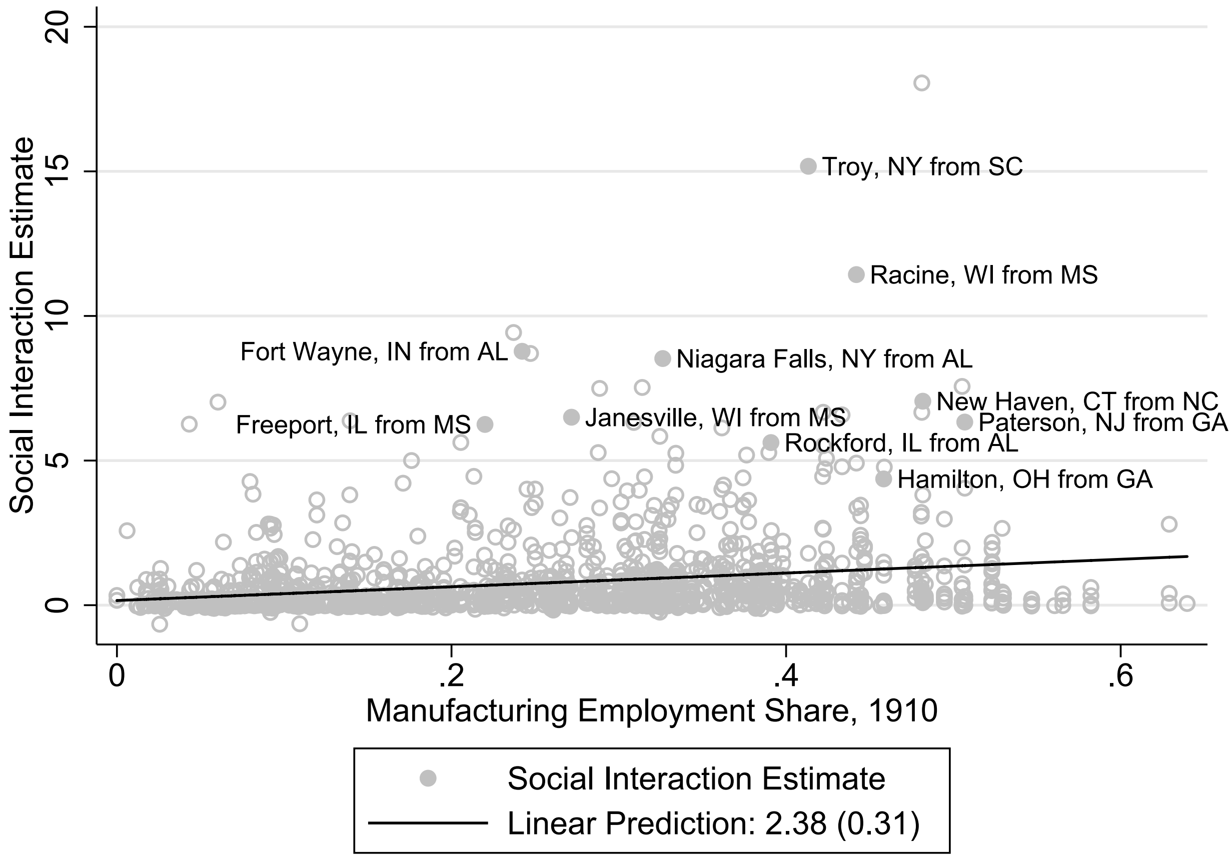
Notes: See note to Figure 2.6.

Figure A.10: Spatial Distribution of Destination-Level Social Interactions Index Estimates, Kansas-born Whites



Notes: See note to Figure 2.7.

Figure A.11: Relationship between Southern Black Destination-Level Social Interactions Index Estimates and 1910 Manufacturing Employment Share



Notes: Linear prediction comes from an OLS regression that includes a constant and 1910 manufacturing employment share. See Table 2.7 for results when including a richer set of covariates. Listed are the cities in Table 2.2.
Sources: Duke SSA/Medicare data and Haines and ICPSR (2010) data

B Appendix for Chapter 3

B.1 Theoretical Details

B.1.1 Proof of Proposition 1

To prove Proposition 3.3.2, we show that the assumptions of a stable equilibrium and non-negative peer effects (i.e., elements of J) imply that the peer effect multipliers m^s, m^n, and m^w are non-negative.

Let $\lambda_1, \lambda_2, \lambda_3$ be the eigenvalues of the 3×3 matrix J. The spectral radius of J is defined as $\rho(J) \equiv \max\{\lambda_1, \lambda_2, \lambda_3\}$. To ensure the equilibrium is stable, we assume that $\rho(J) < 1$. In each peer effect parametrization considered in Table 3.9, all eigenvalues are real and lie in $[0, 1)$, and this condition is satisfied.

The on-diagonal elements of J (J_{11}, J_{22}, J_{33}) are less than one in a stable equilibrium. This follows from the facts that the spectral radius is less than one if and only if $\lim_{k\to\infty} J^k = 0$ and $\lim_{k\to\infty} J^k = 0$ implies that the on-diagonal elements of J are less than one.

In a stable equilibrium, we also have that $\det(I - J) > 0$, where I is the 3×3 identity matrix. This follows from our assumption that $\rho(J) < 1$, the fact that $\det(J) = \lambda_1\lambda_2\lambda_3$, and the fact that $\det(J) = \lambda_1\lambda_2\lambda_3$ if and only if $\det(I - J) = (1 - \lambda_1)(1 - \lambda_2)(1 - \lambda_3)$.

It is straightforward to show that

$$\begin{aligned}
\det(I - J) &= (1 - J_{11})[(1 - J_{22})(1 - J_{33}) - J_{23}J_{32}] && \text{(B1)} \\
&\quad - J_{12}[J_{23}J_{31} + J_{21}(1 - J_{33})] - J_{13}[J_{21}J_{32} + J_{31}(1 - J_{22})] \\
&= (1 - J_{11})m^s - J_{12}m^n - J_{13}m^w, && \text{(B2)}
\end{aligned}$$

where the second equality uses the peer effect multipliers defined in equations (3.7)-(3.9). Because the off-diagonal elements of J are non-negative (by assumption) and the on-diagonal elements of J are less than 1 (as implied by a stable equilibrium), we have that m^n and m^w are non-negative.

As a result,

$$0 < \det(I - J) \leq (1 - J_{11})m^s. \tag{B3}$$

Because $J_{11} < 1$, this implies that m^s is non-negative. QED.

B.1.2 Discussion of Proposition 2

As noted in the text, two jointly sufficient conditions for Proposition 3.3.2 are (a): $d\bar{C}^s/d\text{HHI}^s < d\bar{C}^w/d\text{HHI}^s$ and (b): $d\bar{C}^n/d\text{HHI}^s \leq d\bar{C}^w/d\text{HHI}^s$. Assuming that $\partial F^s/\partial\text{HHI}^s < 0$, conditions (a) and (b) are equivalent to $m^s > m^w$ and $m^n \geq m^w$. Rearranging equations (3.7) and (3.9) shows that condition (a) is satisfied if and only if

$$(1 - J_{22})(1 - J_{33}) > J_{32}(J_{21} + J_{23}) + J_{31}(1 - J_{22}). \tag{B4}$$

The left hand side of inequality (B4) is positive because $J_{22}, J_{33} \in [0, 1)$ in a stable equilibrium with non-negative peer effects. Hence, condition (a) will be true as long as cross-group peer effects, on the right hand side, are small enough.

Similarly, equations (3.8) and (3.9) imply that condition (b) is satisfied if and only if

$$J_{21}(1 - J_{33} - J_{32}) \geq J_{31}(1 - J_{22} - J_{23}). \tag{B5}$$

If blacks with ties to the South have a larger peer effect on blacks without ties to the South than non-blacks, $J_{21} > J_{31} \geq 0$, then inequality (B5) is satisfied if $(J_{22} - J_{33}) + (J_{23} - J_{32}) \geq 0$, which will hold insofar as own-group peer effects among blacks without ties to the South are at least as strong as own-group peer effects among non-blacks ($J_{22} \geq J_{33}$) and an increase in the non-black crime rate leads to a greater increase in the crime rate among blacks without ties to the South than vice versa ($J_{32} \geq J_{23}$), which is plausible because baseline crime rates are higher among blacks than non-blacks.

It is useful to consider the simple case where there are no cross-group peer effects between black and non-black youth, $J_{13} = J_{23} = J_{31} = J_{32} = 0$. In this case, the peer effect multipliers are

$$m^s = \frac{1 - J_{22}}{(1 - J_{11})(1 - J_{22}) - J_{12}J_{21}} \tag{B6}$$

$$m^n = \frac{J_{21}}{(1 - J_{11})(1 - J_{22}) - J_{12}J_{21}} \tag{B7}$$

$$m^w = 0 \tag{B8}$$

In a stable equilibrium, $J_{22} \in [0, 1)$ and $(1 - J_{11})(1 - J_{22}) > J_{12}J_{21}$, ensuring that $m^s > m^w$ and condition (a) holds. Condition (b) additionally requires non-negative peer effects between blacks with and without ties to the South, $J_{21} \geq 0$.

With a 2×2 matrix (i.e., two types of agents), the result holds if and only if $m_{11} = (1 - J_{22})/\det(M) > 0$ and $m_{21} = J_{21}/\det(M) > 0$. Note that $\det(M) = \det(I - J) = 1 - tr(J) + \det(J) = (1 - \lambda_1)(1 - \lambda_2)$, where λ_1, λ_2 are the eigenvalues of J.

Suppose that stability means $\lambda_1, \lambda_2 \in (0, 1)$. [An alternative is $\lambda_1, \lambda_2 \in (-1, 1)$.] This is an intuitive definition when one thinks of eigenvalues as stretching matrices. If the equilibrium is stable, then $\det(M) > 0$.

If $J_{21} \geq 0$ and $J_{22} < 1$, then the result follows. The first assumption states that cross-group peer effects are non-negative. This is plausible, but other possibilities exist. The second assumption is an intuitive stability-like condition, seen in linear-in-means models.

We can calculate the eigenvalues:

$$\lambda_1 = \frac{1}{2}\left[J_{11} + J_{22} + \sqrt{(J_{11} - J_{22})^2 + 4J_{12}J_{21}}\right] \tag{B9}$$

$$\lambda_2 = \frac{1}{2}\left[J_{11} + J_{22} - \sqrt{(J_{11} - J_{22})^2 + 4J_{12}J_{21}}\right] \tag{B10}$$

If the above notion of stability is appropriate, we need to ensure that $\lambda_1 < 1$ and $\lambda_2 > 0$. This will be true for limited values of $(J_{11}, J_{12}, J_{21}, J_{22})$. In other words, the stability assumption imposes restrictions on J. We can figure out what parameter values are consistent with a stable equilibrium using the above inequalities.

In sum, the 2×2 case goes through if

$$\lambda_1, \lambda_2 \in (0, 1) \tag{B11}$$

$$J_{11}, J_{22} \in (0, 1) \tag{B12}$$

$$J_{12}, J_{21} \geq 0 \tag{B13}$$

These are in decreasing order of "bite." But a key takeaway is that, even in the 2×2 case, the result depends on parameter values.

B.2 Estimating a Model of Social Interactions in Location Decisions

Appendix B.2 describes a structural model of social interactions in location decisions. This model allows us to estimate the share of migrants that chose their destination because of social interactions. We include this variable in our regressions to examine whether the effect of social connectedness is driven by variation across cities in unobserved characteristics of migrants.

B.2.1 MODEL OF SOCIAL INTERACTIONS IN LOCATION DECISIONS

Migrants from birth town j are indexed on a line by $i \in \{1, \ldots, N_j\}$, where N_j is the total number of migrants from town j. For migrant i, destination k belongs to one of three preference groups: high (H_i), medium (M_i), or low (L_i). The high preference group contains a single destination. In the absence of social interactions, the destination in H_i is most preferred, and destinations in M_i are preferred over those in L_i.[57] A migrant never moves to a destination in L_i. A migrant chooses a destination in M_i if and only if his neighbor, $i-1$, chooses the same destination. A migrant chooses a destination in H_i if his neighbor chooses the same destination or his neighbor selects a destination in L_i.[58]

Migrants from the same birth town can differ in their preferences over destinations. The probability that destination k is in the high preference group for a migrant from town j is $h_{j,k} \equiv \mathbb{P}[k \in H_i | i \in j]$, and the probability that destination k is in the medium preference group is $m_{j,k} \equiv \mathbb{P}[k \in M_i | i \in j]$.

Migrants with many destinations in their medium preference group will tend to be influenced by the decisions of other migrants. For our empirical work, distinguishing between types of migrants is important because migrants that are more influenced by social interactions might differ along several dimensions. For example, migrants with many destinations in their medium preference group might be negatively selected in terms of earnings ability or be more pro-social, as discussed in the text.

The probability that migrant i moves to destination k given that his neighbor moves there is

$$\rho_{j,k} \equiv \mathbb{P}[D_{i,j,k} = 1 | D_{i-1,j,k} = 1] = \mathbb{P}[k \in H_i] + \mathbb{P}[k \in M_i] \tag{B14}$$

$$= h_{j,k} + m_{j,k}, \tag{B15}$$

where $D_{i,j,k}$ equals one if migrant i moves from j to k and zero otherwise.

The probability that destination k is in the medium preference group, conditional on not being in the high preference group, is $\nu_{j,k} \equiv \mathbb{P}[k \in M_i | k \notin H_i, i \in j]$. The conditional probability definition for $\nu_{j,k}$ implies that $m_{j,k} = \nu_{j,k}(1 - h_{j,k})$. We use $\nu_{j,k}$ to derive a simple sequential estimation approach.

[57] The assumption that H_i is a non-empty singleton ensures that migrant i has a well-defined location decision in the absence of social interactions. We could allow H_i to contain many destinations and specify a decision rule among the elements of H_i. This extension would complicate the model without adding any new insights.

[58] This model shares a similar structure as Glaeser, Sacerdote and Scheinkman (1996) in that some agents imitate their neighbors. However, we differ from Glaeser, Sacerdote and Scheinkman (1996) in that we model the interdependence between various destinations (i.e., this is a multinomial choice problem) and allow for more than two types of agents.

In equilibrium, the probability that a randomly chosen migrant i moves from j to k is

$$
\begin{aligned}
P_{j,k} \equiv \mathbb{P}[D_{i,j,k} = 1] &= \mathbb{P}[D_{i-1,j,k} = 1, k \in H_i] + \mathbb{P}[D_{i-1,j,k} = 1, k \in M_i] \\
&\quad + \sum_{k' \neq k} \mathbb{P}[D_{i-1,j,k'} = 1, k \in H_i, k' \in L_i] && \text{(B16)} \\
&= P_{j,k} h_{j,k} + P_{j,k} \nu_{j,k} (1 - h_{j,k}) + \sum_{k' \neq k} P_{j,k'} h_{j,k} (1 - \nu_{j,k'}) && \text{(B17)} \\
&= P_{j,k} \nu_{j,k} + \left(\sum_{k'=1}^{K} P_{j,k'} (1 - \nu_{j,k'}) \right) h_{j,k}. && \text{(B18)}
\end{aligned}
$$

The first term on the right hand side of equation (B16) is the probability that a migrant's neighbor moves to k, and k is in the migrant's high preference group; in this case, social interaction reinforces the migrant's desire to move to k. The second term is the probability that a migrant follows his neighbor to k because of social interactions. The third term is the probability that a migrant resists the pull of social interactions because town k is in the migrant's high preference group and the neighbor's chosen destination is in the migrant's low preference group.

The share of migrants from birth town j living in destination k that chose their destination because of social interactions equals $m_{j,k}$. As a result, the share of migrants in destination k that chose this destination because of social interactions is

$$
m_k \equiv \sum_j N_{j,k} m_{j,k}, \tag{B19}
$$

where $N_{j,k}$ is the number of migrants that moved from j to k. Our goal is to estimate m_k for each destination.

B.2.2 Estimation

To facilitate estimation, we connect this model to the social interactions (SI) index introduced by Stuart and Taylor (2017*a*). The SI index is the expected increase in the number of people from birth town j that move to destination k when an arbitrarily chosen person i is observed to make the same move,

$$
\Delta_{j,k} \equiv \mathbb{E}[N_{-i,j,k} | D_{i,j,k} = 1] - \mathbb{E}[N_{-i,j,k} | D_{i,j,k} = 0], \tag{B20}
$$

where $N_{-i,j,k}$ is the number of people who move from j to k, excluding person i. A positive value of $\Delta_{j,k}$ indicates positive social interactions in moving from j to k, while $\Delta_{j,k} = 0$ indicates the absence of social interactions. Stuart and Taylor (2017*a*) show that the SI index can be expressed as

$$\Delta_{j,k} = \frac{C_{j,k}(N_j - 1)}{P_{j,k}(1 - P_{j,k})}, \tag{B21}$$

where $C_{j,k}$ is the average covariance of location decisions between migrants from town j, $C_{j,k} \equiv \sum_{i \neq i' \in j} \mathbb{C}[D_{i,j,k}, D_{i',j,k}]/(N_j(N_j - 1))$. We follow the approach described in Stuart and Taylor (2017*a*) to estimate $P_{j,k}$ and $\Delta_{j,k}$ using information on migrants' location decisions from the Duke SSA/Medicare data.[59]

The model implies that $C_{j,k}$ equals[60]

$$C_{j,k} = \frac{2P_{j,k}(1 - P_{j,k}) \sum_{s=1}^{N_j-1} (N_j - s) \left(\frac{\rho_{j,k} - P_{j,k}}{1 - P_{j,k}} \right)^s}{N_j(N_j - 1)}. \tag{B22}$$

Substituting equation (B22) into equation (B21) and simplifying yields[61]

$$\Delta_{j,k} = \frac{2(\rho_{j,k} - P_{j,k})}{1 - \rho_{j,k}}, \tag{B23}$$

which can be rearranged to show that

$$\rho_{j,k} = \frac{2P_{j,k} + \Delta_{j,k}}{2 + \Delta_{j,k}}. \tag{B24}$$

We use equation (B24) to estimate $\rho_{j,k}$ with our estimates of $P_{j,k}$ and $\Delta_{j,k}$.

Equations (B15) and (B18), plus the fact that $m_{j,k} = \nu_{j,k}(1 - h_{j,k})$, imply that

$$\rho_{j,k} = \nu_{j,k} + \frac{P_{j,k}(1 - \nu_{j,k})^2}{\sum_{k'=1}^{K} P_{j,k'}(1 - \nu_{j,k'})}. \tag{B25}$$

We use equation (B25) to estimate $\nu_j \equiv (\nu_{j,1}, \ldots, \nu_{j,K})$ using our estimates of $(P_{j,1}, \ldots, P_{j,K}, \rho_{j,1}, \ldots, \rho_{j,K})$. We employ a computationally efficient algorithm that leverages the fact that equation (B25) is a quadratic equation in $\nu_{j,k}$, conditional on $\sum_{k'=1}^{K} P_{j,k'}(1 - \nu_{j,k'})$. We initially assume that $\sum_{k'=1}^{K} P_{j,k'}(1 - \nu_{j,k'}) = \sum_{k'=1}^{K} P_{j,k'} = 1$, then solve for $\nu_{j,k}$ using the quadratic formula, then construct an updated estimate of $\sum_{k'=1}^{K} P_{j,k'}(1 - \nu_{j,k'})$, and then solve again for $\nu_{j,k}$ using the quadratic formula. We require that each estimate of $\nu_{j,k}$ lies in $[0, 1]$. This iterated algorithm

[59] We use cross validation to define birth town groups. See Stuart and Taylor (2017*a*) for details.

[60] This follows from the fact that the covariance of location decisions for individuals i and $i + n$ is $\mathbb{C}[D_{i,j,k}, D_{i+n,j,k}] = P_{j,k}(1 - P_{j,k}) \left(\frac{\rho_{j,k} - P_{j,k}}{1 - P_{j,k}} \right)^n$.

[61] Equation (B23) results from taking the limit as $N_j \to \infty$, and so relies on N_j being sufficiently large.

converges very rapidly in the vast majority of cases.[62]

We use equation (B18) to estimate $h_{j,k}$ with our estimates of $\rho_{j,k}$ and $\nu_{j,k}$. Finally, we estimate $m_{j,k}$ using the fact that $m_{j,k} = \rho_{j,k} - h_{j,k}$. We use equation (B19) to estimate our parameter of interest, m_k, using estimates of $m_{j,k}$ and observed migration flows, $N_{j,k}$.

B.2.3 RESULTS

Appendix Figure B.2 displays a histogram of our estimates of the share of migrants that chose their destination because of social interactions, m_k, for cities in the North, Midwest, and West regions. The estimates range from 0 to 0.62. The unweighted average of m_k across cities is 0.26, and the 1980 population weighted average is 0.39.

Appendix Table B.10 examines the relationship between log HHI, the log number of migrants, and m_k. The raw correlation between log HHI and m_k is negative, but when we control for the log number of migrants, log HHI and m_k are positively correlated, as expected. This relationship is similar when including state fixed effects.

Appendix Figure B.3 further describes the relationship between log HHI and m_k. Panel A plots the unconditional relationship between log HHI and m_k, while Panel B plots the relationship conditional on the log number of migrants.[63] When we control for m_k in equation (3.12), we identify the effect of social connectedness on crime using variation in the vertical dimension of Panel B.

Conditional on the number of migrants in a destination and the share of migrants that chose their destination because of social interactions, variation in social connectedness continues to arise from concentrated birth town to destination city population flows. To see this, consider two hypothetical cities that each have 20 migrants, one-fourth of whom chose their destination because of social interactions. In the low HHI city, the 20 migrants come from five birth towns. Each town sends four migrants, one of whom moves there because of social interactions. As a result, $\text{HHI}_{\text{Low}} = 0.2$. In the high HHI city, the 20 migrants also come from five birth towns. One town sends 12 migrants, three of whom move there because of social interactions. Two towns each send two migrants, one of whom moves there because of social interactions, and two towns each send two migrants, neither of whom is influenced by social interactions. As a result, $\text{HHI}_{\text{High}} = 0.4$.[64]

[62]For 10 birth towns, the algorithm does not converge because our estimates of $P_{j,k}$ and $\rho_{j,k}$ do not yield a real solution to the quadratic formula. We examined the sensitivity of our results to these cases by (1) dropping birth towns for which the algorithm did not converge, (2) estimating $\nu_{j,k}$ and $\sum_{k'=1}^{K} P_{j,k'}(1-\nu_{j,k'})$ as the average of the values in the final four iterations, and (3) forcing $\hat{\nu}_{j,k}$ to equal zero for any (j,k) observation for which the quadratic formula solution does not exist. The motivation for (3) is that our estimates of $P_{j,k}$ and $\rho_{j,k}$ in these 10 cases were consistent with negative values of $\nu_{j,k}$, even though this was not a feasible solution. All three options yielded nearly identical estimates of our variable of interest, m_k. This is not surprising because these 10 birth towns account for a negligible share of the over 5,000 birth towns used to estimate m_k.

[63]In particular, Panel B plots the residuals from regression log HHI and m_k on the log number of migrants.

[64]Alternatively, suppose that in the high HHI city, the 20 migrants come from three birth towns. One town sends

This example is consistent with Figure 3.2 in that variation in social connectedness arises from the top sending town.

The structural model features local social interactions: each migrant directly influences no more than one migrant.[65] As a result, the model does not distinguish between the case where 12 migrants come from one town, with three migrants influenced by social interactions, and the case where 12 migrants come from three towns, with three migrants influenced by social interactions. Although this simple model does not capture all possible forms of social interactions, we believe that it likely captures the most relevant threats to our empirical strategy for this paper.

B.3 Details on Peer Effect Parametrization

Appendix B.3 provides additional details on the literature that guides our parametrization of peer effects in Section 3.6.

Case and Katz (1991) find that a one percent increase in the neighborhood crime rate leads to a 0.1 percent increase in a Boston youth's self-reported propensity of committing a crime during the last year (Table 10). This implies that a one percentage point increase in the neighborhood crime rate leads to a 0.1 percentage point increase in youth's crime rate, suggesting on-diagonal elements of J close to 0.1.

Glaeser, Sacerdote and Scheinkman (1996) estimate a local social interactions model in which there are two types of agents. Fixed agents are not affected by their peers, and compliers imitate their neighbor.[66] The probability that an agent is a complier thus maps to the on-diagonal elements of J. In Table IIA, the authors report estimates of $f(\pi) = (2-\pi)/\pi$, where π is the probability that an agent is a fixed type. The probability that an agent is a complier is $1 - \pi = 1 - 2/(1 + f(\pi))$. Using FBI UCR data on murders across cities for 1970 and 1985, Glaeser, Sacerdote and Scheinkman (1996) report estimates of $f(\pi)$ between 2 and 4.5, implying on-diagonal elements of J between 1/3 and 2/3. For robbery and motor vehicle theft, the authors estimate $f(\pi)$ in the range of 37-155 and 141-382, suggesting diagonal elements of J very close to 1.

Ludwig and Kling (2007) find no evidence that neighborhood violent crime rates affect violent crime arrests among MTO participants age 15-25 (Table 4). These estimates suggest on-diagonal elements of J close to zero.

Damm and Dustmann (2014) estimate the effect of municipality crime rates on refugees' criminal convictions in Denmark. For males, they find that a one percentage point increase in the municipality crime rate leads to a 7-13 percent increase in the probability of conviction over a

12 migrants, three of whom move there because of social interactions, and two towns each send four migrants, one of whom moves there because of social interactions. As a result, $\text{HHI}_{\text{High}} = 0.44$.

[65] However, a single migrant can indirectly influence several migrants.

[66] Their model is similar to the one described in Appendix B.2.

seven year period from ages 15-21 (Table 3, also see p. 1820). Given an average conviction rate of 46 percent, this translates into a 3-6 percentage point increase in the probability of conviction; we take the midpoint of 4.5. For females, the municipality crime rate has no effect on convictions. Consequently, these estimates imply that a one percentage point increase in the municipality crime rate leads to a $(0.5 \cdot 4.5)/7 \approx 0.32$ percentage point increase in refugees' annual conviction rate. This suggests on-diagonal elements of J close to 1/3. Damm and Dustmann (2014) find that, beyond the impact of the municipality crime rate, the crime rate of co-nationals has an additional impact while the crime rate of immigrants from other countries does not (Table 7). This suggests that cross-group peer effects might be small.

In sum, estimates from Case and Katz (1991) suggest on-diagonal values of J close to 0.1, estimates from Glaeser, Sacerdote and Scheinkman (1996) suggest on-diagonal elements of J close to 0.5 for murder, estimates from Ludwig and Kling (2007) suggest on-diagonal elements of J close to zero, and estimates from Damm and Dustmann (2014) suggest on-diagonal values of J close to 0.3 and off-diagonal elements near zero.

Table B.1: Summary Statistics: Crime and Social Connectedness, 1960-2009

	Mean	SD	First Quartile	Third Quartile	Fraction Zero
Offenses reported to police per 100,000 residents					
Murder	6.7	8.8	1.7	8.7	0.184
Rape	29	28	10	40	0.070
Robbery	215	252	68	270	0.004
Assault	1,134	1,099	287	1,622	0.005
Burglary	1,234	846	670	1,630	0.000
Larceny	3,228	1,785	2,023	4,198	0.000
Motor Vehicle Theft	582	513	260	742	0.000
Population	93,074	94,505	39,476	104,217	-
HHI, Southern Black Migrants	0.020	0.016	0.008	0.028	-
Log HHI, Southern Black Migrants	-4.220	0.781	-4.852	-3.563	-
Top Sending Town Share, Southern Black Migrants	0.061	0.041	0.036	0.074	-
Number, Southern Black Migrants	630	1,315	58	596	-

Notes: Each observation is a city-year. HHI and migrant counts are calculated among all individuals born in the former Confederacy states from 1916-1936. Data on rape is only available starting in 1964. Sample is restricted to cities with less than 500,000 residents in 1980.
Sources: FBI UCR, Duke SSA/Medicare dataset

Table B.2: Summary Statistics: Cities' Average Crime Rates

			Percentile				
	Mean	SD	5	25	50	75	95
Murder	6.7	6.8	1.3	2.7	4.5	8.0	19.2
Rape	29.1	18.3	6.5	16.0	26.3	36.9	65.8
Robbery	212.6	183.1	41.9	93.0	153.0	269.1	611.5
Assault	1,121.6	626.5	326.7	647.5	1,013.1	1,469.5	2,320.4
Burglary	1,233.1	474.0	541.8	891.9	1,185.3	1,510.2	2,095.9
Larceny	3,221.5	1,213.2	1,517.0	2,351.4	3,186.4	3,918.5	5,030.8
Motor Vehicle Theft	576.9	369.8	178.7	309.4	460.6	746.6	1,300.1

Notes: For each city, we construct an average crime rate across years 1960-2009. Table B.2 reports summary statistics of these average crime rates. Sample is restricted to cities with less than 500,000 residents in 1980.
Sources: FBI UCR

Table B.3: Summary Statistics: Cities With and Without 1911-1916 Homicide Rates

	1911-1916 Homicide Rates Observed	
	Yes	No
	(1)	(2)
HHI, Southern black migrants	0.007	0.021
	(0.006)	(0.016)
Number, Southern black migrants	7,999	540
	(16,068)	(2,079)
Population, 1980	549,344	80,839
	(1,099,422)	(170,680)
Percent black, 1980	0.237	0.103
	(0.152)	(0.148)
Percent female, 1980	0.530	0.519
	(0.008)	(0.019)
Percent 25+ with HS, 1980	0.489	0.560
	(0.080)	(0.098)
Percent 25+ with College, 1980	0.118	0.137
	(0.048)	(0.078)
Log area, square miles, 1980	3.886	2.729
	(0.986)	(0.888)
Log median family income, 1979	10.85	11.06
	(0.148)	(0.205)
Unemployment rate, 1980	0.0886	0.0708
	(0.033)	(0.030)
Labor force participation rate, 1980	0.458	0.483
	(0.041)	(0.052)
Manufacturing emp. share, 1980	0.213	0.233
	(0.072)	(0.094)
N (cities)	46	369

Notes: Table reports means and, in parentheses, standard deviations. Column 1 contains cities in the North, Midwest, and West regions that are in our main analysis sample and for which we observe homicide rates for at least one year from 1911-1916. These cities have at least 100,000 residents in 1920 and at least 5 deaths each year. Column 2 contains cities in the North, Midwest, and West regions that are in our main analysis sample but for which we do not observe homicide rates from 1911-1916. Unlike our main analysis sample, we do not restrict to cities with fewer than 500,000 residents in 1980.
Sources: Census (1922, p. 64-65) , Duke SSA/Medicare data, Census city data book

Table B.4: The Relationship between Social Connectedness and City Covariates, 1960-2009, Including African American-Specific Covariates

	Dependent variable: Log HHI, Southern black migrants			
Year covariates are measured:	1970	1980	1990	2000
	(1)	(2)	(3)	(4)
Log number, Southern black migrants	-0.806***	-0.779***	-0.744***	-0.750***
	(0.068)	(0.076)	(0.088)	(0.097)
Log population	-0.006	0.002	-0.022	0.025
	(0.073)	(0.078)	(0.089)	(0.089)
Percent black	-0.018	-0.000	-0.035	-0.075
	(0.059)	(0.077)	(0.073)	(0.066)
Percent female	-0.074	0.025	-0.008	0.018
	(0.060)	(0.079)	(0.089)	(0.076)
Percent age 5-17	-0.080	0.141	0.463*	0.448
	(0.226)	(0.262)	(0.267)	(0.332)
Percent age 18-64	-0.140	0.179	0.500*	0.577
	(0.235)	(0.277)	(0.280)	(0.365)
Percent age 65+	0.007	0.218	0.440**	0.444**
	(0.162)	(0.214)	(0.207)	(0.224)
Percent with high school degree	0.065	-0.131	0.017	-0.015
	(0.132)	(0.107)	(0.091)	(0.101)
Percent with college degree	0.027	0.017	-0.007	-0.016
	(0.073)	(0.054)	(0.082)	(0.086)
Log area, square miles	0.021	-0.028	-0.013	-0.028
	(0.062)	(0.070)	(0.077)	(0.083)
Log median family income	-0.075	-0.011	-0.202**	-0.067
	(0.096)	(0.089)	(0.099)	(0.082)
Unemployment rate	0.176**	-0.025	-0.070	0.029
	(0.083)	(0.087)	(0.092)	(0.058)
Labor force participation rate	0.073	0.007	0.085	-0.035
	(0.052)	(0.088)	(0.105)	(0.056)
Manufacturing employment share	0.203***	0.165***	0.163***	0.191***
	(0.065)	(0.059)	(0.053)	(0.047)
African American-Specific Covariates:				
Percent female	0.040	-0.085	0.012	0.077
	(0.046)	(0.062)	(0.074)	(0.072)
Percent age 5-17	0.122	0.098	0.160	-0.114
	(0.078)	(0.115)	(0.152)	(0.174)
Percent age 18-64	0.130	0.034	0.215	-0.025
	(0.088)	(0.131)	(0.180)	(0.212)

Table B.4: The Relationship between Social Connectedness and City Covariates, 1960-2009, Including African American-Specific Covariates, cont.

	Dependent variable: Log HHI, Southern black migrants			
Year covariates are measured:	1970	1980	1990	2000
	(1)	(2)	(3)	(4)
Percent age 65+	0.044	0.044	0.093	-0.017
	(0.055)	(0.070)	(0.087)	(0.103)
Percent with high school degree	-0.195***	-0.060	-0.112	-0.033
	(0.074)	(0.075)	(0.076)	(0.074)
Percent with college degree	0.160***	0.122*	0.125	0.059
	(0.053)	(0.064)	(0.079)	(0.079)
Unemployment rate	-0.083*	0.065	0.119**	0.101**
	(0.048)	(0.074)	(0.059)	(0.041)
State fixed effects	x	x	x	x
Adjusted R2	0.773	0.757	0.763	0.771
N (cities)	228	228	228	228
p-value: Wald test that parameters equal zero				
Demographic covariates	0.909	0.604	0.434	0.041
Economic covariates	0.023	0.990	0.220	0.521
African American-specific covariates	0.001	0.274	0.389	0.131

Notes: African American-specific covariates are not available for 1960. See note to Table 3.3.
Sources: Duke SSA/Medicare data, Census city data book, NHGIS

Table B.5: The Relationship between Social Connectedness and Measures of Social Capital

	Dependent variable: Log HHI, Southern black migrants							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: All Cities								
Associational density	0.0818	0.0601					0.135	0.109*
	(0.0571)	(0.0489)					(0.0908)	(0.0594)
Social capital index			0.0469	-0.00181			-0.0645	-0.0783
			(0.0558)	(0.0510)			(0.0920)	(0.0633)
Social capital composite index					0.0378	-0.00995		
					(0.0547)	(0.0477)		
Log number, Southern black migrants		-0.850***		-0.852***		-0.852***		-0.851***
		(0.0330)		(0.0324)		(0.0324)		(0.0331)
State fixed effects		x		x		x		x
R2	0.007	0.741	0.002	0.739	0.001	0.740	0.008	0.742
N (cities)	490	490	490	490	490	490	490	490
Counties	227	227	227	227	227	227	227	227
Panel B: Cities with Above Median Black Population Share in 1990								
Associational density	0.309***	0.118					0.514***	0.213**
	(0.0645)	(0.0746)					(0.103)	(0.103)
Social capital index			0.189***	0.0367			-0.264***	-0.149
			(0.0579)	(0.0767)			(0.0957)	(0.0979)
Social capital composite index					0.170***	0.0225		
					(0.0563)	(0.0719)		
Log number migrants		-0.629***		-0.653***		-0.655***		-0.621***
		(0.0600)		(0.0562)		(0.0559)		(0.0595)
State fixed effects		x		x		x		x
R2	0.129	0.598	0.043	0.591	0.034	0.590	0.155	0.603
N (cities)	229	229	229	229	229	229	229	229
Counties	152	152	152	152	152	152	152	152

Notes: All variables are normalized to have mean zero and standard deviation one in the sample used in Panel A. See Rupasingha and Goetz (2008) for definitions of associational density and social capital indices, which are measured at the county level using data from 1988 and 1990. The correlation between the social capital index and the social capital composite index is 0.99. Sample limited to cities with at least 25,000 residents in each decade and which received at least 25 Southern black migrants in the Duke dataset. Standard errors, clustered at the county level, are in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Sources: Duke SSA/Medicare data, Rupasingha and Goetz (2008)

Table B.6: The Effect of Social Connectedness on Crime, 1960-2009, Results for All Explanatory Variables

	Dependent variable: Number of offenses reported to police						
	Murder (1)	Rape (2)	Robbery (3)	Assault (4)	Burglary (5)	Larceny (6)	Motor Vehicle Theft (7)
Log HHI, Southern black migrants	-0.181***	-0.083**	-0.251***	-0.142***	-0.095***	-0.049	-0.163***
	(0.034)	(0.035)	(0.035)	(0.042)	(0.022)	(0.030)	(0.041)
Log number, Southern black migrants	0.150***	0.060**	0.146***	0.075**	0.051***	0.038	0.041
	(0.022)	(0.027)	(0.027)	(0.029)	(0.018)	(0.024)	(0.029)
Log population	0.944***	0.837***	1.118***	0.864***	0.947***	0.871***	1.273***
	(0.053)	(0.042)	(0.052)	(0.049)	(0.030)	(0.042)	(0.053)
Percent black, 1960	2.615***	3.717***	2.703***	3.520***	1.683***	0.588	1.585***
	(0.394)	(0.488)	(0.422)	(0.541)	(0.378)	(0.412)	(0.400)
Percent black, 1970	1.898***	2.512***	1.522***	0.890***	0.904***	0.066	1.204***
	(0.225)	(0.248)	(0.223)	(0.298)	(0.161)	(0.255)	(0.268)
Percent black, 1980	1.598***	1.556***	1.184***	0.592**	0.315**	-0.177	0.872***
	(0.167)	(0.162)	(0.192)	(0.265)	(0.140)	(0.243)	(0.235)
Percent black, 1990	1.544***	0.730***	0.737***	0.183	0.060	-0.085	0.616**
	(0.205)	(0.216)	(0.201)	(0.238)	(0.165)	(0.303)	(0.291)
Percent black, 2000	1.880***	0.117	0.418*	-0.132	0.127	-0.447*	0.890***
	(0.226)	(0.234)	(0.234)	(0.218)	(0.174)	(0.265)	(0.246)
Percent female, 1960	-0.235	2.965	-2.321	1.183	3.846	1.469	1.113
	(3.323)	(3.972)	(4.267)	(4.217)	(2.643)	(2.287)	(3.278)
Percent female, 1970	1.142	2.396	-0.379	-5.374*	-0.069	-0.241	1.260
	(1.880)	(1.971)	(2.195)	(2.880)	(1.258)	(1.451)	(2.595)
Percent female, 1980	-1.743	-1.131	-1.689	-4.141	1.588	-2.773	-0.973
	(2.047)	(2.317)	(2.549)	(3.038)	(1.574)	(2.143)	(3.114)
Percent female, 1990	-3.829	-2.197	0.538	-1.329	1.103	-1.298	4.573
	(2.706)	(2.904)	(3.728)	(2.574)	(2.226)	(2.251)	(4.266)
Percent female, 2000	4.335	1.984	-0.818	3.643*	-1.443	-0.649	-2.015
	(3.008)	(2.383)	(2.603)	(1.959)	(1.611)	(1.809)	(3.010)
Percent age 5-17, 1960	-1.476	-18.408***	0.751	-16.009**	1.816	-7.283**	6.275
	(5.192)	(5.431)	(6.667)	(6.454)	(3.536)	(3.305)	(4.411)
Percent age 18-64, 1960	-1.143	-11.610**	4.168	-8.046	1.531	-6.607***	5.548*
	(4.056)	(4.685)	(5.295)	(4.982)	(2.750)	(2.448)	(3.371)

Table B.6: The Effect of Social Connectedness on Crime, 1960-2009, Results for All Explanatory Variables, cont.

	Dependent variable: Number of offenses reported to police						
	Murder	Rape	Robbery	Assault	Burglary	Larceny	Motor Vehicle Theft
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Percent age 65+, 1960	-2.843	-13.297***	0.545	-14.016***	-0.145	-5.873***	0.182
	(3.270)	(3.851)	(4.903)	(4.282)	(2.601)	(2.092)	(3.248)
Percent age 5-17, 1970	-6.603**	-9.194***	-7.336**	-7.073	-3.975**	-3.004	-0.515
	(2.937)	(2.969)	(3.033)	(4.493)	(1.811)	(2.151)	(3.307)
Percent age 18-64, 1970	-3.771	-4.638*	-3.465	-7.827*	-4.797***	-3.551*	1.888
	(2.705)	(2.514)	(2.751)	(4.153)	(1.588)	(1.913)	(2.775)
Percent age 65+, 1970	-4.117*	-7.088***	-4.046*	-5.228	-3.043**	-2.272	-1.347
	(2.255)	(2.167)	(2.413)	(3.303)	(1.414)	(1.600)	(2.709)
Percent age 5-17, 1980	-8.082***	-10.612***	-3.334	-12.578***	-6.098**	1.356	11.437***
	(2.917)	(2.932)	(4.021)	(4.662)	(2.709)	(4.058)	(4.338)
Percent age 18-64, 1980	-9.361***	-8.200***	-3.751	-11.294***	-5.998***	-0.036	8.985***
	(2.162)	(2.090)	(2.854)	(3.314)	(1.903)	(2.330)	(3.225)
Percent age 65+, 1980	-4.834**	-7.669***	-0.178	-7.982**	-3.899**	2.976	10.184***
	(2.421)	(2.327)	(3.241)	(3.659)	(1.902)	(3.708)	(3.435)
Percent age 5-17, 1990	-17.701***	-9.090**	-7.317*	-8.706*	-4.683*	1.342	6.294
	(4.289)	(4.108)	(4.114)	(4.456)	(2.632)	(3.324)	(5.232)
Percent age 18-64, 1990	-14.688***	-7.455***	-4.407*	-7.640**	-6.078***	0.464	6.159*
	(2.996)	(2.697)	(2.587)	(3.152)	(1.865)	(2.536)	(3.250)
Percent age 65+, 1990	-10.878***	-6.553**	-3.425	-6.599**	-3.676*	2.157	5.563
	(3.419)	(3.059)	(3.106)	(3.335)	(1.923)	(2.183)	(3.845)
Percent age 5-17, 2000	-4.741	-9.525*	-2.977	-0.087	6.760**	2.669	8.752*
	(5.067)	(5.145)	(4.226)	(4.047)	(3.400)	(4.091)	(5.190)
Percent age 18-64, 2000	-5.702	-6.522	-2.049	-1.315	5.537**	2.441	9.519**
	(3.819)	(4.205)	(3.511)	(3.163)	(2.731)	(3.220)	(4.100)

Table B.6: The Effect of Social Connectedness on Crime, 1960-2009, Results for All Explanatory Variables, cont.

	Dependent variable: Number of offenses reported to police						
	Murder	Rape	Robbery	Assault	Burglary	Larceny	Motor Vehicle Theft
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Percent age 65+, 2000	-4.116	-6.737*	-1.575	0.202	6.110**	2.847	7.808**
	(3.921)	(3.900)	(3.226)	(3.061)	(2.590)	(3.131)	(3.827)
Percent with high school degree, 1960	-1.444**	-0.341	0.134	-0.178	0.041	-0.487	-1.186
	(0.631)	(0.651)	(0.878)	(0.806)	(0.572)	(0.663)	(0.722)
Percent with high school degree, 1970	-2.494***	-1.387***	-1.844***	-3.207***	-0.832**	-0.171	-2.596***
	(0.566)	(0.499)	(0.570)	(0.616)	(0.325)	(0.385)	(0.667)
Percent with high school degree, 1980	-2.298***	-0.405	-1.495**	-1.244*	-1.413***	-1.353**	-1.145*
	(0.528)	(0.472)	(0.653)	(0.673)	(0.335)	(0.553)	(0.625)
Percent with high school degree, 1990	-1.893***	1.513***	-1.325**	1.097**	0.841**	0.542	-1.125*
	(0.470)	(0.466)	(0.531)	(0.500)	(0.366)	(0.420)	(0.645)
Percent with high school degree, 2000	-1.397***	2.796***	-0.705	1.561***	1.419***	1.033***	-0.636
	(0.507)	(0.530)	(0.553)	(0.456)	(0.370)	(0.390)	(0.609)
Percent with college degree, 1960	-0.425	1.146	-1.973*	-0.447	0.849	2.421***	0.168
	(1.061)	(1.349)	(1.178)	(1.405)	(0.793)	(0.698)	(1.187)
Percent with college degree, 1970	-0.308	1.252**	-0.221	2.245***	1.548***	1.605***	0.316
	(0.765)	(0.609)	(0.764)	(0.686)	(0.370)	(0.387)	(0.802)
Percent with college degree, 1980	0.420	0.032	0.187	0.244	0.875***	1.434***	-1.306*
	(0.482)	(0.484)	(0.596)	(0.700)	(0.326)	(0.402)	(0.725)
Percent with college degree, 1990	-0.324	-0.574	-0.046	-0.661*	0.725***	0.911***	-1.505***
	(0.414)	(0.376)	(0.373)	(0.361)	(0.281)	(0.285)	(0.548)
Percent with college degree, 2000	0.035	-1.091**	-0.081	-0.320	-0.065	0.615*	-2.208***
	(0.456)	(0.501)	(0.448)	(0.422)	(0.339)	(0.320)	(0.621)
Log area, square miles, 1960	-0.004	0.282***	-0.108	0.080	0.048	0.060	-0.169***
	(0.059)	(0.058)	(0.084)	(0.070)	(0.043)	(0.045)	(0.058)

Table B.6: The Effect of Social Connectedness on Crime, 1960-2009, Results for All Explanatory Variables, cont.

	Dependent variable: Number of offenses reported to police						
	Murder	Rape	Robbery	Assault	Burglary	Larceny	Motor Vehicle Theft
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log area, square miles, 1970	0.042	0.270***	-0.136**	0.127***	0.063***	0.090**	-0.218***
	(0.052)	(0.040)	(0.053)	(0.047)	(0.024)	(0.039)	(0.048)
Log area, square miles, 1980	0.098*	0.272***	-0.105*	0.127***	0.086***	0.133***	-0.186***
	(0.051)	(0.038)	(0.056)	(0.044)	(0.026)	(0.034)	(0.052)
Log area, square miles, 1990	0.092*	0.183***	-0.126**	0.113***	0.081***	0.125***	-0.054
	(0.047)	(0.040)	(0.053)	(0.043)	(0.029)	(0.037)	(0.052)
Log area, square miles, 2000	0.067	0.121***	-0.188***	0.098**	0.061**	0.106***	-0.127***
	(0.049)	(0.040)	(0.048)	(0.042)	(0.029)	(0.040)	(0.044)
Log median family income, 1960	-1.335**	-0.763	-0.736	-0.848	-1.117***	-0.585*	-0.477
	(0.527)	(0.665)	(0.686)	(0.688)	(0.371)	(0.325)	(0.537)
Log median family income, 1970	-0.434	-0.983***	-0.264	-0.049	-0.757***	-0.848***	0.635
	(0.298)	(0.294)	(0.369)	(0.373)	(0.196)	(0.198)	(0.388)
Log median family income, 1980	-0.783***	-1.525***	-0.953***	-0.468	-0.377*	-0.866***	0.028
	(0.216)	(0.241)	(0.342)	(0.361)	(0.217)	(0.235)	(0.355)
Log median family income, 1990	-0.512**	-1.912***	-1.030***	-1.319***	-1.215***	-1.517***	-0.280
	(0.260)	(0.240)	(0.315)	(0.260)	(0.165)	(0.186)	(0.382)
Log median family income, 2000	-1.281***	-2.149***	-1.227***	-1.722***	-1.310***	-1.216***	-0.616***
	(0.189)	(0.197)	(0.152)	(0.174)	(0.153)	(0.160)	(0.229)
Unemployment rate, 1960	-0.628	2.086	6.734**	3.018	2.871	2.433	1.905
	(2.272)	(3.165)	(3.431)	(3.369)	(2.125)	(1.977)	(2.538)
Unemployment rate, 1970	-0.603	-1.855	0.905	1.376	-0.356	-0.128	0.883
	(1.686)	(1.635)	(2.171)	(2.114)	(1.257)	(1.270)	(2.256)
Unemployment rate, 1980	1.473	2.048*	-0.629	2.811*	2.180**	2.787***	1.122
	(1.306)	(1.132)	(1.503)	(1.534)	(0.977)	(0.895)	(1.801)

Table B.6: The Effect of Social Connectedness on Crime, 1960-2009, Results for All Explanatory Variables, cont.

	Dependent variable: Number of offenses reported to police						
	Murder	Rape	Robbery	Assault	Burglary	Larceny	Motor Vehicle Theft
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Unemployment rate, 1990	6.720***	0.768	2.448*	0.672	3.206**	-1.041	2.081
	(2.130)	(1.735)	(1.451)	(1.651)	(1.247)	(1.658)	(2.566)
Unemployment rate, 2000	-1.312	-1.369	-2.271*	0.627	2.313**	2.087*	-0.583
	(1.587)	(1.384)	(1.285)	(0.932)	(1.072)	(1.104)	(1.107)
Labor force participation rate, 1960	4.029*	3.201	5.054**	4.236**	3.114**	2.727***	2.575
	(2.162)	(2.349)	(2.143)	(2.016)	(1.291)	(0.989)	(1.599)
Labor force participation rate, 1970	1.072	1.114	2.498**	3.674***	1.987***	1.827**	0.845
	(1.102)	(0.911)	(1.260)	(1.398)	(0.623)	(0.760)	(1.283)
Labor force participation rate, 1980	2.912***	3.393***	3.105**	3.142**	2.077***	4.067***	1.398
	(1.012)	(0.945)	(1.351)	(1.506)	(0.668)	(1.138)	(1.370)
Labor force participation rate, 1990	2.653***	2.965***	3.234**	2.009**	2.280***	3.077***	1.682
	(0.985)	(1.017)	(1.401)	(0.966)	(0.765)	(0.833)	(1.559)
Labor force participation rate, 2000	0.545	1.144***	1.137***	1.371***	0.223	1.266***	0.238
	(0.429)	(0.372)	(0.388)	(0.300)	(0.302)	(0.325)	(0.482)
Manufacturing employment share, 1960	0.022	0.724	0.969**	1.489***	0.314	-0.069	0.000
	(0.344)	(0.451)	(0.479)	(0.515)	(0.308)	(0.280)	(0.405)
Manufacturing employment share, 1970	0.058	0.476	0.170	0.141	0.062	-0.161	-0.398
	(0.292)	(0.293)	(0.340)	(0.430)	(0.192)	(0.230)	(0.321)
Manufacturing employment share, 1980	0.619**	0.063	0.239	-0.049	-0.300	-0.832**	0.106
	(0.298)	(0.278)	(0.377)	(0.463)	(0.259)	(0.419)	(0.452)
Manufacturing employment share, 1990	0.294	0.209	0.371	0.197	0.370	-0.255	0.002
	(0.350)	(0.360)	(0.381)	(0.425)	(0.320)	(0.423)	(0.465)

Table B.6: The Effect of Social Connectedness on Crime, 1960-2009, Results for All Explanatory Variables, cont.

	Dependent variable: Number of offenses reported to police						
	Murder	Rape	Robbery	Assault	Burglary	Larceny	Motor Vehicle Theft
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Manufacturing employment share, 2000	0.322	0.988**	0.068	0.688	0.641**	0.415	-0.118
	(0.388)	(0.447)	(0.372)	(0.429)	(0.323)	(0.314)	(0.515)
State fixed effects	x	x	x	x	x	x	x
Pseudo R2	0.773	0.838	0.931	0.913	0.938	0.926	0.906
N (city-years)	18,854	17,690	18,854	18,854	18,854	18,854	18,854
Cities	471	471	471	471	471	471	471

Notes: See note to Table 3.4. Sources: FBI UCR, Duke SSA/Medicare data, Census city data book

Table B.7: The Effect of Social Connectedness on Crime, 2000-2009, by Predicted Crimes

	Dependent variable: Number of offenses reported to police						
	Murder (1)	Rape (2)	Robbery (3)	Assault (4)	Burglary (5)	Larceny (6)	Motor Vehicle Theft (7)
All Cities	-0.091	0.078	-0.074	-0.129**	0.002	-0.029	-0.011
	(0.071)	(0.078)	(0.058)	(0.059)	(0.044)	(0.044)	(0.064)
Below Median Predicted Crimes	-0.064	0.171	0.190	-0.041	0.021	0.065	0.285**
	(0.162)	(0.144)	(0.138)	(0.120)	(0.075)	(0.073)	(0.114)
Above Median Predicted Crimes	-0.073	0.044	-0.047	-0.167***	-0.034	-0.042	-0.015
	(0.075)	(0.090)	(0.062)	(0.064)	(0.045)	(0.049)	(0.071)

Notes: Table displays estimates of equation (3.12). Sample restricted to cities with less than 500,000 residents in 1980. Regressions include the same covariates used in Table 3.4. To generate the predicted number of crimes for each city, we estimate equation (3.12) using data from 1995-1999, replacing state-year fixed effects with state-specific linear time trends. We then predict the number of crimes with these coefficients and covariates from 2000-2009, using the average value of log HHI and log number of migrants for all cities when generating the prediction. We estimate regressions using data from 2000-2009. Standard errors, clustered at the city level, are in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$
Sources: FBI UCR, Duke SSA/Medicare data, Census city data book

Table B.8: Negative Selection of Southern Black Migrants into Network Destinations

Sample:	Men and Women			Men			Women		
Dependent variable:	Years of Schooling (1)	Log Income (2)	Log Income (3)	Years of Schooling (4)	Log Income (5)	Log Income (6)	Years of Schooling (7)	Log Income (8)	Log Income (9)
Panel A: Selection into state of residence									
Share of migrants from birth state in state of residence	-1.594***	-0.107***	-0.041	-1.768***	-0.058**	0.019	-1.516***	-0.025	0.090*
	(0.154)	(0.031)	(0.030)	(0.176)	(0.022)	(0.019)	(0.152)	(0.051)	(0.052)
Years of schooling			0.041***			0.044***			0.076***
			(0.002)			(0.001)			(0.005)
N	97,132	77,760	77,760	45,187	42,960	42,960	51,945	34,800	34,800
R2	0.080	0.084	0.099	0.082	0.120	0.147	0.082	0.110	0.150
Panel B: Selection into metropolitan area of residence									
Share of migrants from birth state in metro of residence	-1.990***	-0.182***	-0.108**	-2.057***	-0.118***	-0.036	-1.995***	-0.154***	-0.002
	(0.117)	(0.044)	(0.044)	(0.108)	(0.035)	(0.036)	(0.154)	(0.057)	(0.059)
Years of schooling			0.036***			0.039***			0.070***
			(0.002)			(0.001)			(0.006)
N	66,359	52,958	52,958	30,533	29,201	29,201	35,826	23,757	23,757
R2	0.084	0.070	0.081	0.086	0.102	0.125	0.088	0.096	0.131
Quartic in age	x	x	x	x	x	x	x	x	x
Year of birth fixed effects	x	x	x	x	x	x	x	x	x
Birth state fixed effects	x	x	x	x	x	x	x	x	x
State/metro of residence fixed effects	x	x	x	x	x	x	x	x	x
Year fixed effects	x	x	x	x	x	x	x	x	x

Notes: Sample limited to African Americans born in the South from 1916-1936 who are living in the North, Midwest, or West regions. Standard errors, clustered at the state of residence level, are in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$
Sources: 1960 and 1970 Census IPUMS

Table B.9: The Effect of Social Connectedness on Crime, 1960-2009, Additional Robustness Checks

	Dependent variable: Number of offenses reported to police						
	Murder	Rape	Robbery	Assault	Burglary	Larceny	Motor Vehicle Theft
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Including cities with at least 500,000 residents in 1980							
Log HHI, Southern	-0.168***	-0.159***	-0.187***	-0.194***	-0.139***	-0.120***	-0.235***
black migrants	(0.036)	(0.037)	(0.039)	(0.043)	(0.026)	(0.029)	(0.039)
Pseudo R2	0.935	0.921	0.983	0.947	0.974	0.971	0.968
N (city-years)	19,543	18,324	19,543	19,543	19,543	19,543	19,543
Cities	485	485	485	485	485	485	485
Panel B: Negative binomial model							
Log HHI, Southern	-0.120***	-0.052	-0.129***	-0.079**	-0.039	-0.037	-0.115***
black migrants	(0.032)	(0.032)	(0.039)	(0.036)	(0.027)	(0.029)	(0.043)
Pseudo R2	0.283	0.217	0.187	0.143	0.148	0.123	0.144
N (city-years)	18,854	17,690	18,854	18,854	18,854	18,854	18,854
Cities	471	471	471	471	471	471	471
Panel C: Drop observations if dependent variable is below 1/6 or above 6 times city mean							
Log HHI, Southern	-0.128***	-0.076**	-0.247***	-0.133***	-0.091***	-0.045	-0.158***
black migrants	(0.031)	(0.036)	(0.034)	(0.042)	(0.022)	(0.029)	(0.041)
Pseudo R2	0.766	0.846	0.935	0.902	0.943	0.933	0.910
N (city-years)	15,192	15,695	17,823	15,250	18,712	18,715	18,613
Cities	470	471	471	471	471	471	471
Panel D: Drop observations if dependent variable is below 1/6 or above 6 times city median							
Log HHI, Southern	-0.156***	-0.080**	-0.246***	-0.133***	-0.090***	-0.044	-0.158***
black migrants	(0.032)	(0.036)	(0.034)	(0.042)	(0.022)	(0.029)	(0.041)
Pseudo R2	0.776	0.848	0.935	0.901	0.943	0.933	0.909
N (city-years)	15,711	15,799	17,844	15,246	18,705	18,693	18,652
Cities	471	470	471	471	471	471	471
Panel E: Measure HHI using birth county to destination city population flows							
Log HHI, Southern	-0.154***	-0.053	-0.214***	-0.120***	-0.066***	-0.042	-0.137***
black migrants	(0.033)	(0.032)	(0.038)	(0.039)	(0.023)	(0.032)	(0.041)
Pseudo R2	0.772	0.837	0.930	0.913	0.937	0.926	0.906
N (city-years)	18,854	17,690	18,854	18,854	18,854	18,854	18,854
Cities	471	471	471	471	471	471	471

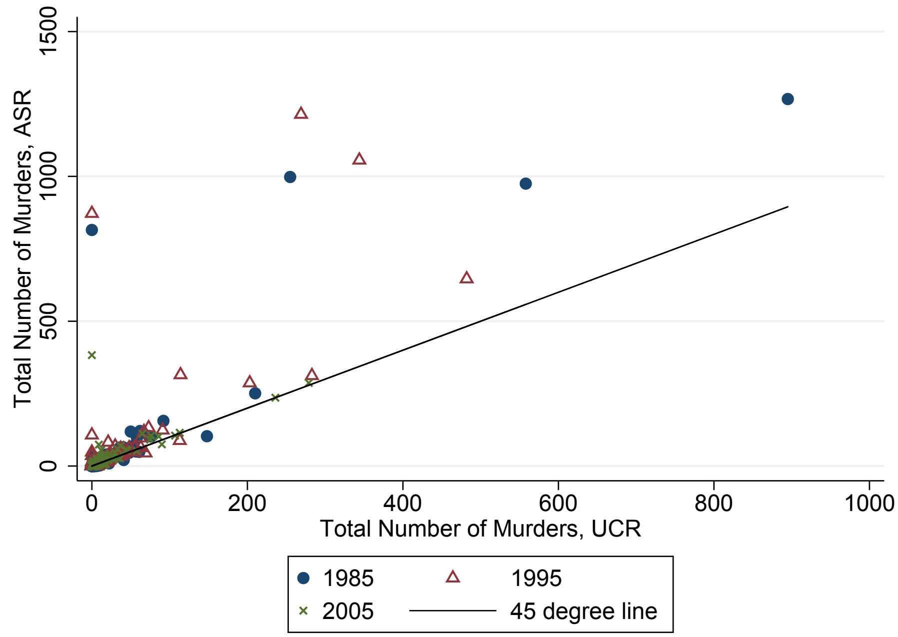
Notes: In Panel B, we estimate a negative binomial model instead of equation (3.12). For Panels C and D, we construct mean and median number of crimes for each city from 1960-2009. Regressions include the same covariates used in Table 3.4. Standard errors, clustered at the city level, are in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$
Sources: FBI UCR, Duke SSA/Medicare data, Census city data book

Table B.10: The Relationship between Social Connectedness, the Number of Migrants, and the Share of Migrants that Chose their Destination Because of Social Interactions

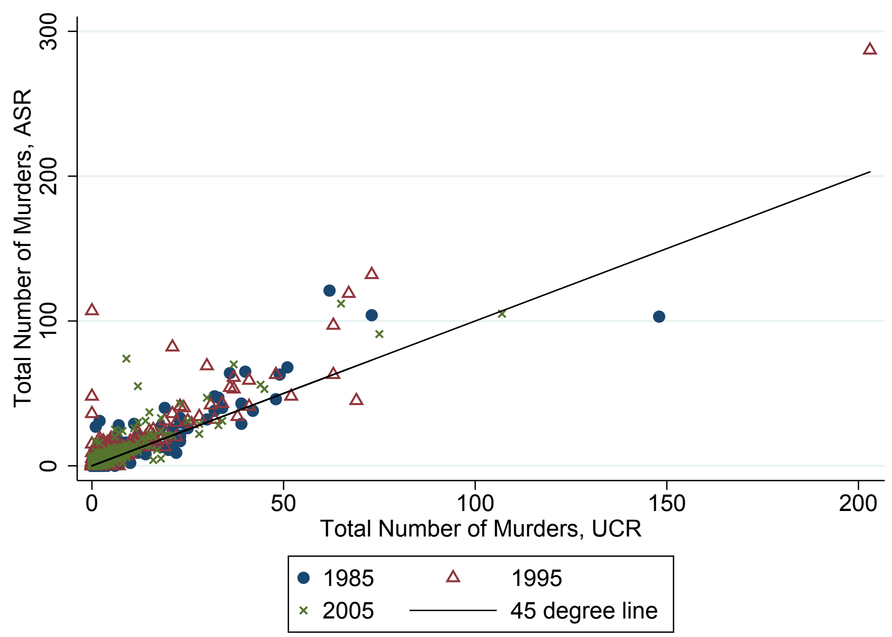
	Dependent variable: Log HHI, Southern black migrants			
	(1)	(2)	(3)	(4)
Log number, Southern black migrants	-0.457***		-0.666***	-0.669***
	(0.014)		(0.021)	(0.023)
Share of migrants that chose destination		-2.423***	2.896***	2.993***
because of social interactions		(0.282)	(0.229)	(0.259)
State fixed effects				x
R2	0.723	0.184	0.834	0.848
N (cities)	471	471	471	471

Notes: Sample restricted to cities with less than 500,000 residents in 1980. We estimate the share of migrants that chose their destination because of social interactions using a structural model, as described in the text.
Sources: Duke SSA/Medicare data,

Figure B.1: The Relationship between Murder Counts from Different FBI Data Sets



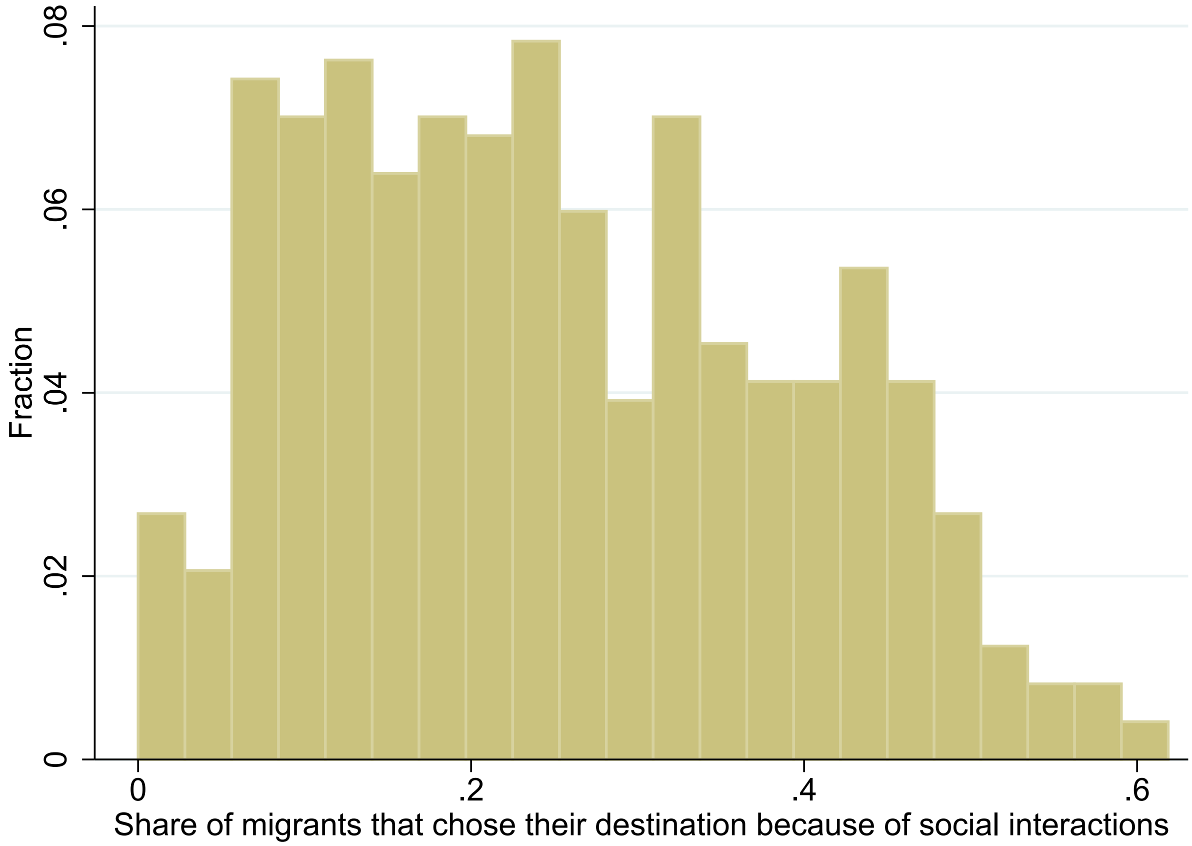
(a) All cities



(b) Cities with less than 500,000 residents in 1980

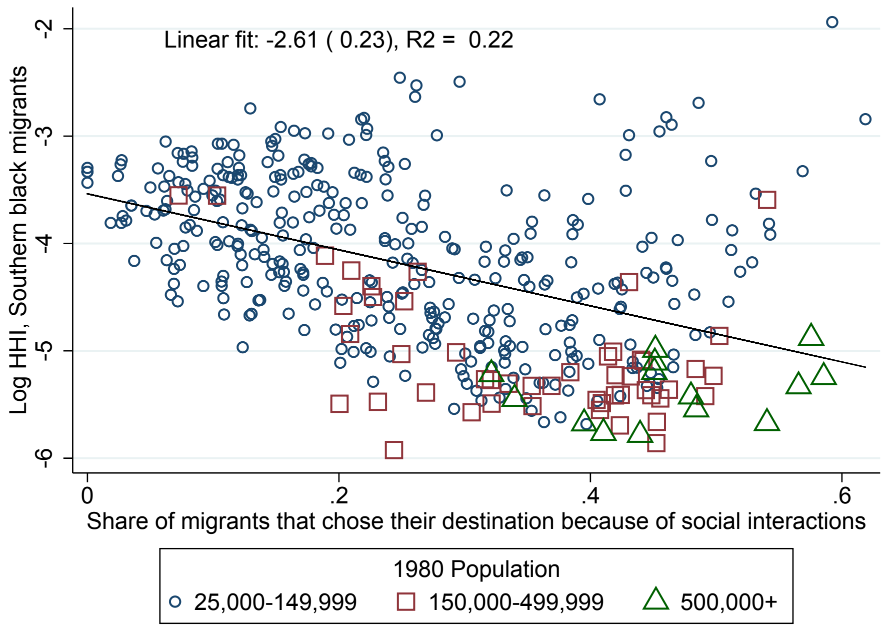
Notes: The UCR data contain the total number of murders per police agency. To construct a similar measure from the ASR data, we calculate the sum of murders committed by adult whites, adult blacks, adult other races, juvenile whites, juvenile blacks, and juvenile other races.
Source: FBI UCR

Figure B.2: Share of Migrants that Chose their Destination Because of Social Interactions

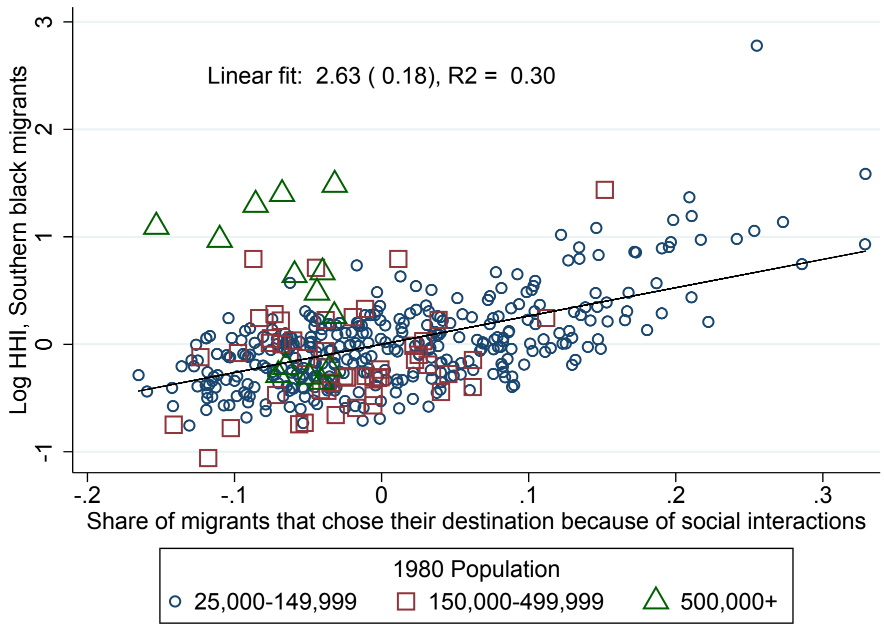


Notes: We estimate the share of migrants that chose their destination because of social interactions using a structural model, as described in the text.
Source: Duke SSA/Medicare data

Figure B.3: The Relationship between Social Connectedness and the Share of Migrants that Chose their Destination Because of Social Interactions



(a) Raw



(b) Conditional on Log Number, Southern Black Migrants

Notes: We estimate the share of migrants that chose their destination because of social interactions using a structural model, as described in the text. Panel B plots the residuals from regressing log HHI and the share of migrants that chose their destination because of social interactions on the log number of migrants.
Source: Duke SSA/Medicare data

BIBLIOGRAPHY

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

Aaronson, Daniel, Bhash Mazumder, Seth G. Sanders, and Evan J. Taylor. 2016. "Estimating the Effect of School Quality on Mortality in the Presence of Migration: Evidence from the Jim Crow South." *Working Paper*.

Albouy, Valerie, and Laurent Lequien. 2009. "Does Compulsory Education Lower Mortality?" *Journal of Health Economics*, 28(1): 155–168.

Alesina, Alberto, and Eliana La Ferrara. 2000. "Participation in Heterogeneous Communities." *Quarterly Journal of Economics*, 115(3): 847–904.

Alesina, Alberto, Reza Baqir, and William Easterly. 1999. "Public Goods and Ethnic Divisions." *Quarterly Journal of Economics*, 114(4): 1243–1284.

Alesina, Alberto, Reza Baqir, and William Easterly. 2000. "Redistributive Public Employment." *Journal of Urban Economics*, 48(2): 219–241.

Almond, Douglas, and Janet Currie. 2011. "Human Capital Development Before Age 5." In *Handbook Labor Economics*. , ed. David Card and Orley Ashenfelter, 1315–1486. Amsterdam:Elsevier.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association*, 91(434): 444–455.

Associated Press. 1983. "Blacks in Pennsylvania Town Recall Southern Past." *The Baytown Sun*.

Barker, David J.P. 1998. *Mothers, Babies and Health in Later Life, 2nd ed.* Edinburgh:Churchill Livingston.

Bartel, Ann P. 1989. "Where do the New U.S. Immigrants Live?" *Journal of Labor Economics*, 7(4): 371–391.

Bauer, Thomas, Gil S. Epstein, and Ira N. Gang. 2005. "Enclaves, Language, and the Location Choice of Migrants." *Journal of Population Economics*, 18(4): 649–662.

Bayer, Patrick, Stephen L. Ross, and Giorgio Topa. 2008. "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes." *Journal of Political Economy*, 116(6): 1150–1196.

Beaman, Lori A. 2012. "Social Networks and the Dynamics of Labour Market Outcomes: Evidence from Refugees Resettled in the U.S." *Review of Economic Studies*, 79(1): 128–161.

Becker, Gary. 1967. *Human Capital and the Personal Distribution of Income.* Ann Arbor Michigan:University of Michigan Press.

Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, 76(2): 169–217.

Beine, Michel, Frédéric Docquier, and Çağlar Ozden. 2011. "Diasporas." *Journal of Development Economics*, 95(1): 30–41.

Bell, Velma Fern. 1933. "The Negro in Beloit and Madison, Wisconsin." Master's diss. University of Wisconsin.

Billings, Stephen B., and Erik B. Johnson. 2012. "A Non-Parametric Test for Industrial Specialization." *Journal of Urban Economics*, 71(3): 312–331.

Black, Dan A., Joonhwi Joo, Robert LaLonde, Jeffrey Smith, and Evan J. Taylor. 2016. "Simple Tests for Selection Bias: Learning More from Instrumental Variables." *IZA Discussion Paper, No. 9346*.

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.

Black, Dan A., Yu-Chieh Hsu, and Lowell J. Taylor. 2015. "The Effect of Early-Life Education on Later-Life Mortality." *Journal of Health Economics*, 44: 1–9.

Blume, Lawrence E., William A. Brock, Steven N. Durlauf, and Yannis M. Ioannides. 2011. "Identification of Social Interactions." In *Handbook of Social Economics*. Vol. 1, , ed. Jess Benhabib, Alberto Bisin and Matthew O. Jackson, 853–964. Elsevier.

Blumstein, Alfred. 2000. "Disaggregating the Violence Trends." In *The Crime Drop in America*. , ed. Alfred Blumstein and Joel Wallman, 13–44. New York:Cambridge University Press.

Boustan, Leah Platt. 2009. "Competition in the Promised Land: Black Migration and Racial Wage Convergence in the North, 1940-1970." *Journal of Economic History*, 69(3): 756–783.

Boustan, Leah Platt. 2011. "Was Postwar Suburbanization 'White Flight'? Evidence from the Black Migration." *Quarterly Journal of Economics*, 125(1): 417–443.

Brinch, Christian N., Magne Mogstad, and Matthew Wiswall. forthcoming, 2016. "Beyond LATE with a Discrete Instrument: Heterogeneity in the Quantity-Quality Interaction of Children." *Journal of Political Economy*.

Brock, William A., and Steven N. Durlauf. 2001. "Discrete Choice with Social Interactions." *Review of Economic Studies*, 68(2): 235–260.

Buckles, Kasey, Andreas Hagemann, Ofer Malamud, Melinda Morrill, and Abigail Wozniak. 2016. "The Effect of College Education on Mortality." *Journal of Health Economics*, 50: 99–114.

Burks, Stephen V., Bo Cowgill, Mitchell Hoffman, and Michael Housman. 2015. “The Value of Hiring through Employee Referrals.” *Quarterly Journal of Economics*, 130: 805–839.

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

Cameron, Stephen V., and Christopher Taber. 2004. “Estimation of education borrowing constraints using returns to Schooling.” *Journal of Political Economy*, 112(1): 132–182.

Card, David. 1994. “Earnings, Schooling and Ability Revisited.” *NBER Working Paper No. 4832*.

Card, David. 1995. “Using Geographic Variation in College Proximity to Estimate the Return to Schooling.” *Aspects of Labour Market Behaviour : Essays in Honour of John Vanderkamp*, 201–222.

Card, David. 1999. “The causal effect of education on earnings.” In *Handbook of Labor Economics, vol 3.* , ed. Orley Ashenfelter, Richard Layard and David Card. Amsterdam:Elservier.

Card, David. 2001. “Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems.” *Econometrica*, 69(5): 1127–1160.

Carneiro, Pedro, James J. Heckman, and Edward J. Vytlacil. 2011. “Estimating the Marginal Returns to Education.” *American Economic Review*, 101(6): 2754–2781.

Carrington, William J., Enrica Detragiache, and Tara Vishwanath. 1996. “Migration with Endogenous Moving Costs.” *American Economic Review*, 86(4): 909–930.

Case, Anne, and Angus Deaton. 2015. “Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century.” *Proceedings of the National Academy of Science*, 112(49): 15078–15083.

Case, Anne C., and Lawrence F. Katz. 1991. “The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths.” *NBER Working Paper 3705*.

Cassar, Alessandra, Luke Crowley, and Bruce Wydick. 2007. “The Effect of Social Capital on Group Loan Repayment: Evidence from Field Experiments.” *Economic Journal*, 117(517): F85âĂŞF106.

Census, United States Bureau of the. 1922. “Mortality Statistics, 1920.” *Twenty-First Annual Report*.

Census, United States Bureau of the. 1979. “The Social and Economic Status of the Black Population in the United States, 1790-1978.” *Current Population Reports, Special Studies Series P-23 No. 80*.

Chalfin, Aaron, and Justin McCrary. 2015. “Are U.S. Cities Underpoliced? Theory and Evidence.”

Chay, Kenneth, and Kaivan Munshi. 2015. "Black Networks After Emancipation: Evidence from Reconstruction and the Great Migration."

Chen, Yuyu, Ginger Zhe Jin, and Yang Yue. 2010. "Peer Migration in China." *NBER Working Paper 15671*.

Clark, Damon, and Heather Royer. 2013. "The Effect of Education on Adult Mortality and Heatlh: Evidence from Britain." *American Economic Review*, 103(6): 2087–2120.

Collins, William J. 1997. "When the Tide Turned: Immigration and the Delay of the Great Black Migration." *Journal of Economic History*, 57(3): 607–632.

Collins, William J., and Marianne H. Wanamaker. 2015. "The Great Migration in Black and White: New Evidence on the Selection and Sorting of Southern Migrants." *Journal of Economic History*, 75(4): 947–992.

Conti, Gabriella, James J. Heckman, and Sergio Urzua. 2010. "The Education-Health Gradient." *American Economics Review: Papers and Proceedings*, 2010(2): 234–238.

Currie, Janet. 2009. "Healthy, Wealthy, and Wise? Socioeconomic status, poor health in childhood, and human capital development." *Journal of Economic Literature*, 47(1): 87–122.

Currie, Janet. 2011. "Inequality at Birth: Some causes and Consequences." *American Economic Review*, 101(3): 1–22.

Currie, Janet, and Enrico Moretti. 2003. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings." *The Quarterly Journal of Economics*, 118(4): 1495–1532.

Curtis White, Katherine J. 2008. "Population Change and Farm Dependence: Temporal and Spatial Variation in the U.S. Great Plains, 1900-2000." *Demography*, 45(2): 363–386.

Cutler, David, Angus Deaton, and Adriana Lleras-Muney. 2006. "The Determinants of Mortality." *Journal of Economic Perspectives*, 20(3): 97–120.

Cutler, David M., and Adriana Lleras-Muney. 2010. "Understand Differences in Health Behaviors by Education." *Journal of Health Economics*, 29(1): 1–28.

Damm, Anna Piil, and Christian Dustmann. 2014. "Does Growing Up in a High Crime Neighborhood Affect Youth Criminal Behavior?" *American Economic Review*, 104(6): 1806–1832.

Deaton, Angus. 2013. "What does the Empirical Evidence Tell Us about the Injustice of Health Inequalities." In *Inequalities in Health: Concepts, Measures and Ethics*. , ed. Nir Eyal, Samia Hurst, Ole Frithof Norheim and Daniel Wikler, 263–281. Oxford:Oxford University Press.

Duranton, Gilles, and Henry G. Overman. 2005. "Testing for Localization Using Micro-Geographic Data." *Review of Economic Studies*, 72(4): 1077–1106.

Epple, Dennis, and Richard E. Romano. 2011. "Peer Effects in Education: A Survey of the Theory and Evidence." In *Handbook of Social Economics*. Vol. 1, , ed. Jess Benhabib, Alberto Bisin and Matthew O. Jackson, 1053–1163. Elsevier.

Evans, William N., Craig Garthwaite, and Timothy J. Moore. 2016. "The White/Black Educational Gap, Stalled Progress, and the Long-Term Consequences of Crack Cocaine Markets." *Review of Economics and Statistics*, 98(5): 832–847.

Feigenberg, Benjamin, Erica Field, and Rohini Pande. 2013. "The Economic Returns to Social Interaction: Experimental Evidence from Microfinance." *Review of Economic Studies*, 80(4): 1459–1483.

Fogel, Robert W. 1997. "New findings on secular trends in nutrition and mortality: Some Implications for population theory." In *Handbook of Population and Family Economics*. Vol. 1A, , ed. Mark Rosenzweig and Oded Stark, 433–481. Amsterdam:Elsevier.

Fogel, Robert W. 2004. *The Escape from Hunger and Premature Death, 1700-2100: Europe, America and the Third World.* Cambridge:Cambridge University Press.

Fox, James Alan. 2000. "Demographics and U.S. Homicide." In *The Crime Drop in America*. , ed. Alfred Blumstein and Joel Wallman, 288–317. New York:Cambridge University Press.

Freeman, Richard B. 1999. "The Economics of Crime." In *Handbook of Labor Economics*. Vol. 3C, , ed. Orley Ashenfelter and David Card, 3529–3571. Amsterdam:North Holland.

Fukuyama, Francis. 2000. "Social Capital and Civil Soceity." IMF Working Paper 00/74.

Gibson, Campbell, and Kay Jung. 2005. "Historical Census Statistics on Population Totals by Race, 1790 to 1990, and by Hispanic Origin, 1790 to 1990, For Large Cities and Other Urban Places in the United States." *U.S. Census Bureau Population Division Working Paper No. 76*.

Giuletti, Corrado, Jackline Wahba, and Yves Zenou. 2014. "Strong versus Weak Ties in Migration."

Glaeser, Edward L., Bruce Sacerdote, and José A. Scheinkman. 1996. "Crime and Social Interactions." *Quarterly Journal of Economics*, 111(2): 507–548.

Glaeser, Edward L., David I. Laibson, José A. Scheinkman, and Christine L. Soutter. 2000. "Measuring Trust." *Quarterly Journal of Economics*, 115(3): 811–846.

Glaeser, Edward L., David Laibson, and Bruce Sacerdote. 2002. "An Economic Approach to Social Capital." *Economic Journal*, 112(483): F437âĂŞF458.

Gottlieb, Peter. 1987. *Making Their Own Way: Southern Blacks' Migration to Pittsburgh, 1916-1930.* Urbana:University of Illinois Press.

Graham, Bryan S. 2008. "Identifying Social Interactions Through Conditional Variance Restrictions." *Econometrica*, 76(3): 643–660.

Gregory, James N. 1989. *American Exodus: The Dust Bowl Migration and Okie Culture in California.* New York:Oxford University Press.

Gregory, James N. 2005. *The Southern Diaspora: How the Great Migrations of Black and White Southerners Transformed America.* Chapel Hill:University of North Carolina Press.

Griliches, Zvi. 1977. "Estimating the Returns to Schooling: Some Econometric Problems." *Econometrica*, 45(1): 1–22.

Grimard, Franque, and Daniel Parent. 2007. "Education and smoking: Were Vietnam war draft avoiders also more likely to avoid smoking?" *Journal of Health Economics*, 26(5): 896–926.

Grossman, James R. 1989. *Land of Hope: Chicago, Black Southerners, and the Great Migration.* Chicago:University of Chicago Press.

Grossman, Michael. 1972. "On the concept of health capital and the demand for health." *Journal of Political Economy*, 80(2): 223–255.

Grossman, Michael. 2000. "The human capital model." In *Handbook of Health Economics*. Vol. 1, , ed. Anthony J. Culyer and Joseph P. Newhouse, 347–408. Amsterdam:Elsevier.

Guerry, André-Michel. 1833. *Essai sur la Statistique Morale de la France.* Paris:Crochard.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2004. "The Role of Social Capital in Financial Development." *American Economic Review*, 94(3): 526–556.

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

Heath, Rachel. 2016. "Why do Firms Hire using Referrals? Evidence from Bangladeshi Garment Factories."

Heckman, James J., and Edward Vytlacil. 1999. "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects." *Proceedings of the National Academy of Sciences*, 96(8): 4730–4734.

Heckman, James J., and Edward Vytlacil. 2005. "Structural Equations, Treatment Effects, and Econometric Policy Evaluation." *Econometrica*, 73(3): 669–738.

Heckman, James J., and Edward Vytlacil. 2007. "Econometric Evaluation of Social Programs, part II." In *Handbook of Econometrics*. Vol. 6B, , ed. James J. Heckman and Edward E. Leamer, 4875–5143. Elsevier.

Heckman, James J., John Eric Humphries, and Gregory Veramendi. 2016. "Dynamic Treatment Effects." *Journal of Econometrics*, 191: 276–292.

Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics*, 24(3): 411–482.

Heckman, James J., Lance Lochner, and Petra E. Todd. 2006. "Earnings functions, rates of return and treatment effects: The Mincer equation and beyond." In *Handbook of the Economics of Education*. Vol. 1, 307–458. Amsterdam:Elsevier.

Heckman, James J., Sergio Urzua, and Edward Vytlacil. 2006. "Understanding Instrumental Variables in Models with Essential Heterogeneity." *The Review of Economics and Statistics*, 88(3): 389–432.

Hellerstein, Judith K., Melissa McInerney, and David Neumark. 2011. "Neighbors and Coworkers: The Importance of Residential Labor Markets." *Journal of Labor Economics*, 29(4): 659–695.

Henri, Florette. 1975. *Black Migration: Movement North, 1900-1920.* New York:Anchor Press/Doubleday.

Hornbeck, Richard. 2012. "The Enduring Impact of the American Dust Bowl: Short- and Long-Run Adjustments to Environmental Catastrophe." *American Economic Review*, 102(4): 1477–1507.

Hornbeck, Richard, and Suresh Naidu. 2014. "When the Levee Breaks: Black Migration and Economic Development in the American South." *American Economic Review*, 104(3): 963–990.

Hurt, Douglas R. 2011. *The Big Empty: The Great Plains in the Twentieth Century.* Tucson:University of Arizona Press.

Ioannides, Yannis M., and Linda Datcher Loury. 2004. "Job Information Networks, Neighborhood Effects, and Inequality." *Journal of Economic Literature*, 42(4): 1056–1093.

Jackson, Blyden. 1991. "Introduction: A Street of Dreams." In *Black Exodus: The Great Migration from the American South*. , ed. Alferdteen Harrison, xi–xvii. Jackson:University Press of Mississippi.

Jamieson, Stuart M. 1942. "A Settlement of Rural Migrant Families in the Sacramento Valley, California." *Rural Sociology*, 7: 49–61.

Johnson, Janna, and Evan J. Taylor. 2016. "The Long-Run Consequences of Urban-Rural Migration." *Working Paper*.

Johnson, Kenneth M., and Richard W. Rathge. 2006. "Agricultural Dependence and Changing Population in the Great Plains." In *Population Change and Rural Society*. , ed. William A. Kandel and David L. Brown, 197–217. Springer.

Karlan, Dean, Markus Mobius, Tanya Rosenblat, and Adam Szeidl. 2009. "Trust and Social Collateral." *Quarterly Journal of Economics*, 124(3): 1307–1361.

Karlan, Dean S. 2005. "Using Experimental Economics to Measure Social Capital and Predict Financial Decisions." *American Economic Review*, 95(5): 1688–1699.

Karlan, Dean S. 2007. "Social Connections and Group Banking." *Economic Journal*, 117(517): F52–F84.

Knack, Stephen, and Philip Keefer. 1997. "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation." *Quarterly Journal of Economics*, 112(4): 1251–1288.

Knowles, Lucas W. 2010. "Beloit, Wisconsin and the Great Migration the Role of Industry, Individuals, and Family in the Founding of Beloit's Black Community 1914 - 1955."

Lange, Fabian, Alan L. Olmstead, and Paul W. Rhode. 2009. "The Impact of the Boll Weevil, 1892-1932." *Journal of Economic History*, 69: 685–718.

La Porta, Rafael, Florencio Lopez de Silanes, Andrei Shleifer, and Robert W. Vishny. 1997. "Trust in Large Organizations." *American Economic Review*, 87(2): 333–338.

Laury, Susan. 1986. "Brownsville Folk Reunite in Decatur." *Herald and Review*.

Lleras-Muney, Adreana. 2005. "The Relationship between Education and Adult Mortality in the United States." *The Review of Economic Studies*, 72(1): 189–221.

Lochner, Lance. 2011. "Non-production Benefits of Education: Crime, Health, and Good Citizenship." In *Handbook of the Economics of Education*. Vol. 4, , ed. Eric A. Hanushek, Stephen J. Machin and Ludger Woessmann. Amsterdam:North Holland.

Ludwig, Jens, and Jeffrey R. Kling. 2007. "Is Crime Contagious?" *Journal of Law and Economics*, 50(3): 491–518.

Malamud, Ofer, and Abigail Wozniak. 2012. "The impact of college education on geographic mobility." *Journal of Human Resources*, 47(4): 913–950.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies*, 60(3): 531–542.

Marks, Carole. 1989. *Farewell, We're Good and Gone: The Great Black Migration.* Bloomington:Indiana University Press.

Marks, Carole. 1991. "The Social and Economic Life of Southern Blacks During the Migrations." In *Black Exodus: The Great Migration from the American South.* , ed. Alferdteen Harrison, 36–50. Jackson:University Press of Mississippi.

Mazumder, Bhashkar. 2008. "Does Education Improve Health? A Reexamination of Evidence from Compulsory Schooling Laws." *Economic Perspectives*, 32(2): 2–16.

Meara, Ellen R., Seth Richards, and David M. Cutler. 2008. "The Gap Gets Bigger: Changes in Mortality and Life Expectancy, by Education, 1981-2000." *Health Affairs*, 27(2): 350–360.

Meghir, Costas, Marten Palme, and Emilia Simeonova. 2012. "Education, Health, and Mortality: Evidence from a social experiment." *NBER Working Paper No. 17932*.

Miguel, Edward, Paul Gertler, and David I. Levine. 2005. "Does Social Capital Promote Industrialization? Evidence from a Rapid Industrializer." *Review of Economics and Statistics*, 87(4): 754–762.

Mincer, Jacob. 1974. *Schooling, Experience, and Earnings.* New York:National Bureau of Economic Research.

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

Moffitt, Robert. 2008. "Estimating Marginal Treatment Effects in Heterogeneous Populations." *Annales d'Economie et de Statistique*, , (91-92): 239–261.

Munshi, Kaivan. 2003. "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market." *Quarterly Journal of Economics*, 118(2): 549–597.

Munshi, Kaivan. 2011. "Labor and Credit Networks in Developing Economics." In *Handbook of Social Economics*. Vol. 1, , ed. Jess Benhabib, Alberto Bisin and Matthew O. Jackson, 1223–1254. Elsevier.

Neal, Derek, and Armin Rick. 2014. "The Prison Boom & The Lack of Black Progress after Smith & Welch."

Putnam, Robert D. 2000. *Bowling Alone: The Collapse and Revival of American Community.* New York, NY:Simon & Schuster.

Quetelet, Adolphe. 1835. *Sur l'Homme et le Developpement De Ses Facultes.* Paris:Bachelier.

Ram, Rati, and Theodore W. Schultz. 1979. "Life Span, Health, Savings, and Productivity." *Economic Development and Cultural Change*, 27(3): 399–421.

Rosenzweig, Mark R., and T. Paul Schultz. 1982. "The behavior of mothers as inputs to child health: The determinants of birth weight, gestation, and rate of fetal growth." In *Economic Aspects of Health.* , ed. Victor R. Fuchs. Chicago:University of Chicago Press, NBER.

Roy, Arthur. 1951. “Some Thoughts on the Distribution of Earnings.” *Oxford Economic Papers*, 3(2): 135–146.

Rubin, Morton. 1960. “Migration Patterns of Negroes from a Rural Northeastern Mississippi Community.” *Social Forces*, 39(1): 59–66.

Ruggles, Steven, J., Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. 2010. “Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database].”

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2015. *Integrated Public Use Microdata Series: Version 6.0 [Machine-readable database].* Minneapolis:University of Minnesota.

Rupasingha, Anil, and Stephan J. Goetz. 2008. *US County-Level Social Capital Data, 1990-2005.* The Northeast Regional Center for Rural Development, Penn State University, University Park, PA.

Rupasingha, Anil, Stephan J. Goetz, and David Freshwater. 2006. “The Production of Social Capital in US Counties.” *Journal of Socio-Economics*, 35(1): 83–101.

Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls. 1997. “Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy.” *Science*, 277(918).

Sapienza, Paola, Anna Toldra-Simats, and Luigi Zingales. 2013. “Understanding Trust.” *Economic Journal*, 123(573): 1313âĂŞ1332.

Schmutte, Ian M. 2015. “Job Referral Networks and the Determination of Earnings in Local Labor Markets.” *Journal of Labor Economics*, 33(1): 1–32.

Scott, Emmett J. 1920. *Negro Migration During the War.* New York:Oxford University Press.

Scroggs, William O. 1917. “Interstate Migration of Negro Population.” *Journal of Political Economy*, 25(10): 1034–1043.

Smith, James P. 1999. “Healthy bodies and thick wallets: The dual relation between health and economic status.” *Journal of Economic Perspectives*, 13(2): 145–166.

Smith, James P., and Finis Welch. 1989. “Black Economic Progress After Myrdal.” *Journal of Economic Literature*, 27(2): 519–564.

Smith, Sheila. 2006. “All-class Reunion Recalls Decatur’s Ties to Brownsville, Tenn.” *Herald & Review*.

Soteriades, Elpidoforos S., and Joseph R. DiFranza. 2003. “Parent’s Socioeconomic Status, Adolescents’ Disposable Income, and Adolescents’ smoking status in Massachusetts.” *American Journal of Public Health*, 93(7): 1155–1160.

Spitzer, Yannay. 2014. "Pogroms, Networks, and Migration: The Jewish Migration from the Russian Empire to the United States 1881-1914."

Stack, Carol. 1970. *All our Kin.* New York:Basic Books.

Steinbeck, John. 1939. *The Grapes of Wrath.* New York:The Viking Press.

Stock, James H., Jonathan W. Wright, and Motohiro Yogo. 2002. "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments." *Journal of Business and Economic Statistics*, 17(3): 518–529.

Stuart, Bryan A., and Evan J. Taylor. 2017*a*. "Social Interactions and Location Decisions: Evidence from U.S. Mass Migration."

Stuart, Bryan A., and Evan J. Taylor. 2017*b*. "The Effect of Social Connectedness on Crime: Evidence from the Great Migration."

Taylor, Evan J. 2016. "Comparing estimation of Marginal Treatment Effects with Local Instrumental Variables and the Separate Approach." *working paper*.

Tibbetts, Stephen G. 2012. *Criminological Theory: The Essentials.* Los Angeles:SAGE Publications.

Tolnay, Stewart E., and E. M. Beck. 1991. "Rethinking the Role of Racial Violence in the Great Migration." In *Black Exodus: The Great Migration from the American South*. , ed. Alferdteen Harrison, 20–35. Jackson:University Press of Mississippi.

Topa, Giorgio. 2001. "Social Interactions, Local Spillovers and Unemployment." *Review of Economic Studies*, 68(2): 261–295.

Topa, Giorgio. 2011. "Labor Markets and Referrals." In *Handbook of Social Economics*. Vol. 1, , ed. Jess Benhabib, Alberto Bisin and Matthew O. Jackson, 1193–1221. Elsevier.

Vytlacil, Edward. 2002. "Independence, Monotonicity, and Latent Index Models: An Equivalence Result." *Econometrica*, 70(1): 331–341.

Walque, Damien De. 2007. "Does Education Affect Smoking Behaviors? Evidence using the Vietnam Draft as an Instrument for College Education." *Journal of Health Economics*, 26(5): 877–895.

Weisburd, David, Gerben J.N. Bruinsma, and Wim Bernasco. 2009. "Units of Analysis in Geographic Criminology: Historical Development, Critical Issues, and Open Questions." In *Putting Crime in its Place: Units of Analysis in Geographic Criminology*. , ed. David Weisburd, Gerben J.N. Bruinsma and Wim Bernasco, 3–31. New York:Springer.

Wilkerson, Isabel. 2010. *The Warmth of Other Suns: The Epic Story of America's Great Migration.* New York:Random House.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA:MIT Press.