ESSAYS IN LABOR AND EDUCATION ECONOMICS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Nolan Andrew Kopkin

May 2013

ESSAYS IN LABOR AND EDUCATION ECONOMICS

Nolan Andrew Kopkin, Ph. D.

Cornell University 2013

In chapter 1, I explore the relationship between discrimination towards blacks and the black-white self-employment rate gap, and provide the first direct empirical evidence that discrimination negatively impacts black self-employment. As a proxy for discrimination, I construct a measure of prejudicial attitudes using responses from the 1993-2010 General Social Survey. After compiling an index of prejudicial attitudes, I estimate the relationship between self-employment and this index of prejudicial attitudes using the 2005-2009 American Community Survey. I find that an amount of prejudice equal to the difference in the least and most prejudiced census divisions increases the black-white self-employment rate gap by 24.2%-34.5%.

In chapter 2, I examine how income tax rates affect the labor migration decisions of NBA free agents. By using a dataset of professional basketball players' free agent contracts from the National Basketball Association (NBA) between the 2001-2002 and 2007-2008 seasons, I am able to identify the effect that changes in income tax rates have on the labor migration decisions of NBA free agents. I find that an increase in the marginal income tax rate faced by NBA basketball players that play for a given team leads to a decrease in the average skill of the NBA free agents that migrate to that team.

In chapter 3, I study how a nutritional improvement in school provided meals affects student outcomes. There has recently been an emphasis on decreasing childhood obesity and increasing the health of schoolchildren in the United States. Improving the nutritional content of school meals is one potential mechanism for achieving these goals. In addition to direct health benefits, these interventions may provide positive effects on students' academic and behavioral outcomes. In 2007, the Buffalo Public School District implemented the *Healthier Options for Public Schoolchildren* (*HOPS*) program, which increased the nutritional content of food provided at school. I find that students in *HOPS* schools experienced a statistically significant increase in standardized math test scores, particularly among low ability, high income, and female students. Additionally, I find that the intervention had no impact on standardized English test scores, attendance, or suspensions.

BIOGRAPHICAL SKETCH

Nolan Andrew Kopkin was an economics Ph. D. student at Cornell University from August 2007 to May 2013. His research interests include labor economics, applied econometrics, and the economics of education. Additionally, Nolan has a strong interest in the economics of discrimination. In August 2013, he will join the faculty in the Africology Department at the University of Wisconsin – Milwaukee. As an undergraduate at the University of Florida, Nolan received a double-major in Economics and Mathematics, and minored in Statistics and Business Administration. Prior to graduate school, Nolan worked for a year at the Bureau of Labor Statistics in Washington, DC.

Dedicated to my mother, Marilyn. We miss you.

ACKNOWLEDGMENTS

I would like to especially thank the members of my graduate committee: Lawrence Kahn, Kevin Hallock, and Michael Lovenheim, who have continued to push me to improve and expand upon my research. I also wish to thank John Abowd, John Bishop, Diane Burton, Damon Clark, Ronald Ehrenberg, Freeha Fatima, George Jakubson, David Just, Matthew Freedman, Genevieve K. Landers, Robert Hutchens, Richard Mansfield, Mirinda Martin, Jordan Matsudaira, Emily Owens, Victoria Prowse, and other attendees of the Cornell University labor economics seminar and economics of education seminar for their many helpful comments. I would also like to thank NORC for the provision of the GSS Sensitive Data Files that are essential to Chapter 1, and the Buffalo Public School District and Bridget Wood for the restricted-use data and information about the *HOPS* program that are essential to Chapter 3.

TABLE OF CONTENTS

CHAPTER 1
DOES DISCRIMINATION AFFECT BLACK ENTREPRENEURSHIP?

1.1 Introduction

Over the last hundred years in the United States, the black self-employment rate has been roughly one-third to one-half of the white self-employment rate (Fairlie and Meyer, 2000). According to the 2010 American Community Survey, while 6.8% of whites own their own businesses, only 3.6% of blacks do.[1] In fact, whites own 83.4% of the businesses in America, compared to only 7.1% for blacks. However, recent survey findings show that there is a strong demand among blacks to start their own businesses. In a survey conducted by Gallup on high-school students' attitudes towards entrepreneurship, 75% of black high-school students stated that they wanted to start their own business in the future, compared to 68% of all other high-school students (Kourlisky and Esfandiari, 1997). Furthermore, a survey of working-age Americans found that blacks are 78 percent more likely to attempt to start new companies, even after controlling for demographic characteristics (Shane, 2008).

If blacks are so interested in starting their own businesses, then why are there so few black entrepreneurs in America? Some of the most common reasons that have been cited in the literature as to why blacks are less likely to start their own businesses are a lack of human capital, both in terms of the education and experience necessary to run a business, a lack of role models or family figures who own their own businesses, a lack of financial capital necessary to facilitate business formation or business success, or the lack of social networks to help one through the process of starting one's own business. For instance, Singh and Crump (2007) suggest that, since blacks with entrepreneurial intentions are more educated than those without entrepreneurial intensions, in order to improve the black self-employment rate, more focus toward improving the educational attainment of blacks is necessary. Fairlie (1999) finds that, while educational attainment explains a small amount of the black-white gap in business formation, racial differences in asset levels and the probability of having a self-employed father account for a much larger share of the gap. Furthermore, Fairlie and Robb (2007) observe that lack of work experience in a family business among blacks negatively affects the outcomes of

[1] These tabulations are based on 2010 American Community Survey 3-year estimates for male and female self-employed workers in own not incorporated business, as a percentage of the civilian employed population 16 years of age and over.

black-owned businesses. Others cite lack of financial capital or lack of access to financial markets as reasons for such large gaps in black-white business formation and business success.

Discrimination against blacks could cause or exacerbate many of these problems facing aspiring black entrepreneurs. Due to discrimination in wage employment, blacks are likely to have lower asset levels than whites (Charles and Guryan 2008). Blacks facing discrimination also may have larger challenges securing business loans or outside investments in their companies (Cavalluzzo, Cavalluzzo, and Wolken 2002; Blanchflower et al 2003; Cavalluzzo and Wolken 2005; Coleman 2008). They are less likely to get an education equal in quality to whites or be able to move into managerial positions to earn the experience necessary for entrepreneurial success (Card and Krueger 1992; Giuliano, Levine, and Leonard 2011). Furthermore, this discrimination may lead to intergenerational effects and lower-quality social networks (Lentz and Laband 1990; Dunn and Holtz-Eakin 2000; Hout and Rosen 2000; Fairlie and Robb 2007). However, no previous empirical studies have explored the relationship between discrimination towards blacks and the black-white self-employment rate gap. This paper explores that relationship, and provides the first direct empirical evidence that discrimination negatively impacts black self-employment.

There are a number of reasons why we should care whether discrimination is negatively affecting black self-employment. Blacks are underrepresented in business ownership despite the strong demand among blacks to start their own businesses, and self-employment may provide a route for economic advancement among blacks when compared to opportunities in wage-employment (Holtz-Eakin, Rosen, and Weathers 2000; Fairlie 2004a; Fairlie 2004b). Moreover, many of the black individuals that fail to become self-employed as a result of discrimination come from the upper half of the skill distribution, and might be successful business owners in the absence of discrimination. Furthermore, the government may be able to intervene to reverse the effect of discrimination if they can determine the mechanisms through which discrimination is most heavily impacting black self-employment.

In this paper, I explore the relationship between discrimination towards blacks and the black-white self-employment rate gap by examining the effect that prejudicial attitudes towards blacks have on the black-white gap in self-employment. While I cannot directly measure discrimination, as a proxy for discrimination, I construct a measure of prejudicial attitudes towards blacks using survey responses from the 1993-2010 General Social Survey Sensitive Data

Files. These are restricted-access data files maintained by the Nation Opinion Research Center (NORC) at the University of Chicago and are obtained under special contractual arrangements. In each year, the survey includes an independently drawn sample of roughly 2,000-3,000 English or Spanish speaking persons 18 years of age or over, living in non-institutional arrangements within the United States. The GSS contains demographic, behavioral, and attitudinal questions, plus topics of special interest, designed to take the "pulse of America." I use some of these questions to compile an index of the prejudicial attitudes of whites at the state level. Then, I merge my index of the prejudicial attitudes of whites with the 2005-2009 American Community Survey (ACS) Public Use Microdata Sample, which contains detailed information about survey respondents. To determine what impact the prejudicial attitudes of whites have on the black-white self-employment rate gap, I estimate regressions of self-employment status, from the ACS, on the state-level index of prejudicial attitudes, controlling for the characteristics of individuals and states that could be correlated with both racial prejudice and with the black-white self-employment rate gap.

I find that a one standard deviation increase in the average prejudice of whites increases the black-white self-employment rate gap by between 0.58-0.83 percentage points, and these results are statistically significant at the 5% level. An amount of prejudice equivalent to the difference in the least prejudiced and most prejudiced census divisions causes the black-white self-employment rate gap to widen by between 1.35-1.93 percentage points. These estimates translate into increases in the black-white self-employment rate gap of 24.2%-34.5%. Additionally, I find that the effect of an increase in average white prejudice on the black-white self-employment rate gap is due almost entirely to a decrease in the black self-employment rate, not an increase in the white self-employment rate. I show that these results are robust to various model and data assumptions.

Additionally, I test whether an increase in average white prejudice widens the black-white self-employment income gap. I find no evidence of a relationship between the average prejudice of whites at the state-level and the black-white self-employment income gap. Thus, while prejudice impacts whether blacks become self-employed, once they are self-employed it does not have an effect on their income.

I also investigate the mechanisms through which prejudice is impacting the black-white self-employment rate gap. Specifically, I consider the role that customer discrimination and financial constraints play in driving the effect of prejudice on the black-white self-employment rate gap. I find suggestive evidence that discrimination is causing blacks to have more difficulty financing their businesses, and that customer discrimination may be playing some smaller role, however these results are mostly inconclusive due to imprecision in the estimates.

Finally, I examine whether prejudice has a stronger effect on the black-white self-employment rate in certain parts of the ability distribution. I find suggestive evidence that the effect of prejudice on the black-white self-employment rate gap is most heavily concentrated in the 4^{th} quintile of the ability distribution, and least heavily concentrated in the 1^{st} quintile. Again, these results are mostly inconclusive because the estimates are imprecise.

In this paper, I explore the relationship between discrimination towards blacks and the black-white self-employment rate gap, and provide the first direct empirical evidence that discrimination negatively impacts black self-employment. However, I find no evidence of an effect on income among self-employed blacks. I find evidence that suggests that discrimination causes increased difficulty for black entrepreneurs in acquiring business financing, and that customer discrimination may be playing some smaller role. I also find suggestive evidence that discrimination has the largest effect on self-employment in the 4^{th} quintile of the ability distribution.

The next section of this paper discusses some background information and reviews the related literature. Section 1.3 describes the data. Sections 1.4 and 1.5 explain the methodology I employ and the results that follow, respectively. Section 1.6 concludes.

1.2 Discussion and Related Literature

My paper most closely relates to literature pertaining to the black-white self-employment rate gap, access to business capital, and lender discrimination. This paper adds to the existing literature by providing empirical support for the role that discrimination plays in explaining why blacks are both less likely to start businesses and less likely to have business success. While there is a sizable literature explaining this gap, no previous empirical studies have analyzed the relationship between discrimination towards blacks and the black-white self-employment rate gap. Some previous studies have used discrimination as a possible explanation for their findings, but none have been able to empirically indentify how discrimination towards blacks affects black entrepreneurship.

Much of the previous literature on the black-white self-employment gap focuses on the role of racial demographic differences. Racial differences in educational attainment (Singh and Crump 2007), asset levels (Fairlie 1999), the probability of having a self-employed father (Fairlie 1999; Hout and Rosen 2000), work experience in family businesses (Fairlie and Robb 2007), and self-employment patterns in previous generations (Lentz and Laband 1990; Dunn and Holtz-Eakin 2000) are all cited as reasons for the black-white gap in self-employment. However, each of these explanations for the black-white self-employment rate gap may be partially driven by discrimination.

Others cite access to business capital as a prominent reason for why there are so few black entrepreneurs. Bates (1997) finds that the startup capital of a new business is partially determined by the human capital of its owner, and since blacks have less human capital they also have less startup capital available for use in new businesses. Fairlie and Robb (2008) report that black-owned businesses have much lower levels of startup capital than white-owned businesses, and these differences persist across industries. Additionally, Robb, Fairlie, and Robinson (2009) show that blacks are much less likely to access external financial markets or to inject additional capital into their businesses in their companies' early years. Moreover, Cavalluzzo, Cavalluzzo, and Wolken (2002), Blanchflower et al (2003), Cavalluzzo and Wolken (2005), and Coleman (2008) all find a measureable black-white gap in loan denial rates. As with demographic differences, racial differences in human capital, startup capital, and access to external financial markets also may be driven in part by discrimination.

Despite the vast amount of research that has been done on the black-white self-employment gap, no previous empirical studies have analyzed the relationship between discrimination towards blacks and the black-white self-employment rate gap. This is a significant oversight in the literature, as discriminatory beliefs could be a primary driving force behind the factors that others have identified as being important in generating this gap. In this paper, I focus on directly investigating the relationship between the black-white self-employment rate gap and discrimination towards blacks, as measured by prejudicial attitudes, in order to explore the potential role of this underlying mechanism.

1.3 Data

The data that I use primarily come from the 1993-2010 General Social Survey (GSS)

Sensitive Data Files and the 2005-2009 American Community Survey (ACS) Public Use Microdata Sample.

1.3.1 General Social Survey Sensitive Data Files

I use the GSS Sensitive Data Files to measure the extent of racial prejudice across states. This data contains state-level geographic identifiers and responses to demographic, behavioral, and attitudinal questions for each individual, and is obtained under special contractual arrangements with NORC.[2] [3] In each year, the survey utilizes an independently drawn sample of roughly 2,000-3,000 English or Spanish speaking persons 18 years of age or over, living in non-institutional arrangements within the United States, using full-probability sampling.[4] Therefore, the GSS is a repeated cross-sectional dataset, and each year is representative of the U.S. population. While the GSS is not designed to be representative at the state level, Brace et al (2002) show that state samples obtained from the GSS are remarkably representative, and that representativeness improves by pooling additional years of the GSS because of the joint effects of adding more primary sampling units and more individual observations. Representativeness can be improved further by adding additional decades, which correspond to different sampling frames (Brace et al 2002). I therefore pool the 1993-2010 survey results to measure prejudicial attitudes across states.

The GSS asks many questions that are particularly focused on determining white individuals' attitudes towards blacks.[5] However, many of these questions focus on issues that are less directly related to racial prejudice, such as government assistance towards blacks or affirmative action. Including only those questions that one-dimensionally focus on attitudes towards blacks, I compile a racial prejudice index for each white individual in each state. Table 1.1 shows the years that each question was asked for those questions that are included in the racial prejudice index. Using the racial prejudice index of each individual, I then estimate the average prejudice of whites towards blacks at the state level, pooling across the 1993-2010

[2]Some of the data used in this analysis are derived from Sensitive Data Files of the GSS, obtained under special contractual arrangements designed to protect the anonymity of respondents. These data are not available from the author. Persons interested in obtaining GSS Sensitive Data Files should contact the GSS at GSS@NORC.org.
[3]The General Social Survey has been used in work by Cutler, Glaeser, and Vigdor (1999), Hout and Rosen (2000), Card, Mas, and Rothstein (2008), and Charles and Guryan (2008).
[4]Spanish speakers were added starting in 2006.
[5]For a complete list of questions that focus on whites attitudes towards blacks asked in the GSS from 1993-2010, see Appendix A.

surveys.

To create a racial prejudice index for each white individual in the GSS, I use the method introduced by Charles and Guryan (2008).[6] First, I make higher values correspond to more prejudiced answers. For instance, for the question that asks each respondent for an opinion to the statement "White people have a right to keep Blacks out of their neighborhoods if they want to, and Blacks should respect that right," I code a response of "Agree Strongly" as three points, a response of "Agree slightly" as two points, a response of "Disagree slightly" as one point, and a response of "Disagree strongly" as zero points. I then normalize the response to each question by the overall mean and standard deviation in 1993, the first year of the analysis.[7] I divide by the 1993 value of the standard deviation so that the relative weight of a particular question does not change over time. Formally, let d_{it}^{q} be respondent i's response to question q in year t. Then respondent i's normalized response to question q in year t, call it $\tilde{d}_{it}^{q}$, is given by

$$\tilde{d}_{it}^{q} = \frac{d_{it}^{q} - E[d_{i,93}^{q}]}{\sqrt{Var(d_{i,93}^{q})}}. \qquad (1)$$

Figure 1.1 shows how the normalized responses to each question have changed over time from 1993-2010. While prejudice has generally fallen over time, there is some variance in this decline across questions. For instance, while many more individuals state that they would vote for a qualified black candidate for president in 2010 than in 1993, the number of individuals that believe that blacks have less motivation, willpower, and in-born ability to learn has declined at a much slower rate.

Once I have normalized the response to each question for each individual in each year, I

[6] Where available, I use the same questions that Charles and Guryan (2008) use in creating my index, with a few notable exceptions. I use two questions that ask "On the average blacks have worse jobs, income, and housing than white people. Do you think these differences are because most blacks have less in-born ability to learn?" and "On the average blacks have worse jobs, income, and housing than white people. Do you think these differences are because most blacks just don't have the motivation or will power to pull themselves up out of poverty?", while Charles and Guryan (2008) exclude these questions from their analysis. Moreover, they include the question "Irish, Italians, Jewish and many other minorities overcame prejudice and worked their way up. Do you believe blacks should do the same without special favors?" I exclude this question, as it also focuses on views pertaining to the appropriate role of government and affirmative action. Another notable difference in the index created by Charles and Guryan (2008) is that they use GSS waves from 1972-2004 and use 1977 as the base year in their analysis, while I use GSS waves from 1993-2010 to more accurately assess a measure of current prejudice. In their index, they also use other questions which were no longer asked post-1991.

[7] Three of the questions were not asked in 1993. For these questions, I normalize by the mean and standard deviation from the first year in which the question was asked.

then calculate the average of the normalized responses for each individual in year *t*. Formally, let $D_{it}^{q} = \sum \tilde{d}_{it}^{q} / Q_t$, where Q_t is the total number of prejudice questions in year *t*. I then estimate a regression of D_{it}^{q} on individual year fixed effects to find the residual $\widetilde{D}_{it}^{q}$, which I will henceforth refer to as the individual-level prejudice index. I use this method to control for the fact that, while discrimination may be decreasing over time, the sample size from each state is changing over time as well, and I wish to separate the state-specific component of the individual-level prejudice index from the time-specific component.[8] Table 1.2 shows the pairwise correlations between each of the normalized individual-level prejudice questions, and the pairwise correlations between each of the normalized questions and the individual-level prejudice index. All of the prejudice questions are positively correlated, and most are highly correlated with other prejudice questions and with the individual-level prejudice index.

Using the individual-level prejudice index and the standardized responses to each question, I estimate a regression of each outcome on age, years of education, and an indicator variable for male, as well as state and year fixed effects, to determine how prejudice varies across the population; by conventional wisdom, older people, the less educated, and men tend to be more prejudiced. Regressions describing the individual-level prejudice index and the standardized responses to each question are shown in Table 1.3. On average, the individual-level prejudice index increases with age, decreases with education, and is larger for males. I cannot reject that the same is true of the standardized responses to each question. These results are consistent with Charles and Guryan (2008). Table 1.2 and Table 1.3 provide confidence that the individual-level prejudice index is an accurate measure of prejudicial attitudes.

Once I have obtained the individual-level prejudice index, $\widetilde{D}_{it}^{q}$, for each individual, I then find the average of $\widetilde{D}_{it}^{q}$ in each state to create an index of the average prejudice of whites at the state level. To do this, I weight by the number of adults in the household of each individual, and I correct the weights for survey non-response where possible.[9] This is done because the full-probability GSS samples are designed to give each household an equal probability of inclusion in the sample, and each adult in the household then has an equal probability of inclusion in the survey. Left unweighted, this would serve to under-represent individuals from large households

[8] As a robustness check, I also include a specification that does not disentangle the state-specific and time-specific components of each individual-level prejudice index. This does not qualitatively change my estimate of the effect of average white prejudice on the black-white self-employment rate gap.

[9] The GSS allows for an area non-response adjustment to the weights from 2004 to the present.

and over-represent individuals from small ones.[10]

Figure 1.2 shows the distribution of prejudice across census divisions. The East South Central division (AL, KY, MI, TN) is the most prejudiced, followed by the West South Central (AR, LA, OK, TX) and South Atlantic divisions (DE, DC, FL, GA, MD, NC, SC, VA, WV). The New England (CT, ME, MA, NH, RI, VT), Mountain (AZ, CO, ID, NM, MN, UT, NV, WY), and Pacific divisions (AK, CA, HI, OR, WA) are the least prejudiced.[11] These tabulations follow closely with conventional wisdom. Figure 1.3 shows how prejudice has changed over time in each census division. In the majority of census divisions, there is not much movement in the prejudice index over time. It is important to note that fluctuations in the prejudice index over time might occur not only because prejudice is changing, but also because the mix of questions asked and the individuals sampled in each year also are changing. Stability of the prejudice index within each census division over time further supports the use of the average state-level prejudice index over this time span.

It is technically accurate that prejudicial attitudes do not necessarily imply discriminatory behavior. However, since Charles and Guryan (2008) find evidence that an increase in the prejudicial attitudes of whites increases the black-white wage gap, which implies that whites are acting upon these attitudes, I will henceforth ignore the distinction.

1.3.2 American Community Survey Public Use Microdata Sample

The ACS is a cross-sectional dataset that gathers detailed information from survey respondents annually. I make use of information about each respondent's location, sex, race, age, educational attainment, self-employment status, self-employment income, industry, occupation, and other demographic characteristics. While the GSS contains similar information about each respondent, I use the ACS to study self-employment outcomes due to the scope of the survey; the GSS samples 2,000-3,000 persons each year, approximately half of whom are males, but I am able to observe over 3 million males in a five year period with the ACS.

Using the ACS, I define someone as self-employed if he reports being self-employed with a job at the time of the survey, and not self-employed if he reports being wage-employed or

[10] As a robustness check, I include a specification where I create the state-level prejudice index using unweighted GSS data. The use of weights has no substantive impact on the estimate of the effect of average white prejudice on the black-white self-employment rate gap.

[11] Based on the contractual arrangements with NORC, the census division is the finest level of geography at which I am permitted to display statistics from the GSS Sensitive Data Files.

unemployed at the time of the survey. In my basic definition of self-employment, I exclude those individuals who report being out of the labor force. However, as a robustness check, I use an alternative definition of self-employment that defines these individuals as not being self-employed. This does not qualitatively change my results.

Table 1.4 shows summary statistics for the full sample of males age 18-64 and for subsamples of males age 18-64 separated by race and self-employment status.[12] In this study I focus on the self-employment decisions of males since female labor supply decisions are typically both more complex than and dissimilar to male labor supply decisions. Furthermore, it has been found in the literature that females are generally much less interested in self-employment than males (Pryor and Reedy, 2009). Approximately 11.4% of the individuals in the sample are self-employed at the time of the survey. Those that are self-employed have more years of experience, are more likely to have a college degree, are more likely to be U.S. citizens, are more likely to be immigrants, are more likely to be married, are less likely to be black, and are less likely to be military veterans. Among the self-employed, blacks have fewer years of experience, are less likely to have a college degree, are less likely to be U.S. citizens, are more likely to be immigrants, are less likely to be married, are more likely to be military veterans, and are less likely to speak English poorly than whites.[13]

1.4 Methodology

In this paper, I seek to identify the effect that racial prejudice has on the self-employment rate gap between blacks and whites using a time-invariant measure of racial prejudice that varies at the state level. Ordinarily, to identify this effect, one would estimate a regression of whether an individual is self-employed on an indicator variable that is equal to 1 if an individual is black, the average prejudice of whites at the state level, and an interaction term between the two, controlling for the characteristics of individuals and states that could be correlated with both racial prejudice and with the black-white self-employment rate gap. Because of the complex sampling design of the American Community Survey, the appropriate standard errors should use a repeated replication method utilizing the replicate weights contained therein (U.S. Census

[12] Experience is defined as max(age-years of education-6, 0).

[13] Hispanic blacks and Hispanic whites are included in counts for blacks and whites, respectively. As a robustness check, I include a specification which excludes all individuals of Hispanic origin. This does not qualitatively change my estimate of the effect of average white prejudice on the black-white self-employment rate gap.

Bureau, 2009).[14] Since I have merged each individual observation in the ACS with state-level data, the standard errors should be clustered at the state level as well. This would indicate that either using repeated replication with the ACS replicate weights alone or clustering the standard errors at the state level alone will not retrieve the appropriate standard errors, and there is no method that allows one to do both simultaneously. To circumvent this problem, I use a two-step feasible generalized least squares (FGLS) estimation method proposed by Hanushek (1974) and used by Borjas (1982) and Lewis and Linzer (2005).

In the first step, I estimate a regression of whether an individual is self-employed on each of his individual-level characteristics separately for each state:

$$P(y_{is} = 1) = f(\varphi_0 + \varphi_{1s} Black_{is} + \varphi_{2s} X_{is} + u_{is}), \tag{2}$$

where y_{is} is equal to 1 if individual i in state s is self-employed, $Black_{is}$ is an individual-specific indicator variable equal to one if individual i in state s is black and equal to zero if individual i in state s is white, X_{is} is a vector of individual-specific characteristics, and u_{is} is an error term. For each individual, I control for the set of demographic characteristics shown in Table 1.4, as well as a quadratic in potential labor market experience. When f is a linear function, φ_{1s} is the state-specific black-white self-employment rate gap controlling for individual-level characteristics. Since each of these regressions include only individual-level characteristics, standard errors can be correctly estimated by repeated replication making use of the ACS replicate weights.

To identify the effect that racial prejudice has on the black-white self-employment rate gap, in the second step I estimate a regression of the state-specific black-white self-employment rate gap on the state-level prejudice index. Specifically, I estimate

$$\widehat{\varphi_{1s}} = \delta_0 + \delta_1 Prej_s + \delta_2 Z_s + v_s , \tag{3}$$

[14] Standard errors estimate the variation in a statistic across multiple samples of a given population, and the true standard error of any characteristic calculated from a single sample can never be known with certainty. Replicate weights allow a single sample to simulate multiple samples and generate more informed standard error estimates that retain all of the information about the complex sampling design of the ACS (U.S. Census Bureau, 2009). There are 80 separate replicate weights in the ACS that allow researchers to derive standard error estimates empirically (U.S. Census Bureau, 2009). The standard error of an estimate using replicate weights in the ACS can be calculated using the formula $SE(\beta) = \sqrt{4/80 \sum_{r=1}^{80} (\beta_r - \beta)^2}$, where β is the estimate using the full-sample weight and β_r is the estimate from the analysis using the rth set of replicate weights (U.S. Census Bureau, 2009).

where $\widehat{\varphi_{1s}}$ is the predicted state-specific black-white self-employment rate gap controlling for individual-level characteristics that is taken from the estimation of equation (2), $Prej_s$ is an index of the average prejudice in state s, Z_s is a vector of state-level characteristics, and v_s is the state-level error term. For each state, I control for the natural log of the fraction of the population that is black, the natural log of the population, the average amount of contact in wage employment for males, the fraction of the black male population and the white male population that has a high school diploma, and the fraction of the black male population and the white male population that had a self-employed father present as a child.[15] The natural log of the fraction of the population that is black, the natural log of the population, the fraction of the black male population that has a high school diploma, and the fraction of the white male population that has a high school diploma are calculated by summing over the relevant subpopulations in the ACS. To determine the fraction of the black male population and the white male population that had a self-employed father present as a child, I use a GSS question that asks "Was your father normally self-employed or did he work for someone else while you were growing up?" and then sum over the relevant subpopulations.

To determine the average amount of customer contact in wage employment for males in each state, I use the Occupational Information Network (O*NET). O*NET is administered by the Employment and Training Administration of the U.S. Department of Labor (USDOL/ETA), and is the nation's primary source for occupational information. O*NET indexes abilities, interests, knowledge, skills, activities, context, and values specific to each occupation in their database. I use an index of the importance of 'Working directly with the Public' as a measure of the amount of customer contact in each occupation. This index ranges from an importance of 98 for public address announcers and sheriffs to an importance of 0 for mathematical technicians. I match the amount of customer contact in each occupation to each male worker in the ACS by the Standard Occupational Classification (SOC) given in both datasets and then estimate the average amount of customer contact in wage employment for males in each state.

Identification of the coefficient δ_1 in equation (3) is based on the covariance between the

[15]The race-specific fraction of the male population that graduated from high school in each state may serve as a measure of the race-specific quality of schooling in that state. It also may help to represent the race-specific quality of potential social networks in that state. The race-specific fraction of the male population that had a self-employed father present as a child may serve as a proxy variable for the probability that an individual in each state had experience in a family business as a child. It may also control for past race-specific self-employment rates in the case that past self-employment affects the current level of the average prejudice of whites in each state.

black-white self-employment rate gap controlling for individual-level characteristics and the index of the average prejudice of whites, controlling for all state-specific characteristics. δ_1 can be interpreted as the increase in the black-white self-employment rate gap brought about by a one standard deviation increase in average white prejudice, controlling for all individual and state-specific characteristics.

Since the dependent variable in equation (3) is estimated, v_s can be decomposed into sampling error, τ_s, and standard regression error, ε_s. Weighting equation (3) by the inverse standard deviation of the error term, $1/\sqrt{(\sigma^2 + \omega_s^2)}$, where σ^2=Var(ε_s) and ω_s^2=Var(τ_s), will produce the most efficient parameter estimates. Since σ^2 is not observed, it must be estimated by feasible generalized least squares; ω_s^2 is approximated using the estimate of the sampling variance of $\widehat{\varphi_{1s}}$ from equation (2). Using FGLS, σ^2 is approximated by estimating equation (3) by unweighted OLS. From the unweighted OLS regression, I retrieve the residuals, $\hat{v}_s$. Then, since equation (2) is estimated using a linear probability model, σ^2 can be estimated by

$$\hat{\sigma}^2 = \frac{\sum_s \hat{v}_s^2 - \sum_s \widehat{\omega}_s^2 + tr\left((X'X)^{-1} X'GX\right)}{S-k}, \tag{4}$$

where $\hat{v}_s$ is the residual for state *s* from the second-step unweighted OLS regression, $\widehat{\omega}_s^2$ is the estimated sampling variance of $\hat{\varphi}_{1s}$, X is the matrix of state-level independent variables, G is the diagonal matrix with ω_s^2 as the diagonal elements, *S* is the number of states, and *k* is the number of state-level independent variables including the constant. Equation (3) is then re-estimated, weighted by $1/\sqrt{(\hat{\sigma}^2 + \widehat{\omega}_s^2)}$.

One detail that is often overlooked in related studies is that, since the state-level prejudice index is an estimated regressor from a relatively small sample, as are the other regressors that come from the GSS, standard errors that do not account for this variability will likely be biased downward.[16] To correct for the imprecision in the estimated regressors that come from the GSS, I bootstrap the standard errors, stratifying on GSS year and primary sampling unit.[17] [18]

[16] While regressors from the ACS are also estimated, the ACS sample is more than one hundred times larger than the GSS sample. Failure to account for the estimation of regressors from the ACS should have no substantial impact on the standard errors.
[17] Each bootstrapped standard error is calculated using 1000 replications.
[18] In some specifications, ordinary heteroskedasticity-robust standard errors may be biased downward by over 25%, highlighting the importance of accounting for imprecision in the regressors.

1.5 Results

I estimate the model using the two-step FGLS procedure discussed in Section 1.4. The distribution of the state-level black-white self-employment rate gap controlling for individual-level characteristics is shown in Figure 1.4. The estimates in Figure 1.4 can be directly compared to those in Figure 1.2. In terms of the magnitude of the black-white self-employment rate gap controlling for individual-level characteristics, the Pacific, which is the least prejudiced census division, has four out of its five states in the lowest twenty; on the other hand, the East South Central, which is the most prejudiced, has only one.

The results from the estimation of equation (3) are shown in Table 1.5. If a coefficient in equation (3) is negative, this refers to a widening of the black-white self-employment rate gap, and if a coefficient is positive, this indicates a narrowing of the gap. I add control variables across columns to illustrate the effect that the additional controls have on the estimate of the effect of average white prejudice. A one standard deviation increase in average white prejudice at the state level widens the black-white self-employment rate gap by between 0.58-0.83 percentage points, all else equal, depending on the specification, and these results are statistically significant at the 5% level in all specifications. An amount of prejudice equivalent to the difference in the least prejudiced and most prejudiced census divisions causes the black-white self-employment rate gap to widen by between 1.35-1.93 percentage points. Since the nationwide black-white self-employment rate gap controlling for observable characteristics is approximately 5.59 percentage points, this represents a 24.2%-34.5% increase. An increase in the average amount of customer contact in wage-employment also widens the black-white self-employment rate gap, and this result is statistically significant at the 5% level.[19] Sequentially adding additional control variables tends to have very little effect on the average white prejudice coefficient.[20]

---

[19] I lack power when estimating state-level regressions which include interaction terms between each individual-level characteristic and the black dummy variable. However, when including these interaction terms in the first step, point estimates of the effect of a one standard deviation increase in average white prejudice on the black-white self-employment rate gap are quantitatively similar to the results presented in Table 1.5.

[20] The results are robust to whether a linear probability model, probit regression, or logistic regression is estimated in the first step, and a linear probability model allows me to interpret the coefficients as average marginal effects; Appendix Table 1.B1, which displays the coefficients from the second step when the first step is a probit regression, shows that qualitatively similar conclusions can be drawn. Borjas and Sueyoshi (1994) and Huber, Kernell, and Leoni (2005) show that this two-step procedure is possible when the first step is a non-linear model. When the first-step is estimated by a probit regression, each second-step regression is weighted by $1/\sqrt{(\sigma^2 + \omega_s^2)}$, and σ^2 is estimated by $\hat{\sigma}^2 = \frac{\sum_s \hat{\epsilon}_s^2 - \sum_s \omega_s^2}{S-k}$, where $\hat{\epsilon}_s$is the residual for state s from a second-step unweighted OLS regression,

The results from Table 1.5 show that a one standard deviation increase in average white prejudice at the state level widens the black-white self-employment rate gap, however, they do not show whether or not the effect of a one standard deviation increase in average white prejudice at the state level lowers the self-employment rate of blacks. To determine the effect of a one standard deviation increase in average white prejudice on the black self-employment rate, I first estimate the state-level black self-employment rate that is not due to observable characteristics. Specifically, I estimate the linear probability model shown in equation (2) separately for each state, but with the black indicator variable omitted. Then, to find the state-level black self-employment rate that is not due to observable characteristics, I calculate the average residual from the first-step regression among blacks in each state. Finally, I use the average residual, which represents the state-level black self-employment rate that is not due to observable characteristics, as the dependent variable in the second-step regression.

The results from this second-step regression are shown in Table 1.6. A one standard deviation increase in average white prejudice at the state level decreases the black self-employment rate by between 0.62-0.75 percentage points, in specifications (2)-(6). These results are statistically significant at the 5% level in specifications (2)-(5) and at the 10% level in specification (6). Specifications (2)-(6) of Table 1.5 show that a one standard deviation increase in average white prejudice at the state level widens the black-white self-employment rate gap by between 0.63-0.83 percentage points. Therefore, the impact of an increase in average white prejudice on the black-white self-employment rate gap is due almost entirely to a decrease in the black self-employment rate, not an increase in the white self-employment rate. Sensibly, an increase in prejudice against blacks should cause black self-employment to decline while having little effect on white self-employment.

It is of some concern that unobservable individual-level characteristics may be able to explain a large portion of the black-white self-employment rate gap that I have attributed to discrimination. It also is of some concern that individuals will react endogenously to observed discrimination; for example, in the presence of discrimination, blacks might acquire less education because they no longer plan to start a business. To provide evidence on these potential sources of bias, I show that observable characteristics of individuals are uncorrelated with the

---

ω_s^2 is the estimated sampling variance of φ_{1s}, S is the number of states, and k is the number of state-level independent variables including the constant. The resultant coefficients in the second-step regression can be interpreted in the same way as coefficients from a probit regression.

average prejudice of whites in each state. If observable characteristics of individuals are uncorrelated with the average prejudice of whites, then unobservable characteristics of individuals can only be correlated with the average prejudice of whites if they are uncorrelated with the observable characteristics of individuals; furthermore, it would rule out that blacks are reacting endogenously to discrimination by changing their observable characteristics.

At the individual level, education, experience, citizenship, immigration, and marriage are positively correlated with self-employment, while military experience and poor English speech are negatively correlated with self-employment. To determine whether observable individual-level characteristics are correlated with the average prejudice of whites, I calculate the state-level black-white gap in each characteristic of interest, and then estimate a bivariate regression of each state-level gap on the average prejudice of whites. Of particular interest is whether observable individual-level characteristics other than race predict a self-employment rate gap that is correlated with the average prejudice of whites. These results are shown in Column (1) of Table 1.7. Together, all of these characteristics imply an increase in the black-white self-employment rate gap of 0.04 percentage points for every one standard deviation increase in average white prejudice, and this result is not statistically significant. Since the combination of observable characteristics that predict self-employment is uncorrelated with the average prejudice of whites, it is unlikely that unobservable characteristics are able to account for the black-white self-employment rate gap that is brought about by the average prejudice of whites, or that blacks are reacting endogenously to prejudice by altering their behavior.

An increase in average white prejudice widens the black-white self-employment rate gap. It is worth asking whether an increase in average white prejudice also widens the black-white self-employment income gap. Prejudice may directly impact the self-employment income of blacks through higher-interest business loans or discrimination from suppliers or consumers; however, since prejudice widens the black-white self-employment rate gap, it also may affect which blacks become or remain self-employed. If average white prejudice widens the black-white self-employment rate gap by making it more difficult for blacks to finance their businesses, then blacks that become or remain self-employed in more prejudiced states will likely be more highly skilled than blacks in less prejudiced states. This might cause estimates of the effect of average white prejudice on the black-white self-employment income gap to be biased toward narrowing the gap, since I cannot control for unobserved individual characteristics. To provide evidence that unobservable individual-level characteristics of self-employed individuals are uncorrelated with the average prejudice of whites, I show that observable characteristics of the self-employed are uncorrelated with the average prejudice of whites in each state.

To determine whether observable characteristics of self-employed individuals are correlated with prejudice levels, I calculate the state-level black-white gap in each characteristic of interest among the self-employed, and then estimate a bivariate regression of each state-level gap on the average prejudice of whites. These results are shown in Column (2) of Table 1.7. Based on observable characteristics, there is no statistically significant evidence that self-employed blacks are more highly skilled in more prejudiced states than less prejudiced states when compared to the self-employed whites in those states. Thus, it is unlikely that self-employed blacks are more highly skilled in more prejudiced states than less prejudiced states when compared to the self-employed whites in those states, based on unobservable characteristics.

To estimate the effect of an increase in average white prejudice on the black-white self-employment income gap, I again use the two-step FGLS procedure discussed in Section 1.4. I use the natural log of self-employment income as the dependent variable in the first-step regressions.[21] The results from the second-step regressions are shown in Columns (1)-(3) of Table 1.8. There is no statistically significant relationship between the average prejudice of whites at the state-level and the black-white self-employment income gap.[22] Since some self-employed workers may pay themselves a wage or salary that should actually be reported as self-employment income, I also estimate the regressions using the natural log of self-employment income plus wage and salary income for all self-employed workers as the dependent variable in the first-step regressions. The results from these second-step regressions are shown in Columns (4)-(6) of Table 1.8 and are not qualitatively different from the results in Columns (1)-(3).

As previously mentioned, estimates of the effect of average white prejudice on the black-white self-employment income gap are biased toward narrowing the gap if the black individuals that become or remain self-employed in more prejudiced states are unobservably more highly

[21] In the regressions I exclude all self-employed workers with self-employment income less than or equal to zero. This includes only 1.12% of self-employed blacks and 1.21% of self-employed whites. The mean self-employment income of blacks with self-employment income below zero is similar to the mean self-employment income of whites with self-employment income below zero.

[22] Using the subset of self-employed workers who have greater self-employment income than wage and salary income, or who have no wage or salary income, does not qualitatively change the results.

skilled than blacks in less prejudiced states.[23] It is not clear if there is no effect of an increase in average white prejudice on the black-white self-employment income gap, or if I fail to find an effect because the black individuals that are observed as self-employed in more prejudiced states are more highly skilled than blacks in less prejudiced states in some unobservable way. However, based on Column (2) of Table 1.7, it is unlikely that the black individuals that are observed as self-employed in more prejudiced states are unobservably more highly skilled than blacks in less prejudiced states.

1.5.1 Role for Customer Discrimination and Business Financing

I have not yet addressed the potential reasons why an increase in average white prejudice widens the black-white self-employment rate gap. One possible explanation could be that a higher level of prejudice among whites widens the black-white self-employment rate gap because white customers discriminate against black-owned businesses. In this case, discriminatory whites only would purchase from black-owned businesses at lower prices than white-owned businesses would charge for identical products, in order to compensate themselves for the disutility that they experience when purchasing from blacks (Becker 1957). In order for customer discrimination to exist, there must be contact between customers and the group that is discriminated against.

I define an individual as being self-employed in a high contact occupation if they are self-employed in an occupation where the importance of 'Working directly with the Public' is above the median level, and I define an individual as not being self-employed in a high contact occupation if they are not self-employed, or if they are self-employed in an occupation where the importance of 'Working directly with the Public' is below the median level. I use an analogous definition for self-employment in a low contact occupation.[24]

To test whether the average prejudice of whites has a greater effect on the black-white self-employment rate gap in occupations with large amounts of customer contact than in occupations with small amounts of customer contact, I use the two-step FGLS procedure discussed in Section

[23] If the black individuals that become or remain self-employed in more prejudiced states are more highly skilled than blacks in less prejudiced states, but this skill difference is observable, then estimates of the effect of average white prejudice on the black-white self-employment income gap are not biased. They are only biased if this skill difference is unobservable.

[24] The median level of the index of the importance of 'Working directly with the Public' is 49.86. The mean of the index is 50.21.

1.4 separately for high contact and low contact occupations. Table 1.9 shows the results from the second-step regressions: a one standard deviation increase in average white prejudice at the state level is shown to widen the black-white self-employment rate gap in high contact occupations by between 0.36-0.48 percentage points, and these results are statistically significant at the 5% level. A one standard deviation increase in average white prejudice at the state level widens the black-white self-employment rate gap in low contact occupations by between 0.32-0.38 percentage points, and these results are only statistically significant (at the 10% level) when the race-specific graduation rates and the race-specific fractions with a self-employed father present as a child are not included. In comparable specifications, the estimate of the effect of prejudice on the black-white self-employment rate gap in low contact occupations is not statistically distinguishable from the effect of prejudice on the black-white self-employment rate gap in high contact occupations, and is between 76%-88% as large in magnitude. While there may be a small effect caused by customer discrimination, it is not a primary driver of the results.

Another potential explanation why an increase in the average prejudice of whites widens the black-white self-employment rate gap is that a more prejudiced environment makes it more difficult for blacks to finance their businesses. One implication of financial market discrimination is that prejudice would have a greater effect on the black-white self-employment rate gap in industries that require a large amount of startup capital. To this end, I estimate the average startup cost of a firm in each industry that corresponds to a two-digit NAICS code. To do this, I use the 2007 Survey of Business Owners, which reports the percent of firms in each industry that had startup costs of less than \$5,000, \$5,000 to \$9,999, \$10,000 to \$24,999, \$25,000 to \$49,999, \$50,000 to \$99,999, \$100,000 to \$249,999, \$250,000 to \$999,999, or \$1,000,000 or more. The survey also allows for an answer of don't know or not applicable. I estimate the average startup cost of a firm in each industry using the mid-point of each range and 1.5 times the top-code, under the assumptions that not applicable means that no startup capital was necessary and the average startup capital of those who did not know their business' startup cost is the same as those who did know their business' startup cost. Then, I divide the ACS sample of firms into two groups by classifying each industry as a high startup cost industry or low startup cost industry. Detailed information about the estimation of average startup costs in each industry can be found in Appendix Table 1.B2.

The race-specific percent of the labor force that is self-employed in each industry is shown in Figure 1.5. The industries that I have identified as low startup cost industries (educational services, services, administrative services, professional services, and construction) contain 53.9% of all self-employed workers, and self-employed workers in each of these industries make up a relatively large fraction of the labor force for both blacks and whites, with educational services being the exception. The largest black-white gaps in self-employment are in the construction and professional services industries. While the number of self-employed workers in high startup cost industries is relatively lower, there tends to be a significant gap in each of these industries, except for transportation. Among high startup cost industries, agriculture has the largest black-white gap in self-employment.[25]

To test whether the average prejudice of whites has a greater effect on the black-white self-employment rate gap in industries with high startup costs than in industries with low startup costs, I use the two-step FGLS procedure discussed in Section 1.4 separately by industry startup cost; Table 1.10 shows the results from the second-step regressions. A one standard deviation increase in average white prejudice at the state level is shown to widen the black-white self-employment rate gap in high startup cost industries by 0.40 percentage points in the full specification shown in Column (3). This result is statistically significant at the 10% level. A one standard deviation increase in average white prejudice at the state level increases the black-white self-employment rate gap in low startup cost industries by only 0.26 percentage points in the full specification, and this result is not statistically significant. Since the magnitude of the effect is much smaller in low startup cost industries, this may be evidence that average white prejudice has a much smaller effect on the black-white self-employment rate gap in industries with low startup costs. This assertion can be strengthened by the fact that 54% of self-employed workers are in industries that I classify as high startup cost industries and 46% are in industries that I classify as low startup cost industries, because the sample could not be split evenly based on two-digit NAICS industries; due to this feature of the data, the coefficients for high startup cost industries are biased downward, and the coefficients for low startup cost industries are biased upward. Prejudice could have a smaller effect in low startup cost industries either because blacks in more prejudiced states have more difficulty acquiring external financing from lenders or investors or because blacks in more prejudiced states have lower wage earnings to use toward

[25] In section 1.5.3, I do a robustness check of my initial results excluding agriculture, and find no substantive difference in my results.

financing their own businesses, as found by Charles and Guryan (2008). Nevertheless, in each specification, the estimate of the effect of prejudice on the black-white self-employment rate gap in low startup cost industries is not statistically distinguishable from the effect in high startup cost industries due to the relatively small number of states in my sample.

I find evidence that suggests that discrimination causes increased difficulty for black entrepreneurs in acquiring business financing, and that customer discrimination may be playing some smaller role. However, the results are mostly inconclusive due to the size of the standard errors.

1.5.2 Heterogeneous Effects Across the Distribution of Ability

The effect of prejudice on the black-white self-employment rate gap is likely concentrated among those black workers that are most likely to try to become self-employed. Individuals with bachelor's or graduate degrees, those that are married, and immigrants are much more likely to enter into self-employment, and are thus susceptible to discrimination that may keep them from becoming self-employed. Those that would never attempt to enter self-employment in the absence of discrimination cannot be prevented from becoming self-employed by discrimination.

To determine how the effect of prejudice on the black-white self-employment rate gap varies across the distribution of ability, I first create an index of ability. To generate this index, I estimate a regression of the natural log of wage income on all of the individual-level demographic characteristics present in X_{is} in equation (2), separately for each state, using only the wage-employed. Then, in each state, I calculate the predicted value from this regression for each worker in the labor market; this is an estimate of the log wage for each individual given his observable characteristics.[26] I use this index to rank workers based on ability.

In each state, I divide each worker into an ability quintile based on his predicted log wage. While only 3.97% in the lowest ability quintile are self-employed, 16.4% are self-employed in the highest ability quintile, and the self-employment rate increases monotonically across quintiles. Thus, the effect of prejudice on the black-white self-employment rate gap is probably most heavily concentrated in the upper tail of the ability distribution.

[26] Of course, the coefficients from this regression may be biased because I only observe wages for the wage-employed; self-employed and unemployed individuals are omitted from the regression. Nonetheless, the predicted value from this regression will give a fairly accurate representation of each person's relative position in the ability distribution.

I use the two-step FGLS procedure discussed in Section 1.4 separately by quintile; Table 1.11 shows the results from the second-step regressions. The magnitude of the effect of average white prejudice on the black-white self-employment rate gap is largest in the 4th quintile of the ability distribution, and is statistically significant at the 5% level. However, while the effects in the 2nd, 3rd, and 5th quintiles are approximately one-third to one-half the size of the effect in the 4th quintile and not statistically significant, and the effect in the lowest quintile is less than one-tenth the size of the effect in the 4th quintile and not statistically significant, none of the differences are statistically significant either due to imprecision in the estimates. However, evidence suggests that the effect of prejudice on the black-white self-employment rate gap is most concentrated in the 4th quintile of the ability distribution, and least concentrated in the 1st quintile.

Theoretically, it should not be surprising if prejudice most heavily affects self-employment in the 4th quintile of the ability distribution. While the prevalence of self-employment is highest in the 5th quintile, individuals in the 5th quintile are those that financial institutions or employers would have the most difficult time discriminating against due to their credentials. A loan officer would likely have a hard time explaining why she didn't give a loan to a black client with an MBA and 25 years of labor market experience, but gave one to a white client with 10 fewer years of experience and only a bachelor's degree.

While I find evidence that suggests that the effect of prejudice on the black-white self-employment rate gap is most heavily concentrated in the 4th quintile of the ability distribution, and least heavily concentrated in the 1st quintile, the results are mostly inconclusive due to the size of the standard errors.

1.5.3 Robustness Checks

I estimate a series of robustness checks to determine if the effect that racial prejudice has on the black-white self-employment rate gap is robust to various model and data assumptions. The results from these regressions are shown in Tables 1.12a and 1.12b. Column (1) of Table 1.12a shows the specification from Column (6) of Table 1.5 for the sake of comparison. To determine if the result is a byproduct of the weighting scheme that I employ, in Column (2), I estimate the second-step regression by unweighted ordinary least squares, and in Column (3), I estimate the second-step regression by weighting each state by the inverse of the standard error of the dependent variable that comes from the first-step regression. In both cases the estimate of the effect of average white prejudice on the black-white self-employment rate gap is very similar to the estimate that uses FGLS weights. In Column (4), I use an alternative definition of self-employment that includes all individuals not in the labor force as not self-employed. This might affect the initial estimates if some individuals drop out of the labor force when they fail to become self-employed, and if it occurs more frequently for blacks in low prejudice states or for whites in high prejudice states. This attenuates the estimate; however, the attenuation is a product of the alternative definition of self-employment; under the alternative definition 18% fewer individuals are defined as self-employed, and the self-employment rate gap is 9.3% smaller.

In Column (5), I calculate the average prejudice of whites at the state level without reweighting each year to receive equal weight. In Column (6), I calculate the average prejudice of whites at the state level by not time-detrending prejudice at the individual level. The adjustments in Columns (5) and (6) have little effect on the sign or magnitude of the estimates. In Column (7), I use a measure of average white prejudice that is calculated without weighting each respondent by the number of individuals in the household or correcting for survey non-response. In Column (8), I exclude all individuals of Hispanic origin, regardless of whether they are black or white. In Column (9), I exclude all immigrants. Each of the adjustments in Columns (7)–(9) has little effect on the sign or magnitude of the estimates.

In Column (10), I present a specification that excludes agriculture from the first-step estimate of the black-white self-employment rate gap. While this attenuates the estimate, 4.3% fewer workers are defined as self-employed when agriculture is excluded; the black-white self-employment rate gap that excludes agriculture is also 9.3% smaller. In Column (11), I time-detrend the fraction of the black male population and the fraction of the white male population that had a self-employed father present as a child to control for the fact that, while the fraction that had a self-employed father present as a child may be changing over time, the sample size from each state is changing over time as well. This also has very little effect on the sign or magnitude of the estimate. In Column (12), I include a measure of the fraction of blacks that are segregated from whites to control for the proximity of blacks and whites. While the inclusion of the fraction of blacks that are segregated from whites increases the magnitude of the coefficient, it likely reacts endogenously to the average prejudice of whites.

In Columns (13) - (18), I check to see if the estimates are robust to particular functional form assumptions in the model. In Column (13), I allow the population and the fraction of the population that is black to enter the second-step regression linearly. In Column (14), I omit the natural log of the fraction of the population that is black. This would be appropriate if the fraction of the population that is black were endogenously determined. In Column (15), I include a quadratic function of both the fraction of the black male population and the fraction of the white male population with a high school diploma, in the case that increases in the fraction with a high school diploma may at some point signal a declining quality of schooling or a lower quality social network if it is particularly easy to obtain a high school diploma. In Column (16), I include a quadratic function of both the fraction of the black male population and the fraction of the white male population with a college degree, in the case that the fraction with a college degree is a better measure of the quality of schooling or quality of social network than the fraction with a high school diploma. In Column (17), in order to allow for a flexible functional form for the fraction of the population that is black, I control for a fourth-degree polynomial in the fraction of the population that is black. In Column (18), in order to allow for a flexible functional form for the fraction of the black male population and the fraction of the white male population that had a self-employed father present as a child, I control for a quartic in the fraction of the black male population and the fraction of the white male population that had a self-employed father present as a child. Each of these functional form assumptions has little effect on the sign or magnitude of the estimates.[27] In Column (19), I test whether average white prejudice, as measured by the GSS prejudice questions that specifically ask about feelings towards blacks, has an effect on the Hispanic-white self-employment rate gap. The measure of average white prejudice towards blacks should have a much smaller impact on the self-employment rate of Hispanics, although it may still have a negative impact since discrimination against blacks and discrimination against Hispanics may be correlated. I estimate this falsification test using only white Hispanics to avoid allowing black Hispanics to bias the results, since prejudice clearly impacts black self-employment. The estimate is positive and statistically

[27] I also run eleven additional robustness checks where I leave one question at a time out of the prejudice index, using the specification from Table 1.5, Column (6). The coefficients on Average White Prejudice range from -0.00496 to -0.00757, and all are statistically significant at the 10% level, except for the one where I leave out the question which asks "On the average blacks have worse jobs, income, and housing than white people. Do you think these differences are because most blacks just don't have the motivation or will power to pull themselves up out of poverty?"

insignificant, which lends additional credibility to my previous results.

1.6 Conclusion

In this paper, I examine the link between prejudicial attitudes towards blacks and the black-white gap in self-employment. I find that a one standard deviation increase in the average prejudice of whites increases the black-white self-employment rate gap by between 0.58-0.83 percentage points. An amount of prejudice equivalent to the difference in the least prejudiced and most prejudiced census divisions causes the black-white self-employment rate gap to widen by between 1.35-1.93 percentage points. These estimates translate into increases in the black-white self-employment rate gap of 24.2%-34.5%. Additionally, I find that the effect of an increase in average white prejudice on the black-white self-employment rate gap is due almost entirely to a decrease in the black self-employment rate, not an increase in the white self-employment rate. It makes practical sense that an increase in prejudice against blacks would cause black self-employment to decline while having little effect on white self-employment. These results are robust to various model and data assumptions.

There are several plausible explanations for these results. Discrimination may exacerbate many of the problems facing potential black entrepreneurs. Due to discrimination in wage employment, blacks are likely to have lower asset levels than whites. Blacks facing discrimination also may have larger challenges securing business loans or outside investments in their companies. They are less likely to get an education equal in quality to whites or be able to move into managerial positions to earn the experience necessary for entrepreneurial success. Furthermore, this discrimination may lead to intergenerational effects and lower-quality social networks. Blacks in prejudiced states may face higher prices from distributors. More prejudiced states may structure their laws and governing practices towards the advantage of whites, possibly without openly discriminating against blacks. Furthermore, perceived discrimination may cause blacks to believe they would be unsuccessful in self-employment, or unsuccessful at obtaining business loans if they were to apply.

This paper presents the first direct empirical evidence that discrimination towards blacks negatively impacts black self-employment. Yet, I find no evidence of an effect of prejudice on black self-employment income. My findings are consistent with prejudicial attitudes towards blacks leading to increased difficulty for black entrepreneurs in acquiring business financing, with a potential smaller role for customer discrimination. Future research might more closely examine the mechanisms through which discrimination is operating, verify whether discrimination against other minority groups, such as Hispanics, negatively impacts self-employment amongst members of those groups, determine what impact racial prejudice against blacks plays in the market for mortgages or home-loans for blacks, or determine if perceived prejudice among blacks follows closely with the prejudicial attitudes of whites that are observed in this study.

Table 1.1: General Social Survey Prejudice Questions 1993-2010, Years Asked

Year	1993	1994	1996	1998	2000	2002	2004	2006	2008	2010
In general, how close do you feel to blacks?			✓	✓	✓	✓	✓	✓	✓	✓
In general, how warm or cool do you feel towards blacks?						✓				
If you and your friends belonged to a social club that would not let blacks join, would you try to change the rules so that blacks could join?	✓	✓								
Blacks have worse jobs, income, and housing than white people. Do you think these differences are because most blacks have less in-born ability to learn?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Do you think most blacks just don't have the motivation or will power to pull themselves up out of poverty?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Do you think there should be laws against marriages between blacks and whites?	✓	✓	✓	✓	✓	✓				
Do you agree that a homeowner should be able to decide for himself whom to sell his house to, even if he prefers not to sell to blacks?	✓	✓	✓				✓	✓	✓	✓
If your party nominated a black for President, would you vote for him if he were qualified for the job?	✓	✓	✓						✓	✓
Do you agree that blacks shouldn't push themselves where they're not wanted?		✓	✓	✓	✓	✓				
Would you yourself have any objection to sending your children to a school where a few/half/most of the children are black?	✓	✓	✓							
Do you agree that white people have a right to keep blacks out of their neighborhoods if they want to, and blacks should respect that right?	✓	✓	✓							

Table 1.2: Pairwise Correlations of Normalized Individual-Level Prejudice Questions

Question	PREJ INDEX	CLOSEBLK	FEELBLKS	RACDIF2	RACDIF4	RACMAR	RACOPEN	RACPRES	RACPUSH	RACSCHL	RACSEG	RACCHNG
PREJ INDEX	1											
CLOSEBLK	0.6879	1										
FEELBLKS	0.8683	0.5322	1									
RACDIF2	0.6591	0.117	0.1482	1								
RACDIF4	0.6797	0.1545	0.2381	0.247	1							
RACMAR	0.643	0.2066	0.1916	0.2862	0.2212	1						
RACOPEN	0.6611	0.1709	.	0.1532	0.1886	0.2113	1					
RACPRES	0.5997	0.2206	.	0.247	0.1976	0.3456	0.2312	1				
RACPUSH	0.7148	0.2612	0.3791	0.2559	0.3634	0.3158	0.2597	0.2781	1			
RACSCHL	0.618	.	.	0.2218	0.2692	0.2529	0.2262	0.2837	0.3216	1		
RACSEG	0.7686	.	.	0.3376	0.3023	0.4123	0.3374	0.3741	0.505	0.3852	1	
RACCHNG	0.6905	.	.	0.2673	0.344	0.3127	0.3537	0.37	.	0.3583	0.4481	1

aCLOSEBLK- In general, how close do you feel to blacks?, FEELBLKS- In general, how warm or cool do you feel towards blacks?, RACCHNG- If you and your friends belonged to a social club that would not let blacks join, would you try to change the rules so that blacks could join?, RACDIF2- On the average blacks have worse jobs, income, and housing than white people. Do you think these differences are because most blacks have less in-born ability to learn?, RACDIF4- On the average blacks have worse jobs, income, and housing than white people. Do you think these differences are because most blacks just don't have the motivation or will power to pull themselves up out of poverty?, RACMAR-Do you think there should be laws against marriages between blacks and whites?, RACOPEN-Do you agree that a homeowner should be able to decide for himself whom to sell his house to, even if he prefers not to sell to blacks?, RACPRES- If your party nominated a black for President, would you vote for him if he were qualified for the job?, RACPUSH- Do you agree that blacks shouldn't push themselves where they're not wanted?, RACSCHL- Would you yourself have any objection to sending your children to a school where a few/half/most of the children are black?, RACSEG- Do you agree that white people have a right to keep blacks out of their neighborhoods if they want to, and blacks should respect that right?

Table 1.3: Determination of Prejudice at the Individual Level

Variables	(1) PREJ INDEX	(2) RACDIF2	(3) RACDIF4	(4) CLOSEBLK	(5) RACMAR	(6) RACOPEN
Age	0.0074***	0.0086***	0.0061***	0.0052***	0.0099***	0.0074***
	(0.0004)	(0.0006)	(0.0006)	(0.0008)	(0.0008)	(0.0009)
Education	-0.0496***	-0.0588***	-0.0713***	-0.0222***	-0.0658***	-0.0168***
	(0.0025)	(0.0050)	(0.0034)	(0.0061)	(0.0062)	(0.0049)
Male	0.102***	0.0434**	0.0771***	0.0764***	-0.0013	0.185***
	(0.0108)	(0.0169)	(0.0213)	(0.0194)	(0.0212)	(0.0286)
Observations	19020	11581	11196	8557	7497	7439
R-squared	0.115	0.072	0.086	0.044	0.158	0.063

Variables	(7) RACPUSH	(8) RACPRES	(9) RACSEG	(10) RACSCHL	(11) FEELBLKS	(12) RACCHNG
Age	0.0134***	0.0027***	0.0135***	0.0053***	0.0061***	0.0086***
	(0.0006)	(0.0008)	(0.0010)	(0.0009)	(0.0010)	(0.0016)
Education	-0.0877***	-0.0437***	-0.0706***	-0.0223***	-0.0442***	-0.0423***
	(0.0042)	(0.0047)	(0.0057)	(0.0061)	(0.0072)	(0.0083)
Male	0.182***	0.0803***	0.0657**	0.0276	0.262***	0.109
	(0.0208)	(0.0213)	(0.0289)	(0.0437)	(0.0479)	(0.0736)
Observations	6061	5224	3221	3117	2151	1212
R-squared	0.188	0.063	0.163	0.054	0.079	0.134

Robust standard errors clustered on state in parenthesis.

*** p<0.01, ** p<0.05, * p<0.1

aAll GSS prejudice questions are standardized by the mean and standard deviation of the first year that each question is asked. Each regression controls for state and year fixed effects.

bCLOSEBLK- In general, how close do you feel to blacks?, FEELBLKS- In general, how warm or cool do you feel towards blacks?, RACCHNG- If you and your friends belonged to a social club that would not let blacks join, would you try to change the rules so that blacks could join?, RACDIF2- On the average blacks have worse jobs, income, and housing than white people. Do you think these differences are because most blacks have less in-born ability to learn?, RACDIF4- On the average blacks have worse jobs, income, and housing than white people. Do you think these differences are because most blacks just don't have the motivation or will power to pull themselves up out of poverty?, RACMAR-Do you think there should be laws against marriages between blacks and whites?, RACOPEN-Do you agree that a homeowner should be able to decide for himself whom to sell his house to, even if he prefers not to sell to blacks?, RACPRES- If your party nominated a black for President, would you vote for him if he were qualified for the job?, RACPUSH- Do you agree that blacks shouldn't push themselves where they're not wanted?, RACSCHL- Would you yourself have any objection to sending your children to a school where a few/half/most of the children are black?, RACSEG- Do you agree that white people have a right to keep blacks out of their neighborhoods if they want to, and blacks should respect that right?

Table 1.4: 2005-2009 American Community Survey, Males Age 18-64

Variable	Full Sample (1)	Black Self-Employed (2)	White Self-Employed (3)	Black Not Self-Employed (4)	White Not Self-Employed (5)
Self-Employed	0.1143				
	(0.3182)				
Black	0.1196				
	(0.3245)				
High School Graduate	0.3000	0.3288	0.2770	0.3686	0.2930
	(0.4583)	(0.4698)	(0.4475)	(0.4824)	(0.4551)
Some College	0.3047	0.3082	0.2786	0.3278	0.3050
	(0.4603)	(0.4617)	(0.4483)	(0.4694)	(0.4604)
Bachelor's Degree	0.1873	0.1516	0.2068	0.1193	0.1948
	(0.3902)	(0.3586)	(0.4050)	(0.3241)	(0.3961)
Graduate Degree	0.0980	0.0884	0.1403	0.0486	0.0995
	(0.2974)	(0.2838)	(0.3473)	(0.2150)	(0.2993)
Experience	20.40	24.14	26.01	18.56	19.85
	(12.34)	(10.74)	(10.57)	(12.23)	(12.38)
Citizen	0.9306	0.9081	0.9427	0.9327	0.9288
	(0.2542)	(0.2889)	(0.2324)	(0.2506)	(0.2572)
Immigrant	0.1086	0.2013	0.1099	0.1253	0.1051
	(0.3111)	(0.4010)	(0.3127)	(0.3311)	(0.3067)
Ever Married	0.6914	0.7305	0.8476	0.5507	0.6899
	(0.4619)	(0.4437)	(0.3594)	(0.4974)	(0.4625)
Military Veteran	0.1650	0.1574	0.1497	0.1789	0.1651
	(0.3712)	(0.3642)	(0.3568)	(0.3833)	(0.3713)
Speaks English Poorly	0.0364	0.0147	0.0271	0.0124	0.0414
	(0.1874)	(0.1204)	(0.1623)	(0.1106)	(0.1993)
Self-Employment Income ($2009)	4333.83	25596.00	34260.85	225.71	578.59
	(24784.22)	(45336.67)	(63455.88)	(3652.678)	(8271.34)
Wage & Salary Income ($2009)	48350.38	20059.46	39386.88	32932.10	52096.26
	(58635.68)	(49518.78)	(80361.73)	(33643.99)	(57407.91)
Observations	3106728	16310	374218	263455	2452745
Population Weight	64896575	425154	6996388	7334233	50140800

[a]Nominal self-employment losses bounded below at $10,000 nominal. Self-employment income is topcoded at the median level above the upperbound.
[b]The full sample includes all individuals in the labor force.

Table 1.5: The Effect of Average White Prejudice on the Black-White Self-Employment Gap

State Self-Employment Gap	(1)	(2)	(3)	(4)	(5)	(6)
Average White Prejudice	-0.0058**	-0.0065***	-0.0063***	-0.0083***	-0.0075***	-0.0065**
	(0.0023)	(0.0024)	(0.0024)	(0.0025)	(0.0028)	(0.0033)
Log Fraction Black		0.0022	0.0020	0.0029	0.0040	0.0040
		(0.0018)	(0.0019)	(0.0021)	(0.0027)	(0.0033)
Log Population			0.0010	0.0016	0.0016	0.0017
			(0.0021)	(0.0024)	(0.0024)	(0.0026)
Average Contact in Wage Employment				-0.0027**	-0.0028***	-0.0033***
				(0.0011)	(0.0010)	(0.0012)
Black Graduation Rate					0.0368	0.0463
					(0.0533)	(0.0652)
White Graduation Rate					0.0207	-0.0016
					(0.0642)	(0.0716)
Fraction of Blacks w/Self-Employed Fathers						0.0061
						(0.0220)
Fraction of Whites w/Self-Employed Fathers						-0.0624*
						(0.0376)
Constant	-0.0441***	-0.0374***	-0.0539	0.0763	0.0325	0.0789
	(0.0055)	(0.0078)	(0.0338)	(0.0537)	(0.0891)	(0.106)
States	46	46	46	46	46	43
R-squared	0.193	0.219	0.222	0.353	0.335	0.428

Bootstrap standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

[a]The table reports the coefficients from a regression of the black-white self-employment gap on the aggregate prejudice index of all whites in the population. The dependent variable in the regression is the coefficient on the dummy variable for black, which comes from state-level linear probability models that control for education categories, a quadratic in experience, citizenship, immigrant status, whether an individual was ever married, whether an individual is a military veteran, whether an individual speaks English poorly, and year fixed effects. The weights for each observation are explained in the text.

[b]Some states are omitted because the General Social Survey did not sample these states in the years for which the prejudice index is calculated, or because the states were not sampled in the years that specific questions were asked.

[c]The first-step regressions use standard errors calculated with successive difference replication using the 80 replicate weights that are included in the American Community Survey.

Table 1.6: The Effect of Average White Prejudice on the Black Self-Employment Rate

Black Self-Employment Rate	(1)	(2)	(3)	(4)	(5)	(6)
Average White Prejudice	-0.0026	-0.0062***	-0.0064***	-0.0075***	-0.0067**	-0.0059*
	(0.0028)	(0.0022)	(0.0022)	(0.0024)	(0.0027)	(0.0032)
Log Fraction Black		0.0096***	0.0098***	0.0104***	0.0116***	0.0094***
		(0.0015)	(0.0016)	(0.0017)	(0.0023)	(0.0028)
Log Population			-0.0012	-0.0009	-0.0003	-0.0009
			(0.0020)	(0.0023)	(0.0021)	(0.0024)
Average Contact in Wage Employment				-0.0014	-0.0017*	-0.0018*
				(0.0010)	(0.0009)	(0.0011)
Black Graduation Rate					0.0412	0.0129
					(0.0465)	(0.0644)
White Graduation Rate					0.0294	-0.0012
					(0.0580)	(0.0645)
Fraction of Blacks w/Self-Employed Fathers						0.0053
						(0.0214)
Fraction of Whites w/Self-Employed Fathers						-0.0600*
						(0.0345)
Constant	-0.0470***	-0.0162**	0.0033	0.0729	0.0195	0.0916
	(0.0065)	(0.0069)	(0.0328)	(0.0494)	(0.0790)	(0.0942)
States	46	46	46	46	46	43
R-squared	0.119	0.153	0.154	0.321	0.322	0.383

Bootstrap standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

[a]The table reports the coefficients from a regression of the black self-employment rate on the aggregate prejudice index of all whites in the population. The dependent variable in the regression is the state-level weighted average of the residuals for blacks, which come from state-level linear probability models that control for education categories, a quadratic in experience, citizenship, immigrant status, whether an individual was ever married, whether an individual is a military veteran, whether an individual speaks English poorly, and year fixed effects. The weights for each observation in this regression are explained in the text.

[b]Some states are omitted because the General Social Survey did not sample these states in the years for which the prejudice index is calculated, or because the states were not sampled in the years that specific questions were asked.

[c]The first-step regressions use standard errors calculated with successive difference replication using the 80 replicate weights that are included in the American Community Survey.

Table 1.7: How Individual-Level Characteristics are Correlated with Average White Prejudice

	Full Sample	Self-Employed
	(1)	(2)
Predicted Self-Employment	-0.0004	0.0110
	(0.0010)	(0.0085)
High School Graduate	-0.0106	-0.102
	(0.0087)	(0.0738)
College Graduate	0.0121	0.0081
	(0.0148)	(0.0239)
Experience	-0.0776	2.639
	(0.317)	(1.778)
Citizen	0.0037	0.0019
	(0.0090)	(0.0056)
Immigrant	-0.0101	-0.0142
	(0.0129)	(0.0172)
Ever Married	-0.0249***	-0.0319
	(0.0085)	(0.0818)
Military Veteran	-0.0131*	0.0168
	(0.0070)	(0.0206)
Speaks English Poorly	0.0069	0.0024
	(0.0047)	(0.0016)

Bootstrap standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
[a]Column (1) and Column (2) show the coefficients from state-level bivariate regressions of the black-white gap in each characteristic on average white prejudice. Each regression is weighted by the inverse of the standard error of the estimate of the dependent variable.
[b]Predicted self-employment comes from a linear probability model that controls for education categories, a quadratic in experience, citizenship, immigrant status, whether an individual was ever married, whether an individual is a military veteran, whether an individual speaks English poorly, and year fixed effects.

Table 1.8: The Effect of Average White Prejudice on the Black-White Self-Employment Income Gap

State Self-Employment Income Gap	Self-Employment Income for Self-Employed Workers			Self-Employment + Wage/Salary Income for Self-Employed Workers		
	(1)	(2)	(3)	(4)	(5)	(6)
Average White Prejudice	0.0138	0.133	-0.0611	0.0450	0.0634	-0.0698
	(0.227)	(0.233)	(0.103)	(0.160)	(0.182)	(0.108)
Log Fraction Black	0.229	0.470	0.222	0.0945	0.140	0.241
	(0.238)	(0.337)	(0.142)	(0.147)	(0.136)	(0.151)
Log Population	0.193	0.207	0.122	0.0877	0.0939	0.127
	(0.170)	(0.191)	(0.119)	(0.139)	(0.149)	(0.141)
Average Contact in Wage Employment	-0.0018	-0.0365	-0.0114	0.0254	0.0189	-0.0116
	(0.0947)	(0.100)	(0.0494)	(0.0749)	(0.0753)	(0.0551)
Black Graduation Rate		9.306	1.372		1.581	1.235
		(7.537)	(2.530)		(2.647)	(2.439)
White Graduation Rate		0.375	0.832		-0.0666	0.439
		(4.090)	(2.400)		(2.370)	(2.505)
Fraction of Blacks w/Self-Employed Fathers			0.996			0.914
			(0.633)			(0.686)
Fraction of Whites w/Self-Employed Fathers			0.498			0.487
			(1.365)			(1.451)
Constant	-2.859	-8.958	-3.167	-2.953	-3.893	-2.801
	(5.871)	(8.373)	(4.352)	(5.558)	(6.586)	(4.893)
States	45	45	42	45	45	42
R-squared	0.152	0.281	0.323	0.098	0.107	0.260

Bootstrap standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

[a]The table reports the coefficients from a regression of the black-white self-employment income gap on the aggregate prejudice index of all whites in the population. The dependent variable in the regression is the coefficient on the dummy variable for black, which comes from state-level regressions that control for education categories, a quadratic in experience, citizenship, immigrant status, whether an individual was ever married, whether an individual is a military veteran, whether an individual speaks English poorly, and year fixed effects. The weights for each observation are explained in the text.

[b]Some states are omitted because the General Social Survey did not sample these states in the years for which the prejudice index is calculated, or because the states were not sampled in the years that specific questions were asked.

[c]One state is omitted because no blacks are observed as being self-employed in that state.

[d]The first-step regressions use standard errors calculated with successive difference replication using the 80 replicate weights which are included in the American Community Survey.

Table 1.9: The Effect of Average White Prejudice on the Black-White Self-Employment Gap in High and Low Contact Occupations

State Self-Employment Gap	High Contact Occupations			Low Contact Occupations		
	(1)	(2)	(3)	(4)	(5)	(6)
Average White Prejudice	-0.0048***	-0.0043***	-0.0036**	-0.0037*	-0.0038	-0.0032
	(0.0012)	(0.0015)	(0.0014)	(0.0020)	(0.0023)	(0.0026)
Log Fraction Black	0.0022	0.0028*	0.0029*	0.0028	0.0029	0.0022
	(0.0014)	(0.0015)	(0.0017)	(0.0019)	(0.0024)	(0.0026)
Log Population	0.0012	0.0016	0.0017	-5.00e-05	-0.0002	0.0003
	(0.0011)	(0.0011)	(0.0012)	(0.0021)	(0.0022)	(0.0022)
Average Contact in Wage Employment	-0.0024***	-0.0026***	-0.0024***	5.63e-05	6.60e-05	-0.0007
	(0.0006)	(0.0005)	(0.0005)	(0.0010)	(0.0011)	(0.0010)
Black Graduation Rate		0.0216	0.0449		0.0042	-0.0003
		(0.0304)	(0.0342)		(0.0429)	(0.0527)
White Graduation Rate		0.0224	0.0328		-0.0164	-0.0384
		(0.0303)	(0.0331)		(0.0593)	(0.0599)
Fraction of Blacks w/Self-Employed Fathers			0.0008			0.0065
			(0.0100)			(0.0164)
Fraction of Whites w/Self-Employed Fathers			0.0099			-0.0700**
			(0.0161)			(0.0307)
Constant	0.0970***	0.0612	0.0214	-0.0258	-0.0118	0.0519
	(0.0303)	(0.0516)	(0.0532)	(0.0430)	(0.0804)	(0.0868)
States	46	46	43	46	46	43
R-squared	0.476	0.491	0.506	0.145	0.145	0.398

Bootstrap standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

[a]The table reports the coefficients from a regression of the black-white self-employment gap in high contact occupations on the aggregate prejudice index of all whites in the population in Columns (1)-(3), and the coefficients from a regression of the black-white self-employment gap in low contact occupations on the aggregate prejudice index of all whites in the population in Columns (4)-(6). The dependent variable in the regression is the coefficient on the indicator variable for black that comes from state-level linear probability models that control for education categories, a quadratic in experience, citizenship, immigrant status, whether an individual was ever married, whether an individual is a military veteran, whether an individual speaks English poorly, and year fixed effects. The weights for each observation are explained in the text.

[b]Some states are omitted because the General Social Survey did not sample these states in the years for which the prejudice index is calculated, or because the states were not sampled in the years that specific questions were asked.

[c]The first-step regressions use standard errors calculated with successive difference replication using the 80 replicate weights that are included in the American Community Survey.

Table 1.10: The Effect of Avg. White Prejudice on the Black-White Self-Employment Gap in High and Low Startup Cost Industries

State Self-Employment Gap	High Startup Cost Industries			Low Startup Cost Industries		
	(1)	(2)	(3)	(4)	(5)	(6)
Average White Prejudice	-0.0046***	-0.0041**	-0.0040*	-0.0033*	-0.0034	-0.0026
	(0.0016)	(0.0018)	(0.0022)	(0.0019)	(0.0021)	(0.0027)
Log Fraction Black	0.0012	0.0016	0.0025	0.0025*	0.0028*	0.0018
	(0.0017)	(0.0022)	(0.0029)	(0.0013)	(0.0017)	(0.0020)
Log Population	0.0016	0.0020	0.0027	0.0006	0.0003	-1.57e-05
	(0.0014)	(0.0015)	(0.0017)	(0.0015)	(0.0016)	(0.0019)
Average Contact in Wage Employment	0.0003	0.0002	-0.0004	-0.0027***	-0.0027***	-0.0026***
	(0.0009)	(0.0010)	(0.0010)	(0.0007)	(0.0008)	(0.0009)
Black Graduation Rate		0.0151	0.0367		0.0040	0.0055
		(0.0474)	(0.0558)		(0.0343)	(0.0444)
White Graduation Rate		0.0241	0.0164		-0.0028	-0.0060
		(0.0409)	(0.0455)		(0.0435)	(0.0500)
Fraction of Blacks w/Self-Employed Fathers			0.0082			0.0004
			(0.0146)			(0.0167)
Fraction of Whites w/Self-Employed Fathers			-0.0384			-0.0150
			(0.0252)			(0.0279)
Constant	-0.0512	-0.0851	-0.0672	0.102***	0.110*	0.111
	(0.0527)	(0.0738)	(0.0843)	(0.0385)	(0.0646)	(0.0756)
States	46	46	43	46	46	43
R-squared	0.271	0.276	0.405	0.299	0.299	0.247

Bootstrap standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

aThe table reports the coefficients from a regression of the black-white self-employment gap in high startup cost industries on the aggregate prejudice index of all whites in the population in Columns (1)-(3), and the coefficients from a regression of the black-white self-employment gap in low startup cost industries on the aggregate prejudice index of all whites in the population in Columns (4)-(6). The dependent variable in the regression is the coefficient on the indicator variable for black that comes from state-level linear probability models that control for education categories, a quadratic in experience, citizenship, immigrant status, whether a respondent was ever married, whether a respondent is a military veteran, whether a respondent speaks English poorly, and year fixed effects. The weights for each observation are explained in the text.

bSome states are omitted because the General Social Survey did not sample these states in the years for which the prejudice index is calculated, or because the states were not sampled in the years that specific questions were asked.

cThe first-step regressions use standard errors calculated with successive difference replication using the 80 replicate weights that are included in the American Community Survey.

Table 1.11: The Effect of Average White Prejudice Across the Distribution of Ability

	1st Quintile	2nd Quintile	3rd Quintile	4th Quintile	5th Quintile
State Self-Employment Gap	(1)	(2)	(3)	(4)	(5)
Average White Prejudice	-0.0013	-0.0088	-0.0061	-0.0177**	-0.0090
	(0.0022)	(0.0054)	(0.0081)	(0.0077)	(0.0090)
Log Fraction Black	-0.0023	0.0121	0.0172**	0.0156	0.0154
	(0.0032)	(0.0088)	(0.0073)	(0.0107)	(0.0096)
Log Population	0.0014	0.0027	0.0081	0.0008	0.0149*
	(0.0020)	(0.0049)	(0.0069)	(0.0075)	(0.0082)
Average Contact in Wage Employment	-0.0004	-0.0038	-0.0044	-0.0033	-0.0067**
	(0.0009)	(0.0026)	(0.0033)	(0.0030)	(0.0034)
Black Graduation Rate	-7.79e-05	0.265	0.105	0.131	-0.0148
	(0.0638)	(0.166)	(0.172)	(0.207)	(0.213)
White Graduation Rate	0.0102	-0.121	-0.0773	-0.133	0.278
	(0.0556)	(0.131)	(0.200)	(0.179)	(0.266)
Fraction of Blacks w/Self-Employed Fathers	-0.0172	0.0052	0.0350	0.0460	0.0165
	(0.0141)	(0.0347)	(0.0491)	(0.0526)	(0.0466)
Fraction of Whites w/Self-Employed Fathers	-0.0178	-0.0005	-0.105	-0.119	-0.0288
	(0.0261)	(0.0732)	(0.0952)	(0.100)	(0.104)
Constant	-0.0291	0.0227	0.0638	0.187	-0.143
	(0.0907)	(0.217)	(0.324)	(0.318)	(0.363)
States	43	43	43	43	43
R-squared	0.279	0.251	0.377	0.363	0.382

Bootstrap standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

[a] The table reports the coefficients from a regression of the black-white self-employment gap on the aggregate prejudice index of all whites in the population. The dependent variable in the regression is the coefficient on the indicator variable for black, which comes from quintile-specific state-level linear probability models that control for education categories, a quadratic in experience, citizenship, immigrant status, whether an individual was ever married, whether an individual is a military veteran, whether an individual speaks English poorly, and year fixed effects. The weights for each observation are explained in the text.

[b] Some states are omitted because the General Social Survey did not sample these states in the years for which the prejudice index is calculated, or because the states were not sampled in the years that specific questions were asked.

[c] The first-step regressions use standard errors calculated with successive difference replication using the 80 replicate weights that are included in the American Community Survey.

Table 1.12a: Robustness Checks of The Effect of Average White Prejudice on the Black-White Self-Employment Gap

State Self-Employment Gap	Original Specification	Unweighted OLS	Weighted by Inverse of SE	Alternative Definition	Prej Not Year Reweighted	Prej Not Time Detrended	GSS Data Not Weighted	Hispanic Origin Excluded	Immigrants Excluded	No Agriculture	Fathers Time Detrended
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Average White Prejudice	-0.0065**	-0.0067*	-0.0067**	-0.0049*	-0.0077**	-0.0075**	-0.0062*	-0.0066*	-0.0070**	-0.0055*	-0.0066*
	(0.0033)	(0.0037)	(0.0033)	(0.0028)	(0.0037)	(0.0034)	(0.0033)	(0.0036)	(0.0035)	(0.0032)	(0.0034)
Log Fraction Black	0.0040	0.0041	0.0047	0.0035	0.0045	0.0047	0.0037	0.0073**	0.0042	0.0032	0.0039
	(0.0033)	(0.0038)	(0.0031)	(0.0029)	(0.0032)	(0.0032)	(0.0033)	(0.0035)	(0.0034)	(0.0031)	(0.0032)
Log Population	0.0017	0.0018	0.0027	0.0004	0.0014	0.0017	0.0017	0.0006	0.0015	0.0013	0.0016
	(0.0026)	(0.0029)	(0.0025)	(0.0021)	(0.0025)	(0.0026)	(0.0026)	(0.0027)	(0.0025)	(0.0025)	(0.0026)
Average Contact in Wage Employment	-0.0033***	-0.0030**	-0.0036***	-0.0022**	-0.0034***	-0.0034***	-0.0032***	-0.0040***	-0.0024**	-0.0041***	-0.0031***
	(0.0012)	(0.0014)	(0.0011)	(0.0010)	(0.0012)	(0.0011)	(0.0012)	(0.0013)	(0.0012)	(0.0011)	(0.0011)
Black Graduation Rate	0.0463	0.0395	0.0633	0.0274	0.0353	0.0447	0.0461	0.0735	0.0504	0.0398	0.0461
	(0.0652)	(0.0736)	(0.0619)	(0.0547)	(0.0650)	(0.0637)	(0.0641)	(0.0701)	(0.0695)	(0.0639)	(0.0637)
White Graduation Rate	-0.0016	-0.0111	0.0141	-0.0323	-0.0145	-0.0031	-0.0082	-0.0079	-0.0044	0.0261	0.0021
	(0.0716)	(0.0837)	(0.0652)	(0.0622)	(0.0704)	(0.0703)	(0.0744)	(0.0815)	(0.0755)	(0.0711)	(0.0703)
Fraction of Blacks w/Self-Employed Fathers	0.0061	0.0057	0.0074	0.0135	0.0096	0.0077	0.0054	0.0073	0.0092	0.0043	0.0065
	(0.0220)	(0.0244)	(0.0210)	(0.0189)	(0.0223)	(0.0219)	(0.0254)	(0.0248)	(0.0226)	(0.0226)	(0.0219)
Fraction of Whites w/Self-Employed Fathers	-0.0624*	-0.0624	-0.0489	-0.0564*	-0.0591	-0.0578	-0.0654*	-0.0507	-0.0637	-0.0415	-0.0558
	(0.0376)	(0.0418)	(0.0361)	(0.0319)	(0.0368)	(0.0375)	(0.0379)	(0.0413)	(0.0389)	(0.0377)	(0.0375)
States	43	43	43	43	43	43	43	43	43	43	43
R-squared	0.428	0.351	0.475	0.374	0.448	0.456	0.413	0.366	0.418	0.369	0.399

Bootstrap standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

[a]The table reports the coefficients from a regression of the black-white self-employment gap on the aggregate prejudice index of all whites in the population. The dependent variable in the regression is the coefficient on the indicator variable for black which comes from state-level linear probability models that control for education categories, a quadratic in experience, citizenship, immigrant status, whether an individual was ever married, whether an individual is a military veteran, whether an individual speaks English poorly, and year fixed effects.

[b]Some states are omitted because the General Social Survey did not sample these states in the years for which the prejudice index is calculated, or because the states were not sampled in the years that specific questions were asked.

[c]The first-step regressions use standard errors calculated with successive difference replication using the 80 replicate weights which are included in the American Community Survey.

[d]Column (1) shows the original FGLS specification found in Column (6) of Table 5. Column (2) shows the coefficients from an unweighted state-level OLS regression. Column (3) shows the coefficients from a state-level OLS regression that is weighted by the inverse of the standard error of the dependent variable from the first-step regression. Column (4) uses an alternative definition of self-employment in the first-step regression which considers those individuals which are not in the labor force as not being self-employed. Column (5) includes a measure of the average prejudice of whites that is calculated without reweighting each year to recieve equal weight. Column (6) includes a measure of the average prejudice of whites that is calculated by not time-detrending prejudice at the individual level. Column (7) includes a measure of the average prejudice of whites that is calculated using GSS data that does not weight each individual by the number of people in each household, nor correct for survey nonresponse. Column (8) excludes all individuals of Hispanic origin, regardless of whether they are black or white. Column (9) excludes all immigrants. Column (10) excludes agriculture from the first-step estimate of the black-white self-employment rate gap. Column (11) seperately time-detrends the self-employment of both black fathers and white fathers to account for changes in self-employment of black fathers and white fathers over time. Columns (4) - (11) are all weighted using the FGLS weights.

Table 1.12b: Robustness Checks of The Effect of Average White Prejudice on the Black-White Self-Employment Gap

State Self-Employment Gap	Segregation Included	Pop/ Fraction Black Linear	Fraction Black Omitted	Grad Rates Squared	College Degree Rates	Fraction Black Quartic	SE Fathers Quartic	Falsification Test
	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
Average White Prejudice	-0.0076**	-0.0065**	-0.0057*	-0.0065*	-0.0065*	-0.0079*	-0.0070*	0.0048
	(0.0035)	(0.0032)	(0.0034)	(0.0038)	(0.0036)	(0.0041)	(0.0039)	(0.0059)
Log Fraction Black	0.0042			0.0039	-9.40e-05		0.0079*	0.0041
	(0.0031)			(0.0037)	(0.0034)		(0.0046)	(0.0052)
Log Population	-0.0001		0.0017	0.0021	0.0021	0.0038	0.0075**	-0.0069
	(0.0028)		(0.0026)	(0.0035)	(0.0025)	(0.0041)	(0.0038)	(0.0047)
Average Contact in Wage Employment	-0.0031***	-0.0037***	-0.0026**	-0.0034***	-0.0031**	-0.0040***	-0.0037**	-0.0008
	(0.0011)	(0.0013)	(0.0012)	(0.0012)	(0.0012)	(0.0013)	(0.0017)	(0.0019)
Black Graduation Rate	0.0420	0.0442	-0.0065	-0.0728		0.0483	0.0883	0.126
	(0.0624)	(0.0696)	(0.0504)	(1.763)		(0.0787)	(0.0823)	(0.140)
White Graduation Rate	-0.0594	0.0175	-0.0105	-1.558		-0.0217	0.0438	-0.0845
	(0.0854)	(0.0775)	(0.0681)	(5.318)		(0.0857)	(0.0898)	(0.114)
Fraction of Blacks w/Self-Employed Fathers	0.0054	0.0099	0.0037	0.0054	0.0060	0.0094	-0.327	0.0031
	(0.0218)	(0.0216)	(0.0214)	(0.0242)	(0.0222)	(0.0237)	(0.350)	(0.0277)
Fraction of Whites w/Self-Employed Fathers	-0.0690*	-0.0535	-0.0724**	-0.0631	-0.0547	-0.0418	-0.160	-0.0391
	(0.0366)	(0.0374)	(0.0368)	(0.0398)	(0.0377)	(0.0437)	(4.617)	(0.0581)
Fraction of Blacks Segregated	0.0345							
	(0.0239)							
Fraction Black		0.0381				-0.143		
		(0.0291)				(0.430)		
Population		3.95e-10						
		(4.52e-10)						
Black Graduation Rate^2				0.0688				
				(1.099)				
White Graduation Rate^2				0.887				
				(3.030)				
Black College Degree Rate					-0.0713			
					(0.0836)			
White College Degree Rate					0.0450			
					(0.0557)			
Fraction Black^2						1.931		
						(4.223)		
Fraction Black^3						-6.867		
						(16.66)		
Fraction Black^4						7.338		
						(23.37)		
Fraction of Blacks w/Self-Employed Fathers^2							1.875**	
							(0.768)	
Fraction of Whites w/Self-Employed Fathers^2							-3.013	
							(8.932)	
Fraction of Blacks w/Self-Employed Fathers^3							-3.362**	
							(1.471)	
Fraction of Whites w/Self-Employed Fathers^3							21.00	
							(37.87)	
Fraction of Blacks w/Self-Employed Fathers^4							1.850**	
							(0.850)	
Fraction of Whites w/Self-Employed Fathers^4							-34.59	
							(51.46)	
States	43	43	43	43	43	43	43	43
R-squared	0.380	0.436	0.394	0.432	0.461	0.476	0.626	0.225

Bootstrap standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

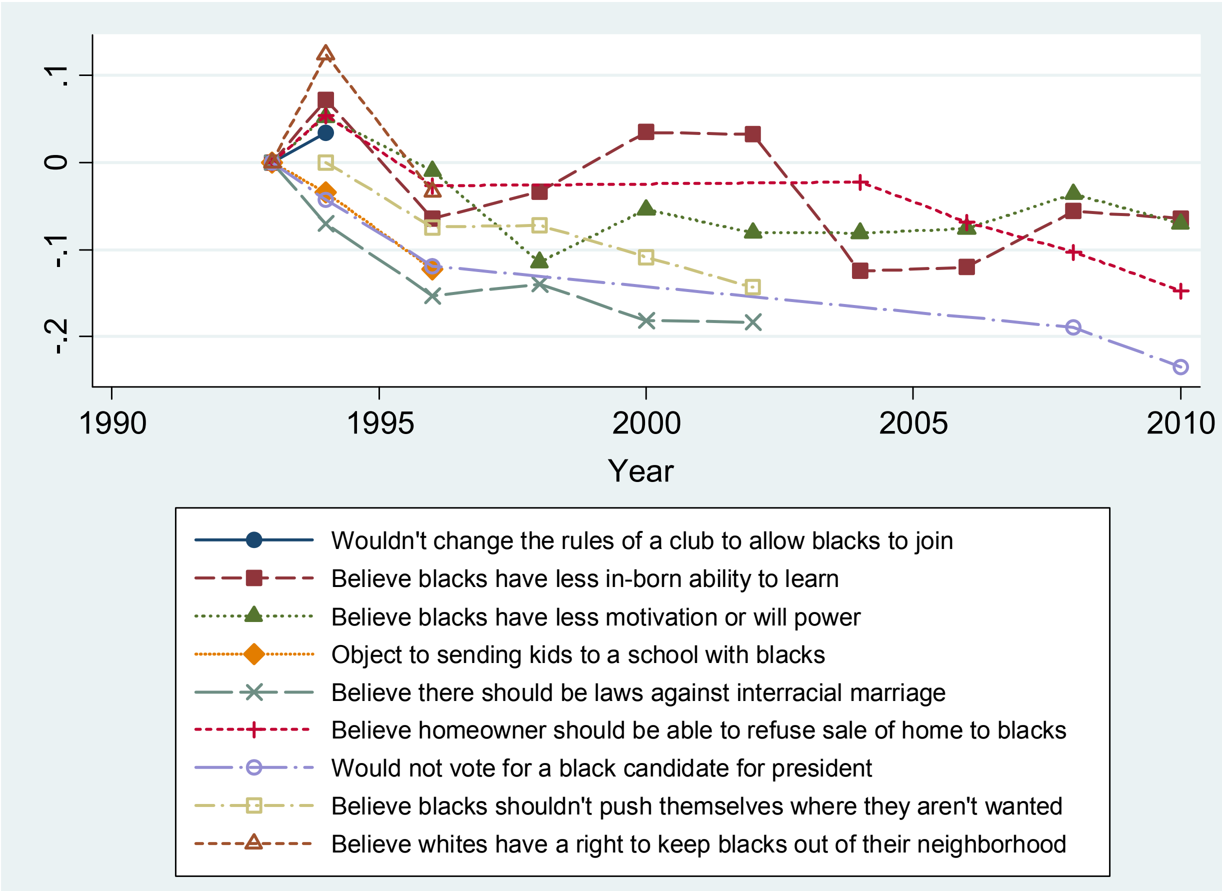
[a]The table reports the coefficients from a regression of the black-white self-employment gap on the aggregate prejudice index of all whites in the population. The dependent variable in the regression is the coefficient on the indicator variable for black that comes from state-level linear probability models that control for education categories, a quadratic in experience, citizenship, immigrant status, whether an individual was ever married, whether an individual is a military veteran, whether an individual speaks English poorly, and year fixed effects.

[b]Some states are omitted because the General Social Survey did not sample these states in the years for which the prejudice index is calculated, or because the states were not sampled in the years that specific questions were asked.

[c]The first-step regressions use standard errors calculated with successive difference replication using the 80 replicate weights that are included in the American Community Survey.

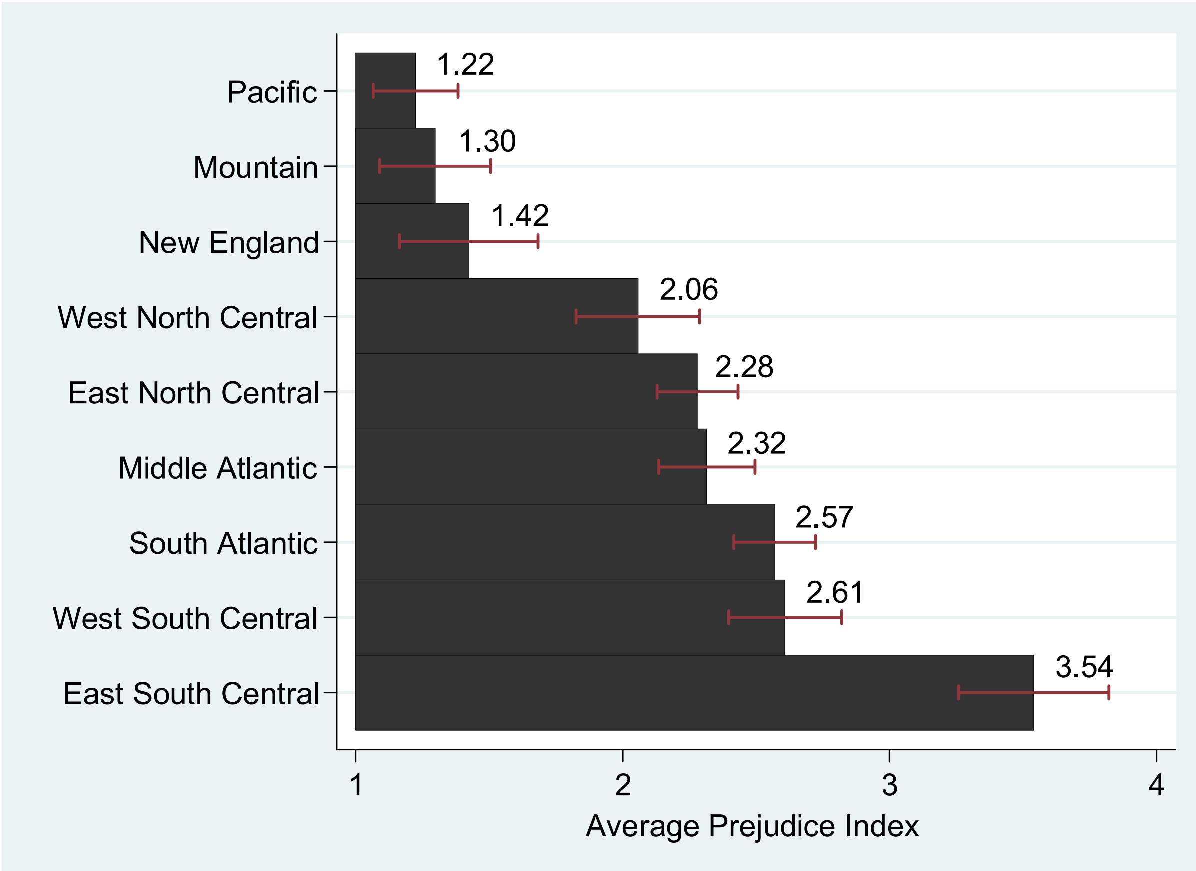
[d]Column (12) includes a segregation index to control for how close blacks are in proximity to whites. Column (13) allows the Fraction Black and the Population to enter the second-stage regression linearly. Column (14) omits Fraction Black. Column (15) uses quadratic terms for fraction of black males and fraction of white males with high school diplomas. Column (16) controls for the fraction of black males and fraction of white males with college degrees instead of high school diplomas. Column (17) controls for a 4th order polynomial in Fraction Black. Column (18) controls for a 4th order polynomial in black and white self-employed fathers. Column (19) presents a falsification test, which presents the results when the dependent variable is the state-level Hispanic self-employment gap, and all blacks are omitted. All are weighted using FGLS weights.

Figure 1.1: How Answers to GSS Prejudice Questions Change over Time



[a]Responses are standardized by the mean and standard deviation of the first year that each question is asked.

Figure 1.2: The Average Prejudice Index across Census Divisions



[a]The error band on each bar represents the 95% confidence interval on the estimate of average prejudice in that census division.

Figure 1.3: How the Average Prejudice Index has changed over Time across Census Divisions

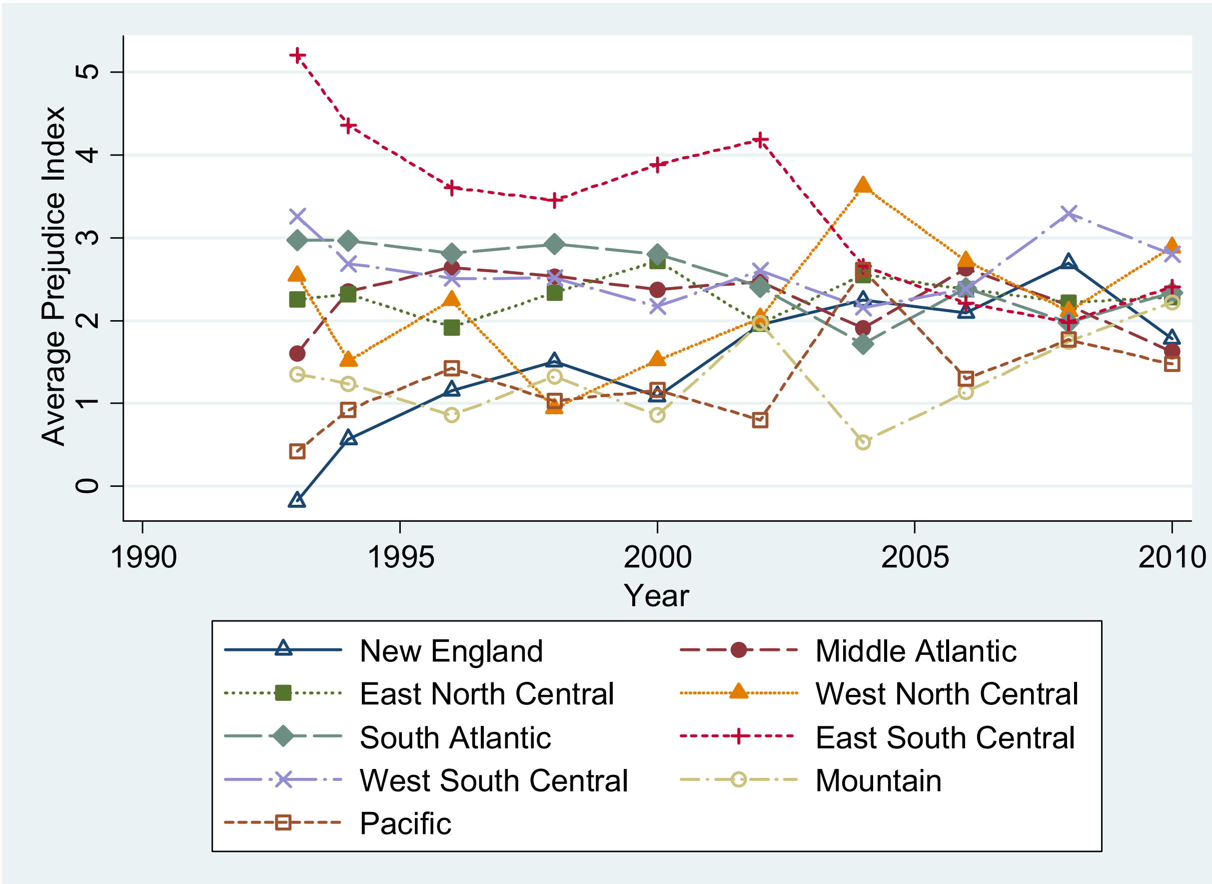
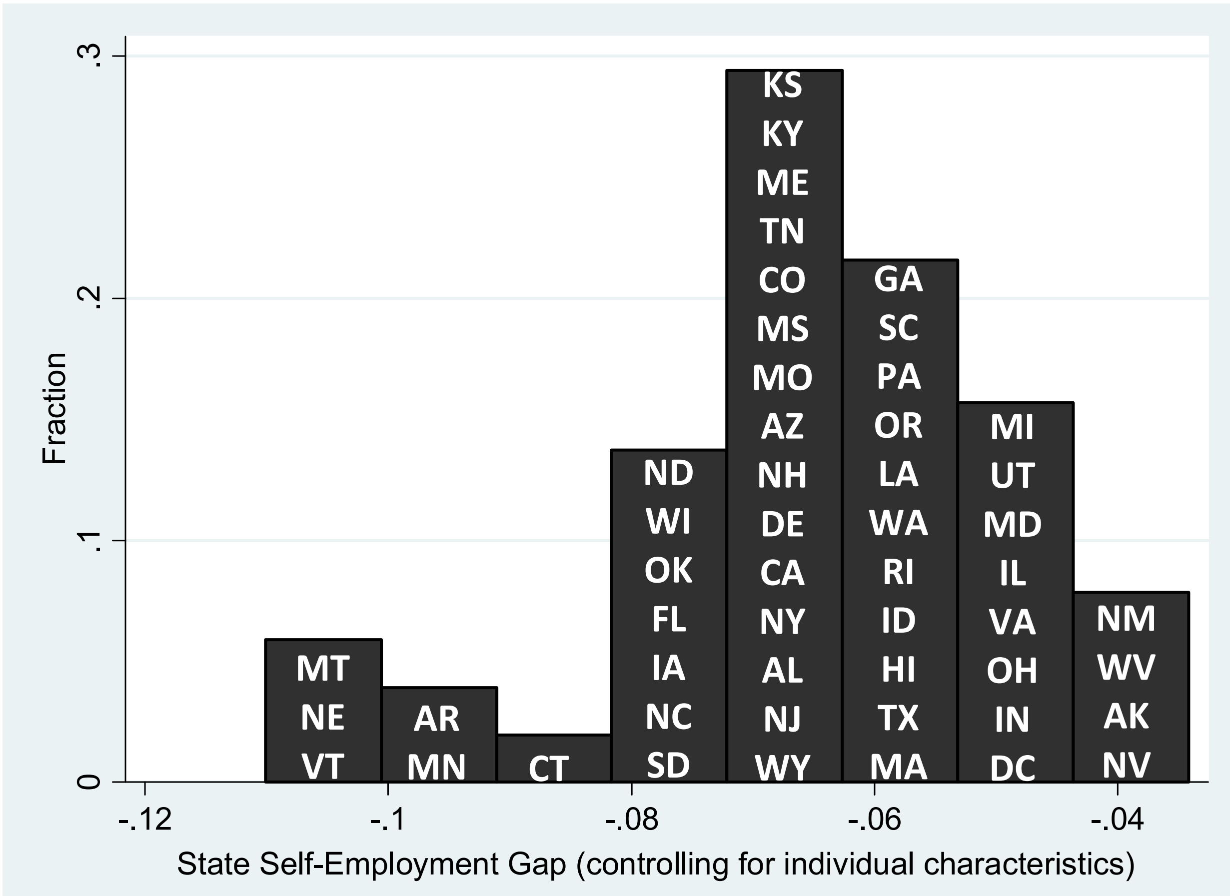
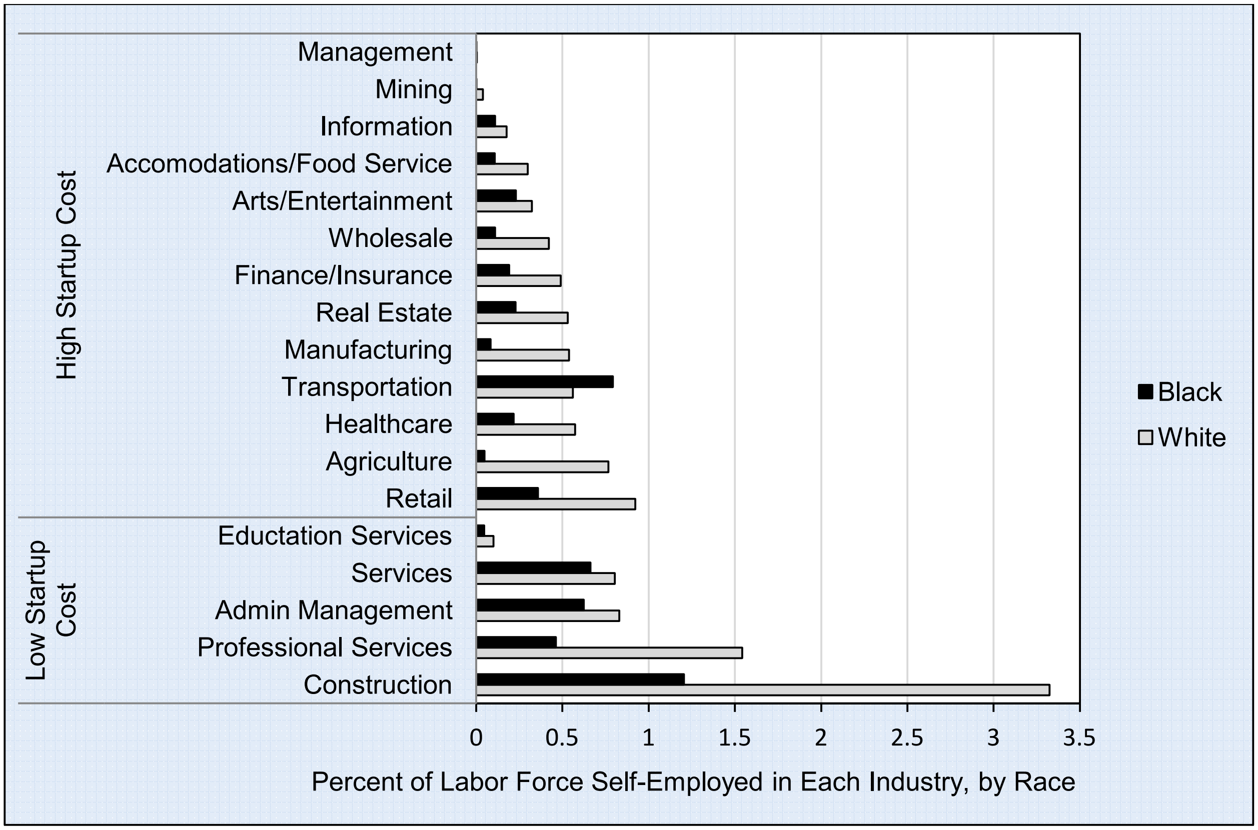


Figure 1.4: Distribution of the State Self-Employment Rate Gap,
Controlling for Individual-Level Characteristics



aBased on American Community Survey Public Use Microdata, 2005-2009 5-year estimates.
bWithin each bar, higher states have larger gaps than those below it.

Figure 1.5: Percent of Labor Force Self-Employed in Each Industry, by Race



Appendix 1.A. Questions from the General Social Survey, 1993-2010

Questions I choose to use because they only deal only with a respondent's views toward blacks:

CLOSEBLK- In general, how close do you feel to blacks? [Very Close=0, 1, 2, 3, Neither one Feeling nor the Other=4, 5, 6, 7, Not at All Close=8]

FEELBLKS- In general, how warm or cool do you feel towards blacks? [Very Warm=0, 1, 2, 3, 4, 5, 6, 7, Very Cool=8]

RACCHNG- If you and your friends belonged to a social club that would not let blacks join, would you try to change the rules so that blacks could join? [Yes=0, No=1]

RACDIF2- On the average blacks have worse jobs, income, and housing than white people. Do you think these differences are because most blacks have less in-born ability to learn? [No=0,Yes=1]

RACDIF4- On the average blacks have worse jobs, income, and housing than white people. Do you think these differences are because most blacks just don't have the motivation or will power to pull themselves up out of poverty? [No=0,Yes=1]

RACMAR-Do you think there should be laws against marriages between blacks and whites? [No=0,Yes=1]

RACOPEN- Do you agree that a homeowner should be able to decide for himself whom to sell his house to, even if he prefers not to sell to blacks. [No=0, Maybe=1, Yes=2]

RACPRES- If your party nominated a black for President, would you vote for him if he were qualified for the job? [Yes=0, No=1]

RACPUSH- Do you agree that blacks shouldn't push themselves where they're not wanted? [Disagree strongly=0, Disagree slightly=1, Agree slightly=2, Agree strongly=3]

RACSCHL (RACFEW/RACHAF/RACMOS)- Would you yourself have any objection to sending your children to a school where a few/half/most of the children are black? [No=0, Yes, but only if Most Black=1, Yes, but only if Half Black=2, Yes, if a Few Black=3]

RACSEG- Do you agree that white people have a right to keep blacks out of their neighborhoods if they want to, and blacks should respect that right? [Disagree strongly=0, Disagree slightly=1, Agree slightly=2, Agree strongly=3]

Questions I choose not to use because they are less directly focused on prejudicial attitudes towards blacks:

AFFRMACT- Do you favor preference in hiring?

BUSING- In general, do you favor or oppose the busing of (Negro/Black/African-American) and white school children from one school district to another?

COLRAC- Consider a person who believes that blacks are genetically inferior. Should such a person be allowed to teach in a college or university?

LIBRAC- Consider a person who believes that blacks are genetically inferior. If some people in your community suggested that a book he wrote which said Blacks are inferior should be taken out of your public library, would you favor removing this book, or not?

HELPBLK- Do you believe the government should be obligated to help blacks?

NATRACE- Are we spending too much on improving the condition of blacks?

NATRACEY- Are we spending too much on assistance to blacks?

RACDIF1- On the average blacks have worse jobs, income, and housing than white people. Do you think these differences are mainly due to discrimination?

RACDIF3- On the average blacks have worse jobs, income, and housing than white people. Do you think these differences are because most blacks don't have the chance for education that it takes to rise out of poverty?

SPKRAC- Consider a person who believes that blacks are genetically inferior. If such a person wanted to make a speech in your community claiming that Blacks are inferior, should he be allowed to speak, or not?

WRKWAYUP- Irish, Italians, Jewish and many other minorities overcame prejudice and worked their way up. Do you believe blacks should do the same without special favors?

Appendix 1.B. Additional Tables

Table 1.B1: The Effect of Average White Prejudice on the Black-White Self-Employment Gap Probit Coefficient

State Self-Employment Gap (Probit Coefficient)	(1)	(2)	(3)	(4)	(5)	(6)
Average White Prejudice	-0.0507***	-0.0616***	-0.0576***	-0.0653***	-0.0675***	-0.0586**
	(0.0177)	(0.0174)	(0.0167)	(0.0171)	(0.0199)	(0.0240)
Log Fraction Black		0.0376***	0.0344**	0.0385**	0.0394*	0.0375
		(0.0146)	(0.0151)	(0.0158)	(0.0223)	(0.0291)
Log Population			0.0246*	0.0288*	0.0258*	0.0254
			(0.0129)	(0.0152)	(0.0157)	(0.0187)
Average Contact in Wage Employment				-0.0107*	-0.0105	-0.0122
				(0.0064)	(0.0070)	(0.0084)
Black Graduation Rate					0.0505	0.0463
					(0.381)	(0.479)
White Graduation Rate					-0.218	-0.431
					(0.433)	(0.486)
Fraction of Blacks w/Self-Employed Fathers						-0.0368
						(0.154)
Fraction of Whites w/Self-Employed Fathers						-0.375
						(0.277)
Constant	-0.289***	-0.186***	-0.590***	-0.0955	0.0942	0.441
	(0.0428)	(0.0586)	(0.213)	(0.316)	(0.613)	(0.742)
States	45	45	45	45	45	42
R-squared	0.243	0.345	0.403	0.444	0.449	0.515

Bootstrap standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

[a]The table reports the coefficients from a regression of the probit coefficient representing the black-white self-employment gap on the aggregate prejudice index of all whites in the population. The dependent variable in the regression is the coefficient on the indicator variable for black that comes from state-level probit regressions that control for education categories, a quadratic in experience, citizenship, immigrant status, whether an individual was ever married, whether an individual is a military veteran, whether an individual speaks English poorly, and year fixed effects. The weights for each observation are explained in the text.

[b]Some states are omitted because the General Social Survey did not sample these states in the years for which the prejudice index is calculated, or because the states were not sampled in the years that specific questions were asked.

[c]One state is omitted because no blacks are observed as being self-employed in that state. The probit coefficient from the first-step probit regression cannot be estimated.

[d]The first-step regressions use standard errors calculated with successive difference replication using the 80 replicate weights that are included in the American Community Survey.

Table 1.B2: Startup Costs by Two-Digit NAICS Industry

Industry	Two-digit Industry NAICS	ACS Percent of Self-Employed	Average Est. Startup Cost	Startup Cost Classification	Less than $5,000	$5,000 to $9,999	$10,000 to $24,999	$25,000 to $49,999	$50,000 to $99,999	$100,000 to $249,999	$250,000 to $999,999	$1,000,000 or more	Don't Know	Not Applicable
Total	00	100%	$69,275.68		30.60%	8.40%	8.90%	5.80%	5.50%	5.30%	3.50%	1.50%	9.20%	21.30%
Agriculture, forestry, fishing, and hunting	11	6.06%	$56,683.25	High	25.00%	8.60%	12.70%	8.30%	6.70%	5.50%	3.00%	0.70%	10.70%	18.60%
Mining, quarrying, and oil and gas extraction	21	0.29%	$140,029.39	High	19.70%	5.90%	9.20%	7.40%	7.90%	7.10%	6.70%	3.50%	15.80%	16.90%
Utilities	22	0.00%	$148,888.36	High	22.00%	4.10%	S	4.90%	3.20%	4.60%	4.00%	5.10%	22.40%	24.60%
Construction	23	26.87%	$34,083.42	Low	34.20%	10.80%	10.00%	5.30%	4.10%	3.10%	1.50%	0.50%	9.70%	20.80%
Manufacturing	31-33	4.25%	$132,138.89	High	22.40%	9.10%	11.30%	7.70%	7.60%	7.60%	5.80%	3.60%	12.70%	12.20%
Wholesale trade	42	3.34%	$106,116.51	High	24.30%	8.40%	10.90%	7.80%	7.90%	7.70%	4.90%	2.50%	11.90%	13.80%
Retail trade	44-45	7.47%	$70,968.64	High	32.50%	8.00%	9.90%	7.50%	7.90%	7.50%	4.00%	1.00%	8.60%	13.30%
Transportation and warehousing	48-49	5.11%	$47,513.71	High	19.90%	8.80%	13.90%	11.10%	8.10%	4.30%	1.70%	0.70%	10.50%	20.90%
Information	51	1.46%	$64,565.52	High	33.30%	8.60%	8.00%	4.50%	4.20%	4.00%	2.60%	1.90%	8.00%	25.00%
Finance and insurance	52	3.98%	$120,296.49	High	27.70%	8.00%	8.80%	5.90%	6.00%	5.60%	4.10%	4.30%	9.00%	20.60%
Real estate	53	4.36%	$174,280.14	High	24.60%	6.50%	7.40%	5.60%	6.40%	8.60%	8.90%	4.80%	12.90%	14.40%
Professional, scientific, and technical services	54	12.34%	$33,578.48	Low	38.00%	9.80%	8.50%	4.40%	3.60%	2.90%	1.60%	0.60%	5.90%	24.70%
Management of companies and enterprises	55	0.03%	534,731.82	High	5.20%	1.90%	2.70%	2.30%	3.70%	5.40%	7.10%	18.40%	37.30%	15.80%
Administrative, support, waste management, etc.	56	7.00%	$30,672.50	Low	33.70%	8.40%	8.20%	4.40%	3.50%	2.70%	1.40%	0.50%	8.20%	29.00%
Educational services	61	0.82%	$27,497.19	Low	32.80%	5.60%	5.10%	3.00%	2.50%	2.30%	1.40%	0.50%	7.50%	39.20%
Health care and social assistance	62	4.64%	$67,591.98	High	25.60%	6.00%	6.60%	5.40%	6.20%	6.40%	4.30%	0.90%	10.40%	28.10%
Arts, entertainment, and recreation	71	2.74%	$39,370.12	High	36.90%	7.00%	6.00%	3.30%	3.00%	2.80%	1.90%	0.90%	7.90%	30.30%
Accommodation and food services	72	2.40%	$190,053.67	High	10.50%	5.30%	9.20%	9.90%	13.00%	16.40%	11.70%	3.30%	12.10%	8.60%
Other services (except public administration)	81	6.85%	$36,807.03	Low	32.80%	9.50%	9.70%	6.00%	4.80%	3.80%	1.90%	0.40%	8.80%	22.40%

[a]All information in the table except the ACS Percent of Self-Employed, the Average Estimated Startup Cost, and the Startup Cost Classification is taken directly from the 2007 Survey of Business Owners. The table shows the percent of respondent firms in each industry that had startup costs of Less than $5,000, $5,000 to $9,999, $10,000 to $24,999, $25,000 to $49,999, $50,000 to $99,999, $100,000 to $249,999, $250,000 to $999,999, or $1,000,000 or more; the survey also allowed for an answer of Don't Know or Not Applicable. An S indicates that the estimate could not be released due to confidentiality.

[b]ACS Percent of Self-Employed is the percent of self-employed individuals that are self-employed in each industry, taken from the ACS.

[c]The Average Estimated Startup Cost for each industry is calculated using the mid-point of each range, under the assumption that Not Applicable means that there was no startup cost to the business, and that the distribution of those that don't know the startup cost is the same as the distribution of those that do know the startup cost.

[d]I derive the Startup Cost Classification for each industry based on the Average Estimated Startup Cost of firms in that industry, doing my best to classify roughly half of the ACS sample of self-employed individuals as High and half as Low.

References: Chapter 1

U.S. Census Bureau. 2009. *American Community Survey Design and Methodology*. Washington, DC: U.S. Government Printing Office. <http://www.census.gov/acs/www/Downloads/survey_methodology/acs_design_methodology.pdf >.

Bates, T. M. 1997. *Race, Self-Employment, and Upward Mobility: An Illusive American Dream*. Washington, DC: Woodrow Wilson Center Press.

Becker, G. 1957. *The Economics of Discrimination*. Chicago, IL: University of Chicago Press.

Blanchflower, D., P. Levine & D. Zimmerman. 2003. Discrimination in the market for small business credit. *The Review of Economics and Statistics* 85: 930–43.

Borjas, G. J. 1982. On regressing regression coefficients. *Journal of Statistical Planning and Inference* 7 (2): 131–37.

Borjas, G. J., and G.T. Sueyoshi. 1994. A two-stage estimator for probit models with structural group effects. *Journal of Econometrics* 64: 165–82.

Brace, P., K. Sims-Butler, K. Arceneaux, and M. Johnson. 2002. Public opinion in the American states: New perspectives using national survey data. *American Journal of Political Science* 46 (1): 173–89.

Card, D., and A. Krueger. 1992. School quality and black-white relative earnings: A direct assessment. *Quarterly Journal of Economics* 107: 151–200.

Card, D., A. Mas, and J. Rothstein. 2008. Tipping and the dynamics of segregation. *Quarterly Journal of Economics* 123 (1): 177–218.

Cavalluzzo, K. S., L.C. Cavalluzzo, and J.D. Wolken. 2002. Competition, small business financing, and discrimination: Evidence from a new survey. *Journal of Business* 75: 641–80.

Cavalluzzo, K.S., and J.D. Wolken. 2005. Small business loan turndowns, personal wealth and discrimination. *Journal of Business* 78 (6): 2153–77.

Charles, K. K., and J. Guryan. 2008. Prejudice and wages: An empirical assessment of Becker's the economics of discrimination. *Journal of Political Economy* 116 (5): 773–809.

Coleman, S. 2008. The use of debt by black-owned firms: Recent evidence from the 2003 survey of small business finances. *Journal of Developmental Entrepreneurship* 13 (1): 39–57.

Cutler, D., E. Glaeser, and J. Vigdor. 1999. The rise and decline of the American ghetto. *Journal of Political Economy* 107 (3): 455–506.

Dunn, T. A., and D.J. Holtz-Eakin. 2000. Financial capital, human capital, and the transition to self-employment: Evidence from intergenerational links. *Journal of Labor Economics* 18(2): 282–305.

Fairlie, R. W. 1999. The absence of the African-American owned business: An analysis of the dynamics of self-employment. *Journal of Labor Economics* 17 (1): 80–108.

Fairlie, R. W. 2004a. Does business ownership provide a source of upward mobility for blacks and Hispanics? *Entrepreneurship and Public Policy*. ed. Doug Holtz-Eakin. Cambridge: MIT Press.

Fairlie, R. W. 2004b. Earnings growth among less-educated business owners. *Industrial Relations* 43 (3): 634-659.

Fairlie, R. W., and A.M. Robb. 2007. Why are black-owned businesses less successful than white-owned businesses? The role of families, inheritances, and business human capital. *Journal of Labor Economics* 25 (2): 289-323.

Fairlie, R. W., and A.M. Robb. 2008. *Race and Entrepreneurial Success: Black-, Asian-, and White-Owned Businesses in the United States.* Cambridge, MA: MIT Press.

Fairlie, R. W., and B.D. Meyer. 2000. Trends in self-employment among white and black men: 1910-1990. *Journal of Human Resources* 35 (4): 643–69.

Giuliano, L., D.I. Levine, and J. Leonard. 2011. Racial bias in the manager-employee relationship: An analysis of quits, dismissals, and promotions at a large retail firm. *Journal of Human Resources* 46 (1): 26–52.

Hanushek, E.A. 1974. Efficient estimators for regressing regression coefficients. *American Statistician* 28 (2): 66–67.

Holtz-Eakin, D., H.S. Rosen H. S., and R. Weathers (2000). Horatio Alger meets the mobility tables. *Small Business Economics* 14 (4): 243–274.

Hout, M., and H.S. Rosen. 2000. Self-employment, family background, and race. *Journal of Human Resources* 35(4): 670–92.

Huber, J. D., G. Kernell, and E. Leoni. 2005. Institutional context, cognitive resources and partisan attachments across democracies. *Political Analysis* 13: 365–86.

Kourilsky, M.L., and M. Esfandiari. 1997. Entrepreneurship education and lower socioeconomic black youth: An empirical investigation. *The Urban Review* 29 (3): 205–15.

Lentz, B. S., and D.N. Laband. 1990. Entrepreneurial success and occupational inheritance among proprietors. *Canadian Journal of Economics* 23: 563–79.

Lewis, J. B., and D.A. Linzer. 2005. Estimating regression models in which the dependent variable is based on estimates. *Political Analysis* 13 (4): 345-64.

Pryor, J. H., and E.J. Reedy. 2009. Trends in business interest among U.S. college students: An exploration of data available from the cooperative institutional research program. Ewing Marion Kauffman Foundation.

Robb, A., R.W. Fairlie, and D.T. Robinson. 2009. Patterns of financing: A comparison between white- and African-American young firms - Fourth in a series of reports using data from the Kauffman firm survey. Ewing Marion Kauffman Foundation.

Shane, S. 2007*. The Illusions of Entrepreneurship. New Haven, CT: Yale University Press.*

Singh, R., & M. Crump. 2007. Educational attainment: A key factor for improving the lagging rate of black entrepreneurship. *The Review of Black Political Economy* 34 (3): 217–29.

CHAPTER 2

TAX AVOIDANCE: HOW INCOME TAX RATES AFFECT THE LABOR MIGRATION DECISIONS OF NBA FREE AGENTS

2.1 Introduction

In the summer of 2010, LeBron James had a decision to make about his NBA playing future. In the week leading up to "the decision", James met with representatives from six NBA teams: the New Jersey Nets, New York Knicks, Miami Heat, Los Angeles Clippers, Cleveland Cavaliers, and Chicago Bulls. In an hour long special which aired on July 8^{th}, 2010 on ESPN, James chose to play the next five years of his career with the Miami Heat, a decision which is likely still being criticized by NBA analysts at the time of publication.

According to many income tax experts, LeBron James made the right decision. "If LeBron James goes to the Miami Heat instead of the Knicks, blame our dysfunctional lawmakers in Albany, who have saddled top-earning New Yorkers with the highest state and city income taxes in the nation, soon to be 12.85 percent on top of the IRS bite," said the New York Post. "On a five-year contract worth $96 million -- what he'd get from the Knicks or the Heat -- LeBron would pay $12.34 million in New York taxes (Gillis, 2010)." Based on a similar contract, "James would pay $5.69 million in state taxes if he re-signed with the Cleveland Cavaliers. If he signed with the New Jersey Nets, James would pay $10.32 million in state taxes (Gillis, 2010)." He would also pay $2.87 million more in state taxes if he signed with the Chicago Bulls as compared to the Miami Heat, and much more if he went to play for the L.A. Clippers in California, where the top income tax rate is 10.55% (Merchak, 2010). According to agent Gregg Clifton of Octagon Sports who represents big-name baseball players such as Tom Glavine and David Wells, "…there's a clear benefit to relocating to tax-free states (Heath, & Crenshaw, 2003)."

2.2 Literature Review

While the causes and effects of labor migration are of central importance in the labor migration literature, little work has been done regarding the effects that changes in state and city income tax rates may have on the labor migration decisions of high income individuals. This gap in the labor migration literature may be due to a lack of data availability. Many data sources, such as the March Current Population Survey and the Panel Study of Income Dynamics, top-code or censor individuals whose income exceeds a certain amount. In addition, these surveys contain only limited information about worker ability. By using a dataset of professional basketball players' free agent contracts from 2001-2007, I am able to circumvent both of these data limitations.

Clark and Hunter (1992) find that working males in their peak earning years are detracted by high income taxes in migration decisions. Wallace (1993, 2002) finds that income taxes are not always entirely borne by labor, and when income taxes are capitalized within wages, that they generally do not affect individual's migration decisions. However, it may be particularly useful to explore the effects that changes in income tax rates may have on the labor migration decisions of high income individuals, since high income individuals bear the largest burden of an income tax, especially when tax rates are progressive. Kirchgassner and Pommerehne (1996) and Feld and Kirchgassner (2001) find that high income individuals in Switzerland, where local tax competition between cantons and cities is relatively high, seem to choose their location based on the amount of income taxes that they must pay. However, their work does not estimate the causal effect of taxation, but rather notes a negative spatial correlation between cities and cantons with high income taxation and the location of high income individuals. Egger and Radulescu (2009) find that personal income tax rates have a negative effect on cross-border flows of skilled workers in OECD countries.

2.3 Taxes and the NBA

The National Basketball Association's regular season lasts from late-October until late-April. Approximately two-fifths of one season takes place at the end of one year, and three-fifths of the season takes place at the beginning of the next year. Thus, two-fifths of the income tax paid by a player in a given season will be taxed in the first year of a season, and three-fifths of the income tax paid by a player in a given season will be taxed in the second year of a season.

Each team in the National Basketball Association plays an eighty-two game regular season schedule. Exactly forty-one of these games are "home games" which are played in the team's home arena and are taxed by the team's home state. For the remaining forty-one "away games" which are played in their opponent's arena, players generally must pay whichever tax rate is higher, the home state's or the away state's (Merchak, 2010).[28] Many teams also play in cities where there are city or county income taxes as well. Of course, in the United States, state income tax is deductible from federal income tax. In Canada, however, this is not the case; provincial income tax is not deductible from Canadian income tax.

Under the collective bargaining agreement of the NBA there are two types of free agents: restricted free agents and unrestricted free agents. A restricted free agent may entertain offers from any team; however the player's current team reserves the right to match any offer in order to keep the player. Any veteran free agent who will have three or fewer years of NBA service following the season in which his contract expires will become a restricted free agent if the player's prior team makes a qualifying offer (the higher of 125% of his previous season's salary or $175,000 plus the league minimum) to him. An unrestricted free agent is a player who is not bound to any particular team, and is free to negotiate a contract with any team ("2005 Collective Bargaining Agreement," 2009). Unrestricted free agents include any veteran free agents with at least four years of service, any veteran free agents with three or fewer years of service that does not receive a qualifying offer from his prior team, and any undrafted rookies ("2005 Collective Bargaining Agreement," 2009). Figure 2.1 shows the average salary of NBA players in each season between 2001 and 2007.

There is a league-wide maximum salary which cannot be exceeded for a single player, and a salary cap which restricts the total amount of a team's payroll. While there are many exceptions to these restrictions which allow teams to exceed the salary cap, the most important of these exceptions fall into the "Larry Bird" family of exceptions. These exceptions, which are named after the famous Boston Celtics forward Larry Bird, allow teams to exceed the salary cap by re-signing their own free agents.

Teams that exceed the salary cap by large amounts are punished with a luxury tax. There is a strong incentive for teams to keep their payroll under the luxury tax threshold; for each dollar by which a team's payroll exceeds the luxury tax threshold, the team must pay to the league one dollar, which is then dispersed evenly among the teams which have not exceeded the luxury tax

[28] There are many tax law s which differ across different jurisdictions. In the years covered by this analysis, athletes were not required to pay income tax when playing on the road at the Toronto Raptors or Washington Wizards. Taxes at the Chicago Bulls were strictly retaliatory. Furthermore, athletes were not required to pay New York City taxes when playing at the New York Knicks. City tax rates for Detroit, Philadelphia, and Indianapolis differ for residents and non-residents. For away games, athletes that play for the Chicago Bulls are taxed 3% on top of the taxes that they pay to the away state. For a more thorough description, see Hoffman and Hodge (2004).

threshold ("2005 Collective Bargaining Agreement," 2009).[29] Table 2.1 shows the top ten NBA payrolls for each season between 2001-2003 and 2007-2008.[30] In the 2002-2003 season, in the first year of the luxury tax, more than half of the teams in the NBA were over the threshold. However, by 2005-2006 season, only one-sixth of the league had a payroll above the luxury tax threshold. But, in recent seasons the number of teams with payroll above the luxury tax threshold has again increased. In the 2007-2008 season, eleven teams had a payroll which exceeded the luxury tax threshold, and in the 2008-2009 season, thirteen teams had a payroll which exceeded the luxury tax threshold. While many teams choose to exceed the salary cap and even the luxury tax threshold, only a very small minority of teams choose to exceed the luxury tax threshold by a large amount.

It appears that most teams are constrained from raising their payroll above what is observed. Because of this, teams may not be able to compensate players for higher tax rates in the form of higher salaries. In fact, of the teams in our sample, 49.8% of them are already over the cap prior to the free agency period. These teams, on average, must sign an average of 4.88 players in free agency in order to get to the league minimum roster size of 14, and will be greatly restricted in the signing of free agents. However, this is an underestimate of the number of teams that are constrained, as an additional 50 teams have less than 20% of total cap space available to sign an average of 5.32 free agents to reach the league minimum roster size.

2.4 Theoretical Framework

Assume that each free agent maximizes his utility through his utility function, $U(W_{kc}(S),X_{kc})$, which is identical for all free agents, and depends only on the skill-dependent after-tax salary and a vector of team and city characteristics, where $W_{kc}(S)$ is the after-tax salary of a free agent with skill level S on team k in city c, and X_{kc} is a vector of characteristics which are specific to team k

[29] It is worth noting that actual luxury tax payments may be somewhat less than 1:1, since luxury tax payments may be deductible as operating expenses for each team's state and federal income taxes (Reilly & Schweihs, 2004).

[30] Team abbreviations: NYK-New York Knicks, DAL-Dallas Mavericks, CLE-Cleveland Cavaliers, DEN-Denver Nuggets, HOU-Houston Rockets, MIA-Miami Heat, BOS-Boston Celtics, SAS-San Antonio Spurs, LAL-Los Angeles Lakers, PHO-Phoenix Suns, GSW-Golden State Warriors, IND-Indiana Pacers, DET-Detroit Pistons, NJN-New Jersey Nets, TOR-Toronto Raptors, LAC-Los Angeles Clippers, CHI-Chicago Bulls, MIL-Milwaukee Bucks, ORL-Orlando Magic, WAS-Washington Wizards, NOH-New Orleans Hornets, NOO-New Orleans/Oklahoma City Hornets, CHA-Charlotte Hornets, ATL-Atlanta Hawks, PHI-Philadelphia 76ers, POR-Portland Trail Blazers, UTA-Utah Jazz, MEM-Memphis Grizzlies, SEA-Seattle SuperSonics, CHR-Charlotte Bobcats, MIN-Minnesota Timberwolves, SAC-Sacramento Kings.

and city c which affect player utility. Furthermore, assume that it is the goal of team k to maximize ΣS subject to some spending constraint B. For simplicity, let $w_{kc}(S)$ be the pre-tax salary of a free agent with skill level S on team k in city c, such that $W_{kc}(S) = w_{kc}(S)*(1- t_{kc})$, where t_{kc} is the state tax rate faced by players on team k in city c.

Suppose that a free agent with skill level $S=s$ receives two pre-tax salary offers; the free agent is offered $w_{11}(s)$ from team 1, and the free agent is offered $w_{22}(s)$ from team 2. If all elements of X_{kc} are identical, such that $X_{11}=X_{22}$, then the free agent will accept the offer from team 1 if and only if the after-tax salary offered by team 1 is greater than (or equal to) the after-tax salary offered by team 2; that is, if $W_{11}(s) \geq W_{22}(s)$ or $w_{11}(s)*(1-t_{11}) \geq w_{22}(s)*(1-t_{22})$.

Suppose for a moment that the state income tax rates faced by players on each team are equivalent, i.e. $t_{11}= t_{22}$. In this case, the free agent will accept the highest pre-tax offer. However, if we now suppose that the tax rate which faces players on team 1 is greater than the tax rate which faces players on team 2, $t_{11} > t_{22}$, then free agents will accept offers from team 1 if and only if the pre-tax salary offer from team 1 is greater than (or equal to) the pre-tax salary offer from team 2; $w_{11}(s) \geq w_{22}(s)*[(1-t_{22})/(1-t_{11})]$, where $(1-t_{22})/(1-t_{11}) > 1$.

Now, let $U(W_{11}(s),X_{11})= U(W_{22}(s),X_{22})$, relaxing the assumption that $X_{11}=X_{22}$ and allowing t_{11} and t_{22} to differ. If t_{11} were to rise relative to t_{22} to a level of t_{11}' and $w_{11}(s)$ and $w_{22}(s)$ were to remain unchanged, then $U(W_{11}(s),X_{11}) < U(W_{22}(s),X_{22})$, and players with offers from team 1 and team 2 would choose to play for team 2. However, if team 1 were to offer a high enough pre-tax wage, $w_{11}'(s)$, such that $w_{11}'(s)*(1- t_{11}') \geq W_{11}(s)$, then it would be possible to for team 1 to sign players with offers from team 2.

The implication of this theory is that for team 1 to sign a free agent which also has an offer from team 2, team 1 must compensate him for the income tax rate differential, holding constant all other team and city specific characteristics which a player cares about. While a free agent may make a counteroffer to team 1 which explicitly considers tax rates, to the extent to which it is not possible for team 1 to compensate a free agent for the income tax rate differential (due to salary cap restrictions, the luxury tax threshold, etc.), it would be expected that team 2 will be able to sign the free agent. Indeed, team 2 will likely be able to sign most of the free agents with offers from both team 1 and team 2, and team 1 will have to target many lower skilled free agents which team 2 has no interest in. In this case, the average skill of the free agents signed by team 2 will be higher than the average skill of the free agents signed by team 1.

In the preceding theoretical analysis, a free agent's skill level *S* has been taken as known, and it has been implicitly assumed that both teams have the same expectation of the free agent's skill level *S*. It may be worth noting that in reality each team may have a different expectation about *S*. There may also be other considerations which may affect each team's evaluation of a given player such as player maturity and off-court behavior, or team cohesiveness.

2.5 The Data

The data used in this paper comes from a multitude of sources. National Basketball Association player contract data, which includes data on team, year, position played, contract length, and total contract dollar amount, is taken from USA Today Salaries Database and the Pro Sports Transactions website. Player performance data is taken from the Database Basketball website. State and federal income tax rate data is taken from the National Bureau of Economic Analysis Taxsim. Canadian central and provincial tax rates are taken from the Canada Revenue Agency. State effective sales taxes and effective property taxes are calculated from the Census of Governments. MSA demographic data, including population, employment, and income per capita, are taken from the Bureau of Economic Analysis. Data pertaining to the relevant exchange rate between the United States and Canada is taken from the Bank of Canada.

Summary statistics for all free agent signees broken down by the relative income tax rate quartiles (deviation from season mean) are available in Table 2.2.[31] Furthermore, all summary statistics are weighted by contract length so that our sample is representative of the entire population of free agent signees.

Inference of summary statistics is by no means rigorous; however, by looking at the summary statistics, a few patterns are noticeable. It appears that income tax rates and properties tax rates have a negative relationship, as do income tax rates and sales tax rates, as one might expect. There does not seem to be much of a trend in total contract amount or average salary across income tax quartiles. However, it does appear that free agents that sign into relatively lower income tax quartiles have more points per game, blocks per game, rebounds per game, assists per game, and a higher free throw percentage, all leading to a higher performance per game measure throughout their career. They are also more likely to have been all-stars, and to

[31] Income tax rates are the effective marginal tax rate for each team deviated from the team level mean in each season. Sales tax and property tax rates are state effective tax rates calculated from the Census of Governments, deviated from the team level mean in each season.

be taken earlier in the NBA draft.

Ideally, to determine the impact of an income tax differential, we would compare two teams which are identical in every aspect except for the income tax rates faced by the two teams, and observe the difference in skill of the free agent signees or the incremental salary increase which is necessary to make a free agent indifferent between the two teams. However, since we cannot observe any team's counterfactual, we must take a different approach. Since we can only observe accepted offers, we must compare the salaries of free agent signees of differing skill levels that sign with different teams, and attempt to control for differences between the players and teams. First, we must determine the relevant comparable team specific income tax rate which faces each NBA free agent signee. Quick inspection of the data shows that every free agent signee in the sample is above the highest tax threshold in their respective state and the majority make at least a few million dollars more, so the highest marginal tax rate in each state is a good approximation for the actual tax rate that each free agent signee will face in that state. State and local tax rates are federal tax deductible, so we must adjust the state and local marginal tax rates to get an effective tax rate. Since forty-one games take place in the home state and forty-one games take place on the road, and each team typically has road games in approximately the same locations each season, we can find a linear combination of effective marginal state and local income tax rates that face each team's players.[32] Because the season takes place over the course of two separate years, the tax is paid over two separate years, possibly at different tax rates. For more about how our relative income tax rate measure is formulated, see the appendix.

Due to the way that our relative income tax rate measure for each year is calculated, we can garner some intuition about when the relative income tax rate will change. Obviously, any change to the highest marginal state or local income tax rate will change the relative income tax rate facing any of that state's home teams, however, it will also alter the relative income tax rate facing any other team in the league, both because their players may pay some tax towards that state, and because it changes the mean tax rate for that year. This makes sense, since a change to the tax rate facing other teams greatly alters the after-tax salary available to each free agent if they had signed with that other team. In the same way, any change to a U.S. federal or Canadian

[32] Each team plays other teams in the same division four times, twice on the road and twice at home. Each team plays all teams in the opposing conference twice, once on the road and once at home. Each team plays all teams in their conference from different divisions three to four times each, split as evenly as possible between home and away games.

central tax rate policy will also change the relative income tax rate facing each team by raising or lowering the relative income tax rate facing free agents that sign with the Toronto Raptors, the only Canadian team in our sample. For example, if the U.S. government decides to increase federal tax rates, the relative income tax rate facing the Toronto Raptors' players will fall since the league's average tax rate will rise relative to the Toronto Raptors' tax rate. Any change in league alignment, such as that which took place prior to the 2004-2005 NBA season, will also alter relative income tax rates since the opponents that teams will face each year will change. It is worth noting that, since provincial tax rates are not tax deductible from Canadian taxes, and Canadian tax rates are typically different from U.S. federal tax rates, provincial tax rates must be adjusted so that they are comparable to any U.S. state in the preceding analysis.[33]

Figure 2.2 shows the relative income tax rate facing each team between the 2001-2002 to 2007-2008 seasons. After looking over Figure 2.2, it should be clear that, while the relative tax rate may vary significantly across teams, they may not vary much within teams over time. However, even changes of a fraction of a percentage point, which occur quite frequently within the sample, may represent a large change in after-tax salary. The largest change in relative income tax occurred for those players on the Toronto Raptors between when the U.S. dropped its federal income tax rate from 39.1% in 2001 to 35% in 2003. Players on the New Jersey Nets also experience a change in income tax rates when the New Jersey increased their state income tax from 6.37% in 2003 to 8.97% in 2004. New York State also increased income tax rates from 6.85% in 2002 to 7.7% in 2003, only to return them to 6.85% in 2006.

Figure 2.3a shows the distribution of relative income tax rates accepted by free agent signees. Figure 2.3b shows the distribution of relative income tax rates accepted by free agent signees weighted by contract length, and Figure 2.3c shows the distribution of relative income tax rates facing each team's players. Since the distributions in Figure 2.3a and Figure 2.3b are nearly identical, we would not expect that there is much correlation between relative tax rates and the length of contract being signed. However, it does appear that those teams with the lowest relative income tax rate are offering shorter contracts. A comparison of Figure 2.3a to Figure 2.3c yields the same conclusion.

[33] The provincial tax in Toronto is not deductible from Candian tax in the same way that state tax is deductible from U.S. federal taxes. While the Canadian government will generally also not give tax credits for income earned in foreign countries, the U.S. government generally will give tax credits for all foreign tax. My analysis has taken this into account. See IRS Publication 514 (Department of the Treasury, 2010).

2.6 Methodology

I will seek to estimate the causal effect of changes in relative income tax rates on the labor migration decisions and compensation of NBA players. It will be necessary to consider those factors which might affect a team's ability to recruit a free agent which are correlated with the relative income tax rate. The factors we will include are the relative sales tax rate, the relative property tax rate, a lag in the natural logarithm of the population (in the MSA), a lag in the natural log of per capita income (in the MSA), a lag in the natural log of employment (in the MSA), the exchange rate (adjusted for purchasing power parity), a lag in the state's crime rate (per 100,000), a lag in the state's student-teacher ratio, the team's wins last season, whether the free agent switches teams, and the amount of money a team can use to attract free agents under the salary cap. Clearly the sales tax and property tax rates ultimately affect a player's amount of discretionary income, and both are correlated with income tax rates. Population, income, and employment are relevant because they affect consumer demand for basketball, and ultimately may affect the salary a team is willing to offer free agents, while at the same time they may be correlated with a change in income tax rate. The hope here is that if taxes change in response to economic problems that also adversely affect an NBA team, that those economic problems can be fully explained by changes in population, income, and employment. The crime rate and student teacher ratio may be reasons why tax rates increase. For example, if a state decides to increase expenditures in response to a high crime rate in order to pay for additional police officers or longer prison terms, they may have to increase tax rates to finance the additional spending. A similar argument would hold for a response to school quality, so both crime rate and student teacher ratio may be correlated with changes in the relative income tax rate, while they may also be correlated with a player's preferences.[34] A team's wins last season might increase labor supply, since players may be more willing to play for a winning team. Whether a player resigns with a team may also ultimately affect the amount of money which a team can offer, through the "Larry Bird" family of exceptions. The amount of money a team has to spend on free agents may also affect whether a given free agent signs with a team, both through the direct effect of increasing the likelihood of a higher salary offer, and through the indirect effect

[34] Crime rates could be correlated with higher tax rates since governments are likely to respond to crime problems by enlarging police forces. See Cornwell and Trumbull (2004) and Loftin and McDowall (1982) for a more thorough description. Figlio and O'Sullivan (2001) present a model which causes tax rates to respond to a high student-teacher ratio.

of being able to improve the team by attracting other top quality free agents. We include lags for some variables because it would be impossible for a free agent to estimate the current value of things such as population counts or crime rates since these data are usually gathered with a lag. We will focus on the differences between a free agent's signing team and the overall average of all other teams, which best represent a free agent's outside option in the available data, since we do not know which other teams have given offers to a free agent.

My empirical model takes the form

$$SKILL_{ijct} = \beta_0 + \beta_1 TAX_{ct} + \beta_2 X_{it} + \beta_3 TEAM\ DUMMY_j + \beta_4 TIME\ DUMMY_t + \varepsilon_{ijct} \quad (1)$$

$$LOG\ SALARY_{ijct} = \gamma_0 + \gamma_1 TAX_{ct} + \gamma_2 Z_{it} + \gamma_3 TEAM\ DUMMY_j + \gamma_4 TIME\ DUMMY_t + \eta_{ijct} \quad (2)$$

where $SKILL_{ijct}$ is the *measured* skill of free agent *i* at time *t* which is hired by team *j* in city *c*, TAX_{ct} is the relative income tax rate of city *c* in year *t*, X_{it} is a vector of the individual specific characteristics of free agent *i* measured at time *t*, $TEAM\ DUMMY_j$ is a dummy variable which takes a value of 1 if a team *j* signs free agent *i* at time *t* and 0 if it does not, and $TIME\ DUMMY_t$ is a vector of dummy variables which take a value of 1 if the free agent *i* was signed in year *t*, and take a value of 0 otherwise.

Note that if the model takes the form shown above, that β_1, the coefficient on TAX_{ct} in (1), is identified by a change in the skill of free agent signees within a team-city pairing brought about by a change in the relative income tax rate facing a team's players over time. In the same fashion, γ_1, the coefficient on TAX_{ct} in (2), is identified by a change in the natural logarithm of annual salaries of free agent signees within a team-city pairing brought about by a change in the relative income tax rate facing a team's players over time. So β_1 represents the average change in the measured skill level of free agent signees which is brought about by a one-percentage point change in the relative income tax rate, and γ_1 represents the average change in the natural logarithm of the annual salary of free agent signees which is brought about by a one percentage point change in the relative income tax rate. If the majority of changes in the tax rate differentials are capitalized in the average salaries of free agent signees, then we would expect little or no effect on the change in skill level of free agent signees brought about by changes in the relative income tax rate (Wallace 2002). However, if changes in the tax rate differential are not capitalized in the average salaries of free agent signees, then we may expect large changes in the skill level of free agent signees brought about by changes in the relative income tax rate (Wallace 2002). Due to the identification strategy that is employed here, if we believe that any factor which we have chosen to ignore is time invariant, then omitting these factors will not change our results.

The problem which remains is that skill is difficult to measure. The many statistics which are measured by the National Basketball Association can be combined to create a proxy for skill. We will use the market value of these statistics to index free agent skill. To determine the market value of each statistic, we run the regression of the natural logarithm of the average annual salary of the contract signed by each free agent on the statistics that each free agent has put up over the preceding years of his career, on a per game or per 48 minute basis. The predicted value of this regression will then represent the measured skill of each free agent signee. The results of this regression are shown in Table 2.3.

Specifications 1-3 show the regression of the natural logarithm of average annual salary on per game statistics over the career of each free agent prior to their contract signing, while specifications 4-6 show the regression of the natural logarithm of average annual salary on statistics which are calculated per 48 minutes (the length of one full NBA game) over the career of each free agent prior to their contract signing. Specifications 4-6 should adjust upward the measured relative skill level of any "highly skilled" free agent which has not played a large number of minutes per game throughout his career and adjust downward the relative skill level of any "low skilled" free agent which has played a high number of minutes per game throughout his career.

From Table 2.3, we get many expected results. For instance, a free agent's annual salary seems to be positively correlated with seasons played, games played per season, and minutes played per game, as well as points, blocks, rebounds, and assists per game (or per 48 minutes), and a free agent's annual salary seems to be negatively correlated with turnovers, fouls, missed field goals, and missed free throw attempts per game (or per 48 minutes). However, steals, which are generally thought to be a positive statistic, tend to be somewhat negatively correlated with salary. One reason why this might be the case is that it is generally the shorter players which get the most steals. To the extent that this height difference is not fully captured by differences in position, it might be captured in the steals statistic in the form of lower salaries since we have not included a height variable in the regression equation. On average, centers make more than both forwards and guards, and forwards make more than guards, although this result is not statistically significant.

Using the predicted values of the regression in Table 2.3 (Column 2), we can estimate the average skill level of free agent signees for each team in the NBA over the seasons in our sample. Figure 2.4a plots the average skill level of free agent signees for each team against the average tax rate for each team relative to the rest of the league. The average skill level of each team's free agent signees is negatively correlated with the relative income tax rate facing each team's players. Figure 2.4b plots the average salary of each team's free agent signees against the average tax rate for each team relative to the rest of the league. The average salary of each team's free agent signees is positively correlated with the relative income tax rate facing each team's players.

We can then standardize the predicted values of the regression from Table 2.3 to put our measure of skill into units that we can more easily interpret. We then run the regression in (1). The regression results are shown in Table 2.4 and Table 2.5.[35] [36] [37] Whether we look at Table 2.4, which uses a performance per game measure as the dependent variable, or Table 2.5, which uses a performance per 48 minutes measure, we find that decreasing the marginal income tax rate by 1 percentage point in our full specification leads to a 0.080-0.088 standard deviation increase in the skill of the average free agent signee for a given team, ceteris paribus. This effect is both economically and statistically significant.[38]

As Wallace (2002) suggests, since the skill of free agent signees responds to changes in the relative income tax rate, it must be the case that teams do not fully compensate all free agents for

[35]While Table 2.4 and Table 2.5 show the regression results using the predicted value of the regression from Column 2 of Table 2.3, and Column 5 of Table 2.3 respectively, use of other specifications from table 2.3 to index skill does not drastically change our results.
[36] Inclusion of team specific linear time trends does not change the point estimate of the coefficient of interest; however standard errors become increasingly large.
[37] For an in-depth look at clustered standard errors, see Angrist and Pischke (2008). For information about two-way clustering, see Cameron et al (2006) and Thompson (2009).
[38]As a robustness test, in order to make sure that these results are not driven by one "outlier" team, I systematically remove one team at a time to see what happens to the coefficient on RELATIVE INCOME TAX RATE. Exclusion of the Toronto Raptors from our sample leads to a coefficient on RELATIVE INCOME TAX RATE of -0.283 when PERFORMANCE PER GAME is the dependent variable; however this reduces both the sample size and the variance of relative income tax rates, leading a much larger standard error of 0.181. Excluding the Houston Rockets from our sample leads to a coefficient on RELATIVE INCOME TAX RATE of -0.028 when PERFORMANCE PER GAME is the dependent variable with a standard error of 0.027; again this reduces both the sample size and the variance of relative income tax rates. Removal of any other team one at a time does not appear to drastically change our results. Using the WIN SCORE PER GAME measure, the coefficient on RELATIVE INCOME TAX RATE with either team removed is negative and statistically significant at the 5% level.

the income tax differential. While it appears that free agents are, on average, reacting to changes in the marginal income tax rate, it is still possible that some teams are able to compensate some free agents for changes in the relative income tax rate in order to sign them. Therefore, it remains to check whether any of the changes in the relative income tax rate are capitalized in the salaries of free agent signees. These results are found in Table 2.6. Theoretically, to fully compensate a player for a 1% increase in tax rates, salaries must increase by approximately 1.01%. Therefore, if any teams are able to compensate some free agents for changes in the relative income tax rate in order to sign them, then we would expect a positive coefficient on the relative income tax rate which is less than 0.0101. However, none of our specifications led to a coefficient on the relative income tax rate which is statistically significant, and all of the specifications lead to a negative coefficient. Therefore, we are able to learn very little about how changes in relative income tax rates affect the salaries of free agent signees. However, it is not hard to believe that teams in the NBA are too financially constrained to be able to compensate free agents for any income tax differential, as we have discussed in previous sections.

To further illuminate the effect of a 1 percentage point fall in the relative income tax rate, we can convert the 0.080 standard deviation increase in the skill of the average free agent signee to a dollar figure which comes from the market price of the incremental increase in skill. In particular, by taking an approximation around the average performance per game measure, we find that a 1 percentage point fall in the relative income tax rate leads to an average free agent signee skill increase worth $231,260 per year. However, if the average free agent were to be taxed an additional 1 percentage point on their average log salary, the tax would amount to only an additional $30,019. This does seem like an extremely large effect for such a relatively small amount of money. We would think that if a team were able to pay an extra $30,019, on average, to a free agent, then they could fully compensate them for a 1% tax differential. However there are a few reasons why this might be plausible. Since the most highly skilled free agents will usually earn maximum contracts with a maximum contract length, it is impossible for relatively high income tax teams to compensate these free agents for any tax differential. Since teams may be strained financially, free agents may opt to take even a little more in after tax dollars to play for teams with lower tax rates. Furthermore, since the skill distribution of free agent signees is very spread out (standard deviation in performance per game is 0.92, which translates to roughly $2,676,123 in terms of market worth), and there are a limited number of free agent signees each year (an average of 93 or approximately 3 per team), a change in relative income tax rates which attract or discourage only one or two free agents from signing may result in large resulting skill differences among free agents signees. Figure 2.5 makes this increasingly clear. Each bar in Figure 2.5 is split into 0.080 standard deviations of the skill variable. If we use 2001 as an example, we can see that in the tails of the distribution there are often one or fewer free agent signees in any 0.080 standard deviation area.

Next, I decide to look at other metrics of NBA performance to see if they show corroborating results. Specifically, I look to the NBA Efficiency, Game Score, and Win Score measures of performance, since they are easily calculated using our dataset. NBA efficiency is a statistic which is reported by the National Basketball Association website, and it is believed that many coaches and executives rely on the NBA efficiency measure to evaluate player performance (Berri, Schmidt, & Brook, 2006). Game Score is a measure created by sports writer John Hollinger in order to measure single game performances. Win score is a measure created by economist David Berri from a regression of team wins on points scored and points allowed per possession (Berri, Schmidt, & Brook, 2006). Despite its name, NBA efficiency measures performance quite inefficiently; it is a linear combination of box score statistics, all with equal weight. Game score gives different weight to the statistics that John Hollinger felt "evaluate players in a fashion consistent with what NBA observers would believe (Berri, 2006)." The win score measure does slightly better by giving more weight to those box score statistics which are more closely correlated with winning or losing games. A major downfall of all three of these measures is that they measure only box score statistics, and thus give no added value to experience. The results are shown in Table 2.7.

Decreasing the relative income tax rate by 1 percentage point leads to a 0.117 standard deviation skill increase in the average free agent signee for a given team, as measured by win score per game, or a 0.182 standard deviation skill increase as measured by win score per 48 minutes. These results are statistically significant at the 2% and 1% level, respectively. The coefficient on relative income tax rate when the dependent variable is NBA efficiency (per game or per 48 minutes) or game score (per game or per 48 minutes) is negative, although not statistically significant. We might believe this is because NBA efficiency and game score are not very accurate measures of skill (Berri, 2006). It does seem that a decrease in relative income tax rates leads to an increase in the skill of free agent signees, regardless of how skill is measured.

As has been argued thus far, we find that decreasing the marginal income tax rate leads to an increase in the skill of the average free agent signee for a given team. If this is the case, then we might expect that team success might be significantly affected by a change in tax rates. This is shown in Table 2.8. We find that decreasing the relative income tax rate by 1 percentage point is associated with 4.074 more regular season wins in that season in the full specification. This result is significant at the 10% level. We also find that decreasing the relative income tax rate by 1 percentage point is associated with 2.422 more playoff wins in that season in the full specification. This result is not significant at the 10% level, although it is extremely close given a small sample size of 206 team-season observations.

2.7 Conclusion

By using a dataset of professional basketball players' contracts from the 2001-2002 to 2007-2008 seasons, which has rich uncensored data on worker productivity, I am able to identify the effect that changes in state and local income tax rates have on the labor migration decisions of high income NBA free agents. I find, after controlling for other observable characteristics of teams, cities, states, and MSAs, that an increase in the relative income tax rates facing players on a given team leads to a decrease in the average skill of the free agents that that team is able to sign.

While it is difficult to measure skill, we determine the value that teams in the National Basketball Association attach to different player attributes by regressing the natural logarithm of annual salary on a free agent's career statistics prior to signing his contract. Using market weights, we can aggregate a free agent's career statistics to determine their measurable skill level. I then find that after controlling for other observable characteristics of teams, cities, states, and MSAs, that a one percentage point increase in the marginal state income tax rates facing players on a given team leads to a 0.080 to 0.088 standard deviation decrease in the average skill of the free agents that that team is able to sign. Other metrics of player performance seem to see corroborate these results.

While the focus of this paper is very specific, it would appear that the results may be generalizable to a larger population. While I find that an increase in the marginal state income tax rates facing players on a given NBA team leads to a decrease in the average skill of the free agents that that team is able to sign, an analogous result might hold for athletes in other sports.

A similar result might also hold for actors, singers, and other entertainers, which are taxed in a similar fashion. Cities are clamoring to improve their culture, and perhaps state income tax rates have some effect on attracting quality in other entertainment areas as well.

Future research might look at the relationship between income tax rates and labor migration in other major sports. Of particular interest might be how the collective bargaining agreement set up in each sport frames this relationship. For instance, in Major League Baseball, where there is no salary cap and the luxury tax is much weaker than in the NBA, are tax rates more likely to be capitalized in salaries? Or, did a hard salary cap in the National Football League prior to the 2010 season prevent teams from compensating free agents for an income tax differential, causing labor migration to be very sensitive to tax rates, and did this change when the salary cap was dropped in the 2010 season?

The current NBA Collective Bargaining Agreement is set to expire on June 30, 2011, and front and center to the issues surrounding the new CBA is the salary cap (Garcia, 2010). The owners are currently pushing for a hard cap for each team which cannot be exceeded, while the union vehemently opposes one. I will not sound off on this debate at the moment, however, based on the results of this paper, it may make sense to allow the salary cap of each team to respond to changes in the relative income tax rate facing each team's players. This would give teams the ability to compensate free agents for changes in the relative income tax rate, and maintain their ability to attract top free agents.

Figure 2.1: Average NBA Player Salary, 2001-02 to 2007-08 Seasons

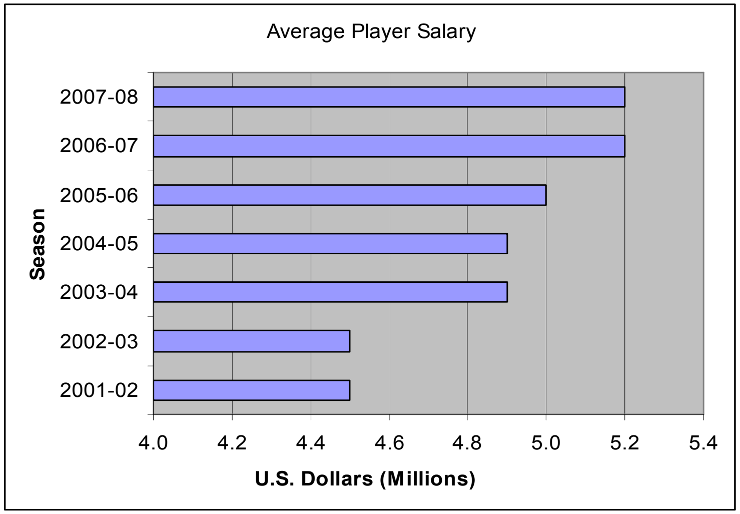


Table 2.1: Top Ten NBA Payrolls in Millions of Dollars, 2001-2007

Season	2001-2002	2002-2003	2003-2004	2004-2005	2005-2006	2006-2007	2007-2008
Salary Cap	$42.5	$40.3	$43.8	$43.9	$49.5	$53.2	$55.6
Luxury Tax Threshold	-----	$52.9	$54.6	-----	$61.7	$65.4	$67.9
Highest Team Payroll	$85.3(NYK)	$104.3(POR)	$84.5(NYK)	$94.1(NYK)	$92.9(NYK)	$82.4(PHO)	$92.8(NYK)
2nd Highest Team Payroll	$84.0(POR)	$93.0(NYK)	$84.3(POR)	$87.4(DAL)	$79.6(SAS)	$81.7(NYK)	$88.4(DAL)
3rd Highest Team Payroll	$57.8(PHI)	$70.8(DAL)	$79.0(DAL)	$80.2(POR)	$68.3(DAL)	$78.2(MIA)	$84.2(CLE)
4th Highest Team Payroll	$56.5(DAL)	$70.4(SAC)	$72.4(MIN)	$70.1(MIN)	$66.6(IND)	$75.5(DET)	$81.2(DEN)
5th Highest Team Payroll	$55.6(MIL)	$65.3(PHI)	$69.6(SAC)	$68.1(ORL)	$64.6(MEM)	$66.7(MIN)	$77.6(HOU)
6th Highest Team Payroll	$54.8(SAC)	$62.6(LAL)	$65.5(LAL)	$66.3(IND)	$61.6(NJN)	$65.6(SAS)	$74.7(MIA)
7th Highest Team Payroll	$54.7(PHO)	$60.5(MEM)	$65.2(PHO)	$65.1(LAL)	$60.5(SAC)	$65.0(DEN)	$73.8(BOS)
8th Highest Team Payroll	$54.5(MIN)	$60.4(NJN)	$63.5(ATL)	$63.8(PHI)	$59.7(PHI)	$64.8(DAL)	$72.5(SAS)
9th Highest Team Payroll	$54.3(UTA)	$60.0(MIL)	$60.3(TOR)	$60.6(SAC)	$59.7(PHO)	$64.8(POR)	$71.3(LAL)
10th Highest Team Payroll	$53.7(DEN)	$59.5(MIN)	$59.1(BOS)	$59.5(MIA)	$59.7(MIA)	$63.8(NJN)	$71.2(PHO)
Teams Above Luxury Tax	-----	16/29	13/28	-----	5/30	6/30	11/30
Teams Above Salary Cap	27/29	29/29	25/28	24/30	24/30	24/30	24/30

[a] Payroll Data comes from the USA Today Salary Database. Salary cap and luxury tax information comes from the SportsCity website

Table 2.2: Summary Statistics by Income Tax Rate Quartile, weighted by contract length

Relative Income Tax Rate	Overall	(1)	(2)	(3)	(4)
Observations	744	209	163	186	186
Relative Income Tax Rate (%)	-0.03 (2.18)	-2.64 (0.13)	-0.61 (0.27)	0.68 (0.52)	2.69 (1.76)
Relative Property Tax Rate (%)	0.06 (0.68)	0.41 (0.67)	-0.09 (0.70)	0.12 (0.80)	-0.45 (1.11)
Relative Sales Tax Rate (%)	0.14 (1.28)	0.74 (0.25)	-0.61 (1.11)	-0.06 (0.89)	0.37 (1.89)
Contract Length (Unweighted)	2.38 (1.77)	2.31 (1.71)	2.44 (1.79)	2.42 (1.87)	2.34 (1.72)
Contract Amount	2.63e07 (3.02e07)	2.67e07 (3.22e07)	2.59e07 (2.82e07)	2.56e07 (2.83e07)	2.70e07 (3.20e07)
Log Contract Amount	16.02 (1.78)	15.92 (1.84)	16.10 (1.73)	16.06 (1.74)	16.02 (1.81)
Average Salary	5.25e06 (4.96e06)	5.32e06 (5.45e06)	5.28e06 (4.72e06)	5.04e06 (4.62e06)	5.38e06 (4.96e06)
Log Average Salary	14.91 (1.18)	14.85 (1.25)	14.97 (1.15)	14.92 (1.13)	14.93 (1.21)
Black	0.783 (0.412)	0.727 (0.447)	0.804 (0.398)	0.809 (0.394)	0.800 (0.401)
Center	0.181 (0.385)	0.154 (0.367)	0.188 (0.392)	0.176 (0.381)	0.204 (0.404)
Forward	0.396 (0.489)	0.424 (0.495)	0.362 (0.482)	0.420 (0.495)	0.369 (0.484)
Seasons	5.58 (3.75)	5.67 (4.14)	5.29 (3.51)	5.72 (3.50)	5.61 (3.81)
Games per Season	55.14 (23.11)	55.13 (25.37)	55.10 (22.55)	56.27 (20.08)	54.02 (23.12)
Minutes per game	22.11 (10.82)	22.87 (11.95)	21.46 (10.18)	22.31 (10.18)	21.65 (10.75)
Blocks per game	0.52 (0.63)	0.585 (0.753)	0.442 (0.477)	0.530 (0.662)	0.522 (0.565)
Steals per game	0.75 (0.48)	0.745 (0.469)	0.781 (0.502)	0.768 (0.515)	0.697 (0.445)
Points per game	9.12 (6.14)	9.79 (6.70)	8.52 (5.68)	9.07 (5.59)	8.95 (6.42)
Offensive Rebounds per Game	1.18 (0.89)	1.20 (0.94)	1.20 (0.84)	1.16 (0.88)	1.21 (0.91)
Defensive Rebounds per Game	2.76 (1.84)	2.89 (2.07)	2.67 (1.72)	2.72 (1.79)	2.75 (1.74)
Assists per Game	2.00 (1.79)	2.07 (1.79)	2.02 (1.84)	2.12 (1.97)	1.78 (1.50)
Turnovers per Game	1.38 (0.84)	1.46 (0.94)	1.35 (0.80)	1.38 (0.79)	1.32 (0.79)
Fouls Per Game	2.04 (0.91)	2.06 (0.99)	2.08 (0.88)	2.00 (0.85)	2.05 (0.92)
Missed FG per Game	4.13 (2.64)	4.30 (2.73)	3.98 (2.53)	4.17 (2.50)	4.05 (2.77)
Missed FT Percentage	0.240 (0.12)	0.231 (0.123)	0.239 (0.114)	0.242 (0.116)	0.250 (0.128)
Performance per Game	14.91 (0.97)	14.96 (1.07)	14.88 (0.91)	14.95 (0.91)	14.86 (0.95)
All-Star Appearances per Season	0.077 (0.193)	0.105 (0.239)	0.054 (0.132)	0.074 (0.181)	0.069 (0.192)
Contract Renewal	0.492 (0.500)	0.505 (0.501)	0.543 (0.500)	0.460 (0.500)	0.466 (0.500)
1/Draft Pick (0 if undrafted)	0.109 (0.199)	0.158 (0.264)	0.067 (0.107)	0.112 (0.213)	0.091 (0.146)

[a] All free agent statistics are career averages prior to contract signing. All statistics are weighted by contract length unless otherwise noted. Standard deviation is in parenthesis.

Figure 2.2: Relative State and Local Income Tax rates in the NBA, 2001-02 to 2007-08 Seasons

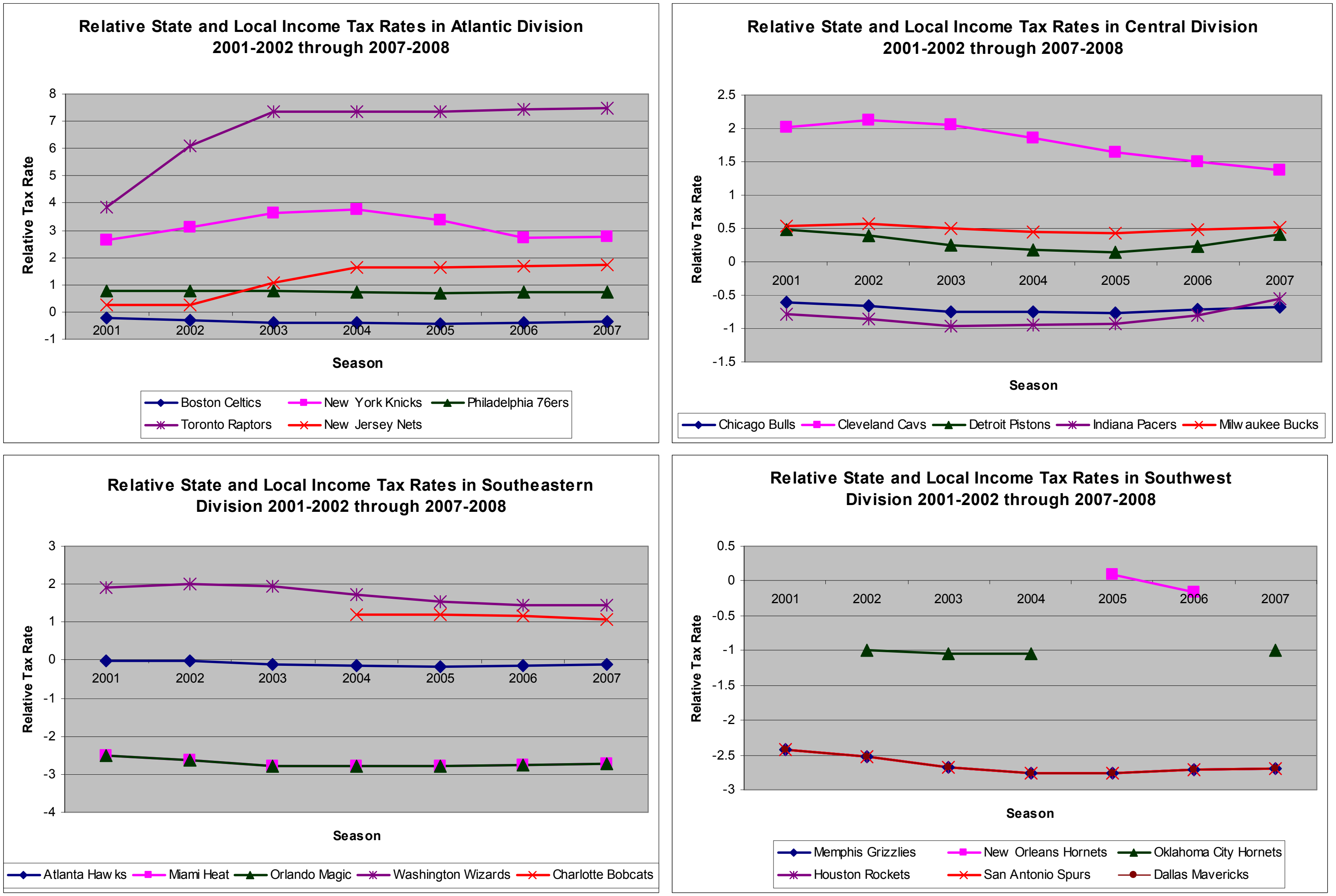
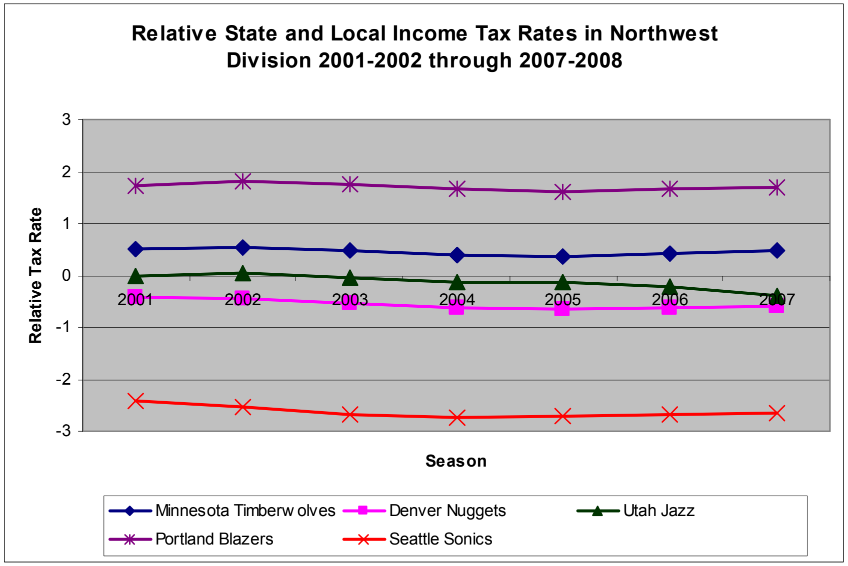


Figure 2.2 continued: Relative State and Local Income Tax rates in the NBA, 2001-02 to 2007-08 Seasons



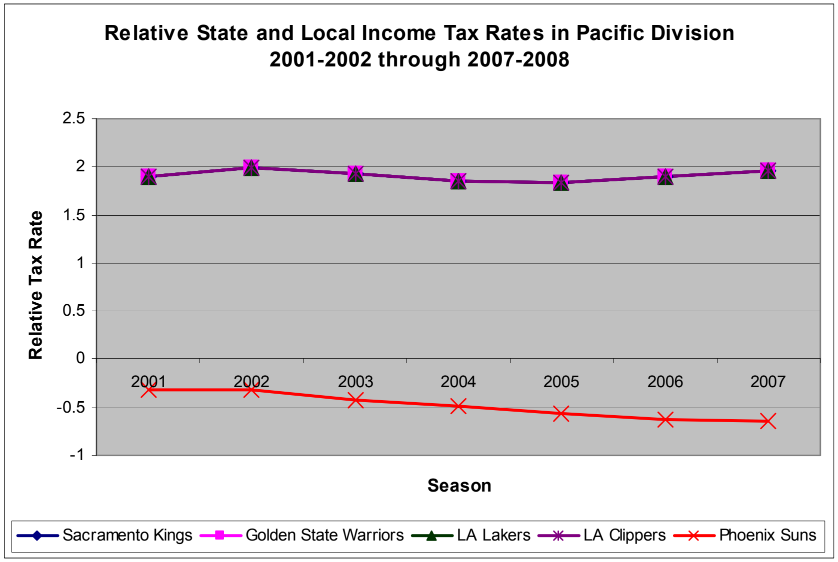
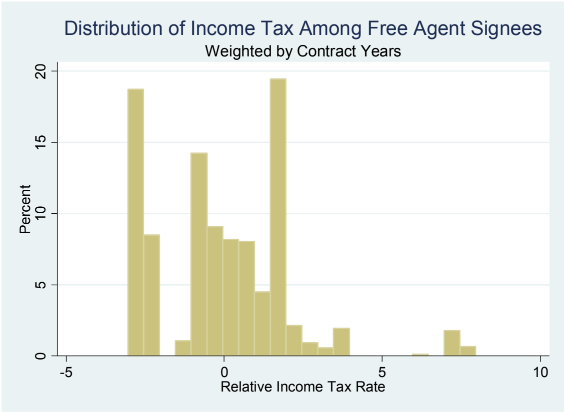
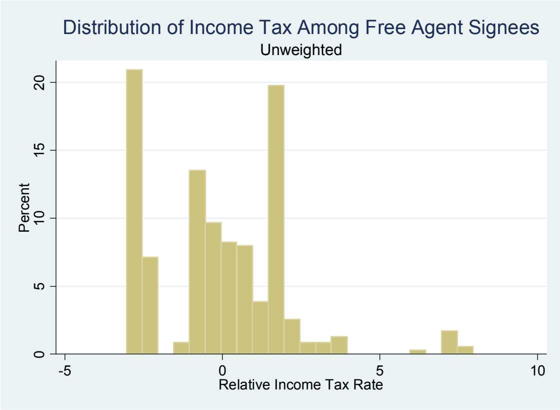


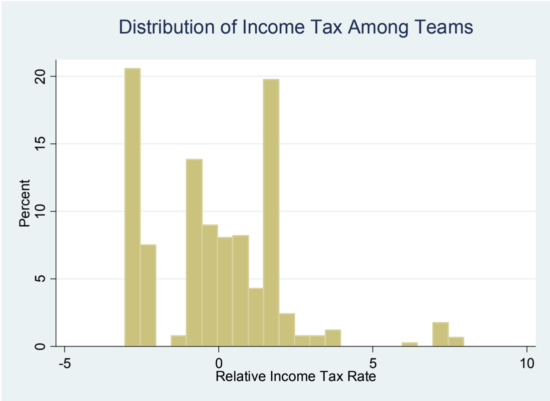
Figure 2.3: Distribution of Income Tax Rates in the NBA, 2001-02 to 2007-08 Seasons



(a)



(b)



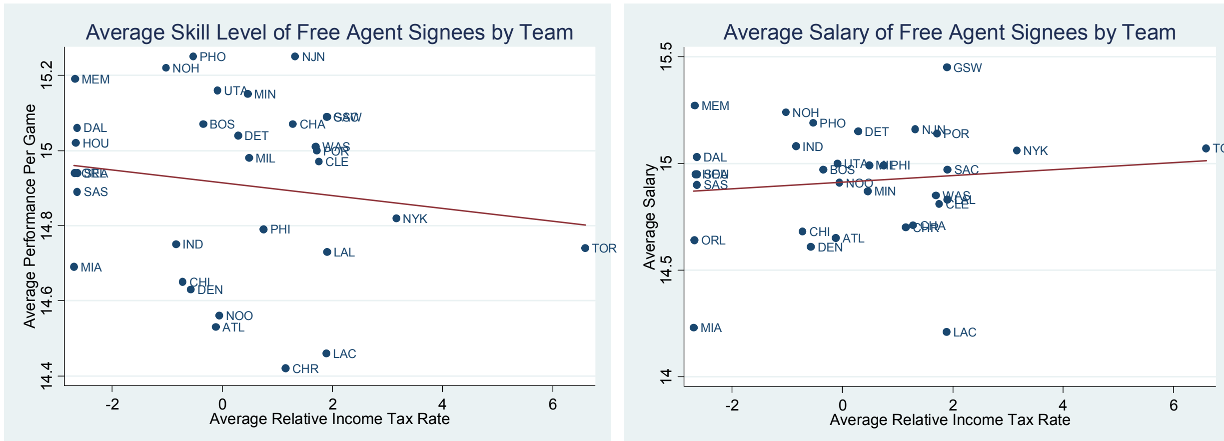
(c)

Table 2.3: Regression of the Logarithm of Yearly Salary Weighted by Contract Length

Explanatory Variable	(1)		(2)		(3)		(4)		(5)		(6)	
CONSTANT	13.152***	(0.075)	13.088***	(0.083)	13.208***	(0.089)	13.158***	(0.080)	13.111***	(0.088)	13.232**	(0.089)
GAMES PER SEASON	0.013***	(0.002)	0.013***	(0.002)	0.014***	(0.003)	0.010***	(0.002)	0.010***	(0.002)	0.013***	(0.003)
MINUTES PER GAME	0.024	(0.013)	0.029*	(0.013)	0.029*	(0.013)	0.061***	(0.006)	0.061***	(0.006)	0.055***	(0.006)
POINTS PER GAME	0.101***	(0.023)	0.100***	(0.023)	0.106***	(0.023)	-----		-----		-----	
BLOCKS PER GAME	0.122	(0.071)	0.075	(0.075)	0.060	(0.077)	-----		-----		-----	
STEALS PER GAME	-0.022	(0.107)	0.017	(0.109)	-0.030	(0.107)	-----		-----		-----	
OFFENSIVE REBS PER GAME	0.183**	(0.077)	0.160*	(0.079)	0.135	(0.079)	-----		-----		-----	
DEFENSIVE REBS PER GAME	0.086	(0.044)	0.084	(0.045)	0.078	(0.045)	-----		-----		-----	
ASSISTS PER GAME	0.142***	(0.037)	0.143***	(0.038)	0.142***	(0.038)	-----		-----		-----	
TURNOVERS PER GAME	-0.300***	(0.107)	-0.288***	(0.107)	-0.294***	(0.107)	-----		-----		-----	
FOULS PER GAME	0.032	(0.069)	-0.008	(0.071)	0.007	(0.072)	-----		-----		-----	
MISSED FG PER GAME	-0.114*	(0.052)	-0.114*	(0.051)	-0.128**	(0.051)	-----		-----		-----	
MISSED FT PERCENTAGE	-0.958***	(0.291)	-1.051***	(0.293)	-0.450	(0.343)	-1.067***	(0.376)	-1.055***	(0.375)	-0.677	(0.379)
BLOCKS PER 48 MINS	-----		-----		-----		0.104***	(0.036)	0.084*	(0.038)	0.086*	(0.038)
STEALS PER 48 MINS	-----		-----		-----		-0.122**	(0.049)	-0.108*	(0.050)	-0.059	(0.050)
POINTS PER 48 MINS	-----		-----		-----		0.039***	(0.009)	0.038***	(0.009)	0.043***	(0.009)
OFFENSIVE REBS PER 48 MINS	-----		-----		-----		0.089***	(0.033)	0.082*	(0.034)	0.075*	(0.033)
DEFENSIVE REBS PER 48 MINS	-----		-----		-----		0.048*	(0.022)	0.045*	(0.022)	0.060***	(0.022)
ASSISTS PER 48 MINS	-----		-----		-----		0.061	(0.018)	0.064***	(0.019)	0.077***	(0.019)
TURNOVERS PER 48 MINS	-----		-----		-----		-0.107**	(0.046)	-0.107**	(0.046)	-0.105*	(0.045)
FOULS PER 48 MINS	-----		-----		-----		-0.076***	(0.025)	-0.081***	(0.025)	-0.015	(0.029)
MISSED FG PER 48 MINS	-----		-----		-----		-0.050***	(0.017)	-0.044**	(0.018)	-0.019	(0.018)
EXPERIENCE	0.077***	(0.028)	0.080***	(0.028)	------		0.054	(0.028)	0.055*	(0.028)	-----	
EXPERIENCE SQUARED	-0.012***	(0.002)	-0.013	(0.002)	------		-0.011***	(0.002)	-0.011***	(0.002)	-----	
CENTER	-----		0.258*	(0.111)	0.221*	(0.109)	-----		0.193	(0.118)	0.119	(0.116)
FORWARD	-----		0.082	(0.077)	0.078	(0.076)	-----		0.063	(0.080)	0.052	(0.078)
Experience dummies	No		No		Yes		No		No		Yes	
Number of Observations	744		744		744		744		744		744	
R^2	0.665		0.667		0.688		0.664		0.666		0.690	

[a] ***=Significant at 0.01, **=Significant at 0.02, *=Significant at 0.05. Standard errors are in parenthesis

Figure 2.4: Average Skill Level and Salary of Free Agent Signees by Team



(a) (b)

Figure 2.5: Skill Distribution of Free Agents, 2001-2002 Season

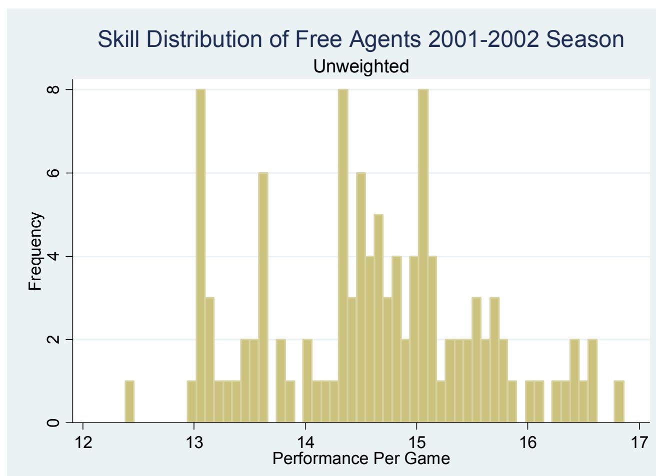


Table 2.4: Regression of the Standardized Market Skill Index on Relative Income Tax Rate, Weighted by Contract Length (Column 2, Table 3)

Explanatory Variable	(1)		(2)		(3)		(4)		(5)	
RELATIVE INCOME TAX RATE	-0.289***	[0.061]	-0.292***	[0.069]	-0.274***	[0.088]	-0.137***	[0.034]	-0.080**	[0.034]
RELATIVE PROPERTY TAX RATE	-----		-0.015	[0.257]	-0.293	[0.470]	-0.372	[0.301]	-0.316	[0.303]
RELATIVE SALES TAX RATE	-----		-0.204	[0.118]	-0.231	[0.166]	-0.096	[0.099]	-0.199	[0.107]
CENTER	-----		0.095	[0.148]	0.067	[0.161]	-0.096	[0.129]	-0.101	[0.110]
FORWARD	-----		-0.242***	[0.093]	0.228*	[0.117]	0.132	[0.122]	0.142	[0.137]
LOG LAG MSA POPULATION	-----		-----		-2.910	[2.166]	-3.265***	[0.879]	-2.343***	[0.467]
LOG LAG MSA EMPLOYMENT	-----		-----		1.528	[2.014]	0.798	[1.047]	0.662	[1.413]
LOG LAG MSA INCOME	-----		-----		-2.856**	[1.196]	-1.880	[1.730]	-0.542	[1.322]
100*[PPP-1] ($COUNTRY/$US)	-----		-----		0.122	[0.137]	0.137	[0.104]	0.178*	[0.081]
TEAM WINS LAST SEASON	-----		-----		0.015	[0.011]	0.010	[0.011]	0.012	[0.011]
LAG CRIME RATE (PER 100,000)	-----		-----		-.0004	[.0004]	-.0003	[.0002]	-.0002	[.0003]
LAG STUDENT-TEACHER RATIO	-----		-----		-0.104	[0.105]	-0.082	[0.078]	-0.075	[0.085]
CONTRACT RENEWAL	-----		-----		-----		0.953***	[0.101]	0.921***	[0.103]
PERCENT OF CAP AVAILABLE	-----		-----		-----		-----		0.009	[0.005]
R^2	0.076		0.089		0.110		0.309		0.325	

[a] ***=Significant at 0.01, **=Significant at 0.02, *=Significant at 0.05
[b] Each regression controls for YEAR and TEAM dummies. [c] Cluster robust standard errors on TEAM and YEAR in brackets.

Table 2.5: Regression of the Standardized Market Skill Index on Relative Income Tax Rate, Weighted by Contract Length (Column 5, Table 3)

Explanatory Variable	(1)		(2)		(3)		(4)		(5)	
RELATIVE INCOME TAX RATE	-0.281***	[0.065]	-0.286***	[0.072]	-0.282***	[0.089]	-0.147***	[0.045]	-0.088***	[0.030]
RELATIVE PROPERTY TAX RATE	-----		-0.025	[0.278]	-0.295	[0.463]	-0.373	[0.309]	-0.316	[0.328]
RELATIVE SALES TAX RATE	-----		-0.251*	[0.114]	-0.261	[0.151]	-0.127	[0.104]	-0.233*	[0.115]
CENTER	-----		0.095	[0.142]	0.067	[0.169]	-0.095	[0.134]	-0.099	[0.118]
FORWARD	-----		0.245***	[0.089]	0.231*	[0.110]	0.136	[0.108]	0.147	[0.120]
LOG LAG MSA POPULATION	-----		-----		-3.290	[2.259]	-3.642***	[1.072]	-2.700***	[0.727]
LOG LAG MSA EMPLOYMENT	-----		-----		2.012	[2.157]	1.290	[1.410]	1.150	[1.758]
LOG LAG MSA INCOME	-----		-----		-2.907**	[1.240]	-1.939	[1.791]	-0.572	[1.348]
100*[PPP-1] ($COUNTRY/$US)	-----		-----		0.097	[0.150]	0.112	[0.112]	0.153	[0.091]
TEAM WINS LAST SEASON	-----		-----		0.013	[0.011]	0.009	[0.010]	0.010	[0.009]
LAG CRIME RATE (PER 100,000)	-----		-----		-.0005	[.0004]	-.0003	[.0003]	-.0002	[.0003]
LAG STUDENT-TEACHER RATIO	-----		-----		-0.100	[0.112]	-0.077	[0.088]	-0.070	[0.094]
CONTRACT RENEWAL	-----		-----		-----		0.945***	[0.115]	0.912***	[0.118]
PERCENT OF CAP AVAILABLE	-----		-----		-----		-----		0.009***	[0.004]
R^2	0.079		0.092		0.110		0.307		0.324	

[a] ***=Significant at 0.01, **=Significant at 0.02, *=Significant at 0.05
[b] Each regression controls for YEAR and TEAM dummies. [c] Cluster robust standard errors on TEAM and YEAR in brackets.

Table 2.6: Regression of the Logarithm of Average Annual Salary on Relative Income Tax Rate, Weighted by Contract Length

Explanatory Variable	(1)		(2)		(3)		(4)	
REL INCOME TAX RATE	-0.016	[0.070]	-0.017	[0.063]	-0.009	[0.062]	-0.010	[0.062]
REL SALES TAX RATE	-0.086	[0.145]	-0.058	[0.143]	-0.050	[0.149]	-0.030	[0.152]
REL PROP TAX RATE	-0.010	[0.266]	-0.048	[0.261]	-0.030	[0.256]	-0.049	[0.257]
CENTER	0.249	[0.185]	-0.036	[0.099]	0.196	[0.171]	-0.038	[0.118]
FORWARD	0.127	[0.137]	0.008	[0.094]	0.107	[0.133]	0.005	[0.095]
PERFORMANCE PER GAME	-----		0.931***	[0.059]	-----		-----	
GAMES PER SEASON	0.012***	[0.003]	-----		0.009*	[0.004]	-----	
MINUTES PER GAME	0.021	[0.021]	-----		0.055***	[0.010]	-----	
POINTS PER GAME	0.113***	[0.038]	-----		-----		-----	
BLOCKS PER GAME	0.027	[0.111]	-----		-----		-----	
STEALS PER GAME	-0.022	[0.144]	-----		-----		-----	
OFFENSIVE REBS PER GAME	0.150	[0.105]	-----		-----		-----	
DEFENSIVE REBS PER GAME	0.074	[0.056]	-----		-----		-----	
ASSISTS PER GAME	0.151***	[0.056]	-----		-----		-----	
TURNOVERS PER GAME	-0.216	[0.210]	-----		-----		-----	
FOULS PER GAME	-0.020	[0.093]	-----		-----		-----	
MISSED FG PER GAME	-0.147***	[0.051]	-----		-----		-----	
MISSED FT PERCENTAGE	-1.000**	[0.405]	-----		-1.014**	[0.431]	-----	
PERFORMANCE PER 48 MINS	-----		-----		-----		0.916***	[0.065]
BLOCKS PER 48 MINS	-----		-----		0.060	[0.061]	-----	
STEALS PER 48 MINS	-----		-----		-0.148	[0.078]	-----	
POINTS PER 48 MINS	-----		-----		0.036*	[0.018]	-----	
OFF REBS PER 48 MINS	-----		-----		0.090***	[0.030]	-----	
DEF REBS PER 48 MINS	-----		-----		0.034	[0.034]	-----	
ASSISTS PER 48 MINS	-----		-----		0.073***	[0.028]	-----	
TURNOVERS PER 48 MINS	-----		-----		-0.095	[0.089]	-----	
FOULS PER 48 MINS	-----		-----		-0.078***	[0.035]	-----	
MISSED FG PER 48 MINS	-----		-----		-0.038	[0.026]	-----	
EXPERIENCE	0.0998	[0.054]	-----		0.067	[0.053]	-----	
EXPERIENCE SQUARED	-0.012***	[0.004]	-----		-0.010***	[0.004]	-----	
LOG LAG MSA POPULATION	-0.698	[2.524]	-0.916	[2.550]	-0.424	[2.424]	-0.613	[2.444]
LOG LAG MSA INCOME	-1.178	[1.787]	-1.351	[1.692]	-1.190	[1.532]	-1.328	[1.503]
LOG LAG MSA EMPLOYMENT	0.664	[2.260]	0.883	[2.022]	0.342	[2.096]	0.454	[1.994]
LAG CRIME RATE	-.0001	[.0002]	-.0002	[.0002]	-.0001	[.0003]	-.0001	[.0002]
LAG STUDENT-TEACHER	-0.030	[0.059]	-0.032	[0.063]	-0.037	[0.062]	-0.036	[0.066]
TEAM WINS LAST SEASON	-0.007	[0.006]	-0.006	[0.005]	-0.005	[0.006]	-0.004	[0.005]
100*[PPP-1] ($COUNTRY/$US)	-0.169***	[0.052]	-0.164***	[0.048]	-0.145***	[0.046]	-0.141	[0.047]
CONTRACT RENEWAL	0.350***	[0.110]	0.342***	[0.104]	0.365***	[0.103]	0.355	[0.102]
PERCENT CAP AVAILABLE	0.004	[0.003]	0.004	[0.003]	0.004	[0.003]	0.004	[0.003]
Number of Observations	744		744		744		744	
R^2	0.731		0.728		0.727		0.724	

[a] ***=Significant at 0.01, **=Significant at 0.02, *=Significant at 0.05
[b] Each regression controls for YEAR and TEAM dummies.
[c] Standard errors clustered on TEAM and YEAR in parenthesis.

Table 2.7: Regression of Different Measures of Skill on Relative Income Tax Rate, Weighted by Contract Length

Explanatory Variable	NBA EFFICIENCY PER GAME		NBA EFFICIENCY PER 48 MINUTES		GAME SCORE PER GAME		GAME SCORE PER 48 MINUTES		WIN SCORE PER GAME		WIN SCORE PER 48 MINUTES	
RELATIVE INCOME TAX RATE	-0.033	[0.050]	-0.016	[0.057]	-0.051	[0.045]	-0.045	[0.062]	-0.117**	[0.048]	-0.182***	[0.047]
RELATIVE PROPERTY TAX RATE	-0.450	[0.272]	-0.356	[0.305]	-0.486	[0.284]	-0.391	[0.322]	-0.413	[0.258]	-0.375	[0.269]
RELATIVE SALES TAX RATE	-0.005	[0.145]	0.124	[0.140]	0.041	[0.162]	0.178	[0.134]	-0.021	[0.135]	0.020	[0.123]
CENTER	-0.080	[0.114]	0.424***	[0.145]	-0.349***	[0.114]	-0.061	[0.140]	0.695***	[0.137]	1.186***	[0.132]
FORWARD	0.128	[0.132]	0.307***	[0.111]	-0.020	[0.134]	0.067	[0.130]	0.589***	[0.128]	0.795***	[0.101]
LOG LAG MSA POPULATION	-3.071	[2.221]	-3.333***	[1.096]	-3.579	[2.425]	-3.715***	[0.752]	-1.478	[2.332]	-1.411	[1.337]
LOG LAG MSA EMPLOYMENT	1.442	[1.674]	2.086	[1.241]	1.690	[1.563]	2.172***	[0.786]	1.090	[2.310]	1.165	[1.789]
LOG LAG MSA INCOME	-1.777	[1.637]	-2.169*	[1.002]	-1.614	[1.837]	-1.923	[1.339]	-1.172	[1.513]	-1.044	[1.011]
100*[PPP-1] ($COUNTRY/$US)	0.219***	[0.068]	0.259***	[0.059]	0.196***	[0.071]	0.197**	[0.087]	0.157***	[0.050]	0.130**	[0.051]
TEAM WINS LAST SEASON	0.017	[0.009]	0.013	[0.009]	0.016	[0.009]	0.013	[0.010]	0.013	[0.008]	0.007	[0.006]
LAG CRIME RATE (PER 100,000)	-.0002	[.0003]	.0001	[.0004]	-.0002	[.0003]	-.00004	[.0004]	-8.94e-6	[.0003]	.0002	[.0003]
LAG STUDENT-TEACHER RATIO	-0.137**	[0.057]	-0.079	[0.055]	-0.142***	[0.055]	-0.091*	[0.044]	-0.145	[0.075]	-0.117*	[0.064]
CONTRACT RENEWAL	0.789***	[0.149]	0.725***	[0.159]	0.789***	[0.147]	0.755***	[0.169]	0.659	[0.107]	0.548***	[0.123]
PERCENT OF CAP AVAILABLE	0.009	[0.005]	0.006	[0.004]	0.008*	[0.004]	0.007*	[0.003]	0.008	[0.004]	0.005	[0.005]
R^2	0.282		0.263		0.2677		0.228		0.327		0.402	

[a] ***=Significant at 0.01, **=Significant at 0.02, *=Significant at 0.05. [b] Each regression controls for YEAR and TEAM dummies.
[c] Cluster robust standard errors on TEAM and YEAR in brackets[11]. [d] All dependent variables are standardized.

Table 2.8: Regression of Team Outcomes on Relative Income Tax Rate, Weighted by Percentage of Free Agents on Each Team

Explanatory Variable	REG SEASON WINS (1)		REG SEASON WINS (2)		REG SEASON WINS (3)		PLAYOFF WINS (4)		PLAYOFF WINS (5)		PLAYOFF WINS (6)	
RELATIVE INCOME TAX RATE	-3.977*	[2.221]	-3.847*	[2.340]	-4.074*	[2.164]	-2.290	[1.415]	-2.251	[1.465]	-2.422	[1.478]
RELATIVE PROPERTY TAX RATE	-----		-45.572	[324.3]	-391.148	[366.5]	-----		-40.673	[188.9]	-118.344	[255.1]
RELATIVE SALES TAX RATE	-----		597.347	[0.114]	290.557	[658.2]	-----		260.584	[219.1]	152.170	[186.4]
LOG LAG MSA POPULATION	-----		-----		-27.747	[37.07]	-----		-----		-37.481	[14.31]
LOG LAG MSA EMPLOYMENT	-----		-----		20.944	[50.25]	-----		-----		28.730	[14.47]
LOG LAG MSA INCOME	-----		-----		13.330	[47.04]	-----		-----		-10.141	[18.18]
100*[PPP-1] ($COUNTRY/$US)	-----		-----		-458.823	[298.0]	-----		-----		-102.007	[109.4]
TEAM WINS LAST SEASON	-----		-----		0.233*	[0.122]	-----		-----		0.046	[0.056]
LAG CRIME RATE (PER 100,000)	-----		-----		-0.002	[0.009]	-----		-----		-.0005	[0.002]
LAG STUDENT-TEACHER RATIO	-----		-----		-0.744	[2.200]	-----		-----		0.531	[1.297]
PERCENT OF CAP AVAILABLE	-----		-----		-0.023	[0.134]	-----		-----		-0.003	[0.037]
Number of Observations	206		206		206		206		206		206	
R^2	0.449		0.453		0.484		0.437		0.443		0.469	

[a] ***=Significant at 0.01, **=Significant at 0.05, *=Significant at 0.10
[b] Each regression controls for YEAR and TEAM dummies. [c] Cluster robust standard errors on TEAM and YEAR in brackets.

Chapter 2: References

2005 Collective Bargaining Agreement. (2009, December 9). Retrieved from http://www.nbpa.org/cba/2005

Angrist, Joshua D., & Pischke, Jörn-Ste¤en. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press. 221-244.

Berri, D. J., Schmidt, M. B., & Brook, S. L. (2006). *The wages of wins: Taking measure of the many myths in modern sport.* Palo Alto, CA: Stanford University Press, 85-138.

Berri, D.J. (2006, November 17). A comment on the player efficiency rating. Retrieved from http://dberri.wordpress.com/2006/11/17/a-comment-on-the-player-efficiency-rating/

Cameron, A.C., Gelbach, J.G., and D.L. Miller (2006). Robust Inference with Multi-way Clustering. Working Paper 09-8, Department of Economics, University of California – Davis.

Clark, David E. and William J. Hunter (1992). The Impact of Economic Opportunity, Amenities and Fiscal Factors on Age-Specific Migration Rates. *Journal of Regional Science, 32*, 349-365.

Cornwell, C., Trumbull, W.N. (1994). Estimating the economic model of crime with panel data. *Review of Economics and Statistics* 76, 360-366.

Department of the Treasury, Internal Revenue Service. (2010). *Foreign tax credit for individuals* (Publication 514). Washington, DC: Government Printing Office. Retrieved from http://www.irs.gov/pub/irs-pdf/p514.pdf

Egger, Peter, & Radulescu, Doina M. (2009). The Influence of Labour Taxes on the Migration of Skilled Workers. *The World Economy*, *32*(9), 5-32.

Feld, L. and G. Kirchgassner (2001). Income Tax Competition at the State and Local Level in Switzerland. *Regional Science and Urban Economics, 31*, 2–3, 181–213.

Figlio, David N., and O'Sullivan, Arthur (2001). "The Local Response to Tax Limitation Measures: Do Local Governments Manipulate Voters to Increase Revenues?" *Journal of Law and Economics* 44, 233–257.

Garcia, A. (2010, November 18). *Cba 101: owners, players brace for negotiations*. Retrieved from http://www.nba.com/2010/news/11/17/labor-primer/index.html

Gillis, K. (2010, July 9). *Lebronomics: could high taxes influence james' team decision?*. Retrieved from http://www.mrc.org/bmi/articles/2010/LeBronomics_Could_High_Taxes_Influence_James_Team_Decision.html

Heath, T., & Crenshaw, A.B. (2003, February 28). The jock tax': how pro athletes are being taxed by the states. *The Washington Post*. Retrieved from http://community.seattletimes.nwsource.com/archive/?date=20030228&slug=jocktax02b

Hoffman, D.K., & Hodge, S.A. Tax Foundation, (2004). *Nonresident state and local income taxes in the united states: the continuing spread of "jock taxes"* (ISSN 1068-0306). Washington, DC: Tax Foundation. Retrieved from http://www.taxfoundation.org/files/fe9737a23595921c41616df11dc3cd48.pdf

Kirchgassner, G. and W. W. Pommerehne (1996). Tax Harmonization and Tax Competition in the European Union: Evidence from Switzerland. *Journal of Public Economics*, 60(3), 351–71.

Loftin, C. and McDowall, D. (1982).The Police, Crime, and Economic Theory. *American Sociological Review,* 47, 393-401.

Merchak, A. (2010, July 8). *State jock taxes: is lebron better off in miami?*. Retrieved from http://www.taxfoundation.org/news/show/26503.html

Reilly, R.F., & Schweihs, R.P. (2004). *The handbook of business valuation and intellectual property analysis*. New York, NY: McGraw-Hill.

Thompson, Samuel Brodsky. Simple Formulas for Standard Errors that Cluster by Both Firm and Time (2009). Available at SSRN: http://ssrn.com/abstract=914002

Wallace, Sally (1993). The Effects of State Personal Income Tax Differentials on Wages. *Regional Science and Urban Economics, 23*, 611-628.

Wallace, S. (2002). The Effect of State Income Tax Structure on Interstate Migration. Fiscal Research Center Working Paper #79, Andrew Young School of Policy Studies, Georgia State University, Atlanta, GA, 21-22.

CHAPTER 3

HOW THE NUTRITIONAL CONTENT OF SCHOOL PROVIDED MEALS AFFECTS STUDENT OUTCOMES

Joint with Mirinda L. Martin

3.1. Introduction

There has recently been an emphasis on decreasing childhood obesity and increasing the health of schoolchildren in the United States. Since many students may consume as many as half their daily calories at school (The White House, 2010), increasing the nutritional content of school provided meals is one of the potential mechanisms recommended for achieving these goals. In 2004, a federal mandate was created that required all school districts participating in the National School Lunch Program to implement a local wellness program by July 2006.

The Buffalo Public School District implemented the *Healthier Options for Public Schoolchildren* (*HOPS*) program in some of their schools during the fall of 2007, after seeing positive results from a similar study implemented in the School District of Osceola County, Florida.[39] *HOPS* increased the nutritional content of the food provided at school during breakfast, lunch, and extended-day snack times. It was expected that *HOPS* would improve health and wellbeing as measured by decreased weight and BMI, as it did in the School District of Osceola County. However, programs like *HOPS* can have other benefits for schoolchildren, such as improved academic outcomes

There are several mechanisms through which improvements in nutrition may serve to enhance students' academic and behavioral outcomes. Sorhaindo & Feinstein (2006) posit several possible explanations: (1) students may be able to learn better, (2) students may be healthier and thus have fewer absences, and (3) students' behavior may improve, causing fewer disruptions in the classroom.

Several studies show that improvements in nutrition can enhance cognitive ability, observed intelligence, and concentration among school-aged children. For example, Politt (1993) shows that iron deficiency, even in early stages, can decrease dopamine transmission, thus negatively impacting cognition. Deficiencies in other vitamins and minerals, specifically thiamine, vitamin E, vitamin B, iodine, and zinc, are shown to inhibit cognitive abilities and mental concentration

[39] See Hollar et al (2010).

(Chenoweth, 2007; Greenbaum, 2007a; Greenbaum, 2007b; Bryan, Osendarp, Hughes, Calvaresi, Baghurst, & van Klinken, 2004; Delange, 2000; Sandstead, 2000). Additionally, amino acid and carbohydrate supplementation can improve perception, intuition, and reasoning (Lieberman, 2003; Frisvold, 2012). There are also a number of studies showing that improvements in nutrient intake can influence the cognitive ability and observed intelligence levels of school-aged children (Benton & Roberts, 1988; Schoenthaler, Amos, Doraz, Kelly, & Wakefield, 1991; Benton & Buts, 1990; Nelson, 1992; Eysenck & Schoenthaler, 1997). Lambert, Agostoni, Elmadfa, Hulsof, Krause, & Livingstone et al. (2004) observe that nutrition can also impact cognitive development in younger children.

If improvements in nutrition make students healthier, then students are likely to have fewer absences and attend class more regularly. The result is that: (1) students will have more time in class, and (2) students will have fewer interruptions in learning over the course of the school year. Additionally, students' behavior may improve and cause fewer disruptions in the classroom, creating a better learning environment for each student in the class. Kleinman, Murphy, Little, Pagano, Wehler, & Regal et al. (1998) focus on student behavior, and show that malnutrition leads to behavior problems. Jones, Borg, Boulware, McCarthy, Sherwin, & Tamborlane (1995) examine the effects of sugar on healthy children, and find that sugar has a negative impact on child behavior; glucose consumption induces "hormonal, symptomatic, and neurophysiologic changes in healthy children." However, these effects are not present for children who consume either a balanced diet or mixed meals that include protein, fat, complex carbohydrates, and fiber; these limit reductions in plasma glucose levels (Jones et al., 1995).

In this paper, we test how the implementation of the *HOPS* program affected students' academic achievement, attendance, and behavior. To estimate the causal effect of the *HOPS* intervention on these outcomes, we estimate a difference-in-differences model using a unique and comprehensive restricted-use dataset from the Buffalo Public School District for the 2003-2004 through 2008-2009 school years. In particular, the Buffalo Public School District is the ideal setting to study this particular issue because over 75% of the students in the district receive free or reduced-price lunch, so the likelihood that students actually eat school lunch in this school district is high.

We find that students in schools that implemented *HOPS* had a statistically significant increase in standardized math test scores compared to students in untreated schools, particularly among low ability students, higher income students, and females. However, inclusion of the students' demographic characteristics attenuates the estimated impact of the *HOPS* invention. Additionally, we find that the intervention had no statistically significant impact on standardized English test scores, or on behavioral outcomes such as attendance or disciplinary suspensions. Although *HOPS* was originally initiated as a way to improve students' health, interventions like *HOPS* may also improve academic outcomes, particularly standardized math test scores, especially among lower ability students, females, and higher income students.

3.2 Background Studies

Advocates of child health have experimented with students' diets in the United States for more than twenty years. Initial studies focused on improving the health of students. However, in more recent years, sociologists and a few economists have looked more closely at the impact of a student's diet and nutrition on academic and behavioral outcomes. For instance, Florence, Asbridge, & Veugelers (2008) find that a higher quality diet is associated with better performance on exams. Programs focused on increasing students' health also show modest improvements in students' academic test scores. Meyers, Sampson, Wietzman, Rogers, & Kayne (1989) find that participation in the School Breakfast Program among low-income children who qualified led to higher test scores and lower tardiness and absence rates.[40] Other studies find that improving the quality of students' diets led to students being on task more often, increased math test scores, possibly increased reading test scores, and increased attendance (Powell, Walker, Chang, & Grantham-McGregor, 1998; Cueto, 2001; Storey, Pearce, Ashfield-Watt, Wood, Baines, & Nelson, 2011; Hollar, Messiah, Lopez-Mitnik, Hollar, Almon, & Agatston, 2010). Additionally, a study by Price (2012) shows that eliminating the sale of soft drinks in vending machines in schools had a positive effect on behavioral outcomes such as tardiness and disciplinary referrals.

Several studies indicate that improving the nutritional content of meals matters much more for schoolchildren who are malnourished, or for those who are economically disadvantaged (Simeon & Grantham-McGregor, 1989; Cueto, Jacoby, & Pollitt, 1998; Hollar et al., 2010). Likely, this is because the economically disadvantaged have experienced a direct impediment to their cognitive abilities due to the lack of nourishment; once their primary physical needs are

[40] See also Kleinman et al (2002).

satisfied, students are then better able to concentrate on learning and thinking through problems (Pertz & Putnam, 1982). Since over 85% of the students in the Buffalo Public School District qualify for free or reduced price meals at school, we should expect a program that focuses on increasing the nutritional content of meals to provide positive effects on students' academic outcomes, as well as health outcomes.

While this study focuses on the long term effect of increasing the nutritional content of school meals, there have been a few studies that look directly at increasing nutritional content and/or calories in the very short term with the intention of having an immediate effect, such as on exam days. Figlio & Winicki (2005) find evidence that schools in Virginia have strategically increased the amount of calories served to students in their school lunches on exam days.[41] In their study, they find evidence that students who received an average of 100 more calories on exam days as compared to similar students performed better on their exams in both math and reading, and the effect was greater for those schools that increased the offered calories the most within the acceptable guidelines. In another study, Simeon and Grantham-McGregor (1989) find that students who received breakfast prior to taking a series of tests performed better than those who only had a cup of hot tea. Cueto (2001) summarizes several studies focused on the effects of fasting or skipping breakfast on student's performance, and shows that fasting does have a negative effect on students' performance, specifically when it comes to memory and visual perception. Correspondingly, Imberman & Kugler (2012), using quasi-random timing in its introduction, find that a program that serves breakfast to students directly in the classroom increased students' math and reading test scores.

There are also a couple of studies that examine, with a more long term focus, the effects of programs that increase the nutritional content of school meals. Belot & James (2011) look at the *Feed Me Better* program initiated in a borough near London, England. They find that the increased nutritional content of school meals led to an improvement in academic outcomes at the individual level, and a decrease in absences at the school level. They also find that the implementation of the program did not decrease take-up rates of school lunch at the schools. Though able to provide credible evidence of their results, our dataset enables us to look at absences at the individual level and control for the effect of individual teachers on the

[41] The National School Lunch Program offers considerable flexibility in the amount of calories served daily provided schools conform to certain nutritional guidelines over the course of the week.

improvement in academic outcomes.

Another study analyzes the initial introduction of the *Healthier Options for Public Schoolchildren* program among those students that received free or reduced lunch in the School District of Osceola County, Florida. Hollar et al. (2010) find reductions in the BMI of students in *HOPS* schools, though these improvements disappeared over the three month summer break from school. They also find evidence of improvements in math and reading scores, although the improvements in reading scores were not statistically significant (Hollar et al., 2010). However, their study was comprised of only four treatment schools and one comparison school, and the results of their analysis rest heavily on the particular control school that they chose. In this study we focus on the effects of the *HOPS* program in the Buffalo Public School District using data from the entire universe of students in grades three through eight.

Due to data limitations, many of the aforementioned studies do not control for lags in student test scores and none control for the influence of individual teachers. We are able to include exam scores prior to the initiation of the *HOPS* program and observe the changes in exam scores for each student in the years before and after the implementation of the program. Additionally, we include teacher fixed effects. Changes in test scores may be highly influenced by individual teachers, and failure to include teacher fixed effects may bias the results if the quality of teachers at treated and untreated schools changes over time.[42] While we are able to include teacher fixed effects, we find ex-post that their inclusion has little effect on our results as it pertains to this particular study.

3.3 The HOPS Program

The goal of the *HOPS* program was to improve the health and nutrition of students in the Buffalo Public School District. The program had three components: increasing the nutritional content of the meals offered at school, teaching the students more about nutrition and focusing on a particular fruit or vegetable each month, and sending home information to students' families encouraging a healthier lifestyle at home. The additional cost of the program was less than $1 per student per month. The Buffalo Public School District increased the nutritional content of school meals by incorporating or substituting foods that were low in fat (particularly foods low in saturated fat and zero trans fats), high in fiber, and contained little added sugar. They included

[42] See Rockoff (2004).

many more fruits and vegetables, and often obtained them from local farmers. *HOPS* schools replaced white bread with whole grain breads, white rice with brown rice, white potatoes with sweet potatoes, and frozen fruits and vegetables with fresh ones. Additionally, these schools incorporated low-fat dairy and cheeses and whole grain sugar cookies into their meals, used whole grain breading for the chicken and whole grain crust for the pizza, and cut back on the amount of chocolate milk served.

The Buffalo Public School District emphasized nutrition education for students and their families through the use of "*HOPS* Foods of the Month", where two foods were chosen each month to focus on. For instance, in September, the Foods of the Month were Whole Grains and Tomatoes. Additionally, schools educated students about the importance of physical activity and incorporated more physical activity during the school day.

While the comparison schools did not receive the *HOPS* program, *HOPS* did have an indirect effect on the nutritional content of food offered at these schools. The comparison schools had access to whole grain breads because all schools in the district used the same bread distributor. The comparison schools also replaced white potatoes with sweet potatoes at times, although not with the frequency of the schools that participated in the program. Therefore, the comparison schools may be partially contaminated with the treatment. Contamination of the comparison schools only happened with regards to the nutritional content of school meals; comparison schools did not adapt any other aspects of the program, such as physical activity or nutrition education. Since the control schools may be contaminated, our estimates of the effect of the *HOPS* program likely understate the true impact of the program.

3.4 Data Description

In order to examine the impact that nutritional improvements in school lunches have on students' academic and behavioral outcomes, we use administrative records from the Buffalo Public School District (BPSD), a school district located in a large urban area that is comprised of more than sixty public schools. This is a unique and detailed restricted-use longitudinal dataset of all students in kindergarten through eighth grade in the Buffalo Public School District between the 2003-2004 and 2008-2009 school years. We are able to observe each student's attendance, recorded disciplinary incidents, teacher, grade, school, and other demographic characteristics: gender, race, free/reduced lunch status, English language proficiency, and disability status.

Additionally, we are able to observe the New York State Assessment Test scores for Mathematics and English Language Arts (ELA) for students in the fourth through eighth grades between the 2005-2006 school year and the 2008-2009 school year.[43] [44]

There are twelve schools in the BPSD that received the *HOPS* intervention during the 2007-2008 and 2008-2009 school years.[45] For our purposes, we will let all other elementary schools in the Buffalo Public School District serve as our comparison group. Table 3.1 demonstrates how the students in *HOPS* schools compare to students in the comparison schools both before and after the start of the *HOPS* program. In the *HOPS* treated schools, there is a slightly greater share of students who are male, disabled, and economically disadvantaged, and a smaller share of students who are white or English proficient. These students also have lower standardized math and English test scores. From looking at the summary statistics, it appears that both math and English test scores are trending down over time for students in the *HOPS* schools, and trending up for students in the comparison schools. It also seems like students in the *HOPS* schools are becoming both absent and suspended more often after the initiation of the *HOPS* program. This is also true of the comparison schools, but to a much smaller extent. Figure 3.1, which shows the trends in standardized math test scores, standardized English test scores, the proportion of days present, and the proportion of days not suspended, echoes these findings.

However, Figure 3.1 also shows that test scores for students in *HOPS* schools seem to be trending downward, while test scores for students in comparison schools seem to be trending upward, even before the introduction of *HOPS.* After the introduction of *HOPS* the decline in test scores seems to be muted for those students in *HOPS* schools, particularly in the first year. From Figure 3.1, it also appears that *HOPS* students go from less likely to be absent and less likely to be suspended to more likely to be absent and more likely to be suspended. This would be difficult to explain based on our understanding of *HOPS.* However, the trends are also similar for those students in the comparison schools.

Figure 3.2 shows the trends in math and English test scores, the proportion of days present, and the proportion of days not suspended, after controlling for demographics characteristics and

[43] We standardize math and English test scores for each student, in each grade, using the mean and standard deviation of each score specific to that student's grade and cohort.
[44] We also have test scores for fourth- and eighth-graders during the 2003-2004 and 2004-2005 school years, which we use to construct lagged test scores for students whose scores were not observed in the 2005-2006 school year.
[45] The schools that received the *HOPS* program in the BPSD are PS #3, #6, #18, #19, #32, #37, #45, #53, #54, #74, #94, and #95.

cohort, grade, school, and teacher fixed effects. The trends for suspensions do not change much, while those for absences show trends that are similar but more pronounced than those in Figure 3.1 for both types of schools. In the 2007-2008 school year, it seems that the *HOPS* program is helping students bring up their math and English test scores. However, most of the gains in math and English scores seem to disappear in the 2008-2009 school year. This may be because the funding for the *HOPS* program was suddenly pulled out early in 2009, around January or February, and students do not take the exams until the end of the spring semester. While some schools tried to keep up with the tenets of the program, there was no longer any external impetus from the funding foundation or district administration; some of the nutritional changes to menus remained, but not most of the ones that distinguished the *HOPS* schools from the comparison schools. It is possible that the effects of the nutrition program had short run effects which lasted only a month or two longer than the actual impetus of the program.[46]

3.4.1 Selection

The schools that participated in the *HOPS* program were not chosen at random. Participation in *HOPS* was voluntary and was determined at a district meeting of school principals. Since test scores for students in *HOPS* schools seem to be trending downward prior to the implementation of *HOPS* while comparison schools seem to be trending upward, an important problem that we face is the possibility that principals self-selected into the program in order to address this downward trend in test scores. However, at the time it seems unlikely that principals realized that nutritional improvements in school lunches would have an impact on test scores. Their decision to implement the program was likely driven by a desire to improve the health and nutritional wellbeing of their schoolchildren. Furthermore, those schools that were willing to volunteer for the program are likely the relevant group of schools for which to estimate the effects of this policy, since these types of policies are generally adopted voluntarily.

Since we are using a difference-in-differences identification strategy, it is particularly important that schools with outcomes that are trending downward are not deliberately self-selecting in order to reverse the trend. To determine if this was the case, we estimate a probit regression of *HOPS* treatment status at the school level on each outcome of interest in the year

[46] Hollar et al (2010) find that the improvements in health and BMI found during the school year were largely undone during the summer months away from school, when the students returned to their typical dietary patterns.

before the intervention was implemented, the trend of each outcome of interest prior to the implementation of the treatment, and school-wide demographic characteristics of students.[47] The marginal effects from these probit regressions are shown in Table 3.2. Most importantly, we find that there are no statistically significant relationships between the treatment and trends in standardized math test scores, standardized English test scores, the proportion of days present, and the proportion of days not suspended. In some of our specifications, we find that *HOPS* schools were more likely to have a greater number of black or other race minority students (Native American, Pacific Islander, Asian) and fewer suspension days per student.[48] [49]

3.5 Methodology

In this paper, we estimate the causal impact of the *Healthier Options for Public Schoolchildren (HOPS)* program on cognitive and non-cognitive outcomes of elementary and middle school students in the Buffalo Public School District in the 2007-2008 and 2008-2009 academic years. The cognitive outcomes of interest include standardized scores from New York State assessment tests in math and English. The non-cognitive outcomes of interest are attendance and disciplinary suspensions.

To estimate the causal effect that the *HOPS* program had on these outcomes, we employ a difference-in-differences model with a single lagged outcome from a year prior to the introduction of the program. Recent evidence has shown that models of this type perform much better than models of gains per year or student fixed effects models in the presence of mean reversion, since models of gains per year or student fixed effects do not allow for decay in the past outcome.[50] All regressions use the lagged outcome from the 2005-2006 academic year.[51]

Formally, let

---

[47] The outcomes of interest are school-wide averages in standardized math test scores, standardized English test scores, the proportion of days not absent, and the proportion of days not suspended, from the 2006-2007 school year. The trend in each outcome is the school-level difference in the 2006-2007 and 2005-2006 school years.

[48] The fraction of black students is statistically significant in column (6), which included only demographic characteristics. The proportion of days not suspended was statistically significant only in column (7), the full specification. The fraction of other race students is statistically significant in columns (6) and (7).

[49] The other race category only comprises about five percent of all students. The two schools with the highest shares of other race students (24% and 31%) participated in the *HOPS* intervention. Without these schools, the distribution of other race minority students is similar between school types.

[50] See Rothstein (2010).

[51] If data from the 2005-2006 school year was not available for a given student, then the lagged outcome from the 2004-2005 school year was used, if available. If data from the 2004-2005 and 2005-2006 school years were not available for a given student, then the lagged outcome from the 2003-2004 school year was used, if available.

$$y_{i,g,t,s,c} = \beta_1 HOPS_s + \beta_3(HOPS_s * POST_t) + \rho y_{i,g',2005,s',c'} +$$
$$\tau_t + \sigma_g + \alpha_{g,t} + \delta_s + \gamma_c + Z\theta_i + \varphi_g * \omega_t * y_{i,g',2005,s',c'} + \varepsilon_{i,g,t,s,c} \quad (1),$$

where $y_{i,g,t,s,c}$ is the outcome variable of interest for student *i*, in grade *g,* in year *t*, in school *s,* with teacher *c,* and $y_{i,g',2005,s',c'}$ is the lagged value of this outcome. $HOPS_s$ is a school-specific indicator variable equal to 1 if a student attends a school that participates in the *HOPS* program, and equal to 0 if a student attends a school that does not participate in the *HOPS* program. $POST_t$ is a time-specific dummy variable equal to 0 in the period before the treatment, the 2006-2007 school year, and equal to 1 in the period after the treatment, the 2007-2008 and 2008-2009 school years. τ_t, σ_g, $\alpha_{g,t}$, δ_s, and γ_c are year, grade, cohort, school, and teacher fixed effects, respectively, and θ_i is a vector of observable characteristics for each student. Additionally, we include $\varphi_g * \omega_t * y_{i,g',2005,s',c'}$, interactions between year and grade fixed effects and the lagged value of the outcome variable, to account for the fact that the lagged outcome may impact the present outcome differentially depending on the grade level of the student and the number of years that have passed since the lagged outcome was observed.

3.6 Results

We present the results of the *HOPS* program on students' standardized math test scores, standardized English test scores, proportion of days present, and proportion of days not suspended in Tables 3.3, 3.4, 3.5, and 3.6, respectively. Table 3.3 shows the impact that the *HOPS* program had on standardized math test scores. The simple difference-in-differences estimate in column (1), which includes a lag in the standardized math test score, shows that the *HOPS* program increased standardized math scores by 0.072 standard deviations, and this result is statistically significant at the 10% level. Inclusion of year, grade, cohort, school, and teacher fixed effects, shown in columns (2)-(6), do not substantially change the point estimate of the impact of the program on standardized math test scores or the standard errors associated with this point estimate. In column (7), inclusion of demographic characteristics tends to attenuate the estimated impact of the *HOPS* program. Although the estimate in column (7) is not statistically significant at the 10% level, it is statistically significant at the 15% level.[52] In column (8), we

[52] In this study, the 15% level may be acceptable since there are only 48 schools, 12 of which are treated.

include interactions between year and grade fixed effects and the 2005 math test score, since the 2005 math test score may impact a student's current math score differently depending on their current year and grade. Once we include year-grade-2005 test score interactions, we no longer have the power to measure a statistically significant impact of *HOPS*. However, the coefficient estimate provides suggestive evidence that there is a positive effect of the program on students' math test scores.[53] [54]

Table 3.4 shows the impact that the *HOPS* program had on standardized English test scores. The simple difference-in-differences estimate in column (1), which includes a lag in the standardized English test score, shows that the *HOPS* program increased standardized English test scores by 0.011 standard deviations, although this result is not statistically significant. Again, inclusion of year, grade, cohort, and school fixed effects, shown in columns (2)-(5), do not substantially change the point estimate of the impact of the program on standardized English test scores or the standard errors associated with this point estimate; none of these estimates are statistically significant. Inclusion of teacher fixed effects in column (6) further attenuates the estimate of the effect of *HOPS* on English test scores, and this point estimate becomes negative in columns (7) and (8) when demographic characteristics are included. While the *HOPS* program had a positive impact on students' standardized math test scores, these estimates provide no evidence that *HOPS* had any impact on students' standardized English test scores. Indeed, most studies show that education policies tend to have a much greater impact on math test scores than reading scores. According to Rothstein (2002), "one likely cause is that students learn math mostly in school, while literacy also comes from habits at home…scores would suffer if students did less out-of-school reading or had a less literate home environment."

Table 3.5 shows the impact that the *HOPS* program had on the proportion of days that students were present. The hypothesis is that if students are made healthier by *HOPS,* then they

[53] Some may have concerns about students switching schools in order to receive or avoid treatment of the *HOPS* intervention. Though it is unlikely that students would actually move or apply for a boundary exception solely to take advantage of or avoid this program, we check to see if it makes a difference. We use the school the student attended the year before the *HOPS* intervention was implemented as an instrument for receiving the treatment, but find no substantive difference in our results.

[54]One potential problem is that we have standardized students' test scores based on the overall mean and standard deviation in each grade and cohort, including *HOPS* treated students. If the *HOPS* treatment has an effect on test scores, then allowing the *HOPS* treatment to affect the overall mean and standard deviation can bias our results. To account for this possible bias, we re-standardize both math and English test scores using the mean and standard deviation of the comparison group only. Re-standardizing by the mean and standard deviation of the comparison group alone has no substantive effect on the results.

will have fewer illnesses and be able to attend school more often. However, depending on the specification, Table 3.5 shows that the *HOPS* program caused students to attend school between 0.61%-0.71% less, approximately one day less each year. However, these results are not statistically significant. Since we do not have information on why students are absent, it may be the case that students miss school for reasons other than illness. Table 3.6 shows the impact that the *HOPS* program had on the proportion of days that students were not suspended. Depending on the specification, students are suspended about one- to two-thirds of a day more per year on average when the *HOPS* program is in effect, although these estimates are not statistically significant. While we believed a priori that *HOPS* could impact the ability for students to learn might come about by reducing absences and improving behavior, we do not find evidence of these mechanisms. This suggests that math scores increase mainly because improvements in nutrition enhance intuition, reasoning, and concentration, but not because students are present more often or are suspended less; our results may also be driven by minor improvements in behavior not captured by observed suspensions.

3.6.1 Heterogeneous Treatment Effects

There is a large amount of evidence that health policies may have a greater effect in economically advantaged children (Perry, Bishop, Taylor, Murray, Mays, & Dudovitz et al., 1998; Müller, Danielzik, & Pusta, 2005; Plachta-Danielzik, Pust, Asbeck, Czerwinski-Mast, Langnase, & Fischer et al., 2007; Belot & James, 2011). Müller et al. (2005) suggest that this is because health policies targeted toward children are most effective when children are treated alongside their parents; in particular, because *HOPS* schools sent pamphlets about healthy eating home with students, children in high-socioeconomic status families are more likely to have healthy eating habits translated into their home environment by their parents. Studies also show that short-term impacts of health policies are stronger among girls than boys, possibly because boys are more reluctant to accept changes in their diet, such as increased intake of vegetables, or less likely to make better food choices despite an increased knowledge of proper nutritional practices (Kelder, Perry, Lytle, & Klepp, 1995; Perry et al., 1998; Müller et al., 2005; Belot & James, 2011). There is also reason to believe that low-ability children will experience greater academic gains from health interventions like *HOPS* if low-ability children are those most in need of improvements in nutrient intake (Schoenthaler & Bier, 1999). We examine the heterogeneous effects of the *HOPS* initiative based on ability, socioeconomic status, and gender.

3.6.1.1 High Ability vs. Low Ability Students

In an effort to compare students who are more similar, we split students by ability based on each student's lagged test score with the hypothesis being that low ability students are more in need of the nutrients provided by the higher nutritional content of *HOPS* meals. This allows us to directly compare low ability students in *HOPS* schools to low ability students in comparison schools, and high ability students in *HOPS* schools to high ability students in comparison schools. A student is grouped as a high ability math student if his/her lagged standardized math test score is greater than or equal to zero, and a student is grouped as a low ability math student if his/her lagged standardized math test score is less than zero.

Table 3.7 shows the impact that the *HOPS* program had on standardized math test scores for high ability and low ability math students. For low ability math students, columns (1)-(8) show that the *HOPS* program increased standardized math scores by 0.076-0.094 standard deviations. The result is statistically significant at the 10% level in columns (1)-(7) and at the 15% level in column (8). For high ability math students, the *HOPS* program also increased their standardized math test scores, though by less than the increase for low ability students. However, the effect of the program for high ability math students, and the difference between the effects for high ability and low ability math students, are both statistically insignificant.[55] No statistically significant effects can be found for standardized English test scores, regardless of ability level.

3.6.1.2 Economically Advantaged vs. Economically Disadvantaged

Table 3.8 shows the effect of *HOPS* on standardized math test scores for economically advantaged and economically disadvantaged students, where economically disadvantaged means that the student qualified for free or reduced price lunches and economically advantaged means that the student did not. *HOPS* had very large effects for economically advantaged students; based on columns (3)-(8), which include grade fixed effects, the *HOPS* initiative increased math test scores by between 0.117-0.150 standard deviations, and these results are statistically

[55] We find that the effect of *HOPS* on high ability students is statistically significant at the 10% level in column (6), when year, grade, cohort, school, and teacher fixed effects are controlled for but demographic characteristics are not.

significant at the 10% level.[56] For economically disadvantaged students, although the magnitude of the effect is positive, *HOPS* had no statistically significant effect on standardized math test scores. However, the difference of the effect for economically advantaged vs. economically disadvantaged students is also statistically insignificant. For standardized English test scores, we find no statistically significant effects, regardless of socioeconomic status.

Economically disadvantaged students have a higher probability of consuming school meals since they receive free and reduced-price lunches. These students are also more likely to suffer from poor eating habits. However, as mentioned earlier, several studies indicate that programs specifically targeted toward decreasing obesity in children have a larger effect for students from higher socioeconomic groups, so it is not surprising that we find that this extends to academic outcomes as well (Muller et al, 2005; Plachta et al, 2007).

3.6.1.3 Male vs. Female

We examine the different effects of the program on both female and male students. The effect of programs similar to *HOPS* that combine teaching students about healthy eating, fruits, and vegetables with the opportunity to consume more nutritious foods for school meals show that female students are much more likely to incorporate what they learn and actually consume more of the healthier food, particularly vegetables (Perry et al., 1998). Table 3.9 reveals the impact of *HOPS* when we allow for different effects of the program on males and females. *HOPS* increased standardized math test scores among females by 0.090-0.098 standard deviations in columns (1)-(5), and the effect is statistically significant at the 10% level. However, in columns (6)-(8), which include teacher fixed effects, the effect of *HOPS* on standardized math test scores for females is not statistically significant, although the size of the coefficients still provide some suggestive evidence of a positive impact on math scores among females. For males, the effect is positive, although generally not statistically significant. However, the difference between the effect for males and females is also not statistically significant. Similar to our other results, we find no statistically significant effects for standardized English test scores, regardless of gender.

[56] The effect is statistically significant at the 5% level in columns (5)-(7), when school fixed effects are controlled for but grade-year-2005 math score interactions are not.

3.7 Conclusion

In this paper, we analyze the effects of the *Healthier Options for Public Schoolchildren* (*HOPS*) program on cognitive and non-cognitive outcomes for students in the Buffalo Public School District. We use a difference-in-differences approach to compare the changes in outcomes of students that attended schools participating in the *HOPS* program with those that attended non-participating schools. We find that students in schools with the *HOPS* program had a statistically significant increase in standardized math scores. However, inclusion of demographic characteristics tends to attenuate the estimated impact of the *HOPS* program and render the point estimates statistically insignificant. The *HOPS* program had no statistically significant impact on standardized English test scores, attendance, or suspensions. Since students are not absent or suspended less often, the evidence suggests that math scores increase largely due to improvements in cognition and concentration, although we cannot rule out minor improvements in behavior not captured by observed suspensions.

We then examine the existence of potentially heterogeneous treatment effects based on ability, socioeconomic status, and gender. We split students into high ability and low ability groups based on each student's lagged test score, and find that the *HOPS* program increased standardized math test scores for low ability math students in our full specification, and this result is generally statistically significant at the 10% level. However, the effect of the program for high ability math students, and the difference between the effects for high ability and low ability math students, are both statistically insignificant. For economically advantaged students, *HOPS* had very large statistically significant effects on standardized math test scores. For economically disadvantaged students the effects are much smaller and not statistically significant. Nevertheless, the difference of the effect for economically advantaged vs. economically disadvantaged students is also statistically insignificant. Similarly, we find larger effects of *HOPS* among females; however, the effect is not statistically different from the effect for males. Furthermore, we find no effect of the *HOPS* program on standardized English test scores, regardless of initial ability, socioeconomic status, or gender.

Our results imply that, although *HOPS* was originally initiated as a way to improve students' health, interventions like *HOPS* may also improve academic outcomes such as standardized math test scores. These effects may be larger for particular groups of students: lower ability students, higher income students, and females. The *HOPS* program cost less than an additional $1 per student per month. When considering whether to implement such a program, schools may consider the potential benefits that programs such as *HOPS* might have on academic outcomes such as test scores in addition to the intended health benefits associated with them.

Table 3.1: Summary Statistics by Year and Treatment Status

	Pre-*HOPS*			Post-HOPS		
	HOPS Untreated	*HOPS* Treated	Difference	*HOPS* Untreated	*HOPS* Treated	Difference
Number Student-Year Observations	11141	4374		21691	8186	
Male	0.5089	0.5194	-0.0105	0.4983	0.5214	-0.0231
	(0.4999)	(0.4997)	[0.0089]	(0.5000)	(0.4996)	[0.0065]
Disabled	0.2444	0.2526	-0.0082	0.2316	0.2363	-0.0046
	(0.4298)	(0.4346)	[0.0077]	(0.4219)	(0.4248)	[0.0055]
Economically Disadvantaged	0.7745	0.8455	-0.0709	0.8371	0.9037	-0.0666
	(0.4179)	(0.3615)	[0.0067]	(0.3692)	(0.2950)	[0.0041]
Black	0.5759	0.5947	-0.0188	0.5642	0.5828	-0.0186
	(0.4942)	(0.4910)	[0.0088]	(0.4959)	(0.4931)	[0.0064]
Hispanic	0.1447	0.1861	-0.0414	0.1524	0.1945	-0.0421
	(0.3518)	(0.3892)	[0.0068]	(0.3594)	(0.3958)	[0.0050]
Other Race	0.0211	0.0629	-0.0418	0.0217	0.0900	-0.0683
	(0.1437)	(0.2428)	[0.0039]	(0.1458)	(0.2862)	[0.0033]
Not English Proficient	0.0682	0.1511	-0.0829	0.0662	0.1713	-0.1051
	(0.2521)	(0.3582)	[0.0059]	(0.2486)	(0.3768)	[0.0045]
Standardized Math Score	0.0550	-0.1393	0.1943	0.0705	-0.1862	0.2567
	(1.0037)	(0.9763)	[0.0177]	(0.9927)	(0.9948)	[0.0130]
Standardized ELA Score	0.0599	-0.1535	0.2133	0.0706	-0.1936	0.2642
	(1.0018)	(0.9782)	[0.0179]	(1.0001)	(0.9731)	[0.0131]
Proportion of Days Present	0.9167	0.9213	-0.0046	0.9127	0.9086	0.0041
	(0.0806)	(0.0739)	[0.0014]	(0.0829)	(0.0790)	[0.0010]
Proportion of Days Not Suspended	0.9973	0.9980	-0.0007	0.9878	0.9858	0.0020
	(0.0077)	(0.0053)	[0.0001]	(0.0481)	(0.0469)	[0.0006]
Lag Standardized Math Score	0.0355	-0.1121	0.1476	0.0155	-0.1931	0.2087
	(1.0064)	(0.9700)	[0.0200]	(1.0041)	(0.9978)	[0.0185]
Lag Standardized ELA Score	0.0372	-0.1374	0.1746	0.0155	-0.1945	0.2100
	(1.0187)	(0.9385)	[0.0205]	(1.0098)	(0.9515)	[0.0188]
Lag Proportion of Days Present	0.9295	0.9254	0.0041	0.9278	0.9215	0.0063
	(0.0638)	(0.0626)	[0.0012]	(0.0619)	(0.0648)	[0.0009]
Lag Proportion of Days Not Suspended	0.9937	0.9929	0.0008	0.9965	0.9952	0.0012
	(0.0212)	(0.0210)	[0.0004]	(0.0153)	(0.0175)	[0.0002]

[a] Standard deviation in parenthesis. The difference reported is the mean of the untreated group minus the mean of the treated group. Standard error from a two sample t-test with unequal variance between groups is reported in brackets.

[b] Pre-*HOPS* period is the 2006-2007 school year. Post-*HOPS* period is the 2007-2008 and 2008-2009 school years.

Table 3.2: Marginal Effects from Probit Regressions Describing Selection into the HOPS Program

HOPS Treatment Status	(1)	(2)	(3)	(4)	(5)	(6)	(7)
2006-2007 Math Score	-0.179				-0.0405		-0.0531
	(0.127)				(0.524)		(0.427)
Math Score Trend	-0.172				-0.658		-0.253
	(0.338)				(0.559)		(0.587)
2006-2007 ELA Score		-0.146			-0.188		-0.268
		(0.124)			(0.455)		(0.525)
ELA Score Trend		0.0617			-0.145		-0.691
		(0.460)			(0.741)		(0.890)
2006-2007 Proportion of Days Present			-2.282		-3.244		1.097
			(2.272)		(3.596)		(3.267)
Proportion of Days Present Trend			4.811		5.267		2.650
			(3.862)		(4.754)		(3.903)
2006-2007 Proportion of Days Not Suspended				18.48	47.66		74.24**
				(29.37)	(32.35)		(30.91)
Proportion of Days Not Suspended Trend				-3.244	-9.917		6.778
				(2.790)	(6.080)		(13.29)
Economically Disadvantaged						0.200	0.500
						(0.207)	(0.594)
Black						0.568*	0.432
						(0.290)	(0.487)
Hispanic						0.537	0.463
						(0.856)	(1.052)
Other Race						4.623*	4.833*
						(2.384)	(2.681)
Not English Proficient						0.0749	-0.142
						(0.959)	(1.183)
Disabled						-0.482	-1.356
						(0.388)	(1.320)
Male						0.611	-0.253
						(1.213)	(1.956)
Observations	47	47	47	47	47	47	47

[a]*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses.

[b]Each trend is the average of each student's difference between the value in the 2006-2007 school year and value in the 2005-2006 school year for each school. If the value from the 2005-2006 school year is missing, then a value from the first previous school year is substituted.

Table 3.3: The Impact of *HOPS* on Standardized Math Test Scores

Standardized Math Scores	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HOPS Treated*Post-Treatment	0.0723*	0.0722*	0.0765*	0.0696*	0.0746*	0.0786*	0.0659	0.0577
	(0.0424)	(0.0425)	(0.0430)	(0.0414)	(0.0414)	(0.0432)	(0.0427)	(0.0405)
Math Score from 2005-06	0.680***	0.680***	0.686***	0.691***	0.653***	0.591***	0.558***	0.487***
	(0.0207)	(0.0207)	(0.0205)	(0.0198)	(0.0175)	(0.0162)	(0.0143)	(0.0361)
Economically Disadvantaged							-0.0542***	-0.0514***
							(0.0168)	(0.0164)
Black							-0.142***	-0.141***
							(0.0139)	(0.0138)
Hispanic							-0.0771***	-0.0752***
							(0.0229)	(0.0233)
Other Minority							0.0221	0.0251
							(0.0275)	(0.0283)
Not English Proficient							-0.102**	-0.109***
							(0.0383)	(0.0369)
Disabled							-0.188***	-0.186***
							(0.0219)	(0.0217)
Male							0.0102	0.0100
							(0.0140)	(0.0140)
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade Fixed Effects			Yes	Yes	Yes	Yes	Yes	Yes
Cohort Fixed Effects				Yes	Yes	Yes	Yes	Yes
School Fixed Effects					Yes	Yes	Yes	Yes
Teacher Fixed Effects						Yes	Yes	Yes
Year-Grade-Lag Math Score Interaction								Yes
Observations	26723	26723	26723	26723	26723	26723	26723	26723
R-squared	0.490	0.490	0.494	0.497	0.519	0.602	0.610	0.613

[a]*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered on school in parentheses.

[b]All specifications include an indicator variable for HOPS treatment and year fixed effects.

Table 3.4: The Impact of *HOPS* on Standardized English Test Scores

Standardized English Scores	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HOPS Treated*Post-Treatment	0.0113	0.0113	0.0126	0.0100	0.0176	0.000363	-0.0127	-0.0287
	(0.0295)	(0.0295)	(0.0298)	(0.0293)	(0.0287)	(0.0361)	(0.0368)	(0.0325)
ELA Score from 2005-06	0.683***	0.683***	0.687***	0.691***	0.645***	0.565***	0.525***	0.474***
	(0.0212)	(0.0212)	(0.0210)	(0.0207)	(0.0126)	(0.00975)	(0.00945)	(0.0284)
Economically Disadvantaged							-0.0716***	-0.0708***
							(0.0129)	(0.0130)
Black							-0.108***	-0.110***
							(0.0174)	(0.0173)
Hispanic							-0.0325	-0.0338
							(0.0226)	(0.0227)
Other Minority							-0.0402	-0.0400
							(0.0317)	(0.0319)
Not English Proficient							-0.185***	-0.189***
							(0.0419)	(0.0416)
Disabled							-0.184***	-0.185***
							(0.0166)	(0.0167)
Male							-0.0256**	-0.0259**
							(0.00992)	(0.00989)
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade Fixed Effects			Yes	Yes	Yes	Yes	Yes	Yes
Cohort Fixed Effects				Yes	Yes	Yes	Yes	Yes
School Fixed Effects					Yes	Yes	Yes	Yes
Teacher Fixed Effects						Yes	Yes	Yes
Year-Grade-Lag ELA Score Interaction								Yes
Observations	25111	25111	25111	25111	25111	25111	25111	25111
R-squared	0.526	0.526	0.528	0.530	0.546	0.595	0.602	0.605

[a]*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered on school in parentheses.

[b]All specifications include an indicator variable for HOPS treatment and year fixed effects.

Table 3.5: The Impact of *HOPS* on the Proportion of Days Present

Proportion of Days Present	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HOPS Treated*Post-Treatment	-0.00613	-0.00611	-0.00673	-0.00702	-0.00696	-0.00660	-0.00685	-0.00709
	(0.00576)	(0.00575)	(0.00596)	(0.00587)	(0.00640)	(0.00636)	(0.00630)	(0.00626)
Prop. Days Present 2005-06	0.650***	0.650***	0.668***	0.671***	0.633***	0.629***	0.623***	0.737***
	(0.0236)	(0.0236)	(0.0249)	(0.0254)	(0.0190)	(0.0180)	(0.0180)	(0.0547)
Economically Disadvantaged							-0.00808***	-0.00789***
							(0.00128)	(0.00127)
Black							0.00648***	0.00678***
							(0.00128)	(0.00127)
Hispanic							-0.00439***	-0.00444***
							(0.00162)	(0.00163)
Other Minority							0.00310	0.00304
							(0.00339)	(0.00325)
Not English Proficient							0.00243	0.00277
							(0.00332)	(0.00334)
Disabled							-0.00604***	-0.00593***
							(0.000777)	(0.000832)
Male							-0.000713	-0.000508
							(0.000916)	(0.000927)
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade Fixed Effects			Yes	Yes	Yes	Yes	Yes	Yes
Cohort Fixed Effects				Yes	Yes	Yes	Yes	Yes
School Fixed Effects					Yes	Yes	Yes	Yes
Teacher Fixed Effects						Yes	Yes	Yes
Year-Grade-Lag Prop. Present Interaction								Yes
Observations	47640	47640	47640	47639	47639	47639	47639	47639
R-squared	0.281	0.281	0.301	0.303	0.351	0.381	0.384	0.395

[a]*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered on school in parentheses.

[b]All specifications include an indicator variable for HOPS treatment and year fixed effects.

Table 3.6: The Impact of *HOPS* on the Proportion of Days Not Suspended

Proportion of Days Not Suspended	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HOPS Treated*Post-Treatment	-0.00273	-0.00284	-0.00306	-0.00337	-0.00324	-0.00240	-0.00239	-0.00173
	(0.00225)	(0.00225)	(0.00223)	(0.00210)	(0.00206)	(0.00144)	(0.00143)	(0.00122)
Prop. Days Not Suspended 2005-06	0.423***	0.432***	0.399***	0.412***	0.404***	0.365***	0.349***	1.074***
	(0.0625)	(0.0629)	(0.0624)	(0.0620)	(0.0603)	(0.0577)	(0.0568)	(0.263)
Economically Disadvantaged							-0.00142**	-0.00134**
							(0.000560)	(0.000535)
Black							-0.00402***	-0.00410***
							(0.000880)	(0.000881)
Hispanic							-0.00174**	-0.00183**
							(0.000733)	(0.000738)
Other Minority							6.91e-05	0.000315
							(0.00151)	(0.00147)
Not English Proficient							0.000170	0.000345
							(0.000662)	(0.000632)
Disabled							-0.000947**	-0.000826**
							(0.000396)	(0.000402)
Male							-0.00316***	-0.00295***
							(0.000392)	(0.000370)
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade Fixed Effects			Yes	Yes	Yes	Yes	Yes	Yes
Cohort Fixed Effects				Yes	Yes	Yes	Yes	Yes
School Fixed Effects					Yes	Yes	Yes	Yes
Teacher Fixed Effects						Yes	Yes	Yes
Year-Grade-Lag Prop. Not Sus. Interaction								Yes
Observations	47880	47880	47880	47879	47879	47879	47879	47879
R-squared	0.054	0.074	0.081	0.091	0.102	0.175	0.179	0.208

[a]*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered on school in parentheses.

[b]All specifications include an indicator variable for HOPS treatment and year fixed effects.

Table 3.7: The Impact of the *HOPS* Program on Math Test Scores, by Math Ability

Standardized Math Scores	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HOPS Treated*Post-Treatment	0.0788*	0.0787*	0.0834*	0.0761*	0.0826*	0.0936*	0.0833*	0.0770
	(0.0462)	(0.0462)	(0.0470)	(0.0453)	(0.0468)	(0.0500)	(0.0487)	(0.0480)
HOPS Treated*Post-Treatment*High Math	-0.0368	-0.0368	-0.0390	-0.0362	-0.0392	-0.0462	-0.0502	-0.0464
	(0.0436)	(0.0437)	(0.0441)	(0.0430)	(0.0432)	(0.0417)	(0.0395)	(0.0388)
Post-Treatment*High Math	-0.0910***	-0.0910***	-0.104***	-0.0969***	-0.0981***	-0.0972***	-0.0965***	0.00351
	(0.0281)	(0.0281)	(0.0286)	(0.0286)	(0.0272)	(0.0221)	(0.0216)	(0.0295)
High Math	0.252***	0.252***	0.259***	0.257***	0.245***	0.242***	0.232***	0.171***
	(0.0339)	(0.0339)	(0.0333)	(0.0314)	(0.0247)	(0.0239)	(0.0240)	(0.0252)
Math Score from 2005-06	0.615***	0.615***	0.621***	0.625***	0.587***	0.523***	0.494***	0.424***
	(0.0255)	(0.0255)	(0.0251)	(0.0243)	(0.0207)	(0.0201)	(0.0182)	(0.0373)
Economically Disadvantaged							-0.0505***	-0.0485***
							(0.0174)	(0.0169)
Black							-0.140***	-0.139***
							(0.0136)	(0.0136)
Hispanic							-0.0725***	-0.0712***
							(0.0228)	(0.0233)
Other Minority							0.0255	0.0273
							(0.0276)	(0.0283)
Not English Proficient							-0.110***	-0.115***
							(0.0362)	(0.0352)
Disabled							-0.183***	-0.181***
							(0.0224)	(0.0222)
Male							0.00927	0.00899
							(0.0139)	(0.0139)
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade Fixed Effects			Yes	Yes	Yes	Yes	Yes	Yes
Cohort Fixed Effects				Yes	Yes	Yes	Yes	Yes
School Fixed Effects					Yes	Yes	Yes	Yes
Teacher Fixed Effects						Yes	Yes	Yes
Year-Grade-Lag Math Score Interaction								Yes
F-Test: HOPS Treated*Post-Treatment + HOPS Treated*Post-Treatment*High Math	1.68	1.67	1.9	1.61	2.05	2.84*	1.79	1.63
Observations	26723	26723	26723	26723	26723	26723	26723	26723
R-squared	0.495	0.495	0.499	0.502	0.523	0.606	0.614	0.615

[a]*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered on school in parentheses.

[b]All specifications include an indicator variable for HOPS treatment, year fixed effects, and an interaction between High Math Ability and HOPS treatment.

Table 3.8: The Impact of the *HOPS* Program on Math Test Scores, by Socioeconomic Status

Standardized Math Scores	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HOPS Treated*Post-Treatment	0.120	0.120	0.137*	0.129*	0.150**	0.138**	0.127**	0.117*
	(0.0718)	(0.0718)	(0.0729)	(0.0736)	(0.0630)	(0.0599)	(0.0585)	(0.0587)
HOPS Treated*Post*Economically Disadvantaged	-0.0660	-0.0660	-0.0817	-0.0794	-0.0977	-0.0746	-0.0734	-0.0676
	(0.0798)	(0.0798)	(0.0798)	(0.0799)	(0.0713)	(0.0699)	(0.0676)	(0.0689)
Post-Treatment*Economically Disadvantaged	0.0689*	0.0686*	0.0802**	0.0738**	0.0921***	0.0433*	0.0512**	0.0134
	(0.0348)	(0.0347)	(0.0350)	(0.0346)	(0.0265)	(0.0221)	(0.0215)	(0.0267)
Economically Disadvantaged	-0.215***	-0.215***	-0.223***	-0.219***	-0.159***	-0.106***	-0.0866***	-0.0627***
	(0.0495)	(0.0495)	(0.0499)	(0.0476)	(0.0314)	(0.0184)	(0.0178)	(0.0184)
Math Score from 2005-06	0.665***	0.665***	0.670***	0.676***	0.649***	0.588***	0.558***	0.487***
	(0.0174)	(0.0174)	(0.0172)	(0.0166)	(0.0169)	(0.0156)	(0.0142)	(0.0362)
Black							-0.142***	-0.141***
							(0.0140)	(0.0140)
Hispanic							-0.0774***	-0.0751***
							(0.0231)	(0.0234)
Other Minority							0.0222	0.0253
							(0.0274)	(0.0279)
Not English Proficient							-0.102**	-0.109***
							(0.0384)	(0.0370)
Disabled							-0.188***	-0.186***
							(0.0219)	(0.0217)
Male							0.0100	0.00993
							(0.0140)	(0.0140)
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade Fixed Effects			Yes	Yes	Yes	Yes	Yes	Yes
Cohort Fixed Effects				Yes	Yes	Yes	Yes	Yes
School Fixed Effects					Yes	Yes	Yes	Yes
Teacher Fixed Effects						Yes	Yes	Yes
Year-Grade-Lag Math Score Interaction								Yes
F-Test: HOPS Treated*Post-Treatment + HOPS Treated*Post*Economically Disadvantaged	0.83	0.82	0.92	0.77	0.95	1.23	0.63	0.57
Observations	26723	26723	26723	26723	26723	26723	26723	26723
R-squared	0.494	0.494	0.498	0.501	0.521	0.603	0.610	0.613

[a]*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered on school in parentheses.

[b]All specifications include an indicator variable for HOPS treatment, year fixed effects, and an interaction between Economically Disadvantaged and HOPS treatment.

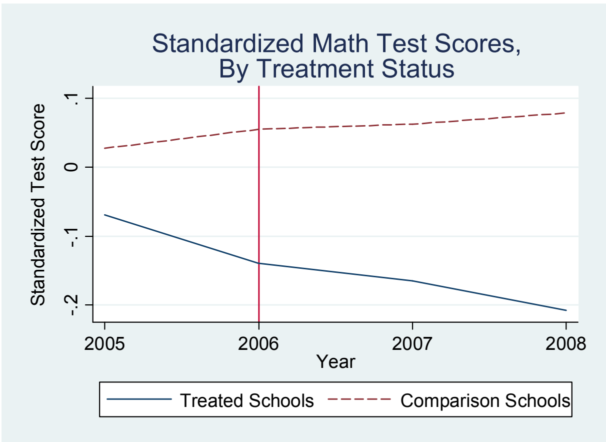
Table 3.9: The Impact of the *HOPS* Program on Math Test Scores, by Gender

Standardized Math Scores	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HOPS Treated*Post-Treatment	0.0937*	0.0935*	0.0975*	0.0898*	0.0924*	0.0866	0.0763	0.0650
	(0.0515)	(0.0516)	(0.0522)	(0.0503)	(0.0501)	(0.0517)	(0.0500)	(0.0477)
HOPS Treated*Post-Treatment*Male	-0.0373	-0.0373	-0.0366	-0.0350	-0.0300	-0.0130	-0.0176	-0.0118
	(0.0415)	(0.0415)	(0.0422)	(0.0419)	(0.0423)	(0.0386)	(0.0376)	(0.0364)
Post-Treatment*Male	-0.0335**	-0.0335**	-0.0296*	-0.0316**	-0.0393**	-0.0484***	-0.0456***	-0.0426***
	(0.0156)	(0.0158)	(0.0154)	(0.0148)	(0.0155)	(0.0123)	(0.0126)	(0.0130)
Male	-0.0294	-0.0294	-0.0332*	-0.0345**	-0.0286	0.0205	0.0360*	0.0345*
	(0.0177)	(0.0177)	(0.0175)	(0.0171)	(0.0172)	(0.0181)	(0.0191)	(0.0188)
Math Score from 2005-06	0.681***	0.681***	0.686***	0.692***	0.653***	0.591***	0.558***	0.487***
	(0.0207)	(0.0207)	(0.0204)	(0.0198)	(0.0174)	(0.0162)	(0.0143)	(0.0362)
Economically Disadvantaged							-0.0546***	-0.0516***
							(0.0168)	(0.0164)
Black							-0.142***	-0.141***
							(0.0139)	(0.0139)
Hispanic							-0.0773***	-0.0754***
							(0.0229)	(0.0233)
Other Minority							0.0223	0.0252
							(0.0275)	(0.0283)
Not English Proficient							-0.102**	-0.109***
							(0.0382)	(0.0368)
Disabled							-0.188***	-0.186***
							(0.0219)	(0.0217)
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade Fixed Effects			Yes	Yes	Yes	Yes	Yes	Yes
Cohort Fixed Effects				Yes	Yes	Yes	Yes	Yes
School Fixed Effects					Yes	Yes	Yes	Yes
Teacher Fixed Effects						Yes	Yes	Yes
Year-Grade-Lag Math Score Interaction								Yes
F-Test: HOPS Treated*Post-Treatment + HOPS Treated*Post-Treatment*Male	1.47	1.46	1.56	1.33	1.43	1.84	1.36	1.23
Observations	26723	26723	26723	26723	26723	26723	26723	26723
R-squared	0.491	0.491	0.495	0.498	0.520	0.602	0.61	0.613

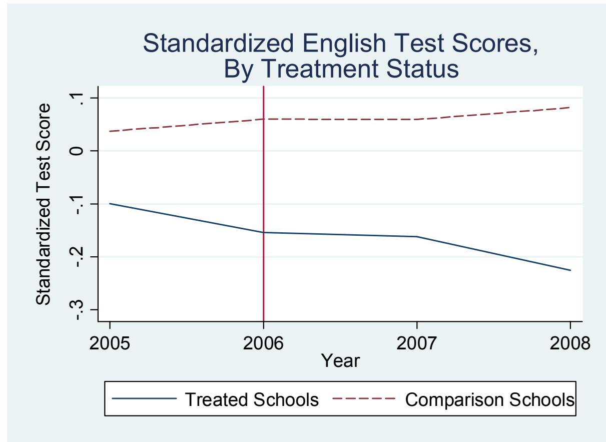
[a]*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered on school in parentheses.

[b]All specifications include an indicator variable for HOPS treatment, year fixed effects, and an interaction between Male and HOPS treatment.

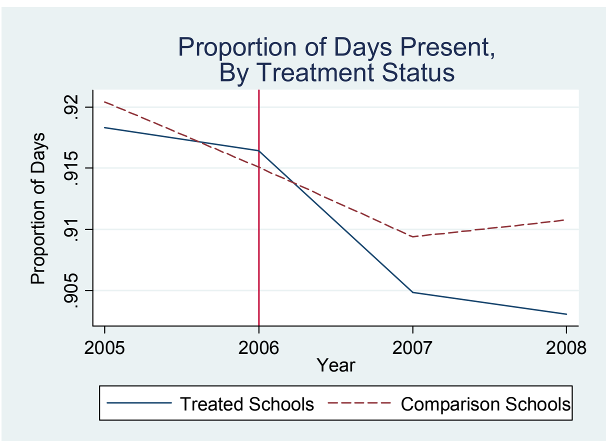
Figure 3.1: Outcome Trends by Treatment Status



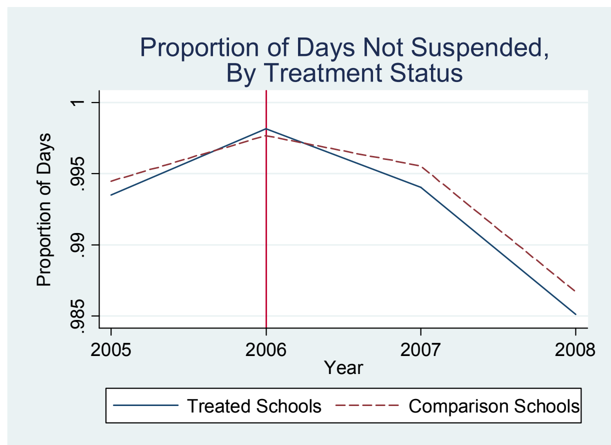
(a)



(b)



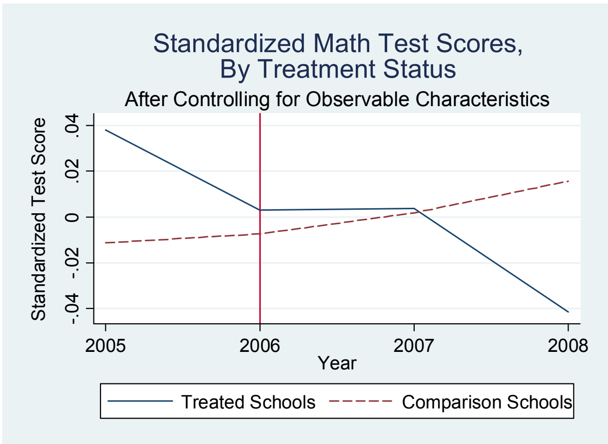
(c)



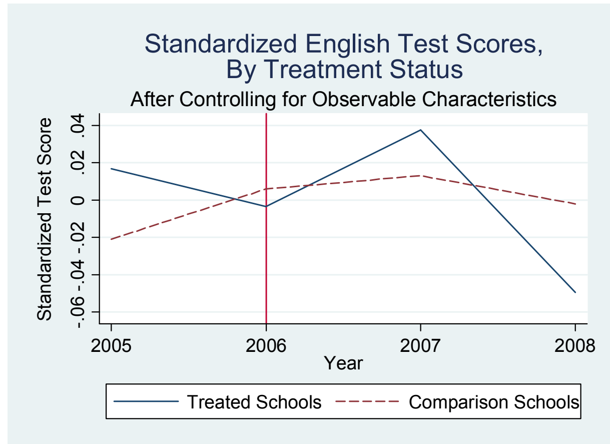
(d)

Figure 3.1a, Figure 3.1b, Figure 3.1c, and Figure 3.1d show the trends in standardized math test scores, standardized English test scores, the proportion of days present, and the proportion of days not suspended, respectively, for *HOPS* treated schools and control schools.

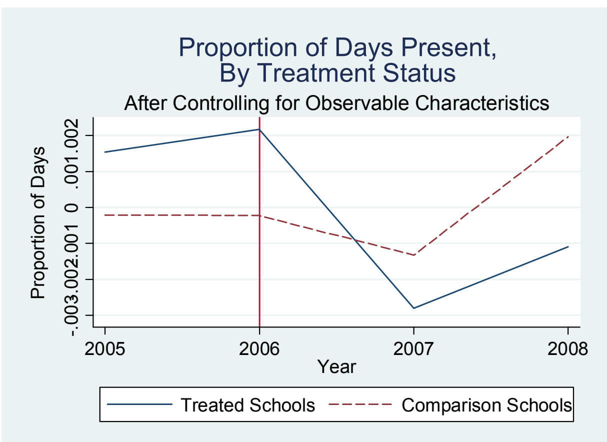
Figure 3.2: Outcome Trends by Treatment Status, Controlling for Observable Characteristics



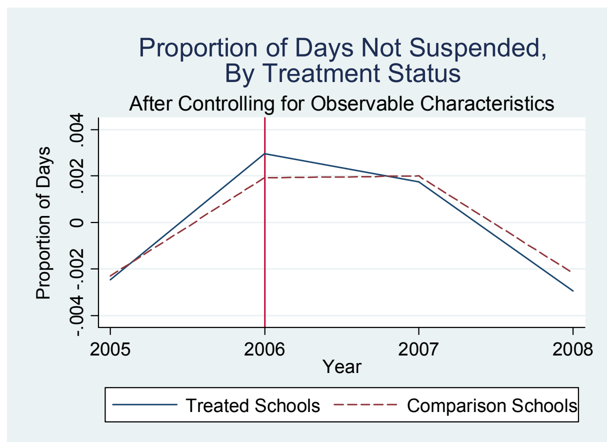
(a)



(b)



(c)



(d)

Figure 3.2a, Figure 3.2b, Figure 3.2c, and Figure 3.2d show the trends in standardized math test scores, standardized English test scores, the proportion of days present, and the proportion of days not suspended, respectively, for *HOPS* treated schools and control schools; each controls for demographic characteristics and cohort, grade, school, and teacher fixed effects.

References: Chapter 3

Belot, M., & and James, J. (2011). Healthy school meals and educational outcomes. *Journal of Health Economics*, *30*(3), 489–504.

Benton, D. & Buts, J. (1990). Vitamins/mineral supplementation and intelligence. *The Lancet*, *335*, 1158–1160.

Benton, D. & Roberts, G. (1988). Effect of vitamin and mineral supplementation on intelligence in a sample of schoolchildren. *The Lancet, 1*, 140–143.

Bryan, J., Osendarp, S., Hughes, D., Calvaresi, E., Baghurst, K. & van Klinken, J. (2004). Nutrients for cognitive development in school-aged children. *Nutrition Reviews, 62*(8), 295–306.

Chenoweth, W. (2007). Vitamin B complex deficiency and excess. In R. Kliegman, H. Jenson, R. Behrman, & B. Stanton (Eds.), *Nelson Textbook of Pediatrics, 18th edition*. Philadelphia: Saunders.

Cueto, S., Jacoby, E., & Pollitt, E. (1998). Breakfast prevents delays of attention and memory functions among nutritionally at risk boys. *Journal of Applied Developmental Psychology*, *19*(2), 219–233.

Cueto, S. (2001). Breakfast and dietary balance: The enKid study. *Public Health Nutrition*, *4*, 1429–1431.

Delange, F. (2000) The role of iodine in brain development. *Proceedings of the Nutrition Society, 59*, 75–79.

Eysenck, H., & Schoenthaler, S. (1997). Raising IQ level by vitamin and mineral supplementation. In R. Sternberg and E. Grigorenko (Eds.), *Intelligence, heredity and environment* (pp. 363 – 392). Cambridge: Cambridge University Press.

Figlio, D., & Winicki, J. (2005). Food for thought: The effects of school accountability plans on school nutrition. *Journal of Public Economics, 89*, 381–394.

Florence, M., Asbridge, M., & Veugelers, P. (2008). Diet quality and academic performance. *Journal of School Health*, *78*, 209–215.

Frisvold, D. (2012). Nutrition and cognitive achievement: An evaluation of the school breakfast program. Working Paper, Emory University.

Greenbaum, L. (2007a). Vitamin E deficiency. In R. Kliegman, H. Jenson, R. Behrman, & B.

Stanton (Eds.), *Nelson Textbook of Pediatrics, 18th Edition*. Philadelphia: Saunders.

Greenbaum, L. (2007b). Micronutrient mineral deficiencies. In R. Kliegman, H. Jenson, R. Behrman, & B. Stanton (Eds.), *Nelson Textbook of Pediatrics, 18th Edition*. Philadelphia: Saunders.

Hollar, D., Messiah, S., Lopez-Mitnik, G., Hollar, T., Almon, M., & Agatston, A. (2010). Effect of a two-year obesity prevention intervention on percentile changes in body mass index and academic performance in low income elementary school children. *American Journal of Public Health*, *100*(4), 646–653.

Imberman, S. & Kugler, A. (2012). The effect of providing breakfast on student performance: Evidence from an in-class breakfast program. *NBER Working Paper 17720*. Retrieved from http://www.nber.org/papers/w17720.

Jones, T., Borg, W., Boulware, S., McCarthy, G., Sherwin, R., Tamborlane, W. (1995). Enhanced adrenomedullary response and increased susceptibility to neuroglygopenia: Mechanisms underlying the adverse effect of sugar ingestion in children. *Journal of Pediatrics*, *126*, 171–177.

Kelder, S., Perry, C., Lytle, L., & Klepp, K. (1995). Community-wide youth nutrition education: Long-term outcomes of the Minnesota Heart Health program. *Health Education Research: Theory and Practice, 10*(2), 119–131.

Kleinman, R., Murphy, J., Little, M., Pagano, M., Wehler, C., Regal, K., & Jellinek, M. (1998) Hunger in children in the United States: Potential behavioral and emotional correlates. *Pediatrics, 101*(1), e3.

Lambert, J., Agostoni, C., Elmadfa, I., Hulsof, K., Krause, E., Livingstone, B., Socha, P., Pannemans, D., Samartins, S., 2004. Dietary intake and nutritional status of children and adolescents in Europe. *British Journal of Nutrition 92*(2), 147–211.

Lieberman, H. (2003). Nutrition, brain function, and cognitive performance. *Appetite, 40*, 245–254.

Meyers, A., Sampson, A., Wietzman, M., Rogers, B., & Kayne, H. (1989). School breakfast program and school performance. *American Journal of Diseases of Children, 143*, 1234–1239.

Müller, M., Danielzik, S., & Pusta, S. (2005). School- and family-based interventions to prevent overweight in children. *Proceedings of the Nutrition Society, 64*, 249–254.

Nelson, M. (1992) Vitamin and mineral supplementation and academic performance in schoolchildren. *Proceedings of the Nutrition Society, 51*, 303–313.

Perry, C., Bishop, D., Taylor, G., Murray, D., Mays, R., Dudovitz, B., Smyth, M., Story, M. (1998). Changing fruit and vegetable consumption among children: The 5-a-Day Power Plus Program in St. Paul, Minnesota. *American Journal of Public Health, 88*(4), 603–609.

Pertz, D., & Putnam. L. (1982). An examination of the relationship between nutrition and learning. *The Reading Teacher, 35*(6), 702–706.

Plachta-Danielzik, S., Pust, S., Asbeck, I., Czerwinski-Mast, M., Langnase, K., Fischer, C., Bosy-Westphal, A., Kriwy, P., & Mueller, M. (2007). Four-year follow-up of school-based intervention on overweight children: The KOPS study. *Obesity, 15*(12), 3159–3169.

Pollitt E. (1993). Iron deficiency and cognitive function. *Annual Review of Nutrition, 13*, 521–537.

Powell, C., Walker, S., Chang, S., & Grantham-McGregor, S. (1998). Nutrition and education: A randomized trial of the effects of breakfast in rural primary school children. *American Journal of Clinical Nutrition, 68*, 873–879.

Price, J. (2012). De-fizzing schools: The effect on student behavior of having vending machines in schools. *Agricultural and Resource Economics Review, 41*(1), 92–99.

Rockoff, J. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review, 94*(2), 247–252.

Rothstein, J. (2010). Teacher quality in educational production: tracking, decay, and student achievement. *The Quarterly Journal of Economics, 125*(1), 175–214.

Rothstein, R. (2002, August 28). LESSONS; Sums vs. summarizing: SAT's math-verbal gap. *New York Times*. http://www.nytimes.com/2002/08/28/nyregion/lessons-sums-vs-summarizing-sat-s-math-verbal-gap.html.

Sandstead, H. (2000). Causes of iron and zinc deficiencies and their effects on brain. *Journal of Nutrition, 130*, 347–349.

Simeon, D., & Grantham-McGregor, S. (1989). Effects of missing breakfast on the cognitive functions of school children of differing nutritional status. *American Journal of Clinical Nutrition, 49*, 646–653.

Schoenthaler, S., Amos, S., Doraz, W., Kelly, M., & Wakefield, J. (1991). Controlled trial of vitamin – mineral supplementation on intelligence and brain function. *Personality and Individual Differences, 12*, 343–350.

Schoenthaler, S., & Bier, I. (1999). Vitamin – mineral intake and intelligence: A macrolevel analysis of randomized controlled trials. *Journal of Alternative and Complementary Medicine, 5*, 125–134.

Sorhaindo, A., & Feinstein, L. (2006). What is the relationship between child nutrition and school outcomes. *Wider Benefits of Learning Research Report No.18*. Centre for Research on the Wider Benefits of Learning.

Storey, H., Pearce, J., Ashfield-Watt, P., Wood, L., Baines, E., & Nelson, M. (2011). A randomized controlled trial of the effect of school food and dining room modifications on classroom behaviour in secondary school children. *European Journal of Clinical Nutrition, 65*, 32–38.

The White House, Office of the First Lady. (2010). First Lady Michelle Obama Launches Let's Move: America's Move to Raise a Healthier Generation of Kids [Press release]. Retrieved from http://www.whitehouse.gov/the-press-office/first-lady-michelle-obama-launches-lets-move-americas-move-raise-a-healthier-genera.