

Moving North and Into Jail? The Great Migration and Black Incarceration

Katherine Eriksson

December 31, 2016

Abstract:¹ Black incarceration rates in the U.S. grew relative to white incarceration rates throughout the first half of the 20th century despite substantial convergence in education levels and wages between the two groups. This paper considers the First Great Migration prior to 1940 as a factor which increased black male incarceration rates. I construct an individual-level dataset of all southern-born male prisoners and non-prisoners in the 1940 US Census; both groups are matched to their childhood household in the 1920 Census in order to control for across-household selection using household fixed effects. I estimate that migrating to the North roughly doubled an individual's chance of being incarcerated, increasing the probability of incarceration by 1.55 percentage points. I estimate that the Great Migration was responsible for about 7 percent of the increase in black incarceration rates between 1920 and 1940.

¹Eriksson: kaeriksson@ucdavis.edu. Department of Economics, University of California, Davis, Stellenbosch University, and National Bureau of Economic Research. I thank FamilySearch.org and Ancestry.com for access to the census indexes from 1920 and 1940. I also appreciate help from undergraduate research assistants at UCLA and Cal Poly. I am grateful for feedback from Ran Abramitzky, Leah Boustan and participants at the UC-Berkeley Complete Count Census Workshop.

I. Introduction

Incarceration rates of black men rose sharply after emancipation at the end of the Civil War. By 1880, black men were over 3 times more likely than white men to be incarcerated. By 1923, the ratio had grown to 4.3 and by 1940 it was 4.8. As a proportion of the population, 0.85 percent of black men were incarcerated in 1920, while by 1940 over 2.1 percent were in jail or prison. This growth in the racial incarceration gap occurred despite substantial convergence between black and white education and income levels during the same period. While there has been much interest in recent racial gaps in incarceration (Alexander, 2012; Raphael and Stoll, 2007; Lochner and Moretti, 2004), little is known about the origins and evolution of the racial incarceration gap in the early 20th century United States. This paper explores one factor: migration of southern blacks to northern cities during the Great Black Migration.

Between 1900 and 1970, approximately 6 million blacks migrated from the South to the North. Jim Crow laws and the boll-weevil infestation in the South reduced labor market opportunities and pushed men north in search of better opportunities—real incomes were 75 percent higher in the North than the South (Collins and Wanamaker, 2014). During World War I, labor shortages caused partly by a slow-down in immigrant inflows meant that agents from large factories in the North recruited southern men to work in the industrializing North (Grossman, 1991; Lemann, 1991; Wilkerson, 2010; Collins, 1997); this initial migration set off chain migration in the subsequent decades. While incomes were higher in the north, not all outcomes were better—infant mortality was almost twice as high in northern cities than the rural south (Eriksson and Niemesh, 2017) and migrants themselves faced lowered longevity (Black et al., 2015). Muller (2012), using mostly aggregate statistics, argues that the Great Migration increased black

incarceration rates. I add to this analysis by using individual level data to enable me to control for pre-migration characteristics.

This paper estimates the impact of the Great Black Migration between 1920 and 1940 on black incarceration rates. By 1940, 18.4 percent of southern-born black men lived in the North, over 92 percent of them in cities. At the same time, urbanization rates within the South increased from 20.5 percent in 1900 to 38.7 percent in 1940. I show that the causal impact on an individual's propensity to be incarcerated of migrating to the North or within the South was positive and large. I hypothesize that higher incarceration rates in the North are due to a combination of discrimination, in the job application process and in the criminal justice system, and lower wages as a result of lower skills.

I use the male full population from the 1940 United States Census, and match individuals back to their childhood households in 1920 to enable me to estimate family fixed effects models. This enables me to control for any selection at the household-level. By looking at pre-migration characteristics of childhood households, I also show that there is little selection into migration, consistent with other literature in this time period (Collins and Wanamaker, 2014; Boustan, 2017).

I find that black men who migrate to the North are 1.55 percentage points more likely to be incarcerated in 1940, roughly double the probability than if they stayed in the South. A similar pattern holds for white migrants, but of smaller magnitude. The pattern for black migrants is driven largely by those who migrated within the last five years who have a three times larger effect of migrating to the North than those who migrated more than five years ago. Unsurprisingly, recent migrants are younger, less likely to be working, and have lower incomes, even accounting for education and age differences. I also show that migrating *within* the South increases a man's probability of incarceration but that this effect is not as large as migrating to the North.

Finally, I show that northern-born black men had higher incarceration rates than southern-born men, despite having higher levels of education and being less likely to reside in urban areas; I conclude from this that either the inflows of migrants from the South depressed labor market opportunities for Northern blacks or that Southern-born blacks were less likely to turn to crime as a way of handling low incomes. In line with work on a later part of the Great Migration (Stuart and Taylor, 2016), I also find suggestive evidence that larger migrant networks decreased an individual's likelihood of being incarcerated.

This paper contributes to two literatures. First, it adds a negative aspect to the literature on convergence in black-white outcomes throughout the 20th century. While education and income levels converged substantially by 1940 (Margo, 1990; Smith and Welch, 1989), partly due to investments in southern education (Aaronson and Mazumder, 2012; Heckman et al., 2000), incarceration rates were persistently increasing for black men. Second, this paper contributes to our knowledge about crime and cities. The movement of a large portion of the population from rural to urban areas during this time period contributed to increasing incarceration rates, consistent with more recent work which shows that crime rates are higher in cities (Glaeser and Sacerdote, 1999).

The paper proceeds as follows. In the next section, I give an overview of incarceration in the early part of the 20th century and then lay out a conceptual framework for thinking about migration and incarceration. In Section 3, I describe the census data. Section 4 lays out the estimation strategy, Section 5 describes the results, and Section 6 concludes.

II. Background and Conceptual Framework

In the United States, incarceration rates for blacks have always been higher than those of whites. Figure 1 graphs the number of incarcerated individuals per capita (times 1000) by race and region from 1890 to 1980. In 1890, 3 out of 1000 blacks were incarcerated; the black incarceration rate was 3.1 times as high as the white incarceration rate. The black-white incarceration ratio grew to 4.8 in 1940 before falling back to 3.1 by 1950. Thereafter, the ratio grew through 1980. Rates for blacks living in the North were higher than for those living in the South throughout the period—blacks in the more urban North were between two and three times more likely to be incarcerated than those in the rural South.

The figure's numbers divide by the relevant population which includes both men and women of all ages. This is the only possible denominator which is consistent over time is total population due to the fact that in some years data is reported for men and women combined and age of prisoners is not available. Given that about 90% of prisoners were male, multiplying by 1.8 would give the rate for men. For example, the incarceration rate of black males in the South in 1923 was about 0.43. Incarceration rates for the most often incarcerated ages of 18 to 45 are even higher; indeed, for men aged 23-35, the ages used in this paper, the incarceration rate was 0.85 in 1920.

Historical evidence suggests that the initial racial gap in incarceration rates (circa 1890) may have been, in part, the result of a discriminatory system that was set up to incarcerate black men in the South. Following the Civil War, many Southern states passed a series of laws, referred to as "Black Codes", designed to control the mobility and restrict the economic opportunities of black freedmen. One subset of these laws criminalized vagrancy and allowed prisons to lease out their inmates as low-cost labor to local farms (Naidu 2010). Parchman Farm in Mississippi is known as one of more brutal examples of large scale farming using free labor (Oshinsky 1996).

In the North, contemporary observers accused authorities of discriminatory arrest and sentencing practices (Sellin, 1928; Muhammed, 2010). Race riots were common through the 1910's and 1920's in the migrant destination cities of Chicago, Detroit, and Philadelphia. After one riot in Chicago, the ratio of black to white men arrested was 20:1, which newspapers pointed out was unlikely the ratio of actual infractions. Over the 1930's and 1940's, crime statistics moved from being broken down by native and foreign-born to dropping the foreign-born distinction to focus on black versus white.² A recent popular literature (Muhammed, 2010; Alexander, 2012) traces more recent racial discourse on incarceration back to the "condemnation of blackness" following the influx of black men into cities during Great Migration. W. E. B. Dubois and others blamed black criminality on a combination of discrimination, both in getting jobs and the justice process, overcrowded and segregated housing, and the disproportionate availability of liquor stores and saloons in black neighborhoods (Du Bois, 1899).

III. Data

A. Constructing a Matched Sample

Previous papers have been hampered by the lack of disaggregated data on crime in this time period. Other papers (Muller, 2012) have used aggregate incarceration statistics by state, race, and census year to estimate the impact of migration on incarceration.³ Other papers use city-level crime reports for the selected group of cities for which they are available before the introduction of the FBI's Uniform Crime Reports in the 1940's (Feigenbaum and Muller, 2016). I aim to

² Moehling and Piehl (2009) show that in the early 20th century, immigrants were no more likely to be incarcerated for violent crimes at most ages, counter to the political rhetoric at the time.

³ Muller (2012) also briefly considers a micro-approach using the 1940 IPUMS 1% census sample, but uses 1935 location to determine pre-migration location; this paper expands on this approach with a larger dataset, different estimation strategy, and longer time horizon.

estimate the causal impact of migrating on the individual propensity to be incarcerated. Lacking individual data about crime or arrest rates, I use individual-level information about being incarcerated in the US census as my measure of incarceration.

I collect the full universe of southern-born prisoners and non-prisoners from the 1940 census. I then link these individuals back to themselves in their childhood household in 1920. This enables me to control for household background and childhood county fixed effects. My primary sample comprises southern-born black men who are between 23 and 35 years of age in the 1940 census; the age restriction is required so that I can find full households in the census 20 years previously.⁴ I also consider white men and northern-born black men in order to examine possible explanations for my main findings and create similar linked samples for these groups.

I identify prisoners using the group quarters and relationship to household head variables in the full count 1940 census.⁵ I use the restricted data available on the NBER server because I require names to be able to link individuals to 1920.⁶ I then use all non-institutionalized men in the relevant age range as the comparison group.

To match individuals backwards from 1940 to 1920, I follow an iterative procedure used in Abramitzky, Boustan, and Eriksson (2012). I first standardize first and last names using the NYSIIS algorithm (Atack and Bateman 1992) which spells names with the same phonetic sound identically. Individuals are then matched by first name, last name, state of birth, race, and age from

⁴ 3% of black 15 year olds and 6.3% of black 16 year olds are found outside their household in 1920.

⁵ Specifically, I classify as prisoner anyone with *gqtype* = 2. I then remove men who report a relationship to household head that is not “prisoner”, “convict”, “inmate”, “boarder”, “lodger” or missing. I add to this set anyone with a different group quarters value who reports “prisoner” or “convict” as relationship to household head. Finally, I looked up by hand approximately 1,000 images for which the relationship to household head was missing for all individuals and that I suspected were either prisons or hospitals. This process is necessary because the full count census group quarters variable has known inconsistencies (Ruggles et al., 2016).

⁶ A more complicated method is used in Eriksson (2015) which used data collected before the full count 1940 data was cleaned and included a group quarters variable; incarceration rates of southern-born black men are similar across the two datasets.

the 1940 census to the 1920 census. I allow individuals to misreport their age by up to two years in either direction. Inherent in any matching procedure is a trade-off between sample size and accuracy. Furthermore, accuracy may come at the cost of representativeness if the uniqueness of an individual's name is correlated with his socio-economic status, so there is a tradeoff between representativeness of the matched sample and accuracy.

To prioritize accuracy in order to minimize measurement error from placing individuals in the incorrect household or county of birth, in a second robustness sample I go one step further and require individuals to be unique by name/birth state/race within a five year age band. That is, in each census, I drop individuals who have an exact match based on name, state of birth, and race whose age is within plus or minus two years of the individual's age. The match rates in my study are consistent with the literature, averaging around 30 percent. Match rates fall to 12 percent using the second, more restrictive, matching procedure. I reweight my data to account for differences in match rates across incarceration status.

I construct two additional robustness samples where I do not standardize names but use the raw reported name. This lowers the chance of picking the wrong person, but potentially at a cost of missing matches that have names spelled slightly differently.⁷ The two samples use the iterative method and five year age band above.

In Table 1, I examine the representativeness of the matched samples based on both adult and childhood characteristics. I find, consistent with other literature that the matched sample is slightly positively selected from the population based on socio-economic characteristics which are observable. Looking first at characteristics in 1940, men in the matched sample made on average \$221.5 in 1939, about 3 percent of the mean. They are slightly more likely to report making wage

⁷ Recent work (Bailey et al., 2017) suggests that standardizing names results in more false positive matches than using raw names. Therefore, I present results without name standardization as well as the standard method.

income (0.2 percentage points). Men in the matched sample are also slightly more educated (0.3 years) and older (0.22 years). They are less likely to be living in urban areas (1 percentage point) and are equally likely to be migrants either over 20 years or the past five years. The matched sample is mostly balanced in terms of 1920 household head characteristics. Men in the matched sample have slightly less literate household heads (0.02 percentage points) but the magnitude is not large. The largest differences come from the fact that the household heads of the matched sample are less likely to be tenant farmers (and therefore more likely to be in the left-out urban occupation category) and are less likely to own their own home. Overall, the patterns aren't consistent with any one direction of bias. In addition, I control for 1920 household characteristics in the regressions and this does not change the main results.

B. Incarceration Rates and Summary Statistics

I first examine incarceration rates in the north and south by race in Table 2. The first (3) columns present incarceration rates based on region of residence from the published census bureau statistics. The incarceration rate for black men is 1.9 percentage points relative to 0.441 for white men. These rates are higher in the north: 3.24 percentage points for black men in the north versus 1.69 for black in men in the south. White men in the north are 0.8 percentage points more likely to be incarcerated than white men in the south: 1.38 percentage points versus 0.5 percentage points.

These published statistics do not tell us anything about the experience of specific migrants, because place of birth is not available. Therefore, I use my matched sample in Columns (4) through (6) to calculate incarceration rates for southern-born men living in either region. Rates are similar to the first three columns: Northern black incarceration rates are 2.94 versus 1.74 in the south; white rates are 1.43 in the north versus 0.56 in the south. The different rates between the north and

south using both methods suggest that the causal impact of moving north will be positive and large. For southern-born blacks, the difference is 1.2 percentage points.

Table 3 presents summary statistics for the matched sample of southern-born black men. I separate prisoners from non-prisoners. First, we see that prisoners are more likely to be migrants than non-prisoners: 25.3 percent versus 16.1 percent. Prisoners are also more likely to be recent migrants who moved north in the past five years: 5.4 percent relative to 2.1 percent. Prisoners are less educated than non-prisoners (1.5 to 2 years) and migrants are more educated than non-migrants (1.4 to 2 years). Overall levels of education are low, between 4.9 and 7.9 years of completed schooling. The average age in my sample is around 28.5 years. I look at urban residence status for non-prisoners only because prisons could be in urban or rural areas and prisoners do not choose where to be incarcerated. Among non-prisoners, 90.9 percent of migrants live in urban areas with a population of more than 2,500; 41.7 percent of non-prisoners remaining in the south live in urban areas. The summary statistics yield two preliminary conclusions. First, there seems to be positive selection into migration based on education. Second, given that most migrants move to urban areas, higher incarceration rates in cities are likely part of the explanation for the regional differences.

IV. Estimating the causal impact of migration on incarceration

My main estimation estimates the effect of moving from the South to the North for southern-born black men between 1920 and 1940. I use a linear probability model as my primary estimation strategy:

$$(1) \quad \text{Prisoner}_i = \beta_0 + \beta_1 \text{Migrant}_i + \beta_2 \text{education}_i + X_i \beta_3 + u_i$$

where *Prisoner* is a binary variable equal to one if the individual is incarcerated in 1940, *Migrant* is equal to one if the individual migrated between 1920 and 1940, *education* is years of school completed, and X_i is a vector of controls, including age fixed effects and, for some regressions, household-level controls from the 1920 census.

In the presence of positive selection into migration, we would expect the OLS estimate of β_1 to be biased downwards: if higher ability individuals choose to move North in search of higher wages and if higher ability individuals are less likely to be incarcerated, we will underestimate the true β_1 . I address this in multiple ways. First, I show that controlling for observable characteristics, particularly education, is important. Second, I consider whether migrants are selected at a state or county level by including state and then county-level fixed effects. Next, I control for household-level observable characteristics from 1920: occupation of household head, literacy of both parents, and home ownership. Finally, I control for household fixed effects where the household is defined based on the child's family in 1920. This final strategy allows me to control for any across-family unobservable characteristics which are correlated with migration and incarceration. Any selection remaining in the estimates will be due to within-family unobservable characteristics (Collins and Wanamaker, 2013; Abramitzky et al. 2012).

In Table 4, I look at the determinants of migration in my sample to determine to what extent individuals are selected into migration based on childhood family characteristics. Each column regresses migrant status on variables defined in the 1920 household: urban status, household head literacy, household head home ownership status, and dummies for four different occupational categories: tenant farmer, owner-occupied farmer, farm laborer, general laborer. These categories make up over 80% of the occupations of black men in the south in 1920; other occupations are mainly urban-based occupations such as carpenter and blacksmith.

Column (1) adds only age fixed effects. In Column (2), I add state fixed effects, and in Column (3) I add county fixed effects. Across the three specifications, men who grew up in urban areas are more likely to migrate than those who grew up in rural areas. In the final column, the difference is 2.86 percentage points. When controlling for county fixed effects, men whose household head is literate are 1 percentage point more likely to migrate. Surprisingly, men from households who owned their homes are 1.5 percentage points more likely to migrate than men whose households did not own their home. This is likely due to higher home-ownership in farming areas which are more rural. Finally, men whose households work in farming or as laborers are less likely to migrate than men whose household heads worked in the comparison category. Within farming, those who own their house are the least likely to migrate while those from a farm laborer background are most likely to migrate. Nonetheless, when controlling for county fixed effects, the differences are small, around 1 percentage point. Selection into migration seems weak, if anything.

V. Results

In this section, I estimate the effect of migration on the probability of incarceration for southern-born black and white men. I then show that most of the effect is driven by recent migration. I show that recent migrants have lower-paying occupations and are less education than long-term migrants. Finally, I show that southern-born black men are actually *less* likely to be incarcerated than northern-born black men, despite lower income and education levels. I also find some evidence that larger migrant networks lower incarceration rates but this is not necessarily a causal association.

A. Effects of Migration on the probability of being incarcerated

Table 5 presents estimates of Table 1 for the primary estimation sample of southern-born black men. With no controls, the effect of migrating is 1.23 percentage points, the same as the raw difference in Table 2. I then add education in the remaining columns, as well as adding state, county, and then household fixed effects in Columns (3)-(5); all fixed effects are defined based on childhood residence in 1920. The main effect increases to 1.56 percentage points when including education due to the fact that migrants have more education than non-migrants. This difference remains stable when adding state and county of origin fixed effects. Finally, the estimate falls to 1.04 percentage points with household fixed effects. In Table A.1, I show that these patterns are consistent across the four matching methods described above. I struggle with statistical power in the matches using the five year age band, but the point estimates are similar.

We might worry that the sample which identifies the main coefficient in the household fixed effects model is different in some way than the full sample. Therefore, Table 6 restricts to only men for whom at least one brother is also in the matched sample. Further, I restrict to households which have at least one migrant and one non-migrant. The patterns between Table 5 and Table 6 are remarkably similar: the effect is 1.5 percentage points with county fixed effects and 1.04 percentage points with household fixed effects. Surprisingly, this suggests *negative* selection across households. The likely explanation is that migrants come from more urban areas and urban areas have higher incarceration rates even in the south.

Table 7 turns to estimating the effect of migration on southern-born white men. While incarceration rates are lower for white men than black men, the effect of migrating north is 0.8 percentage points when not controlling for education and between 0.89 and 0.95 when controlling for education and state or county fixed effects. Finally, the coefficient falls to 0.6 percentage points when including household fixed effects.

B. Effects of Recent versus Long-term Migration on Incarceration

One reason incarceration rates are higher in the north could be that migrants are separated from their families and job networks. While real income was higher in the north *on average*, recent migrants likely struggled to find jobs in an environment in which they were competing with immigrants for jobs. If recent migrants held lower paying jobs and were less likely to be employed, we would expect them to be incarcerated more often than long-term migrants.

Table 8 splits migrants into two types: those who arrived between 1935 and 1940 (recent migrants) and those who arrived before 1935 (long-term migrants). This is possible to do in the 1940 census because the census asked for the first time in 1940 where individuals lived five years ago. In Column (1), we see that recent migrants are about three times more likely to be incarcerated than long-term migrants (3.47 versus 1.25 percentage points more likely than non-migrants). This pattern persists when adding state and county of origin fixed effects. The difference in the coefficients is significant until adding household fixed effects.⁸ Nonetheless, in all specifications recent migrants are more likely to be incarcerated than long-term migrants.

Looking at the differences between recent and long-term migrants in Table 9, we see that it is indeed the case that recent migrants earn less than long-term migrants. Average wage income is \$80.10, or 15 percent, lower for recent migrants. Recent migrants are younger with slightly less education (0.2 years) but these differences do not account for the full wage difference which remains significant at \$48.31 when controlling for age and education. Recent migrants are *more likely* to report wage income, likely because long-term migrants are more likely to be self-employed. An alternative measure of income is the *occscore* variable created by the census which

⁸ The household fixed effects estimates are imprecise here, likely because to identify these coefficients requires variation in migration timing within households.

measures median occupational income based on the 1950 census. The advantage of looking at this variable is that it is defined for self-employed as well as wage workers. Occupational income is \$114.9 lower for recent migrants than non-migrants. Overall, the evidence is consistent with recent migrants being poorer and less likely to be working.

C. Effects of Migration within the South versus to the North on Incarceration

Incarceration rates in the north are likely higher because migrants move to urban areas. A large literature has explored the reasons that crime, and therefore incarceration, rates are higher in cities (Glaeser and Sacerdote, 1999). Therefore, I split individuals into three groups: those who remain in their birth state (non-migrants), those who migrate out of their birth state by 1940 but who are still in the South, and those who migrate North. Most migrants within the south also migrate to cities such as Atlanta, Baltimore, and Washington D.C. which have higher incarceration rates than rural areas.

Table 10 shows that men who migrate north have a higher probability of being incarcerated than those who migrate within the south but that migration within the south raises the probability of incarceration. Men who migrate north are 1.8 percentage points more likely to be incarcerated than those remaining in their birth state; men migrating within the south increase their probability of incarceration by 1.07 percentage points. The difference in the two coefficients is statistically significant and these patterns remain with state or county of origin fixed effects.

D. Comparing Northern-born black men to Southern-born black men

Next, in Table 11, I compare incarceration rates of Northern- and Southern-born black men. Not controlling for education or age, Northern-born men are 0.41 percentage points *more*

likely to be incarcerated than Southern-born black men. Controlling for age, this coefficient falls slightly to 0.35 percentage points. Northern-born men have on average 1.5 years more education than Southern-born men, so when controlling for education, the gap increases to 1.02 percentage points. The two groups have approximately the same income levels, despite the higher levels of education of Northern-born suggesting that Southern-born men are positively selected along an unmeasured margin relate to the Northern-born. Southern-born men are also more likely to live in cities in the North, which would tend to go against the pattern of being less likely to be incarcerated.

E. Suggestive evidence that larger networks were associated with lower incarceration

Finally, I consider whether migrant networks helped or hurt migrants in terms of incarceration in Table 12. I calculate a variable equal to the proportion of a state's black population that was born in the same state as the migrant. The mean of this new variable is 11 percent, suggesting large clustering based on birth state. Given that migrants tended to migrate to the city served by the closest railroad, this makes sense. For example, almost all Tennessee migrants went to Illinois or Michigan, while migrants from North Carolina went to Pennsylvania, New Jersey, and New York.

I run a simple regression of individual incarceration status on this variable, and then add controls for age and education, and finally birth state fixed effects which are identified off of variation in state of residence within a state of birth. With no controls, the coefficient is -0.0836, meaning a 10 percentage point increase in the proportion of a state's black population from your own state decreases your propensity to be incarcerated by 0.836 percentage points. A one standard deviation (6.2 percentage points) change in the network size decreases the propensity to be

incarcerated by 0.51 percentage points. Adding controls for education and age increases the coefficient slightly in Column (2). Finally, controlling for state of birth fixed effects lowers the coefficient to -0.0541. A one standard deviation increase in network size then would reduce the probability of incarceration by 0.33 percentage points.

F. Estimating the Effect of the Great Migration on Black Incarceration

Migrating to the North between 1920 and 1940 roughly doubled the probability of incarceration for a black man, increasing the probability by between 1 and 1.5 percentage points. In this same time period, the black male incarceration rate increased from 0.895 to 1.962, a 1.07 percentage point increase. The Great Migration increased the proportion of southern-born black men between ages 23 and 35 from 14.3 percent in 1920 to 18.4 percent in 1940.

A back of the envelope calculation to estimate the effect of the great migration on the overall black incarceration rate assumes that incarceration rates in the south would have remained the same in 1940 if those 5.1 percent of black men had not left. One can therefore calculate the change in the incarceration rate caused by the great migration as the change in the black population in the north times the effect on an individual's probability of incarceration from migrating. This is between 0.051 and 0.0764 percentage points. That is, the black incarceration rate would have been 0.051-0.0764 percentage points lower in the absence of the part of the Great Migration which took place between 1920 and 1940. This is between 4.8 and 7.2 percent of the change in the black incarceration rate between 1920 and 1940. This effect is much lower than, for example, the effect of Rosenwald schools on black incarceration (0.85 percentage points) found in Eriksson (2015).

VI. Conclusion

Incarceration rates of black men increased faster than those of white men in the first half of the 20th century despite some convergence in wages and education levels between white and black men. This paper estimates the contribution of a major event during this time period, the Great Black Migration which moved millions of men from the rural South to Northern cities. While migrants moved North in hopes of better opportunities and less discrimination, contemporary observers argued that black men faced worse discrimination from the criminal justice system than they did in the South as many struggled to find jobs. This paper illustrates another negative dimension of the Great Migration for the migrants themselves.

References

- Aaronson, Daniel and Bhashkar Mazumder. 2011. "The Impact of Rosenwald Schools on Black Achievement," *Journal of Political Economy*, 119(5): 821-888.
- Abramitzky, Ran, Leah Boustan, and Katherine Eriksson. 2012. "Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration". *American Economic Review*. 102(5): 1832-1956.
- Alexander, Michelle. 2012. *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. The New Press: New York, New York.
- Bailey, Martha, Morgan Henderson, and Catherine Massey. 2017. "How do Automated Linking Methods Perform? Evidence from the LIFE-M Project". Manuscript.
- Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach". *Journal of Political Economy* 76.
- Black, Dan A., Seth G. Sanders, Evan J. Taylor, and Lowell J. Taylor. 2015. "The Impact of the Great Migration on the mortality of African-Americans: Evidence from the Deep South". *American Economic Review* 105(2): 477-503.
- Borjas, George. 1987. "Self-selection and the Earnings of Immigrants". *American Economic Review* 77(4): 531-53.
- Boustan, Leah Platt. 2007. "Competition in the Promised Land: Black Migration and Racial Wage Convergence in the North, 1940-1970". *Journal of Economic History* 69(3): 755-82.
- Bowler, Alida. 1931. "Recent Statistics on Crime and the Foreign Born." Pp. 83–194 in Report on Crime and the Foreign Born, edited by Edith Abbott. Washington, D.C.: U.S. Government Printing Office.
- Cahalan, Margaret Werner. 1986. "Historical Corrections Statistics in the United States, 1850-1984." Rockville, MD: U.S. Department of Justice, Bureau of Justice Statistics. <http://www.ncjrs.gov/pdffiles1/pr/102529.pdf>.
- Chay, Kenneth and Kaivan Munshi. 2011. "Slavery's Legacy: Black Mobilization in the Postbellum South" Manuscript, Brown University.
- Collins, William J. 1997. "When the Tide Turned: Immigration and the Delay of the Great Black Migration." *Journal of Economic History* 56:607–32.
- Collins, William and Marianne Wanamaker. 2014. "Selection and Economic Gains in the Great Migration of African Americans: New Evidence from Linked Census Data." *American Economic Journal: Applied Economics* 6(1): 220-252.

- Donohue, John, James J. Heckman and Petra E. Todd. 2002. "The Schooling of Southern Blacks: The Roles of Legal Activism and Private Philanthropy, 1910-1960," *The Quarterly Journal of Economics*, 117(1): 225-268.
- Du Bois, W.E.B. 1899. *The Philadelphia Negro*. Schocken Books Inc., New York, New York.
- Eriksson, Katherine. 2015. "Access to Schooling and the Black-White Crime Gap in the Early 20th Century US South: Evidence from Rosenwald Schools". NBER Working Paper No. 21727.
- Eriksson, Katherine and Gregory Niemesh 2016. "Death in the Promised Land: The Great Migration and Black Infant Mortality". Manuscript.
- Fajnzylber, Pablo, Daniel Lederman and Norman Loayza. 2002. "Inequality and Violent Crime." *Journal of Law and Economics*.
- Feigenbaum, James and Christopher Muller. 2016. "Lead Exposure and Violent Crime in the Early Twentieth Century." *Explorations in Economic History* 62: 51-86.
- Ferrie, Joseph. 1996. "A New Sample of Males Linked from the Public Use Micro Sample of the 1850 U.S. Federal Census of Population to the 1860 U.S. Federal Census Manuscript Schedules." *Historical Methods* 29: 141-56.
- Glaeser, Edward L. and Bruce Sacerdote. 1999. "Why is there more crime in cities?" *Journal of Political Economy* 107(S6): S225-S258.
- Glaeser, Edward L., Bruce Sacerdote, and Jose A. Scheinkman. 1996. "Crime and Social Interactions." *Quarterly Journal of Economics*, 111(2): 507-548
- Gould, Eric D., Bruce A. Weinberg, and David B. Mustard. 2002. "Crime Rates and Local Labor Market Opportunities in the United States: 1979-1997." *Review of Economics and Statistics* 84(1): 45-61.
- Grossman, James, 1989. *Land of Hope: Chicago, Black Southerners, and the Great Migration*. Chicago: University of Chicago Press.
- Langan, Patrick. 1991. *Race of Prisoners Admitted to State and Federal Institutions, 1926-86*. Department of Justice: Washington, D.C.
- Lemann, Nicholas, 1991. *The Promised Land : The Great Black Migration and How It Changed America*. New York: Vintage Books.
- Margo, Robert A., 1990. *Race and Schooling in the South, 1880-1950*. Chicago: University of Chicago Press.
- Moehling, Caroline, and Anne Morrison Piehl. 2014. "Immigrant Assimilation into US prisons, 1900-1930" *Journal of Population Economics* 27(1): 173-200.

- Mossell, Sadie. 1921. "The Standard of Living among One Hundred Negro Migrant Families in Philadelphia." *The Annals of the American Academy of Political Science* 98: 173-218.
- Muhammed, Khalil Gibran. 2010. *The Condemnation of Blackness: Race, Crime, and the Making of Modern Urban America*. Harvard University Press: Cambridge, Massachusetts.
- Muller, Christopher. 2012. "Northward Migration and the Rise of Racial Disparity in American Incarceration, 1880-1950." *American Journal of Sociology* 118(2): 281-326.
- Naidu, Suresh. 2010. "Recruitment Restrictions and Labor Markets: Evidence from the Post-Bellum U.S. South" *Journal of Labor Economics*. April.
- Oshinsky, David M. 1996. *"Worse Than Slavery": Parchman Farm and the Ordeal of Jim Crow Justice*. New York: The Free Press.
- Raphael, Stephen, and Michael A. Stoll. 2007. "Why Are so Many Americans in Prison?" in Raphael and Stoll, eds *Do Prisons Make us Safer? The Benefits and Costs of the Prison Boom*. Russell Sage Foundation.
- 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). Minneapolis: University of Minnesota.
- Sellin, Thorsten. 1928. "The Negro Criminal: A statistical note". *Annals of the American Academy of Political and Social Science*. November: 52-64.
- Smith, James and Finis Welch. 1989. "Black Economic Progress after Myrdal" *Journal of Economic Literature* 27(2): 519-64.
- U.S. Department of Commerce. 1904. *Occupations at the Twelfth Census*. Washington, D.C.
- U.S. Department of Commerce. 1914. *Prisoners and Juvenile Delinquents 1910*. Washington, D.C.:
- U.S. Department of Commerce. 1926. *Prisoners 1923*. Washington, D.C.
- U.S. Department of Commerce. 1943. *Sixteenth Census of the United States: 1940. Population. Special Report on institutional population, 14 years and over, characteristics of inmates in penal institutions, and in institutions for the delinquent, defective, and dependent*. U.S. Government: Washington, D.C.
- U.S. Department of Commerce. 1953. *United States Census of Housing: 1950, Parts 2-6*. Washington, D.C.

U.S. Department of Commerce. 1963. *United States Census of Population: 1960, Subject Reports: Inmates of Institutions*. Washington, D.C.: Government Printing Office.

U.S. Department of Commerce. 1973. *United States Census of Population: 1970, Subject Reports: Persons in Institutions and Other Group Quarters*. Washington, D.C.

U.S. Department of Commerce and Labor. 1907. *Special Reports: Prisoners and Juvenile Delinquents in Institutions: 1904*. Washington, D.C.

U.S. Department of Commerce and Labor, Bureau of the Census. 1906. *Wealth, Debt and Taxation*. Washington D.C.

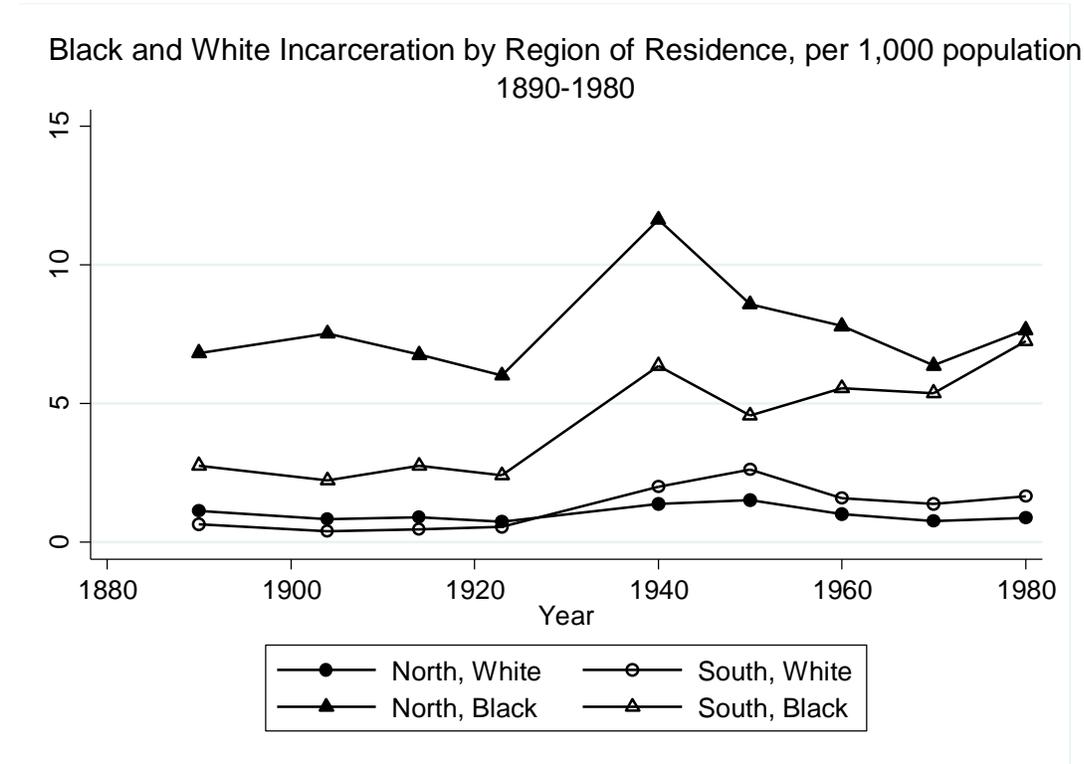
U.S. Department of the Interior. 1895. *Report on Crime, Pauperism, and Benevolence in the United States at the Eleventh Census: 1890, Part II*. Washington, D.C.

Western, Bruce, Jeffrey R. Kling, and David Weiman. 2001. "The Labor Market Consequences of Incarceration". *Crime and Delinquency*. 47(3): 410-427.

Wilkerson, Isabel. 2010. *The warmth of other suns: the epic story of America's great migration*. New York: Random House.

Figures

Figure 1: Black and White Incarceration rates, 1890-1980



Notes: Incarceration figures taken from US Department of Interior (1895), U.S. Department of Commerce and Labor (1907), US Department of Commerce (1914, 1926, 1943, 1955, 1963, 1973, 1983). Population (denominator) taken from IPUMS (Ruggles et al 2010). Figure depicts total number of prisoners by race/census region/year divided by relevant population, where population is interpolated between census years for non-census years. Men and women included—multiply figure numbers by 1.8 to calculate male incarceration rates.

Tables

Table 1: Incarceration rates in the Census vs. published statistics

	Census Data—Southern-born black men			Published Statistics—All black men		
	Overall	North	South	Overall	North	South
Overall	0.998	2.129	0.868	0.577	0.469	0.769
Black	1.962	2.943	1.739	1.900	3.245	1.691
White	0.628	1.437	0.558	0.441	1.387	0.499

Notes: Incarceration rates are based on Southern-born men only in the first three columns. Incarceration rates in columns (4)-(6) are calculated based on the published population totals from U.S. Department of Commerce (1943) Table 8.

Table 2: Characteristics of migrants and non-migrants, 1940

	<u>Non-prisoners</u>			<u>Prisoners</u>	
	<i>All</i>	<i>Migrants</i>	<i>Non-migrants</i>	<i>All</i>	<i>Non-migrants</i>
Migrant	0.161 (0.367)			0.253 (0.435)	
Recent Mig	0.021 (0.143)			0.054 (0.225)	
Education		7.976 (3.026)	5.919 (3.295)	6.353 (3.109)	4.957 (3.116)
Income		446.96 (474.74)	222.618 (289.313)	--	--
Age		29.09 (3.615)	28.44 (3.624)	28.76 (3.511)	28.19 (3.541)
Urban		0.909 (0.287)	0.417 (0.493)	--	--

Notes: N= 410,247. Income measured in 1940 dollars; Yearly income is only reported in 1940 for wage earners and is capped at \$5,000.

Table 3: Comparing matched sample to the population

	Population	Matched Sample	Difference
<i>Panel A: Characteristics of Sample in 1940</i>			
Income	5630.48 (6240.95)	5940.16 (6259.28)	221.50*** (16.99)
=1 if Income>0	0.708 (0.454)	0.716 (0.451)	0.002* (0.001)
Education	5.618 (3.425)	6.074 (3.498)	0.301*** (0.091)
=1 if Urban	0.483 (0.499)	0.481 (0.499)	-0.010*** (0.001)
Age	28.53 (4.810)	28.76 (3.870)	0.218*** (0.011)
<i>Panel B: Migration Choices, 1920-1940</i>			
=1 if migrate to North	0.161	0.166	0.005* (0.003)
=1 if recent migrant to North	0.024	0.023	-0.001** (0.001)
<i>Panel C: Characteristics of Household Head in 1920</i>			
=1 if Literate	0.620 (0.485)	0.619 (0.485)	-0.002* (0.001)
=1 if Owns home	0.758 (0.427)	0.724 (0.446)	-0.034*** (0.001)
=1 if Owner-Occupier Farmer	0.131 (0.337)	0.132 (0.338)	0.001* (0.001)
=1 if Farm Laborer	0.089 (0.280)	0.085 (0.281)	-0.003*** (0.001)
=1 if Laborer	0.092 (0.287)	0.091 (0.289)	-0.001 (0.002)
=1 if Tenant Farmer	0.456 (0.498)	0.434 (0.495)	-0.021** (0.001)

Notes: Income and urban status are only defined for non-prisoners. Income is measured in 2010 dollars and is only reported for wage earners in the 1940 census. I restrict to men aged 23-35 in the 1940 Census or aged 3-15 in the 1920 Census.

Table 4: Selection into Migration to the North

	Outcome =1 if Migrated to the North		
	(1)	(2)	(3)
=1 if Urban	0.0495*** (0.0091)	0.0231*** (0.0034)	0.0286*** (0.0028)
=1 if Head is literate	-0.0115*** (0.0035)	0.0108*** (0.0013)	0.0100*** (0.0013)
=1 if Head Owns home	-0.0739*** (0.0081)	-0.0156*** (0.0025)	-0.0151*** (0.0024)
<i>Household Head's Occupation:</i>			
=1 if Tenant Farmer	-0.0716*** (0.0072)	-0.0177*** (0.0032)	-0.0145*** (0.0030)
=1 if Owner-occupier Farmer	-0.0187*** (0.0034)	-0.0169*** (0.0022)	-0.0165*** (0.0021)
=1 if Farm Laborer	-0.0357*** (0.0042)	-0.0089*** (0.0025)	-0.0064*** (0.0024)
=1 if Laborer	-0.0324*** (0.0041)	-0.0162*** (0.0027)	-0.0163*** (0.0026)
Controls	Age	+State FE	+County FE
R-squared	0.0125	0.0521	0.0608
N	410,247	410,247	410,247

Notes: All characteristics measured in 1920 in the childhood home. Urban is defined as a city with more than 2500 residents. Household's head occupation's left out category is all other occupations, mostly more urban occupations such as carpenter, blacksmith, etc.

Table 5: Effect of migration on the probability of being incarcerated, Southern-born black men

<i>Outcome:</i>	(1)	(2)	(3)	(4)	(5)
	In Prison	In Prison	In Prison	In Prison	In Prison
=1 if migrant	0.0123*** (0.0010)	0.0156*** (0.0011)	0.0154*** (0.0011)	0.0155*** (0.0011)	0.0104*** (0.0035)
Education		-0.0019*** (0.0001)	-0.0023*** (0.0001)	-0.0024*** (0.0001)	-0.0017*** (0.0004)
FE	None	None	State	County	Household
Sample mean	0.0161	0.0161	0.0160	0.0161	0.0191
R ²	0.0012	0.0034	0.0058	0.0134	0.8671
N	410,247	399,682	399,682	399,682	399,682

Notes: Regressions control for age and age-squared; ages are restricted to between 23 and 35 in 1940. Men are matched to their childhood census locations in the 1920 census. State, county, and household fixed effects are defined in 1920. Standard errors are clustered at the childhood county level.

Table 6: Effect of migration on the probability of being incarcerated, matched brother-pairs of Southern-born black men

<i>Outcome:</i>	(1)	(2)	(3)	(4)`
	In Prison	In Prison	In Prison	In Prison
=1 if in the North in 1940	0.0148*** (0.0016)	0.0148*** (0.0016)	0.0150*** (0.0016)	0.0104*** (0.0022)
Education	-0.0020*** (0.0001)	-0.0022*** (0.0001)	-0.0023*** (0.0002)	-0.0017*** (0.0002)
FE	None	State	County	Household
Sample mean	0.0162	0.0162	0.0162	0.0162
R ²	0.0037	0.0061	0.0229	0.6137
N	154,278	154,278	154,278	154,278

Notes: Regressions control for age and age-squared; ages are restricted to between 23 and 35 in 1940. Men are matched to their childhood census locations in the 1920 census. State, county, and household fixed effects are defined in 1920. Standard errors are clustered at the childhood household level.

Table 7: Effect of migration on the probability of being incarcerated, Southern-born matched white brother-pairs

<i>Outcome:</i>	(1)	(2)	(3)	(4)	(5)
	In Prison	In Prison	In Prison	In Prison	In Prison
=1 if migrant	0.0082*** (0.0005)	0.0089*** (0.0005)	0.0095*** (0.0006)	0.0095*** (0.0006)	0.0060*** (0.0016)
Education		-0.0012*** (0.0001)	-0.0012*** (0.0001)	-0.0012*** (0.0001)	-0.0008*** (0.0001)
FE	None	None	State	County	Household
Sample Mean					
R ²	0.0009	0.0036	0.0040	0.0080	0.8653
N	1,550,358	1,510,822	1,510,822	1,510,822	1,510,822

Notes: Regressions control for age and age-squared; ages are restricted to between 23 and 35 in 1940. Men are matched to their childhood census locations in the 1920 census. State, county, and household fixed effects are defined in 1920. Standard errors are clustered at the childhood census county level.

Table 8: Recent vs. long-term migrants, black men

<i>Outcome:</i>	(1) In Prison	(2) In Prison	(3) In Prison	(4) In Prison
=1 if Recent Migrant	0.0347*** (0.0031)	0.0345*** (0.0031)	0.0347*** (0.0031)	0.0152* (0.0087)
=1 if Long-term Migrant	0.0125*** (0.0010)	0.0124*** (0.0011)	0.0124*** (0.0011)	0.0096*** (0.0040)
Education	-0.0020*** (0.0001)	-0.0022*** (0.0001)	-0.0024*** (0.0001)	-0.0017*** (0.0004)
p-value	0.000	0.000	0.000	0.541
FE	None	State	County	HH
Sample Mean	0.0161	0.0161	0.0161	0.0161
R ²	0.0039	0.0063	0.0139	0.8672
N	399,682	399,682	399,682	399,682

Notes: Regressions control for age and age-squared; ages are restricted to between 23 and 35 in 1940. Men are matched to their childhood census locations in the 1920 census. State, county, and household fixed effects are defined in 1920. Standard errors are clustered at the childhood census county level. The final column restricts to households in which at least two sons are matched. P-values are reported for the null hypothesis that the coefficient for recent migrant is equal to the coefficient for long-term migrant.

Table 9: Differences between recent and long-term migrants

	Recent Migrant	Long-term Migrant	Difference
Age	28.35 (3.590)	29.57 (3.697)	-1.225*** (0.024)
Income	548.10 (433.51)	628.21 (493.23)	-80.10*** (3.175)
=1 if Income>0	0.796 (0.402)	0.768 (0.421)	0.029*** (0.003)
Occupational Income	17.01 (8.845)	18.16 (8.610)	-1.149*** (0.060)
Education	7.498 (3.321)	7.691 (3.058)	-0.192*** (0.022)
=1 if Urban	0.851 (0.355)	0.922 (0.266)	-0.071*** (0.002)

Table 10: Effect of interstate vs. interregional migration, matched brother pairs of black men

<i>Outcome:</i>	(1) In Prison	(2) In Prison	(3) In Prison	(4) In Prison
=1 if North	0.0174*** (0.0011)	0.0180*** (0.0011)	0.0181*** (0.0011)	0.0127*** (0.0035)
=1 if outside birth state	0.0096*** (0.0010)	0.0107*** (0.0010)	0.0107*** (0.0011)	0.0100*** (0.0036)
Education	-0.0020*** (0.0001)	-0.00231*** (0.0001)	-0.0024*** (0.0001)	-0.0018*** (0.0004)
p-value	0.000	0.000	0.000	0.5574
Sample Mean	0.0161	0.0161	0.0161	0.0161
FE	None	State	County	Household
R ²	0.00400	0.00649	0.01417	0.86730
N	399,682	399,682	399,682	399,682

Notes: Regressions control for age and age-squared; ages are restricted to between 23 and 35 in 1940. Men are matched to their childhood census locations in the 1920 census. State, county, and household fixed effects are defined in 1920. Standard errors are clustered at the childhood census county level. I drop individuals in federal prisons for this analysis. P-values are associated with the null hypothesis that the coefficients for migration within the South and migrating to the North are equal.

Table 11: Northern-born versus Southern-born Black male incarceration

<i>Outcome:</i>	(1) In Prison	(2) In Prison	(3) In Prison
=1 if Northern-Born	0.0041*** (0.0006)	0.0035*** (0.0006)	0.0102*** (0.0007)
Education			-0.0045*** (0.0001)
FE	None	Age	Age
Sample Mean, Southern-Born	0.0161	0.0161	0.0161
R ²	0.0039	0.0063	0.0139
N	283,876	283,876	275,197

Notes: Regressions control age fixed effects; ages are restricted to between 23 and 35 in 1940. Table uses unmatched men living in the North in 1940. Standard errors are clustered at the state of residence level.

Table 12: Association between Migrant Network Size and Incarceration, Southern-born Black Men

<i>Outcome:</i>	(1) In Prison	(2) In Prison	(3) In Prison
Proportion of population from birth state	-0.0836*** (0.0060)	-0.1014*** (0.0059)	-0.0541*** (0.0073)
Education		-0.0043*** (0.0001)	-0.0045*** (0.0001)
FE	Age	Age	State of birth
Sample Mean	0.0295	0.0295	0.0295
R ²	0.0016	0.0083	0.0095
N	190,995	185,660	185,660

Notes: Regressions control for age fixed effects; ages are restricted to between 23 and 35 in 1940. Men are matched to their childhood census locations in the 1920 census. Table uses unmatched men living in the North in 1940. Standard errors are clustered at the state of residence level.

Table A.1: Robustness of Main Results in Table 5 to Matching Procedure

<i>Outcome:</i>	(1)	(2)	(3)	(4)	(5)
	In Prison				
<i>Panel A: Match 1</i>					
=1 if migrant	0.0123*** (0.0010)	0.0156*** (0.0011)	0.0154*** (0.0011)	0.0155*** (0.0011)	0.0104*** (0.0035)
N	410,247	399,682	399,682	399,682	399,682
<i>Panel B: Match 2</i>					
=1 if migrant	0.0123*** (0.0019)	0.0151*** (0.0019)	0.0142*** (0.0019)	0.0137*** (0.0019)	0.0111 (0.0196)
N	102,008	99,375	99,375	99,375	102,008
<i>Panel C: Match 3</i>					
=1 if migrant	0.0123*** (0.0011)	0.0149*** (0.0011)	0.0151*** (0.0011)	0.0151*** (0.0011)	0.0116** (0.0050)
N	280,372	273,307	273,307	273,307	273,307
<i>Panel D: Match 4</i>					
=1 if migrant	0.0123*** (0.0015)	0.0149*** (0.0016)	0.0151*** (0.0016)	0.0148*** (0.0016)	0.0126 (0.0112)
N	123,454	120,355	120,355	120,355	120,355
Controls	Age	+Education	+Education	+Education	+Education
FE	None	None	State	County	Household

Notes: All regressions control for age fixed effects; ages are restricted to between 23 and 35 in 1940. Men are matched to their childhood census locations in the 1920 census. State, county, and household fixed effects are defined in 1920. Standard errors are clustered at the childhood census county level.