Examining the North-to-South migration of Black and White people in the United States from 1960 to 2020

Qiong Wu*

August 19, 2024

Abstract

This paper explores the reasons behind the migration of three groups of people from the American North to the American South from 1960 to 2020: Black people, White people, and all races. Using Census and survey data from 1960 to 2020, we attempt to determine a relationship between the median family income of each group of people and the percentage of that group of people who migrated from the North to the South. We find that the strongest correlation between median family income and migration rate exists among Black people, although further, more detailed analyses of the other factors underlying this migration pattern are needed to tell a more complete story.

1 Introduction

Migration is one phenomenon that has existed for as long as people have existed. Studying migration patterns can be a rewarding process that reveals underlying economic motivations and explains human behavior. Migration is an essential part of history; researching and understanding migration can help us to learn from the past. In the United States, specifically, migration has long been a widely discussed topic. Though much of this discussion centers on international migration, it is also important to take notice of migrations that have taken place within the country.

In this paper, we attempt to examine the migration of large numbers of White and Black people from the North to the South from 1960 [1] to 2020, a pattern that is particularly interesting given that the larger and much more well-known migration trend only a few decades prior to this selected period saw people moving in the complete opposite direction, from the South to the North. This earlier trend, also known as the Great Migration, was a notable occurrence in Black American history, and it is not immediately apparent why significant numbers of Black people would move back to the South

*Syracuse University. Email: qwu102@syr.edu.

[1] Originally, our instructions from Prof. Pearson (refer to Appendix B) were to examine data from 1950 to 2020, but due to issues with the 1950 data, as explained on our GitHub page (refer to Appendix B), we were unable to include that sample year in our analysis or report.

so shortly after. Uncovering their motivations forms the foundation for our own motivation for this project.

Many researchers have attempted to explore the reasons behind this second, reversed migration pattern, known as the "New Great Migration" (Frey, 2022), listing possible economic and social factors. We focus on one factor from Hunt et al. (2008), suggesting that income may have played a role in the overall increase in the rates of Black and White migration from the North to the South in the latter half of the 20th century. Our data work attempts to determine whether or not there is a relationship between median family income and migration rate for Black people, White people, and all races, and we found evidence to suggest that there could be. We also found that the relationship among median family income and migration rate is stronger among Black people than among White people.

In this paper, we attempt to situate ourselves within the body of research that exists for this topic. We briefly introduce existing papers on the subject, discuss our own goals, and determine what could be done to improve our data work and analysis. In our Literature Review section, we explore in more detail ideas from previous researchers who have suggested reasons for the observed North-to-South migration from 1960 to 2020, and at its conclusion present our own hypothesis (briefly introduced above). In the Data section, we proceed to describe the data we needed to test our hypothesis, as well as the steps we took to access, extract, and work with the data. The Results section shows what we found in more detail, and the Discussion and Conclusion sections acknowledge the gaps in our work and propose some next steps a person could take to build upon this project.

2 Literature Review

The Great Migration, which lasted from the early to mid-1900s (approximately 1910 to 1970), saw large numbers of people leaving the American South and moving to the North in hopes of creating a better life. A significant number of these migrants were Black, which makes sense in historical context: in the early 1900s, the Jim Crow era (1870s to 1960s) was at its peak in the South, whereas the North was a somewhat less racially discriminatory environment and therefore a somewhat better option for people of color.

Around the 1970s, however, this migration pattern seemed to reverse itself, as large numbers of people, a significant proportion of whom were Black, began to leave the Northern United States and move back to the South in a "New Great Migration" (Frey, 2022). Because it showed a complete turnaround in migration patterns from previous decades, the New Great Migration has been studied by many researchers who have hoped to find the reasons behind this significant shift. In particular, Cromartie and Stack (1989) point out that a majority of the Black people who migrated from the North to the South as part of the Sunbelt phenomenon were return migrants (people who were born or have previously lived in an area) to the South or at least had a previous acquaintance with the region. Why would so many Black people suddenly move back into the region many of them had vacated only decades prior?

Hunt et al. (2008) suggest a general answer to this question: For Black people in the 1970s, the South simply became a socially and economically better place for them to live while the North became a worse option. This idea is also supported by the work of other researchers.

Some of the factors underlying the improved economic and social prospects for Black people in the South in the 1970s included an increase in the demand for labor in the South, caused by growth in

the service and manufacturing sectors (Cromartie & Stack, 1989). Proportionally more Black people in the 1970s were able to take advantage of this Southern economic boom due to "marked declines in the Black/White wage gap" (Hunt et al., 2008) and improvements in Black education (Smith & Welch, 1989). In addition, Newman (1983) claims that the state corporate income tax was lower in the South during the 1970s, creating a "more favorable business climate."

On the other side of the country, it was becoming more difficult for Black workers to make a living in the North. Although the North had industrialized much earlier than had the South, this perhaps became a weakness for the region in the 1970s, when globalization and deindustrialization began to take place (Hunt et al., 2008). In addition, issues such as redlining were also prevalent in the North during that period, leading to increased racial segregation. Eventually, according to Hunt et al. (2008), "de-facto" racial segregation" became lower in the South than in other regions of the United States, including the North. Combined, all of these factors contributed to the North becoming a worse environment for Black people to live in than the South. This decline, combined with perceptions of better economic opportunities in the South due to the reasons listed above, contributed to the migration of Black people from the North to the South.

The above factors may have contributed to the reasons why people, especially Black people and workers, would have wanted to move to the South, but Greenwood & Hunt (1984) provide a reason for why such a large number of people participated in the New Great Migration. According to Greenwood & Hunt (1984), migration is a "self-reinforcing phenomenon." This means that as more and more people migrate, the destination (where people migrated to) becomes a more and more attractive place for even more migrants to move to. Greenwood & Hunt (1984) attribute this to the fact that migrants may influence prices and profitability, as well as further market growth and catalyze increased investment in their destinations. The idea that migration paves the way for further migration to take place may also partially explain why a larger number of Black people moved from the North to the South over time.

Our research is influenced by one specific claim made by Hunt et al. (2008), that a possible cause of the New Great Migration was the decrease in the racial wage gap between Black and White people. This claim suggests that the migration was tied to financial means, which was an idea implied throughout the existing literature, although we had not found it stated outright. We decided to test whether this idea had any merit by examining whether or not a higher median family income was related to a higher rate of North-to-South migration among Black people as compared to among White people. We hypothesized that an increase in the median family income of each group of people would correspond with an overall increase in the percentage of that group of people that moved from the North to the South, then tested data to verify this claim.

3 Data

Our data is drawn from IPUMS USA, a website that provides census and survey data from 1790 to the present. The data we collected from this website is anonymized and on an individual level. This means that we do not have the means to track specific people over time; each observation is only a snapshot of a person's life in a specific year and does not tell their whole story.

We used seven data samples, one from each decade from 1960 to 2020: (1) the 5% sample from 1960, (2) the 1% metro fm1 sample for 1970, (3) the 1% metro sample for 1980, (4) the 1% metro sample for 1990, (5) the American Community Surveys (ACS) sample for 2000, (6) the ACS sample

for 2010, and (7) the ACS sample for 2020.

In order to test the relationship between the median family income in a population and the tendency of members in that population to participate in North-to-South migration, we chose to use the following variables: race (RACE), birthplace (BPL), state (FIPS code) (STATEFIP), total family income (FTOTINC), and a consumer price index multiplier (CPI99). Variables such as race and birthplace are self-explanatory. State (FIPS code) records the Federal Information Processing Standards (FIPS) code for the state in which the respondent's household was located. FIPS codes are identifying codes assigned alphabetically to states and other US territories. Total family income is defined as the "total pre-tax money income earned by one's family from all sources for the previous year" (IPUMS USA, n.d.). The consumer price index multiplier converts all of the dollar values across each year to 1999 dollars, therefore allowing for the direct comparison of dollar amounts from different years.

Since our research attempts to determine a relationship between median family income and migration rates among Black and White people, the race variable was essential to our research. As stated in our Literature Review section, we focused on Black people and White people.

We used the birthplace and state FIPS code variables to obtain a general idea of migration patterns in each decade. Our original plan was to attempt to work with both return and nonreturn migration to the South. Return migration simply means that an individual was born in the South, moved to the North at some point in their life, then eventually migrated back to the South. Nonreturn migration describes individuals who were not born in and have not lived in the South, but who migrated to the region.

Due to the limitations of our resources, we only focused on nonreturn migrants in our two areas of interest (the North and the South). We therefore collected data only from individuals who were born (birthplace) in the North and ended up living in the South (state FIPS code). It may also be important to note here that due to prior familiarity and potential family ties in the South, return migrants may have had more than purely economic reasons to move to the South, whereas nonreturn migrants would not have shared these motivations. Therefore, even without any limits in our resources, it would most likely have been appropriate to analyze the motivations and migration rates for each group separately regardless.

The North we defined as Connecticut, Delaware, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont; the South included Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, North Carolina, Oklahoma, Tennessee, Texas, and Virginia.

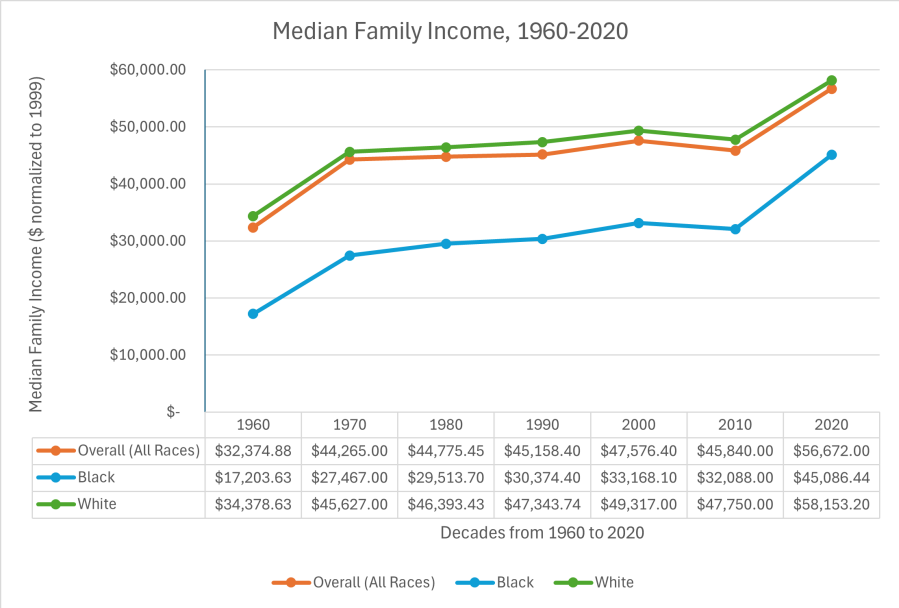
Since our goal was to determine how family income affected migration rates among Black and White people, we needed to examine the median family income over time within each group. To accomplish this, we found that we needed to combine the total family income and consumer price index multiplier variables by multiplying them together. This enabled us to look at adjusted income data without worrying about the effect of inflation.

These samples and variables describe the nature of the raw data we downloaded from IPUMS. Our data collection processes consisted of writing lines of code in Stata to obtain the information we needed (median family income and migration rates), and is detailed in Appendix A. We then used Excel to record, analyze, and visualize this data; details of this arrangement are noted in Appendix B.

4 Results

We hypothesized that a higher median family income would lead to a higher rate of North-to-South migration, and that this would be true among both Black and White people. The reasoning was that a higher income gives people a greater financial ability to migrate in search of even better economic opportunities. In the 1970s especially, the South was economically a much better place to work than the North due to a variety of reasons covered in our Literature Review section, and would therefore have been an economically attractive place for people to migrate to if they had had the means to do so. As a result, our data work focused on attempting to determine whether or not there was a relationship between the median family income of Black and White people in the North and their inclination to migrate to the South. Our expectation was to find that a higher median income in each group would correspond to a higher rate of migration for that group of people.

Through our data work (described in the Data section and detailed in Appendix A), we found the median income and North-to-South migration rates for Black people, White people, and all races. An overview of our results are shown in Figures 1 and 2.



	1960	1970	1980	1990	2000	2010	2020
Overall (All Races)	$32,374.88	$44,265.00	$44,775.45	$45,158.40	$47,576.40	$45,840.00	$56,672.00
Black	$17,203.63	$27,467.00	$29,513.70	$30,374.40	$33,168.10	$32,088.00	$45,086.44
White	$34,378.63	$45,627.00	$46,393.43	$47,343.74	$49,317.00	$47,750.00	$58,153.20

Figure 1: Median Family Income, 1960-2020

Although the conclusions that follow are not drawn directly from these two graphs, they provide some necessary context for the results of the data work.

It is important to note that the correlation coefficient between the median family income for all races and for Black people is 0.9893, and the correlation coefficient between the median family income for all races and for White people is 0.9993. This may be intuitive, because White people were the largest American racial population from 1960-2020, with Black people being the second largest racial population.

However, the correlation coefficient between the migration rates for all races and for Black people is 0.9389, and the correlation coefficient between the migration rates for all races and for White

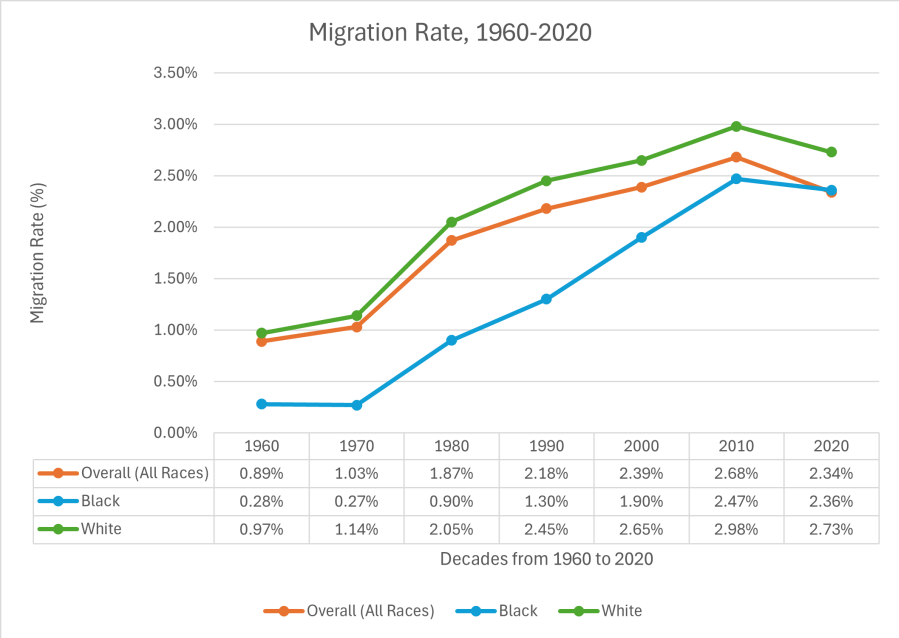


Figure 2: Migration Rates, 1960-2020

people is 0.9981. The latter may also be intuitive; the median family incomes for White people were higher than for all races, so they would have the most ability to migrate. However, the (slightly) lower correlation coefficient for all races and Black people suggests that people of other races had a higher migration rate, though this is not a part of our investigation.

One last thing to consider in the overall picture is that for each of the three populations, migration rates, which had been increasing from 1970-2010, decreased from 2010-2020. This could be reflective of the Great Recession in the late 2000s, which affected global family income, leading to serious financial losses for many families. Details like these validate our data and establish their ties to the real world.

These contextual conclusions, although not directly related to the hypotheses, set the stage for understanding the more specific results that follow. Figures 3 and 4 compare the median family incomes and migration rates for Black and White people.

In this group, the data is displayed in two different graphs: one for Black people and one for White people. The left axis shows the median family income, and the right axis shows the migration rate. The horizontal axis represents time and includes each of the sample years from 1960 to 2020.

Our results showed that, overall, there seems to be a relationship between median family income and migration rates; both variables seem to increase over time for each of the two populations. Upon further analysis, we also found the correlation coefficients between median family income and migration rates for each population. The correlation coefficient is 0.7870 for Black people and 0.7426 for White people. That the correlation coefficient between median family income and migration rates is higher for Black people is indicative that among the two populations, median family income is more important to Black people in giving them the ability to migrate to the South.

Through our data analyses, we were also able to corroborate Hunt et al.'s (2008) claim that one of

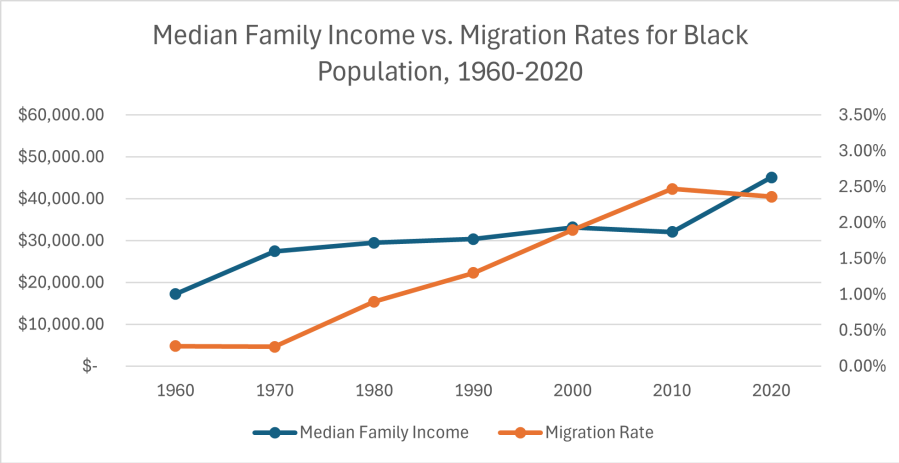


Figure 3: Median Family Income vs. Migration Rates for Black People, 1960-2020

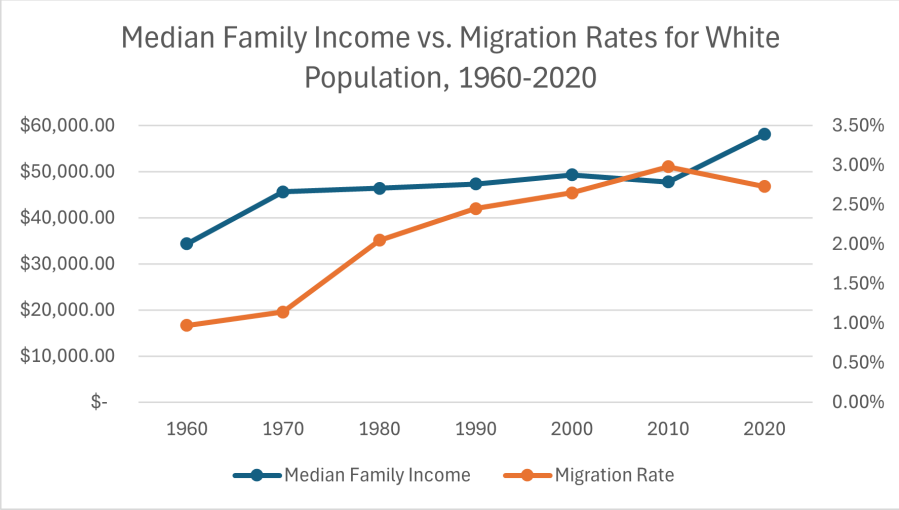


Figure 4: Median Family Income vs. Migration Rates for White People, 1960-2020

the reasons for increased North-to-South migration for Black people was the decrease in the income differential between Black and White people, as shown in Figure 5.

This final graph compares the median family income differential between Black and White people and the migration rates of Black people from 1960 to 2020. On the left axis is the median family income differential between Black and White people, and on the right axis is the migration rate of Black people. As in the other four graphs, the horizontal axis represents time.

Visually, the graph shows that the median family income differential decreases slightly from 1960 to 2020 and the migration rate increases slightly from 1960 to 2020. The correlation coefficient between the two variables is -0.8117, which suggests a very strong negative relationship. As the difference in median family incomes between Black and White people decreases, the migration rates for Black people increases. This supports Hunt et al.'s (2008) claim that a decrease in the income differential relates to an increase in migration.

There are a few implications from this data. The first is that family income is a more prohibitive

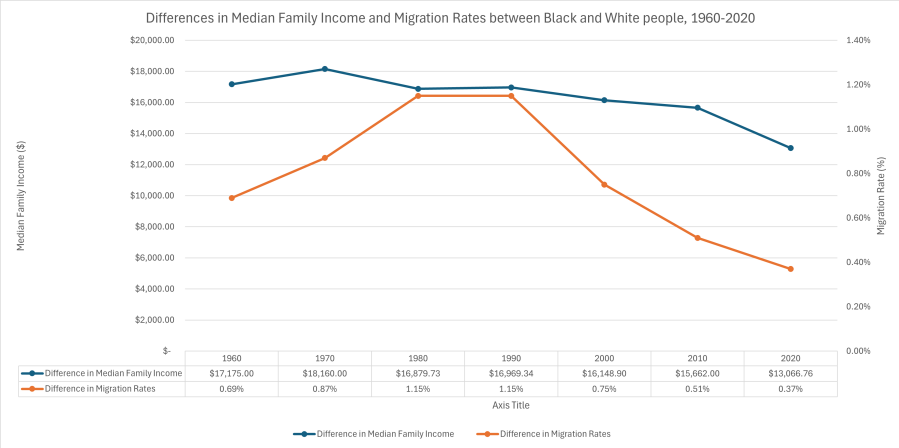


Figure 5: Median Family Income Differential between Black and White People vs. Black Migration Rates, 1960-2020

barrier in migration for Black people than for White people in the United States from 1960 to 2020. This is perhaps because Black people had a lower median family income to begin with than White people, and that many more Black families than White families were earning below some "benchmark" income required for a family to migrate successfully.

Although the data implies a correlation between median family income and migration rate, it does not prove any causal relationship between the two variables. It is therefore equally possible to draw the conclusion from our analyses that migration leads to median income, which would reverse the narrative of this project. If the underlying theory is that people migrate in search of a higher income, it could be true that migration helped them to find that higher income. It would not be prudent to make a definitive causal statement from our analyses alone.

Lastly, although there is a "strong" correlation between median family income and migration rates among each population, there is not a "very strong" (above +/-0.8) or "perfect" (+/-1) correlation. This implies the existence of other factors influencing people's decisions to migrate from the North to the South.

5 Discussion

Our data work is imperfect because there are a few aspects of our approach that leave something to be desired. Our limited resources prevent a more complete analysis—our conclusions should not be taken as absolute, comprehensive truth.

The central issue with our data is that it is anonymous, meaning that we are not able to track the information of specific people over time. This alters how our "migration rate" should be interpreted and prevents our analysis of return migration.

When we state a migration rate in a specific sample year, it does not directly refer to the percentage of people who moved from the North to the South in that decade. Instead, it should be interpreted as the number of people surveyed that year who had moved from the North to the South at some point in their lives. For example, the 2.18% overall migration rate in 1990 (see Figure 2) does not mean that 2.18% of the overall population migrated from the North to the South between 1980 and 1990. It

means, instead, that in 1990, 2.18% of the overall population had moved from the North to the South at some point in their lives. This is a small but significant distinction required to ensure that the data and its resulting conclusions are not misconstrued.

Our inability to track individuals over time using the IPUMS data also prevented our analysis of all migrants; we were, as mentioned in the Data section of this paper, only focused on nonreturn migration. The reason for this is that if a person was born in the South and lived in the South during each sample year, we had no way to know if they had lived in the North at some point in their life unless it was one or five years ago (there are variables in IPUMS that show a person's place of residence 1 and 5 years ago). For example, if a person had been born in the South in 1900, then migrated to the North in 1920, then migrated back to the South in 1950, none of this information would be apparent in the 1960 data. Our analysis, therefore, only included a subgroup of North-to-South migrants.

Another issue with the data we used is the inconsistency of sampling methods over time. For example, we used the 5% population sample for 1960 and the 1% metro sample for 1980. Not only are the percentages of the overall population used in each sample year different, but the 1980 sample also specifically focuses on people who lived in metropolitan areas. Across different years, therefore, data collection and sampling methods are varied. In order to have the most accurate portrayal of the data possible, we would have needed to work with the entire US population. Since that is unrealistic, consistent sampling and data collection methods would at least have been preferable to what we did have.

Further exploration of this research topic would first require obtaining a data set that fixes all of the above issues. Other steps include determining and examining other factors that affect North-to-South migration from 1960 to 2020, as well as examining the differences between the factors that influence return and nonreturn migrants. A comprehensive portrait of the causes and effects of North-to-South migration from 1960 to 2020 would include much more detailed analysis than we were able to accomplish.

6 Conclusion

Through our analysis of the median income and migration rates of Black and White people over time, we found that there is a strong relationship between the two variables for each population, and that the relationship between the two variables is strongest among Black people. Unfortunately, our data and tools for analysis were very limited, and it is unclear which of the two variables influences the other, if such a causal relationship exists at all. We believe our work contributes to the existing body of research on North-to-South migration from 1960 to 2020 as it corroborates a claim made by a previous body of research (Hunt et al., 2008) and raises questions for further exploration.

Bibliography

Cromartie, J., & Stack, C. B. (1989). Reinterpretation of Black return and nonreturn migration to the South 1975-1980. *Geographical Review,79*(3), 297-310. https://www.jstor.org/stable/215574

Frey, W. H. (2022, September 12). *A "New Great Migration" is bringing Black Americans back to the South.* Brookings. https://www.brookings.edu/articles/a-new-great-migration-is-bringing-black-americans-back-to-the-south/

Greenwood, M. J., & Hunt, G. L. (1984). Migration and interregional employment redistribution in the United States. *The American Economic Review, 74*(5), 957-969. https://www.jstor.com/stable/555

Hunt, L. L., Hunt, M. O., & Falk, W. W. (2008). Who is headed South? U.S. migration trends in Black and White, 1970-2000. *Social Forces, 87*(1), 119-195. https://api.semanticscholar.org/CorpusID:143467060

IPUMS USA. (n.d.). *FTOTINC*. IPUMS USA. https://usa.ipums.org/usa-action/variables/FTOTINC#description_section

Newman, R. J. (1983). Industry migration and growth in the South. *The Review of Economics and Statistics, 65*(1), 76-86. https://www.jstor.org/stable/1924411

Smith, J. P., & Welch, F. R. (1989). Black economic progress after Myrdal. *Journal of Economic Literature, 27*(2), 519-564. https://www.jstor.org/stable/2726688

Appendix A. Codes

An analysis of the results from our data work would not be complete without providing clarity on the exact processes by which the data was obtained. Below are the codes we added to the end of each Stata command file and with which we generated the data we needed. They are numbered for clarity in the explanations below; the numbered list is not part of the code.

1. * Family Income
2. gen realinc=ftotinc*cpi99
3. sort race
4. sum realinc, detail
5. by race: sum realinc, detail
6. * Migration Rate
7. gen north=0
8. replace north=1 if bpl==09
9. replace north=1 if bpl==10
10. replace north=1 if bpl==23
11. replace north=1 if bpl==24
12. replace north=1 if bpl==25
13. replace north=1 if bpl==33
14. replace north=1 if bpl==34
15. replace north=1 if bpl==36
16. replace north=1 if bpl==42
17. replace north=1 if bpl==44
18. replace north=1 if bpl==50
19. gen south=0
20. replace south=1 if statefip==01
21. replace south=1 if statefip==05
22. replace south=1 if statefip==12

23. replace south=1 if statefip==13

24. replace south=1 if statefip==22

25. replace south=1 if statefip==28

26. replace south=1 if statefip==37

27. replace south=1 if statefip==40

28. replace south=1 if statefip==47

29. replace south=1 if statefip==48

30. replace south=1 if statefip==51

31. gen migsouth=0

32. replace migsouth=1 if north==1 & south==1

33. tab migsouth

34. by race: tab migsouth

Lines 1-5 describe the codes used to generate median family income. Line 1 is simply a section heading to make the code more organized. Line 2 generates a new variable within Stata, real income (realinc), which combines family income and the consumer price index multiplier to allow for the comparison of income data from different decades. Line 3 sorts the data by race. Lines 4-5 summarize the raw real income data, listing the percentile amounts overall and by race. We recorded the 50th percentile amounts as the median family incomes.

Lines 6-34 describe the codes used to generate the migration rates, with Line 6 labeling the section. Line 7 creates a dummy variable *north*. By default, north=0. Lines 8-18 of the code change the value of *north* to 1 for each birthplace identified as a Northern state using its FIPS code. This identifies and marks those individuals who were born in the north.

Lines 19-30 are similar to Lines 7-18. Line 19 creates a dummy variable *south*. By default, south=0. Lines 20-30 of the code change the value of *south* to 1 for each state identified as a Southern state using its FIPS code. This identifies and marks those individuals who lived in the South at the time of the Censuses and surveys.

Lines 31-32 combine the previous two sections (Lines 7-18 and Lines 19-30) to identify non-

return migrants. Line 31 creates a dummy variable *migsouth* (migration to the south). By default, migsouth=0. Line 32 of the code changes the value of *migsouth* to 1 for each person who was born in the North (north=1) and now lives in the South (south=1). Therefore, the percentage of people who have a *migsouth* value of 1 is the percentage of nonreturn migrants to the South. Lines 33-34 run the tabulate function, creating frequency tables for the *migsouth* table.

Appendix B. Project Materials

A central record of all the materials and processes for this project exists in a GitHub repository located at the following link: https://github.com/ecn310/course-project-migration. The GitHub repository contains all the information and resources necessary for a person unfamiliar with this project to recreate or expand upon it.

The GitHub repository also contains links to a working folder, which contains the original Excel and Word documents used to create this revised version of the project. The link (https://sumailsyr-my.sharepoint.com/:f:/g/personal/qwu102_syr_edu/ElFttFZHsVlEp0PZXKwhwTgBf4TW_41Wh42sCX_AVBQrvg?e=3Szf2H) is accessible to anyone with a working Syracuse University Microsoft account.