Essays in the Economics of Education

by

Marie Claire Hull

Department of Economics
Duke University

Date: ______________________
Approved:

____
Peter Arcidiacono, Supervisor

____
Hugh Macartney

____
Arnaud Maurel

____
Seth Sanders

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Economics
in the Graduate School of Duke University
2015

Abstract

Essays in the Economics of Education

by

Marie Claire Hull

Department of Economics
Duke University

Date: ______________________
Approved:

Peter Arcidiacono, Supervisor

Hugh Macartney

Arnaud Maurel

Seth Sanders

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Economics in the Graduate School of Duke University
2015

Abstract

This dissertation presents three essays in the economics of education. In the first essay, I examine the importance of family shocks for student learning. In the second, I describe how the evolution of the Hispanic-white test score gap varies by immigrant generation. The last essay explores the racial divide in education and the labor market through evidence from interracial families.

To my parents, for believing in me

Contents

List of Tables

List of Figures

Acknowledgements

First and foremost, I would like to thank my advisor, Peter Arcidiacono, for his help and guidance while completing my Ph.D. I am also deeply grateful to the rest of my dissertation committee, Hugh Macartney, Arnaud Maurel, and Seth Sanders, for valuable feedback and encouragement. This work has benefitted greatly from comments from participants in the Duke Labor Lunch Group. I would also like to thank the North Carolina Education Research Data Center and the North Carolina State Center for Health Statistics for granting access to the data used in Chapters 2 and 3.

For Chapter 4, my co-authors and I would like to thank Peter Gottschalk, Lisa Kahn, Kevin Lang, Derek Neal, and seminar participants at Duke University, Boston College, Brigham Young University, UC Davis, UNC Greensboro, Yale University, the 2011 Institute for Research on Poverty Summer Research Workshop, the 2014 AEA Winter Meetings, and the BU/BC Green Line Labor Seminar. This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis.

My family and friends have been an invaluable source of support during this process. I am especially grateful to my mother, Marilyn Hull, for her words of wisdom and my father, Rob Hull, for his perspective. Finally, I would like to thank my friends, Brian Clark, Kate Duch, and Kristoph Kleiner, for seeing me through the best and the worst.

1

Introduction

This dissertation presents three essays in the economics of education. A common theme in each essay is how different family backgrounds lead to inequality in educational outcomes.

In Chapter 2, I examine the importance of family shocks for student learning. Disruptions in family life can take many forms, but all have the potential to impact student learning. With school administrative data matched to birth records, I estimate the effect of unexpected changes in the home environment, or family shocks, on student achievement. I identify family shocks from siblings observed in the same year. Family shocks are an important determinant of achievement: A one standard deviation family shock leads to a 0.13 standard deviation change in math score and a 0.15 standard deviation change in reading score. These magnitudes are similar to recent estimates of the effect of a one standard deviation change in teacher quality. I also find that family shocks have a relatively larger impact on students from affluent families. This heterogeneity is likely due to a greater involvement of affluent parents in their children's learning, as is evidenced by the students' time use reports.

In Chapter 3, I describe how the evolution of the Hispanic-white test score gap varies by immigrant generation. Past research has shown that Hispanic students make test score gains relative to whites as they age through school; however, this finding stands in contrast to the experience of blacks, who show little change in their relative position over the same time frame. Distinguishing Hispanic students by immigrant generation, I find that the children of immigrants (first- and second-generation Hispanics) drive the improvement in Hispanic test scores. Later-generation Hispanics consistently perform slightly below whites, perhaps due to negative selection into ethnic identification. Thus, previous estimates vastly understate the progress of first- and second-generation Hispanic immigrants. From a negative gap in 3rd grade, these students surpass socioeconomically similar whites in math and reading by middle school and end 8th grade as much as a quarter of a standard deviation ahead.

Chapter 4, which is co-authored with Peter Arcidiacono, Andrew Beauchamp, and Seth Sanders, explores the racial divide through evidence from interracial families. We examine gaps between minorities and whites in educational and labor market outcomes, controlling for many covariates including maternal race. Identification comes from different reported races within the family. Estimates show two distinct patterns. First, there are no significant differences in outcomes between black and white males with white mothers. Second, large differences persist between these groups and black males with black mothers. The patterns are insensitive to alternative measures of own race and school fixed effects. Our results suggest that discrimination is not occurring based on child skin color but through mother-child channels, such as dialect or parenting practices.

2

Family Shocks and Academic Achievement

Disruptions in family life can take many forms: parental job loss, divorce, illness or death in the family. Not all changes are negative. Parents may alter their work schedules to accommodate their children, or family income may unexpectedly increase through a bequest. Each of these changes qualifies as a shock to the family environment that may impact student learning. However, the causal effects that these family shocks have on educational outcomes have proven difficult to establish.[1] Furthermore, after a shock takes place, it is unclear which mechanisms are behind any change in outcomes. Depending on how shocks translate into student outcomes, we might expect students from some families to suffer more than others after experiencing the same size shock to their home environment. Recent debates in the economics of education have focused on what school inputs and interventions are most effective, while changes in family life and their effect on education have received less attention. Understanding how families contribute to academic achievement is

[1] For the effect of parental job loss on student outcomes, see Kalil and Ziol-Guest (2008) and Stevens and Schaller (2011). For reviews on the effect of divorce, see Amato and Keith (1991) and Amato (2001), and for changing family arrangements, see Tillman (2007). The effect of income shocks on children's schooling has been examined in the context of developing countries, but to my knowledge, no paper using U.S. data has tackled this issue.

important in its own right, but it also deserves attention since families and schools may play complementary roles in student learning.

In this paper, I abstract from specific types of family shocks and study how unexpected changes in the family environment as a whole impact student test scores. I estimate a model of test score production that controls for student ability, past inputs, and school quality. Then, I separate the residual into a family-year-specific component, or family shock, and an error. I identify family shocks from sibling pairs observed in the same year. Rich administrative school data from North Carolina matched to birth records from the state yields a sample of 5 million student-year observations with a family link. Given my formulation, the family shock parameter is best understood as a net family shock, or the total effect of various family shocks. By abstracting from specific shocks, I can speak more generally about how changes in the family environment go on to affect student learning and delve into the underlying mechanisms. My framework also lets me analyze why children from some families are more susceptible to shocks than others.

I find that family shocks play an important role in a student's academic achievement. A one standard deviation shock to the family environment leads to a 0.13 standard deviation change in a student's math score and a 0.15 standard deviation change in reading score. These estimates are on par with, if not larger than, recent estimates of teacher value-added from Chetty, Friedman, and Rockoff (2014a). They find that a one standard deviation change in teacher quality moves math scores by 0.14 standard deviations and English scores by 0.1 standard deviations.[2] One can think of both teacher assignment and family shocks as yearlong events in a student's life. In this paper, I show that a one standard deviation change in family inputs is

[2] These estimates are typical in this literature. Rockoff (2004), Rivkin, Hanushek, and Kain (2005), Aaronson, Barrow, and Sander (2007), and Kane and Staiger (2008) all put the impact of a one standard deviation better teacher on test scores between 0.08 and 0.15 standard deviations. Rothstein (2014) replicates Chetty, Friedman, and Rockoff's (2014a) results with the school data used in this paper and reproduces all key results on teacher value-added.

at least as important as a one standard deviation change in teacher quality in terms of its impact on test scores.

The development of the teacher quality literature offers a useful analogy for the methods, interpretation, and importance of the results in this paper. From Hanushek (1971) to Kane, Rockoff, and Staiger (2008), studies have found inconsistent evidence that observable teacher characteristics have an impact on student learning. However, with the rise of large, administrative data sets and increased computing power, our understanding of the impact of teacher quality has changed. Studies such as Rockoff (2004) and Rivkin, Hanushek, and Kain (2005) were among the first to leverage the fact that we observe a teacher teach many students over time in order to estimate the total impact of having an effective teacher.[3] They revealed that observable characteristics only explain a small fraction of the total variation in teacher quality.

Similarly, past studies have only found small impacts of observable family shocks on student outcomes. For example, in their preferred specification, Stevens and Schaller (2011) find that parental job loss increases the probability of grade repetition by 0.008 with a standard error of 0.004. Furthermore, the causation for some observable family shocks is so difficult to establish that only descriptive studies exist. Take the case of divorce. Amato and Keith (1991) demonstrate a negative association between divorce and a host of cognitive and social outcomes. Within the context of family shocks, major fights or the initial separation may be more disruptive than the finalization of the divorce. However, these potentially more important events are likely unobserved. In this study, I leverage the fact that I observe multiple siblings in a year to infer the total impact of family shocks on test scores. A unique data set from North Carolina that merges school administrative data and birth records

[3] Earlier studies such as Hanushek (1971) and Murnane (1975) employ a similar logic but are limited by data availability.

makes this study possible.[4] As with teachers, I find that the total variation in family shocks is quite large relative to the impacts of observed family shocks.

After establishing the net impact of a family shock, I explore whether children from some families are more vulnerable to family shocks than others. In other words, does a one standard deviation change in family inputs have a larger impact on some students than others? The way that differences among families might play out is not clear. Consider a negative family shock to student achievement. In disadvantaged families, parents might have limited ability to shield their children from the shock, while affluent parents may be better able to substitute other inputs to offset it. On the other hand, parents from affluent families might be more involved in their children's learning, and so their children might suffer more from a change in their time spent together. If the parents are already uninvolved, the child has less to lose when parental resources are stretched thin. Another factor is the type of shock that leads to the change in test scores since some families are more likely to experience certain shocks than others. Here, I focus on the net impact of a one standard deviation shock.

I take advantage of two features of my data to shed light on the heterogeneity in family shocks and the mechanisms behind them. The first is demographic information in the education and birth data sets that lets me analyze family shocks by family characteristics. The second is time use variables collected at the time of testing: time spent using a computer, free reading, doing homework, and watching television. The free reading and homework variables are particularly informative since educational activities like these are the most productive uses of time for cognitive development (Fiorini and Keane, 2014). Taken together, these variables tell me how family shocks

[4] The only other similarly linked data set comes from Florida. Figlio et al. (2014) uses this data to study the effect of neonatal health on educational outcomes. A handful of recent papers also use the matched North Carolina data, mostly to study the impact of early childhood education, but no other paper to my knowledge uses the mother identifier.

affect students' home activities and how family shocks to time use move with family shocks to achievement.

The evidence supports the hypothesis that shocks to affluent families lead to bigger changes in parental involvement, and therefore, shocks in these families have a bigger impact on the children's achievement. First, I show that family shocks to achievement often have a larger impact on children from families with a higher socioeconomic status (SES). For example, the impact of a one standard deviation family shock on math achievement increases significantly in mother's education, father's education, and income. Second, I show that time spent engaged in educational activities also responds more to family shocks when a student has affluent parents. Last, I demonstrate that family shocks to educational time use are more closely related to family shocks to achievement, as compared with family shocks to other uses of time.

In analyzing the net impact of family shocks, I lose the ability to point to a specific event and say precisely how it affects student achievement. However, a more general treatment helps us see the bigger picture of how disruptions to family life influence a child's learning in school. This work serves as a complement to studies on specific family shocks—it is still important to understand the link between specific changes in the family environment and educational outcomes. But in light of the difficulty in establishing causation, a study of the aggregate impact of shocks has value. It establishes basic relationships between changes in the home environment and performance at school, along with some of the intervening mechanisms.

The findings in this paper are connected to several strands of literature in the economics of education. Much work has focused on how permanent family characteristics, like parents' education, contribute to inequality in educational outcomes.[5] Here, I emphasize that changes in family inputs also play a role in educational out-

[5] See Björklund and Salvanes (2011) for a review.

comes as well educational inequality. Negative family shocks decrease the distance between high-SES and low-SES students while positive shocks increase it. In some ways, these findings are more policy-relevant since there is scope to manipulate parents' current inputs whereas permanent family characteristics are by nature fixed.

Another question of interest for policymakers is what interventions are most successful at reducing outcome gaps between poor and rich students, or minority and white students. There is some evidence that disadvantaged students respond more to improvements in school inputs than affluent students.[6] My results also suggest that schools and teachers play a greater role in the learning of disadvantaged students since their test scores respond less to changes in family inputs. Even in the event that the home circumstances for an affluent child change, the school might still help mitigate any negative effects. One implication of this study is that teachers and principals should communicate information about a child's home environment to the teacher and principal of his sibling. If school employees look out for students in this way, the impact of negative shocks may not be so detrimental. Finally, while more involved parents may leave their children more vulnerable to negative shocks, they also may provide an extra boost to their children's learning in the event of a positive family shock. This research suggests that home interventions that help parents guide their children toward educationally enriching activities will produce achievement gains at school.

Finally, this paper is related to research in other disciplines on the impact of stress at home on a child's cognitive development. In the medical literature, the Adverse Childhood Experiences (ACE) Study has shown that childhood struggles,

[6] For example, Aaronson, Barrow, and Sander (2007) show that black students and students with initially lower achievement benefit more from a higher quality teacher. Krueger (1999) finds that small class sizes have a larger effect on minorities and poor students, and Neal (1997) finds that urban minorities benefit more from attending Catholic school. In recent work, Jackson, Johnson, and Persico (2015) show that the positive effects of increases in school funding on adult outcomes are larger for low-income students.

like abuse, neglect, and other family dysfunction, influence a host of adult outcomes across physical, psychological, behavioral, and economic dimensions.[7] This study affirms the importance of parental inputs but emphasizes that children from affluent families are also vulnerable to disruptions at home.

In the next section, I offer a more in-depth discussion of the relationship between family shocks and academic achievement. I describe the matched education and birth data sets in section 2.2 and go over descriptive results in section 2.3 to motivate my econometric model in section 2.4. In section 2.5, I present my results on family shocks and discuss the mechanisms that lead to changes in academic outcomes. Section 2.6 concludes.

2.1 From family shock to academic achievement

In this section, I discuss the process by which a shock to the family environment could affect a student's performance at school. I address some of the challenges in establishing causation and the mechanisms that connect a family shock to academic achievement. Finally, I develop two hypotheses that explain why some children may be more vulnerable to family shocks than others.

In previous studies of specific family shocks, the process that links family shocks and educational outcomes is essentially a black box. While the researchers can speculate about mechanisms, they have little to no evidence to support or refute their hypotheses. Part of the difficulty lies in the snowball nature of family shocks—rarely does one shock occur by itself. Take the example of parental job loss. When a parent loses his job (precipitating event), several other events (follow-on events) may occur simultaneously or shortly thereafter. Household income may drop, affecting the resources available to invest in the child's human capital. The unemployed parent

[7] See Felitti et al. (1998) for an early overview. Also see the Center for Disease Control's webpage on the study: www.cdc.gov/violenceprevention/acestudy/.

may spend more time at home, potentially increasing his time investment in the child. There also may be some psychological stress associated with the job loss, which could affect a child's academic performance. After the parent loses his job, it is not clear which, if any, of these events would lead to a change in academic performance. It is also unclear which effects would dominate.

Here, I lump the precipitating event and all follow-on events under the term "family shock." While some might argue that I lose the link to the precipitating event, that link was never direct from the start. By reformulating the problem, I can analyze the net effect of a family shock without delving into the tangle of events that changed the home environment for a student.

Separate from the net family shock are all the changes in family inputs that occur as a result of the shock. For example, a family shock might affect the time that a student spends working on homework, reading, using a computer, or watching television; then any of these changes at home could go on to affect performance at school. While those are the time use variables available in my data set, we can think of each as a signal of what has changed in the home environment after a shock. In other words, the causal impact of any of these time use variables may be small, but together they shed light into the black box of mechanisms. With my reformulation, I am able to focus on how family shocks change inputs at home and then study the link between these mechansims and academic achievement.

The importance of family shocks and the mechanisms behind them need not be the same for all families. In fact, it is not clear from the outset which families might be more vulnerable to family shocks and why. One possibility is that disadvantaged students respond more to changes at home. Their parents may be less able to shield them from a negative shock by substituting other inputs. For example, an affluent family may have savings or extended family to rely on. If the parents do not have money to buy books, they may take the time to check out books from the public library. Disadvantaged parents may not be able to make these substitutions, or they may be backed into a corner where they have few substitution options left. In this case, we would expect a one standard deviation change in family inputs to have a larger effect on low-SES students.

A second possibility is that affluent students are more vulnerable to family shocks. Their parents may be more involved in their learning, which could lead to bigger changes in home inputs after a family shock. Suppose affluent parents spend more time helping their children complete their homework. After a negative shock, the parents may not be able to devote as much time to this activity. If disadvantaged parents already do not spend much after-school time with their children, then parental involvement may not change much after a family shock. In a sense, affluent students may be more vulnerable because they have farther to fall.

The test score data lets me determine which types of students are more vulnerable to family shocks, and the time use data helps me understand why.

2.2 Data

The education data for this study is provided by the North Carolina Education Research Data Center (NCDERC). I present summary statistics in Table 2.1. These records cover all students in 3rd through 8th grade attending public school in North Carolina from 1997 to 2013. While test scores are available for all years, data on computer use and free reading was only collected 1999-2011, data on homework was collected through 2011, and data on TV watching was only collected through 2006. Since my econometric model requires a lagged outcome, I omit grade 3 and the first year an outcome was collected (1997 for test scores, homework, and TV watching, 1999 for computer use and free reading). These restrictions leave a sample of 8.3 million test score observations from 2.4 million students, though the sample size is less for the time use outcomes. The education data include other student demographics:

sex, race/ethnicity, and subsidized lunch status. The North Carolina Department of Public Instruction (NCDPI) stopped requiring that schools report subsidized lunch status for individual students after 2006. Finally, each student-year record indicates the school attended.

The NCERDC matched students to the birth records of children born in North Carolina from 1987 onward. The North Carolina State Center for Health Statistics provided the raw birth record files. The match rate of student-year records to birth records for eligible students is 64%.[8] The birth records contain mother and father education, mother marital status, mother and father age, and other characteristics of the pregnancy and birth, such as alcohol and tobacco consumption. For births in 1988 or later, the NCERDC provided a unique mother identifier, which links the birth records and education records of siblings. The sample contains 723,362 unique mothers.

I make some adjustments to the outcome variables for ease of interpretation. The test scores are reported on a developmental scale that changed twice during the sample period. I normalize test scores to be mean zero, standard deviation one, by grade and year. In the raw data, the response options for the time use variables are given in ranges. In Appendix Table A.1, I report the original categorical responses, their frequencies, and the conversion scale. I convert the ranges to a continuous scale using the midpoint of the range. For the top option, I use a value close to the lower bound. For analysis with the computer use variable, I condition on the student having a computer at home; i.e., I exclude an observation if the student indicates that his family does not own a computer. Similarly, I condition on whether the student reports that his teacher(s) assigned homework for the homework variable. For the main part of the analysis, I also standardized the time use variables by grade

[8] Some students were not matched to birth records because of the years of available data. For example, a student in 8th grade in 1997 was likely born in 1982 or 1983.

Table 2.1: Descriptive statistics

Sample:	Full		Matched		Sibling-pair	
Student-year observations	8,301,355		4,890,464		1,426,970	
Unique students	2,403,870		1,200,317		549,989	
Unique schools	2,255		2,244		2,211	
Match rate to birth data for eligible observations	0.642					
Unique mothers			723,362		236,881	
Math score (std devs)	0	(1)	-0.009	(0.988)	-0.029	(1.007)
Reading score (std devs)	0	(1)	-0.012	(0.988)	-0.065	(1.007)
Homework (std devs)	0	(1)	-0.012	(0.990)	-0.028	(0.988)
Free reading (std devs)	0	(1)	-0.045	(0.969)	-0.058	(0.968)
Computer use (std devs)	0	(1)	-0.025	(0.978)	-0.033	(0.970)
TV watching (std devs)	0	(1)	0.022	(1.002)	0.010	(1.016)
Homework (hrs/wk)	2.53	(2.42)	2.49	(2.41)	2.44	(2.39)
Free reading (hrs/day)	0.81	(0.62)	0.78	(0.60)	0.77	(0.60)
Computer use (days/mo)	3.47	(6.29)	3.26	(6.09)	3.17	(5.99)
TV watching (hrs/day)	2.57	(1.99)	2.56	(2.00)	2.53	(2.02)
White	0.580		0.611		0.597	
Black	0.281		0.297		0.311	
Hispanic	0.078		0.040		0.035	
Other race	0.035		0.028		0.030	
Multiracial	0.025		0.024		0.027	
Female	0.493		0.497		0.497	
Subsidized lunch	0.435		0.429		0.511	
Mother education			12.6	(2.4)	12.6	(2.5)
Mother age			25.9	(5.9)	25.3	(5.6)
Mother married			0.670		0.665	
Mother immigrant			0.061		0.052	
No father information			0.133		0.150	
Father education			12.7	(2.4)	12.8	(2.5)
Father age			28.9	(6.5)	28.5	(6.3)
Father immigrant			0.073		0.069	
First born			0.441		0.323	
Birth weight			7.31	(1.31)	7.20	(1.38)
Alcohol when pregnant			0.014		0.013	
Tobacco when pregnant			0.179		0.177	

Sample: Students in grades 4-8, years 1998-2013. Standard deviations in parentheses.

and year, as students' activities likely change as they age and as techonology changes over the years. The normalization facilitates comparison with test scores since all outcomes are in standard-deviation units.

Means and standard deviations for the time use variables before standardization variables are in Table 2.1. In this sample, the average student uses a computer at home for school 3.5 days per month, reads in his free time 49 minutes per day, spends 2.5 hours per week on homework, and watches television 2.6 hours per day. I also report summary statistics for the sample matched to birth data and for the sample of student-year observations matched to a sibling in that year. I identify family shocks from this last subsample. While the matched sample looks overall similar to the full sample, the sibling-pair sample appears negatively selected. In contrast to the mean-zero test scores for the full sample, the sibling-pair sample has a mean math score of -0.029 standard devations and a mean reading score of -0.065 standard deviations. The differences between the mean time use variables for the full sample and sibling-pair sample are equally or less substantial, relative to their standard deviations. Students in the sibling-pair sample spend less time using a computer for school, doing homework, and free reading, and more time watching television.[9] The racial composition is overall similar across samples, though the matched sample is slightly less Hispanic. The rate of subsidized lunch eligibility is higher in the sibling-pair sample, indicating that these students are poorer.

Finally, Table 2.1 gives summary statistics for variables only in the birth records. At the student-year observation level, the average mother was 26 years old at the time of the child's birth and had completed high school. Sixty-seven percent of mothers were married. The probability a student's mother was born outside of the U.S. is 6.1% in the matched sample but 5.2% in the sibling-pair sample. Since

[9] Before standardization, students in the sibling-pair sample appear to watch less TV. After I standardize by grade and year, we see that they watch more television.

the NCERDC matched siblings through the mother, two siblings could in fact be half-siblings. Then the father information, if given, would not necessarily match. For 15% of observations in the sibling-pair sample, no information on the father is present. For the rest, the average father was 29 years old at the time of birth and had completed high school.

2.3 Descriptive results

In this section, I discuss descriptive results on the outcomes and what they tell us about the relationship between the home environment and outcomes for different types of families.

In Table 2.2, I show that mean academic and time use outcomes vary substantially by family traits. The relationship between family background and achievement is well documented in the test score literature—stark differences exist by parent's education attainment, income, and race. The estimated differences here are typical: The black-white test score gap is three-quarters of a standard deviation, and the difference in test scores between children with a mother who graduated college versus high school is similar.

I also observe variation in time use by family background, though these differences are less stark. I report all the time use estimates in standard deviation units. Homework time, free reading, and computer use are all increasing in mother's and father's education; TV watching decreases with parental education. The biggest spread between the children of high school educated and college educated mothers is in TV watching at 0.39 standard deviations, followed by homework time at 0.27 standard deviations, free reading at 0.15 standard deviations, and computer use at 0.08 standard deviations. The other indicator of socioeconomic status, eligibility for subsidized lunch, follows the same pattern except for computer use. Poor families use a computer more days a month for school work, though the difference is small at

Table 2.2: Mean outcomes by demographics

	Math score	Reading score	Homework	Free reading	Computer use	TV watching
All	-0.009	-0.012	-0.012	-0.045	-0.025	0.022
	(0.988)	(0.988)	(0.990)	(0.969)	(0.978)	(1.002)
Mother's education						
< high school	-0.475	-0.492	-0.137	-0.123	-0.040	0.134
	(0.873)	(0.927)	(0.938)	(0.957)	(1.022)	(1.061)
High school	-0.168	-0.155	-0.070	-0.087	-0.051	0.107
	(0.914)	(0.932)	(0.954)	(0.956)	(0.974)	(1.021)
Some college	0.164	0.179	0.034	-0.011	-0.032	-0.018
	(0.917)	(0.899)	(0.997)	(0.975)	(0.951)	(0.958)
College	0.673	0.629	0.196	0.066	0.033	-0.287
	(0.895)	(0.834)	(1.065)	(0.973)	(0.959)	(0.836)
Grad school	0.859	0.815	0.257	0.179	0.077	-0.408
	(0.882)	(0.815)	(1.099)	(1.016)	(0.991)	(0.785)
Father's education						
Missing	-0.537	-0.521	-0.146	-0.123	0.021	0.317
	(0.867)	(0.922)	(0.935)	(0.965)	(1.089)	(1.117)
< high school	-0.335	-0.362	-0.115	-0.115	-0.079	0.058
	(0.876)	(0.925)	(0.942)	(0.952)	(0.965)	(1.008)
High school	-0.060	-0.047	-0.044	-0.074	-0.063	0.046
	(0.911)	(0.921)	(0.963)	(0.956)	(0.947)	(0.988)
Some college	0.267	0.277	0.061	0.016	-0.036	-0.081
	(0.906)	(0.881)	(1.006)	(0.983)	(0.934)	(0.920)
College	0.700	0.653	0.203	0.065	0.038	-0.311
	(0.882)	(0.818)	(1.065)	(0.965)	(0.956)	(0.809)
Grad school	0.924	0.864	0.286	0.193	0.097	-0.451
	(0.864)	(0.790)	(1.109)	(1.015)	(1.005)	(0.755)
Subsidized lunch	-0.454	-0.455	-0.131	-0.106	0.006	0.231
	(0.864)	(0.931)	(0.940)	(0.975)	(1.070)	(1.089)
No sub. lunch	0.315	0.308	0.070	0.005	-0.031	-0.127
	(0.946)	(0.903)	(1.015)	(0.963)	(0.932)	(0.907)

Table 2.2 Mean outcomes by demographics

	Math score	Reading score	Homework	Free reading	Computer use	TV watching
Mother's race						
Black	-0.523	-0.501	-0.135	-0.128	0.073	0.419
	(0.860)	(0.908)	(0.935)	(0.951)	(1.125)	(1.123)
Hispanic	-0.181	-0.333	-0.096	-0.049	0.008	-0.044
	(0.895)	(0.919)	(0.926)	(0.926)	(1.001)	(0.916)
White	0.239	0.238	0.046	-0.009	-0.069	-0.169
	(0.950)	(0.933)	(1.010)	(0.977)	(0.904)	(0.878)
Mother's nativity						
Native	-0.017	-0.010	-0.014	-0.051	-0.030	0.027
	(0.987)	(0.988)	(0.989)	(0.969)	(0.974)	(1.005)
Immigrant	0.114	-0.042	0.025	0.051	0.075	-0.146
	(0.990)	(0.984)	(1.001)	(0.976)	(1.043)	(0.906)
Observations	4,882,823	4,862,999	3,869,819	3,663,761	3,218,859	2,445,804

Sample: students in birth data. Standard deviations in parentheses.

0.04 standard deviations. Children of black and Hispanic mothers also report higher frequencies of computer use relative to white mothers. They also spend less time on homework and free reading, and more time watching television.

Since the academic and time use outcomes are determined jointly, I examine the correlations between student-year outcomes to understand how they move together.[10] These results are in Table 2.3. I use standardized outcomes in the correlation tables, meaning all outcomes are measured relative to the distribution for a grade and year. First, math and reading scores are highly correlated with a correlation coefficient of 0.75. Test scores are also positively related to homework time and free reading. Homework time is more strongly related to math scores (0.20) relative to reading scores (0.17), while the opposite is true for free reading, which correlates more highly with reading scores (0.22) than math scores (0.15). These relationships

[10] Using the North Carolina data, Clotfelter, Ladd, and Vigdor (2006) include the time use variables as regressors in value-added models of math and reading achievement. They find that achievement is generally increasing in homework time (but more so for math) and free reading time (but more so for reading). The patterns for computer use and TV watching are not monotonic.

Table 2.3: Correlations between outcomes

	Math score	Reading score	Homework	Free reading	Computer use	TV watching
Math score	1.000					
Reading score	0.752	1.000				
Homework	0.195	0.165	1.000			
Free reading	0.153	0.216	0.141	1.000		
Computer use	-0.033	-0.043	0.070	0.057	1.000	
TV watching	-0.162	-0.156	-0.025	-0.045	0.010	1.000

Sample: students in birth data. All outcomes standardized by grade and year.

suggest that math is a more homework-intensive subject, while a student can improve his reading skills by reading on his own. In contrast, the test scores are almost as negatively correlated with TV watching at -0.16 for both subjects. When students watch television, they may substitute away from more educationally enriching activities, like doing homework and reading. Finally, test scores are negatively correlated with computer use for school work, though this relationship is the weakest at -0.03 for math and -0.04 for reading. Although students ostensibly spend this time studying, more computer use could lead to lower test scores for several reasons. Students might get easily distracted by social media or other websites while they work on homework, or the learning they do on a computer could be less enriching than learning they do out of a book.

Table 2.3 also shows the correlations in how students spend their time. Free reading and homework time are the most highly correlated among these variables (0.14). These two variables are also negatively correlated with TV watching, which suggests that these activities may be substitutes. Given that time spent using a computer for school work is a subset of homework time, it is not surprising that these two variables are correlated. However, free reading and computer use are also positively correlated, though they correlate in different directions with test scores. These descriptives indicate that the relationship between computer use and test

Table 2.4: Correlations between own outcome and sibling outcome

	Math score	Reading score	Homework	Free reading	Computer use	TV watching
Math score	0.488					
Reading score	0.440	0.465				
Homework	0.113	0.103	0.069			
Free reading	0.082	0.097	0.037	0.107		
Computer use	-0.015	-0.022	0.012	0.002	0.076	
TV watching	-0.163	-0.154	-0.043	-0.033	0.009	0.214

scores may not be so straightforward.

Next, I calculate the correlations between own outcomes and sibling outcomes and present the results in Table 2.4. Each sibling-year pair represents one observation. These correlations give insight into the relationship between family environment and academic and time use outcomes. In general, the sibling correlations have the same sign as, but a lower magnitude than, the own outcome correlations. The correlation between own math score and sibling math score, as well as own reading score and sibling reading score, is close to 0.5. This correlation is the first evidence in this paper that characteristics shared within a family, whether genetics, permanent inputs, or transitory inputs, influence achievement. The rest of the diagonal shows the within-family correlations of time use outcomes. The association for TV watching is especially strong at 0.21, relative to the other time use variables. Television sets often function as a public good within a household, and siblings may watch TV together. TV watching by one child may even distract his sibling from other activities. In fact, sibling television use is just as closely correlated with outcomes as own television use.

In Table 2.5, I examine whether the relationship between achievement and time use varies by family characteristics. In other words, I examine whether these correlations are uniformly positive or negative for all family types, as well as whether they are equally strong across families. I arrange family traits from least to most advantaged. Although computer use and achievement are negatively correlated for most families, they have a small, positive correlation for the most affluent families (parent went to graduate school, mother is white). Furthermore, the correlations are increasing in socioeconomic status, meaning that computer use is the most negatively associated with achievement in the least affluent families (parent dropped out of high school, student eligible for subsidized lunch). The pattern suggests that the context of computer use is key. Affluent parents may be better able to monitor computer use, or homework activities on the computer may be more effective for children of these parents.

Another activity that appears highly contextual is TV watching. The achievement of children of black mothers is positively correlated with time spent watching televsion, while it is negatively correlated for all other groups shown. For other less affluent groups (children eligible for subsidized lunch, with a Hispanic mother, or with a high school dropout parent), the correlation is negative but relatively close to zero. Again, the correlation patterns between TV watching and achievement are monotonic in socioeconomic status. They are also consistent with the mainstreaming hypothesis, which contends that low-SES children receive more cognitive stimulation from watching TV than they would otherwise. Among high-SES children, the opposite is true, so high levels of TV viewing is especially detrimental for these students (Morgan and Gross, 1980).

Finally, for time spent free reading and on homework, the intensity of the relationship between these activities and achievement varies by demographics. The correlation between homework time and test scores is almost uniformly increasing in socioeconomic status; however, the evidence for free reading is mixed. For math, the correlation between free reading and test scores is higher for low-SES students. There is no clear pattern for reading test scores, and all the correlations by family characteristics are clustered around the overall correlation. These estimates suggest that time

Table 2.5: Correlations between outcomes by demographics

	Homework		Free reading		Computer use		TV watching	
	Math	Read	Math	Read	Math	Read	Math	Read
All	0.195	0.165	0.153	0.216	-0.033	-0.043	-0.162	-0.155
Mother's education								
< high school	0.129	0.099	0.151	0.204	-0.087	-0.095	-0.047	-0.048
High school	0.158	0.127	0.135	0.197	-0.073	-0.079	-0.103	-0.100
Some college	0.177	0.145	0.125	0.202	-0.036	-0.045	-0.151	-0.148
College	0.189	0.155	0.107	0.206	0.005	-0.008	-0.205	-0.198
Graduate school	0.181	0.150	0.116	0.222	0.034	0.019	-0.220	-0.218
Father's education								
Missing	0.127	0.096	0.157	0.205	-0.085	-0.093	-0.015	-0.014
< high school	0.136	0.106	0.140	0.199	-0.073	-0.082	-0.046	-0.046
High school	0.161	0.131	0.129	0.197	-0.061	-0.067	-0.103	-0.101
Some college	0.175	0.142	0.119	0.200	-0.030	-0.039	-0.132	-0.133
College	0.185	0.151	0.101	0.201	0.004	-0.008	-0.185	-0.181
Graduate school	0.186	0.152	0.101	0.210	0.035	0.014	-0.203	-0.198
Subsidized lunch	0.126	0.096	0.161	0.208	-0.099	-0.101	-0.038	-0.036
No sub. lunch	0.198	0.165	0.143	0.212	-0.019	-0.025	-0.163	-0.157
Mother's race								
Black	0.130	0.102	0.144	0.192	-0.062	-0.069	0.021	0.023
Hispanic	0.161	0.129	0.141	0.205	-0.078	-0.095	-0.026	-0.030
White	0.194	0.162	0.137	0.215	0.012	0.001	-0.123	-0.123

Sample: students in birth data. All outcomes standardized by grade and year.

spent free reading is equally productive for all types of students. For disadvantaged students, poor reading skills could inhibit performance in math, making the time they spend improving their reading comprehension especially beneficial. This evidence suggests that the mechanism through which family inputs affect achievement is not the same for all families.

Several stylized facts emerge from this descriptive analysis. One, student achievement is positively correlated with homework time and free reading time, negatively correlated with TV time, and barely correlated with computer time for school work.

Two, there are relatively strong within-family correlations in test scores and TV watching. Three, the correlation between achievement and time use depends on family background. Even the signs of the correlations for computer use and TV watching depend on family characteristics. For all time use outcomes, the strength of the association depends on socioeconomic status of the family. And four, the outcomes themselves vary by family background, though not as much for time use outcomes as for academic ones. Students from high-SES families generally score better on tests, spend more time free reading and on homework, and spend less time watching television. The evidence on the relationship between computer use and affluence is mixed.

While this descriptive analysis establishes basic relationships between outcomes and student demographics, it only hints at how these variables may co-move within a family. Within-family correlations pick up permanent family characteristics as well as transitory ones. My econometric model identifies the impact that family shocks have on outcomes so that I can analyze how and why changes at home affect achievement.

2.4 Econometric model

2.4.1 Test score production

I adopt a model of test score production that is common in this literature:

$$y_{ifst} = \gamma y_{it-1} + x_i\beta + \delta_{st} + \varepsilon_{ifst} \tag{2.1}$$

where y_{ifst} is the test score for student i from family f attending the school-grade pair s in year t. This year's test score is a function of last year's test score y_{it-1}, individual characteristics x_i, a school-grade-year fixed effect δ_{st}, and an idiosyncratic error ε_{ifst}. In the teacher quality literature, this specification is often called a value-added model since controlling for the previous year's score isolates the teacher's contribution from other factors known to predict test scores. Next, I decompose the idiosyncratic shock into a family-year-specific component, or family shock, and an error.

$$\varepsilon_{ifst} = \xi_{ft} + \nu_{ifst} \tag{2.2}$$

Like other papers in the test score literature, I take the lagged test score as a sufficient statistic for all previous inputs as well as endowed ability. Many other studies have developed the assumptions inherent in the lagged test score specification; Todd and Wolpin (2003) have perhaps the most thorough discussion. Importantly for this paper, the lagged test contains the previous year's family shock. In other words, the lagged score fully captures the effects of any past family shocks on the current year's test score. I also assume that the effects of inputs and ability, including the family shock, all decay at a constant rate γ. These assumptions are standard in the test score production literature.[11]

The school-grade-year fixed effects control for unobserved variation across schools, within a school over time, and within different grades in the same school. They also control for other geographical and temporal variation, such as neighborhood characteristics and local labor market conditions. Including the grade component of these fixed effects is important because siblings often attend the same school at the same time. Without it, a school shock might confound my estimate of a family shock.[12]

To estimate equation 2.1, I assume that the idiosyncratic shock ε_{ifst} is orthogonal to the regressors y_{it-1}, x_i, and δ_{st}. The family shock is one component of the total shock to a student's test score in a year and thus must also be orthogonal to the regressors. This assumption is key for identification of the family shock parameters.

[11] See Clotfelter, Ladd, and Vigdor (2006), Todd and Wolpin (2007), Aaronson, Barrow, and Sander (2007), and Chetty, Friedman, and Rockoff (2014a) for examples of similar specifications.

[12] Note that school shocks could still confound the estimate of a family shock when siblings are in the same grade at the same school. This possibility arises for twins or multiples; however, non-singletons only make up 1.4% of the sample.

For the interpretation of ξ_{ft} as a family shock to hold, it must not be contaminated by any other transitory trait that is shared by siblings but occurs outside of the family. For example, if siblings had the same math teacher, the estimate of the family shock parameter could reflect their teacher's value-added in addition to the family shock.

2.4.2 Application to time use outcomes

I use the same specification (equation 2.1) and error decomposition (equation 2.2) for the time use outcomes as for the test score outcomes. With the same model, the results for academic outcomes and time use outcomes are easier to compare. However, the interpretation of the parameters changes. While the lagged test score controls for ability and past inputs, the lagged outcome for time use variables controls for habit persistence. With both types of outcomes, the interpretation of the family shock is as an unexpected *change* in the outcome that is shared by siblings. The school-grade-year fixed effects still control for geographical and temporal trends, which could still include school characteristics (e.g., teachers' propensity to assign large amounts of homework) but might also include trends in technology adoption (in the case of the computer use outcome).

2.4.3 Calculation of the variation in family shocks

The parameter of interest in the paper is the standard deviation of ξ_{ft}. It measures the effect of a one standard deviation family shock on an outcome. Since the estimates of the ξ_{ft} come from a small number of observations, computation with the traditional variance formula introduces a substantial amount of estimation error. Instead, I create a sample of sibling pairs and compute the covariance of residuals for siblings in the same year.

$$Cov(\varepsilon_{ifst}, \varepsilon_{i'fs't}) \quad \text{for } i \neq i' \tag{2.3}$$

The variance of the family shock is equal to the covariance of residuals for sibling-year pairs as long as the family shock is uncorrelated with the error, i.e. $Cov(\xi_{ft}, \nu_{ifst}) = 0$, and the errors for sibilings in the same year are unrelated, i.e. $Cov(\nu_{ifst}, \nu_{i'fs't}) = 0$ for $i \neq i'$. I report the variation in family shocks as a standard deviation for ease of interpretation.

2.5 Family shock results

In this section, I briefly summarize the regression results before presenting my main results on the impact of family shocks on academic achievement and time use outcomes. I also discuss the relative importance of family shocks on different outcomes and how family shocks are related across outcomes.

2.5.1 Regression results

I present results from the outcome regressions in Table 2.6. Estimates from the test score models in columns 1 and 2 will be familiar to many readers. All coefficients have the expected sign, with the exception of alcohol consumption during pregnancy. The coefficients for parent's education and race/ethnicity are generally larger in magnitude than the estimates in Clotfelter, Ladd, and Vigdor (2006), which uses a similar specification to study the impact of teacher characteristics on achievement. The estimated effects of lagged score are also typical.

Perhaps less familiar with readers are the results for the time use outcomes, which are in columns 3-6. All outcomes are in standard-deviation units, normalized by grade and year. Habit persistence is highest for free reading and TV watching at 0.4 standard deviations; the corresponding coefficients for computer use and homework time are about half of that. The patterns of coefficients for mother's and father's education and race/ethnicity generally follow the same pattern as the differences in means, though the marginal differences are much smaller. The biggest differences

Table 2.6: Outcome regression results

	Math score	Reading score	Homework	Free reading	Computer use	TV watching
Lagged outcome	0.788	0.756	0.195	0.397	0.178	0.418
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mother's education						
High school	0.046	0.051	0.017	0.018	-0.006	0.006
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Some college	0.080	0.089	0.046	0.042	-0.004	-0.011
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
College	0.127	0.126	0.084	0.054	0.002	-0.067
	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)
Graduate school	0.145	0.144	0.100	0.092	0.006	-0.098
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)
Father's education						
High school	0.044	0.049	0.028	0.018	0.004	-0.028
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Some college	0.073	0.080	0.054	0.049	0.007	-0.054
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
College	0.109	0.108	0.083	0.048	0.017	-0.109
	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)
Graduate school	0.137	0.137	0.113	0.100	0.031	-0.145
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Black	-0.152	-0.184	-0.126	-0.073	0.081	0.359
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	-0.121	-0.189	-0.093	-0.101	0.066	0.078
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Other race	0.047	-0.056	0.027	0.034	0.123	0.027
	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)
Multiracial	-0.049	-0.049	-0.043	0.005	0.014	0.144
	(0.001)	(0.001)	(0.003)	(0.002)	(0.003)	(0.003)
Female	0.023	0.052	0.063	0.188	0.069	-0.075
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Birth data	-0.130	-0.097	0.060	0.231	0.036	-0.067
	(0.052)	(0.057)	(0.100)	(0.099)	(0.116)	(0.108)
Mother age	0.001	0.002	0.002	-0.000	0.002	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mother married	-0.005	-0.003	0.003	0.000	0.003	-0.009
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Mother immigrant	0.060	0.057	0.014	0.054	0.015	-0.023
	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)

Table 2.6 Outcome regression results

	Math score	Reading score	Homework	Free reading	Computer use	TV watching
No father information	0.015	0.020	0.020	-0.002	0.017	-0.010
	(0.002)	(0.002)	(0.004)	(0.004)	(0.005)	(0.004)
Father age	-0.000	-0.000	0.000	0.000	0.001	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Father immigrant	0.042	0.049	0.010	0.047	0.014	-0.006
	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)
First born	0.021	0.042	0.034	0.043	0.023	-0.006
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Birth weight	0.006	0.004	0.001	-0.001	0.001	0.005
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Alcohol when pregnant	0.008	0.003	-0.001	-0.008	-0.008	-0.033
	(0.002)	(0.002)	(0.004)	(0.004)	(0.005)	(0.004)
Tobacco when pregnant	-0.033	-0.026	-0.018	0.018	-0.001	0.020
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Observations	8,287,683	8,247,200	6,682,110	5,848,259	5,127,654	4,464,194
R^2	0.706	0.651	0.118	0.190	0.098	0.260
F-stat for FE	11.64	4.97	7.38	2.78	6.01	2.72

Standard errors in parentheses.

occur between black and white children, particularly for time spent watching television. For homework time, free reading, and TV watching, most estimates for the parent education categories are less than 0.1 standard deviations in magnitude. For computer use, they are all within 0.03 standard deviations.

2.5.2 Family shock results

The family shock results are in Table 2.7. To generate these estimates, I take the residuals from the test score model and calculate the covariance for sibling-year pairs. I present the results in standard-deviation units.

I find that the net impact of a family shock is 0.13 standard deviations on math scores and 0.15 standard deviations on reading scores. For context, Chetty, Friedman, and Rockoff (2014a) find that teacher value-added is 0.14 standard devations in math and 0.1 standard deviations in English. Like Chetty, Friedman, and Rockoff (2014a),

Table 2.7: Standard deviation of family shocks by demographics

	Math score	Reading score	Homework	Free reading	Computer use	TV watching
All	0.133	0.149	0.144	0.200	0.178	0.247
Mother's education						
< high school	0.127	0.155	0.087	0.162	0.177	0.248
High school	0.131	0.149	0.141	0.178	0.174	0.253
Some college	0.131	0.145	0.151	0.222	0.175	0.245
College	0.147	0.144	0.196	0.251	0.187	0.230
Graduate school	0.150	0.148	0.205	0.272	0.197	0.239
Father's education[a]						
< high school	0.132	0.159	0.129	0.177	0.177	0.245
High school	0.130	0.145	0.137	0.189	0.167	0.242
Some college	0.137	0.150	0.161	0.229	0.157	0.244
College	0.152	0.149	0.185	0.250	0.188	0.240
Graduate school	0.152	0.143	0.210	0.260	0.203	0.239
Subsidized lunch	0.128	0.157	0.119	0.168	0.190	0.261
No sub. lunch	0.140	0.149	0.165	0.203	0.157	0.231
Marital status[b]						
Unmarried	0.129	0.151	0.110	0.164	0.213	0.263
Married	0.135	0.149	0.155	0.212	0.168	0.241
Mother's race						
Black	0.121	0.148	0.122	0.166	0.211	0.278
Hispanic	0.151	0.165	0.193	0.183	0.217	0.240
White	0.137	0.148	0.152	0.215	0.162	0.226
Mother's nativity						
Immigrant	0.155	0.165	0.200	0.241	0.208	0.275
Native	0.132	0.148	0.142	0.198	0.177	0.247
Sibling-pair obs.	825,830	818,228	590,258	602,517	485,927	368,832

All outcomes in standard-deviation units.

[a] For sibling pairs with the same father's education.

[b] For siblings with the same marital status for the mother at the time of birth.

I use a value-added specification, which facilitates the comparison. Thus, a one standard devation change in the family inputs has close to the same impact on math scores as a one standard deviation change in teacher quality. For reading scores, the impact of a one standard deviation change in family environment is 50% larger than a one standard deviation change in teacher quality.

Turning to the family shock results by demographics, I can evaluate the two hypotheses on why the same size family shock might have a larger impact on some families than others. One possibility is that children from disadvantaged families are more vulnerable to family shocks. Their parents may be less able to insure against shocks, or they might be less able to substitute other inputs. Another possibility is that children from advantaged families are more vulnerable. If their parents spend a lot of time with them, making their achievement production process more family intensive, they might suffer more with a change in the family environment. If the parents were less involved from the outset, there is less scope for family changes to affect learning.

In math, the impact of a family shock is increasing in affluence, whether measured by mother's education, father's education,[13] or subsidized lunch status. Because of the large sample size, most of these differences are statistically significant. Parents' education and income are the more straightforward measures of socioeconomic status. For other family characteristics, the patterns still generally support the same hypothesis. Family shocks have a larger impact on children of white mothers compared to black mothers, but they are largest for children of Hispanic mothers. They are also larger for children of married mothers, where mother's marital status is

[13] I only calculate these statistics for siblings with the same level of father's education. When it differs, the children are more likely to be half-siblings. Even when it is the same, the children could still be half-siblings. Regardless, the interpretation is more clear when siblings have the same level of father's education, whether their fathers are different or not.

measured at birth.[14] Finally, children of immigrant mothers are more vulnerable to family shocks compared to native mothers. By and large, the evidence for math scores is consistent with the hypothesis that family shocks have a bigger effect on children from affluent families.

For reading scores, the evidence is mixed. There is no consistent relationship between family shock impacts and parents' education. Poor students are more vulnerable to family shocks, and the impacts for children of black mothers and children of white mothers are the same. Children of Hispanic mothers and children of immigrant mothers are relatively more susceptible to family shocks. Children of single mothers are relatively more susceptible.

Comparing the math and reading results, we see differences in both the magnitude of the effects (intercept) and how the vary across socioeconomic status (slope). The impact of a family shock on reading scores is generally higher than the impact on math scores. These results are in line with the idea that math scores are determined more at school and reading scores are determined more at home. However, this story is more true for the children with parents with low levels of education. For children with highly educated parents, the impact of a family shock on math scores and on reading scores is more similar. This pattern could arise if highly educated parents have a higher math ability themselves and are better equipped to help their children with the children's math homework.

Overall, these results offer more support for the hypothesis that family shocks have a bigger impact on children with affluent parents. However, the test score results alone give little indication of the mechanisms at work in the learning process. For this, I turn to the results for the time use variables.

From the top row of Table 2.7, family shocks have a larger effect on students'

[14] In calculating these statisitics, I condition on mothers having the same marital status at the birth of both children.

time use than test scores when I compare all outcomes in standard-deviation units. In contrast to the 0.13 effect in math and 0.15 effect in reading, a one standard deviation family shock moves homework use by 0.14 standard deviations (0.33 hours per week), free reading by 0.20 standard deviations (0.12 hours per day), computer use for school work by 0.18 standard deviations (1.07 days per month), and TV watching by 0.25 standard deviations (0.50 hours per day).

Examining these results by demographics, we see that the impact of a family shock on free reading and homework increases with the family's socioeconomic status. The relationship between mother's education and the effect of a family shock on homework time is particularly striking. In response to a one standard deviation family shock, a student with a mother that dropped out of high school only sees his homework time change by 0.09 standard deviations (0.21 hours per week). However, a student with a mom that went to graduate school sees his homework time change by 0.20 standard deviations (0.49 hours per week). The same patterns for math scores hold up for homework time and free reading, but the differences for the time use variables are larger.

Going back to the process of family shock to academic achievement, it seems that first a family shock affects how much time a child spends on educationally enriching activities, like homework and reading for pleasure. When the parents are more affluent, the family shock has a bigger impact on these activities. Then, these changes in time use at home, and likely other family inputs that are positively related to learning, go on to impact achievement. However, since achievement is a function of family inputs and school inputs, and the school inputs are held constant, the effect of the family shock on achievement is lessened but still present.

In contrast to the free reading and homework time variables, the pattern for the effect of family shocks on TV watching and computer use is mixed. Recall that these variables are negatively correlated with achievement. The impact of a family shock on computer use is often increasing in parents' education but is decreasing in income as measured by subsidized lunch eligibility. The impact of a family shock on TV watching is more often decreasing in affluence. The differences by socioeconomic status are also smaller for these variables. For example, the range of impacts by mother's education is 0.02 standard deviations for both of these variables, while it is 0.12 standard deviations for homework time. While family shocks have a large impact on TV time, they seem to affect all children's TV time about equally. The same is not true for the educational activities.

We can learn more about the relationship between test scores and the family shock mechanisms by analyzing the correlations between family shocks to achievement and family shocks to time use. Similar to my calculation of the variation in family shocks, I modify the traditional correlation formula. Let k and ℓ represent two different outcomes. Normally, the formula to find the correlation in family shocks to these two different outcomes would be

$$Corr(\xi_{ft}^k, \xi_{ft}^\ell) = \frac{Cov(\xi_{ft}^k, \xi_{ft}^\ell)}{\sqrt{Var(\xi_{ft}^k)Var(\xi_{ft}^\ell)}}$$

I replace both variances in the denominator with the covariance of sibling residuals, which is the same modification I made to calculate the variation in family shocks. I replace the numerator with covariance of sibling residuals for different outcomes, i.e. $Cov(\varepsilon_{ifst}^k, \varepsilon_{i'fs't}^\ell)$ for $i \neq i'$. For this substitution to be valid, the family shock to one outcome must be uncorrelated with the individual error for any other outcome, i.e. $Cov(\xi_{ft}^k, \nu_{ifst}^\ell) = 0$, and the siblings' individual errors for different outcomes must also be unrelated, i.e. $Cov(\nu_{ifst}^k, \nu_{i'fs't}^\ell) = 0$ for $i \neq i'$.

Table 2.8 reports results on the correlations between family shocks to different outcomes. The correlation between family shocks to math achievement and family shocks to reading achievement is 0.82, indicating that family disruptions have a

Table 2.8: Correlations between family shocks

	Math score	Reading score	Homework	Free reading	Computer use	TV watching
Math score	1.000					
Reading score	0.820	1.000				
Homework	0.327	0.279	1.000			
Free reading	0.224	0.343	0.276	1.000		
Computer use	-0.095	-0.100	0.091	0.014	1.000	
TV watching	-0.118	-0.055	-0.101	-0.024	-0.060	1.000

Sample: students in birth data.

similar effect on a student's cognitive development across subjects. I also find a positive relationship between family shocks to achievement and family shocks to educational time use. This means that a family shock that increases homework time or free reading time also increases test scores. The relationship between family shocks to free reading and family shocks to reading scores ($\rho = 0.34$) is stronger than the relationship between family shocks to free reading and family shocks to math scores ($\rho = 0.22$). Conversely, family shocks to homework time are more strongly associated with family shocks to math scores ($\rho = 0.33$) compared to reading scores ($\rho = 0.28$).

Turning to the relationship between family shocks to achievement and family shocks to non-educational time use, we see that the magnitudes are all lower compared to the correlations with educational uses of time. These correlations are also also all negative, ranging from -0.06 to -0.12. The correlations between family shocks to different uses of time are generally weak, with the exception of free reading and homework time.

Let us return to the process through which a family shock affects academic achievement. Of the various potential mechanisms, educational time use appears the most important. The evidence suggests that family shocks that change the time a student spends on educational activities (like free reading and homework) are the ones that have a bigger impact on the student's learning. These results also help us understand why family shocks have a bigger impact on children from affluent families. In the face of a family shock, the larger changes in free reading and homework time for these children likely cause the larger changes in test scores.

2.6 Conclusion

In this paper, I estimate a model of test score production and analyze the family-year-specific component of the residual to determine the impact of family shocks on student test scores. I find that a one standard deviation family shock moves math scores by 0.13 standard deviations and reading scores by 0.15 standard deviations. These estimates are similar in magnitude to recent estimates of the effect of having a one standard deviation better teacher (Chetty, Friedman, and Rockoff, 2014a). Chetty, Friedman, and Rockoff's (2014b) estimate on the long-term impact of teacher effectiveness offers some insight on how a yearlong change to a student's inputs translates into later life outcomes. They determine that a one standard deviation improvement in teacher quality increases lifetime income by almost $40,000. While recent work in the economics of education has sought to understand which school interventions are most effective and why, here I emphasize the importance of the family side of achievement production. Transitory family characteristics, in addition to permanent family characteristics, influence student learning.

The second contribution of this paper is to determine which families are more vulnerable to family shocks, and why. A same-size family shock has a larger impact on students from affluent families, likely because these parents are initially more involved in their children's learning. These results suggest that successful home interventions are ones that help parents guide their children to educationally enriching activities. They also support the idea that school interventions are most helpful to disadvantaged students—changes to affluent students' achivement comes from changes at home. Furthermore, they suggest that school employees should communicate information about changes in a child's home environment so that the impact of a shock on the child's sibling might also be mitigated.

This paper represents a first step in understanding the relationship between family shocks and academic outcomes, and several unanswered questions remain. First, much of the evidence in this paper comes from correlations. For example, I look to the correlation between the effect of a family shock on free reading time and the effect of a family shock on reading test scores to understand the relationship between the time use mechanism and the outcome. Future work on this topic must ascertain the direct causal link between various mechanisms and outcomes. The data in this paper is not well-suited to address this issue since all outcomes are measured simultaneously. A second question of interest is whether family shocks are more important at certain ages. Is the cognitive ability of younger children more sensitive to changes in parental investments, as found Cunha, Heckman, and Schennach (2010)? My framework cannot answer this question directly since I use sibling pairs to calculate the impact of family shocks, and siblings are almost always different ages. A final issue is the persistence of family shocks. Do children bounce back quickly after a family shock, or does the impact of the shock carry through to adulthood? In the teacher value-added literature, teacher effects fade out quickly when persistence is evaluated in terms of future years' scores (Kane and Staiger, 2008; Rothstein, 2010). However, Chetty, Friedman, and Rockoff (2014b) find that teacher value-added impacts a host of later life outcomes including college attendance, teenage pregnancy, and earnings.

3

The Academic Progress of Hispanic Immigrants

At 16.3% of the population, Hispanics are now the largest and fastest growing minority in the United States, yet they have lower levels of human capital than whites by several measures.[1] Furthermore, due to high fertility rates and high rates of migration into the country, Hispanics are disproportionately young, implying that a high proportion are currently receiving their education in our nation's schools. Given that they will be a key part of our future workforce, ensuring that Hispanic youth leave school with the skills to succeed has never been more important, yet Hispanic-white differences in human capital are almost always discussed as an afterthought to black-white differences in research on racial and ethnic inequality.

In recent work, Clotfelter, Ladd, and Vigdor (2009) document that of the three largest racial/ethnic groups (white, black, and Hispanic), Hispanic students are the only one that change their relative position in test scores over time. They score about 0.1 standard deviations below observably similar whites in 3rd grade, equal them in 5th, and outscore them by 0.1 standard deviations in 8th. In contrast,

[1] According to the definition used in the 2010 Census, "Hispanic or Latino" refers to a person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin regardless of race. There is some variation in this definition by source and over time.

black students' covariate-adjusted reading gap stands at a constant -0.5 standard deviations; their math gap is similar in size, though it does narrow about a tenth of a standard deviation between 3rd and 8th grades. Since IQ is relatively stable by age 10, changes in underlying ability cognitive ability cannot explain the rise in Hispanic achievement over this time period, leaving noncognitive inputs and environmental factors as potential explanations (Cunha et al., 2006).

What, or rather, who, accounts for the relative gains in Hispanic test scores? Most papers that note an upward trend in Hispanic test scores attribute it to the assimilation of immigrant Hispanics over the sample period.[2] But in reality, only 13% of the school-age Hispanic population are immigrants, 42% have immigrant parents but were born in the U.S., and the remaining 45% come from families that have been in the U.S. for generations (Ruggles et al., 2010). Any measurement of human capital accumulation for Hispanics is a weighted average of these three groups. Moreover, any representation of how Hispanic students fare as they age blends together both the intercepts and slopes of each of these groups. In this light, lumping all Hispanics together or referring to all Hispanics as immigrants may be misleading. Although assimilation, whether it be through language acquisition or the adoption of native customs, might explain convergence in Hispanic test scores, it does not explain why Hispanic students outscore their native peers.

Several plausible scenarios could give rise to the upward trend in Hispanic test scores. I illustrate three of these in Figure 3.1, using the immigrant status of the parents to divide the Hispanic population into two groups. While each immigrant generation likely has its own intercept, they might share a common slope if Hispanic families are uniformly better at producing academic achievement. I show this

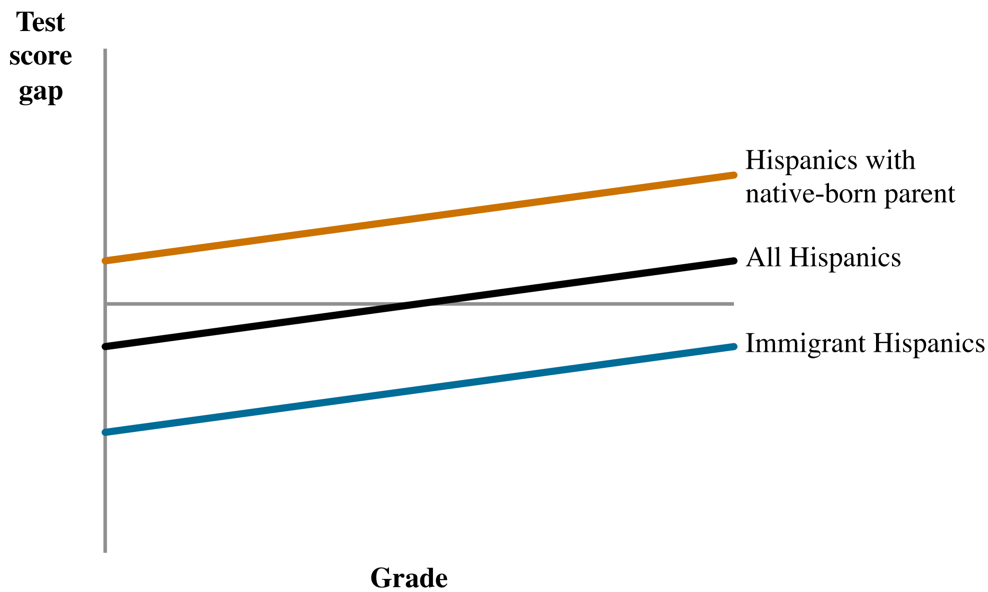
[2] In this paper, "assimilation" takes on two slightly different meanings. Here, I refer to the cultural process of assimilation, whereby immigrants adopt the native language and customs. Assimilation in test scores, like assimilation in wages, occurs when immigrant and native test scores are indistinguishable.

scenario in Panel A. This case leaves no room for differential gains for immigrant Hispanics through the assimilation process.

It could also be the case that Hispanics from established families exhibit almost no change in test scores, like their white and black peers. This scenario would leave immigrant Hispanics responsible for all of the gains. Even here, the overall trendline could mask different levels of academic achievement for immigrant and non-immigrant Hispanics. In Panel B, I depict the case in which positively selected immigrant ancestors give later-generation Hispanics a lingering advantage in school. Then, the progress made by more recently arrived immigrants, though substantial, would not be enough to completely close their gap with whites. In Panel C, later-generation Hispanics appear slightly worse off compared to whites due to negative selection into identification as Hispanic, but more recent arrivals catch up quickly and go on to outscore whites. (I provide more detail on how each of these possibilities could come about in the following section.) In these latter two scenarios, the overall Hispanic trendline understates the gains made by immigrant Hispanic students. In the final one, their test scores are also much higher than previously thought, especially as the students age.

Recognizing the important role of generational status, I decompose the Hispanic-white test score gap across grades. I find that the last scenario, depicted in Panel C of Figure 3.1, holds true. To estimate these gaps, I make use of a rich, administrative data set from North Carolina public schools that contains reading and math scores, as well as socioeconomic background information, for several cohorts of students as they progress from 3rd to 8th grade. Information on immigrant generation for students born in-state comes from matched birth records, and I use a finite mixture distribution to model test score production for unmatched students. Longitudinal data is critical in this type of study so as not to confuse changes in cohort quality with the speed of assimilation, a point first made by Borjas (1985).

(A) Same slope, different intercepts



(B) Positively selected immigrant ancestors

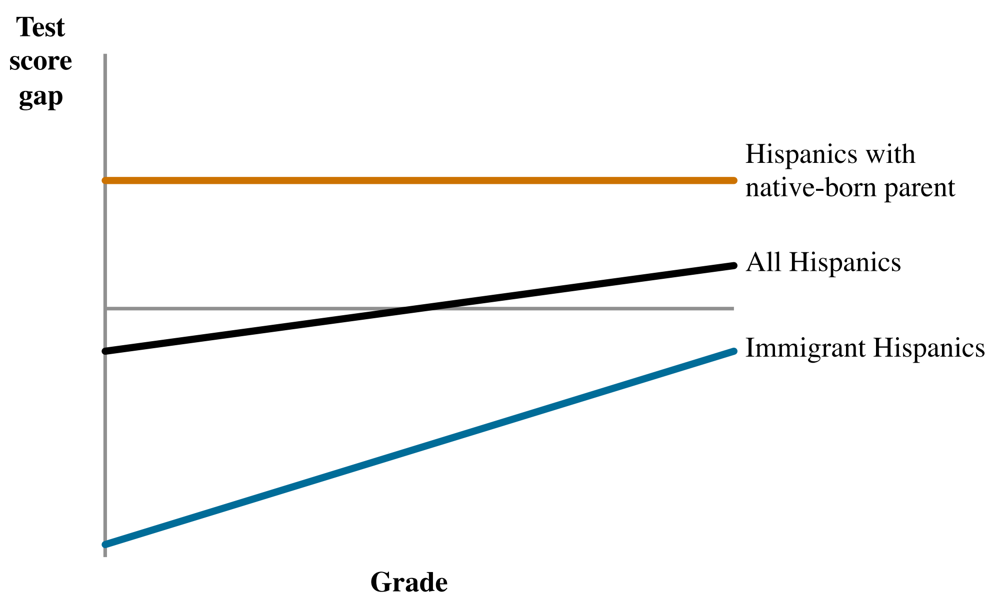


FIGURE 3.1: Three scenarios generating the overall Hispanic trend

(C) Negative selection into ethnic identification

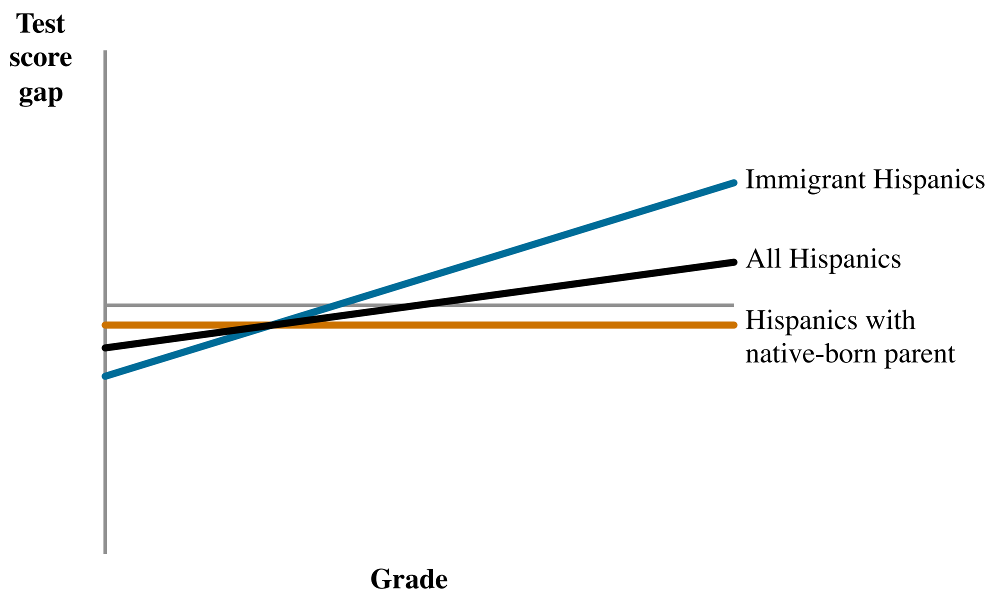


Figure 3.1: Three scenarios generating the overall Hispanic trend

Figure provided for illustrative purposes only.

To the best of my knowledge, no other paper has pinned down trends in test score gaps by immigrant generation with the level of precision found here. The contribution of this paper is threefold: First, I confirm that newly arrived Hispanic immigrants are responsible for the downward trend in Hispanic scores found with repeated cross sections. Previous work could not rule out the possibility that all Hispanic students are more sensitive to moves. Second, I find that later-generation Hispanics, like whites and blacks, consistently score slightly below socioeconomically similar whites. Although there is no evidence that their scores catch up to whites, this finding is of no real economic concern: An accurate measure of ethnic background would put them on par with whites. Most importantly, I show that the children of immigrants drive the growth in Hispanic test scores and that the speed of convergence for immigrant Hispanics is quite rapid. First-generation Hispanics that arrive by 3rd grade equal similar whites within a few years. The second generation makes similarly impressive gains to finish 8th grade 0.11 standard deviations ahead of whites in reading and 0.24 standard deviations ahead in math. Lumping all Hispanic students together understates the considerable progress of these immigrant students. While Borjas (1985) found assimilation in wages to be quite slow, my analysis shows that young immigrants, as well as the children of immigrants, catch up to native whites within a few years.

Two features of my results indicate that cultural assimilation is not the only force at work in the growth in Hispanic test scores. First, second-generation students make the same gains in math and close to the same gains in reading as the first generation. Though there is likely some room for the language skills of these second-generation students to catch up with whites, especially if Spanish is spoken in the home, they still have more years of U.S. experience overall. Second, the test scores of whites and immigrant Hispanics more than converge—second-generation Hispanics perform significantly better than whites in reading and math in 8th grade, the first generation for math only. If Hispanic immigrants adopting native customs was the sole factor behind the growth in Hispanic test scores, we would expect the growth to taper off as Hispanic scores approached average white scores.

Achievement test scores are determined by cognitive inputs, noncognitive inputs, and environmental factors. In terms of the rise in Hispanic immigrant achievement found here, Cunha et al. (2006) rule out changes in IQ after age 10, and controls for school quality rule out the school environment. Thus, noncognitive skills (e.g., motivation and self-discipline) and the home environment must drive the test score gains for children of immigrants. Although I do not have direct evidence, an explanation that is consistent with my findings is that immigrant parents create a home environment that values achievement. This environment gives their children a boost beyond any effects of assimilation. In any case, Hispanic immigrant students fare quite well under the current system, even better than previously suggested. While these students are at a severe socioeconomic disadvantage, their home environment and noncognitive abilities combine to lift them up above the achievement of similar whites.

In the following section, I discuss in greater detail the processes of immigration, assimilation, and selection that could generate the overall Hispanic trend. I also provide an overview of previous work on the Hispanic-white achievement gap and the role of generational status in educational outcomes for immigrants. In Section 3.2, I describe the data sets used in this paper, which motivate the development of my econometric framework in Section 3.3. I report results in Section 3.4, detailing trends in raw and adjusted Hispanic-white test score gaps by immigrant generation as students move through school. Section 3.5 concludes.

3.1 Theoretical and empirical background

3.1.1 Immigration, assimilation, and selection

In their paper on minority achievement gaps, Clotfelter, Ladd, and Vigdor (2009) were the first to find that Hispanic students impove relative to whites with continuous enrollment over the course of grades 3 to 8. In fact, Hispanics score the same as whites in math and reading in 5th grade after adjusting for covariates, and they outperform them by about a tenth of a standard deviation in 8th grade. That Hispanics perform more poorly than whites of comparable socioeconomic status early in school but outscore them in later grades runs counter to the experience of black students. As shown by Clotfelter, Ladd, and Vigdor (2009), among others, black children also have lower test scores in the early grades, but there is no evidence of convergence as the students age.[3] Furthermore, the enormous improvement in Hispanic achievement in these grades cannot be due to changes in cognitive ability, in accordance with the

[3] Fryer and Levitt (2004, 2006) find that the test scores of black Kindergarteners are no different from socioeconomically similar whites, but the difference opens up to commonly observed levels by 3rd grade.

human development literature cited in the review article by Cunha et al. (2006), which states that there are critical periods of development and that IQ becomes stable by age 10.

How do we interpret the relative success of Hispanic children? One interpretation is that Cunha et al. (2006) are wrong and that the underlying cognitive ability of Hispanic children improves relative to their non-Hispanic peers well beyond age 10. Another interpretation, however, recognizes that Hispanics are a heterogeneous group, most importantly heterogeneous in how recently their families migrated to the U.S. Some Hispanic children are themselves immigrants, some are U.S.-born but have immigrant parents, and some are children of Hispanic ancestry whose families have been in the U.S. for three or more generations. Thus, in every grade, the average test score of Hispanic students is a mixture of the performance of these three groups. Differences in the home environments and noncognitive skills among these groups of Hispanics could combine to generate the overall test score pattern.

Distinguishing the performance of Hispanic students by their generation helps shed light on the anomalous finding of steady gains relative to white and black peers. There are arguments for why Hispanics that have been in the U.S. for three or more generations might perform worse, the same, or better than white students. Most scholars believe that across generations there is regression towards the mean in most socioeconomic outcomes. Hispanics have been in the U.S. for fewer generations than whites on average, and it may be that third-generation and higher Hispanics today are the offspring of positively selected immigrant grandparents or great-grandparents. This type of selection would imply a moderately positive but constant Hispanic-white gap for later-generation Hispanics. On the other hand, Duncan and Trejo (2011) show that native-born Hispanics that self-identify as such tend to be negatively selected largely because more successful Hispanics are less likely to retain their Hispanic identity and report it in surveys. Although this kind of selection would give rise to negative gap, it would be no real economic cause for concern; a true measure of Hispanic ethnicity would show that these Hispanics track whites.[4]

There are also arguments for how rapidly and how completely immigrant children can assimilate. While we know that Hispanic students as a whole surpass the performance of similar white students in later grades, whether this stems from immigrant children surpassing white students in later grades has not been established. Although immigrant students surely come in at a disadvantage, this disadvantage might be small enough that rapid growth in their test scores could put them on par with whites in a few years. Any growth beyond this point must come from factors beyond cultural integration since full assimilation only implies that average native achievement and average immigrant achievement are equal. Alternatively, immigrant students could enter so far behind or have such a slow growth rate that I never observe complete convergence with whites during the ages in my sample. A slow convergence rate would be consistent with the slow rate of wage assimilation found in Borjas (1985); however, immigrant children likely catch up to natives faster than immigrant adults.

Given the overall upward trend in Hispanic test scores, I identify three of the more plausible scenarios and illustrate them in Figure 3.1. In Panel A, I show the case where Hispanic families across the board are better at improving their children's achievement scores, but immigrant Hispanics start at a lower level. The overall trendline represents an average of achievement levels but not of slopes. For the next two panels, I assume that Hispanics with native-born parents have flat trendlines relative to whites, as do black students. In Panel B, I depict the case where their achievement lies above whites due to positive selection of their immigrant grandparents and great-grandparents. To generate the overall trendline, recent immigrants must im-

[4] To clarify, a true measure of Hispanic ethnicity would capture whether the student or any of his ancestors (parents, grandparents, etc.) was from Cuba, Mexico, Puerto Rico, South or Central America, or other Spanish culture.

prove rapidly, but they would still score below whites in later grades. This scenario leaves open the possibility that the growth in Hispanic immigrant test scores tapers off as they approach whites. Finally, Panel C shows what happens when there is negative selection into ethnic identification for established Hispanic families. Their test scores then fall slightly below whites in each grade. To make up for the lower scores of later-generation immigrants still identifying as Hispanic, recent immigrants must assimilate quickly and go on to outscore whites by a wide margin. Then, the overall pattern would understate the relative performance of immigrant Hispanic youth in later grades as well as the speed of their convergence. In this final scenario, a factor besides assimilation must drive the growth in Hispanic test scores in these later grades.

3.1.2 Related empirical work

Despite the fact that Hispanics recently overtook blacks as the largest minority group, the Hispanic-white test score gap is almost always discussed as an afterthought to the black-white gap in papers on minority achievement. Outside of sample size concerns when Hispanics were a smaller part of the population, this feature of the literature on racial/ethnic inequality partly came about because of the difficulty in analyzing Hispanic-white differences due to heterogeneity among Hispanics. As they emerge as an important demographic group in their own right, researchers have begun to tackle the particular issues that Hispanics face.

Past studies usually estimate the raw Hispanic-white test score gap as smaller than the black-white gap but vary in how much smaller. For example, Phillips and Chin (2004) place it at 78% of the size of the black-white gap in math in 4th grade and 84% in reading using the National Assessment of Educational Progress (NAEP) data. With a more recent wave of the same data set, Reardon and Galindo (2009) estimate the size of the Hispanic-white achievement gaps in reading and math to be about three quarters of a standard deviation in 4th and 8th grade. Unfortunately, the NAEP is more useful for studying trends across time than how Hispanic-white differences evolve as students progress through school since the data set contains a limited number of grades and samples repeated cross-sections.

The Early Childhood Longitudinal Study-Kindergarten Class (ECLS-K) provides estimates on achievement gaps for some of the youngest ages. While this data is useful for analyzing some types of achievement inequality, language proficiency rules for administering assessments lower its value for studying high-immigrant populations.[5] Nevertheless, Fryer and Levitt (2004, 2006) and Reardon and Galindo (2009) estimate Hispanic-white differences to bounce between -0.3 and -0.5 standard deviations in reading, with no clear trend. In math, they find that the raw gap in the fall of Kindergarten is around -0.8 standard deviations but shrinks to -0.5 standard deviations by the spring of 5th grade.

Crucial to the analysis of the evolution of the Hispanic-white gap for cohorts of students is whether the sample is restricted to continuous enrollees. Without this restriction, researchers find that the size of the gap is constant or even grows as recently arrived immigrants enter the sample. With it, they find progress. In addition, controlling for key background characteristics explains much of the Hispanic-white gap. Clotfelter, Ladd, and Vigdor (2009) condition on continuous enrollment in North Carolina public schools and find that between 3rd and 8th grade, Hispanic students close their reading and math test score gaps with similar whites in 5th

[5] For one, students were not given the reading test if they were not proficient in oral English, meaning that there is severe sample selection into this assessment on the order of 29% for all Hispanics and 77% for first-generation Hispanics in the early waves of the survey. Also, students not proficient in oral English but proficient in oral Spanish were given the math assessment in Spanish. While these math scores do not suffer from the normal language bias, one cannot ascertain how improvement in English language skills contributes to any narrowing of the Hispanic-white math gap. Furthermore, some evidence indicates that a number of Hispanic students were misclassified as being proficient in oral English when parents later reported Spanish as the primary language spoken at home (Reardon and Galindo, 2009). Thus, one cannot be certain that math scores suffer from no language bias.

grade and then finish 8th grade about 0.1 standard deviations ahead. Expanding on this paper, Clotfelter, Ladd, and Vigdor (2012) demonstrate that each group of late arrivers makes progress in the time it stays enrolled in the school system.

Past literature has also examined the impact of immigrant generation and time of arrival on a variety of educational outcomes, but there remains a lack of consensus on this topic as conclusions vary depending on the outcome studied and the racial/ethnic composition of the immigrant pool (see Kao and Tienda, 1995; Rong and Grant, 1992; Chiswick and DebBurman, 2004; Cortes, 2006). With the problematic ECLS-K data, Reardon and Galindo (2009) examine raw test score differences for Mexican students by immigrant generation. In math, they find the magnitude and pattern of raw gaps to be similar for first- and second-generation students. These students begin Kindergarten with math scores 1.1 standard deviations lower than whites and close the gap to around -0.8 standard deviations over the next two years, but it remains relatively constant through 5th grade. The third-generation Mexican students display a similar pattern, though they start school only 0.5 standard deviations below whites. The profile of reading scores resembles that of the math scores for the second and third generations; the authors do not report estimates for the first generation since so few were proficient enough in English to take the reading assessment.

This paper improves upon past work by breaking down trends in the Hispanic-white test score gap for an intact cohort. With this data set and econometric model, I establish which Hispanics are responsible for the trend reversal, the size of the test score gap for each generation, and how they evolve as the students age. In the process, I determine how quickly test scores for immigrant Hispanics converge with whites of a similar socioeconomic status. By focusing on a single ethnic group, I acknowledge that the experience of immigrant Hispanics might be quite different from, for example, immigrant Asians.

3.2 Data

The North Carolina education data is a rich, longitudinal, administrative data set that links information on students, teachers, and public schools over time. The database is maintained by the North Carolina Education Research Data Center (NCERDC), which is housed at Duke University. Every year, the data center processes files received by the North Carolina Department of Public Instruction (NCDPI) and adds them to the existing data. Since the availability of some parts of the data varies over time, I restrict my analysis to certain years. I draw information on students from the End of Grade (EOG) files, which contain reading and math test scores for 3rd through 8th graders as well as student background characteristics such as race/ethnicity, parent's education, and free/reduced price lunch status. Encrypted identifiers follow students over time, even as they change schools, as long as they stay within the universe of North Carolina public schools.

My sample consists of two cohorts of students in 3rd grade in 2000 and 2001. For the 2000 cohort, normal grade progression means that these students would be in 4th grade in 2001, 5th grade in 2002, and so on, up to 8th grade in 2005 (and similarly for the 2001 cohort). Since late arrival into and early exit out of the sample are important variables in my analysis, I make some adjustments so that I do not confuse retention with entering or leaving the sample. Specifically, I follow students who were held back in 3rd grade or any subsequent grade, and I drop any students that enter a cohort in 4th grade or above due to grade repetition.[6] This procedure

[6] With this procedure, I miss students that start attending North Carolina public schools at the same time that they repeat a grade. To check the extent to which this might be a problem, I also estimate my model restricting the sample to students who were born between October 1990 and September 1991 for the 2000 cohort and between October 1991 and September 1992 for the 2001 cohort. The cutoff for entering Kindergarten was October 16 for these cohorts, and the months in these ranges are also the most frequent months of birth, covering 77% of the sample. While the overall pattern of my results remains the same, I do find that all generations of Hispanic students perform worse relative to whites, as compared to the sample that does not drop students who were not the appropriate age for grade.

ensures that my results are valid for the cohorts of students attending 3rd grade in 2000 and 2001 and that late arrival and early exit variables only capture students that enter and exit the universe of North Carolina public schools. Note that for the rest of the paper "grade" means imputed grade, or the grade that a student would be in if he had been promoted each year. Finally, I normalize test scores within each cohort-grade pair.[7] To maintain comparability across time given changes in the racial and ethnic composition of students, I use the mean and standard deviation of the white test score distribution in each cohort-grade pair for normalization.

The NCERDC has also matched students in the education files to birth records supplied by the North Carolina State Center for Health Statistics. These birth records provide even more background information on students, including the location of birth for each of the parents, but unfortunately, I have no records for students born outside of North Carolina. Furthermore, the match rate of 3rd grade education data to birth data is 66% for all students but 19% for Hispanic students. I expect a lower match rate for Hispanics since first-generation immigrants do not have an in-state birth record by definition. Furthermore, some native-born Hispanics in North Carolina moved to the state between birth and the end of 3rd grade. If all Hispanic students born in North Carolina were matched, the education to birth data match rate would be 32%.[8] I make use of the birth certificate information when possible and let my econometric model take care of the rest. When a student's birth certificate indicates that at least one parent was born outside the country, I classify him as a second-generation immigrant.[9] Students with both parents born in the U.S. are categorized as third-generation, but in reality, their roots could reach back even further.

[7] Test scores are reported on the same developmental scale regardless of grade, so I can include retained students in the normalization.

[8] Calculated from the Census/ACS. See Appendix B for the full details.

[9] This definition of second-generation is consistent with other social science research.

For the group of Hispanic students without birth records, I rely on the Census and the American Community Survey (ACS) to fill in the distribution of immigrant generations for school-age Hispanic youth not born in North Carolina (Ruggles et al., 2010). To ensure the samples are comparable, I use school-age Hispanics living in North Carolina and attending public school from the 2000 Census and 2001-2006 waves of the ACS. I choose the 2000 and 2001 3rd grade cohorts so that the Census/ACS and NCERDC data line up in terms of available information. Earlier cohorts leave me without corresponding years in the Census/ACS. In later ones, the NCERDC does not provide free/reduced lunch status and parent education, the key variable to predict immigrant generation.

While the North Carolina education data is not nationally representative, there are several reasons that make the state interesting to use as a case study. First, North Carolina is in fact representative of the country in a few key ways. It features a diversified economy with a mixture of urban, suburban, and rural areas spread across the Appalachian Mountains, the Piedmont, and the Atlantic coast. Hispanics reside in all parts of the state, drawn to the demand for workers in such industries such as agriculture, construction, service, and retail (Gill, 2010). Second, while Hispanics make up a smaller proportion of North Carolina's population than they do nationally (8.4% versus 16.3% as of 2010), the state has experienced astounding growth in the number of Hispanics residing there in the past two decades. Therefore, the state could serve as a model for what will happen in other parts of the nation.

3.3 Econometric model

If the immigrant generation for every Hispanic student was known, I could estimate generational test score gaps by ordinary least squares, taking the following equation as the model:

$$y_{ik} = g_i'\alpha_k + x_{ik}'\beta_k + \varepsilon_{ik} \tag{3.1}$$

where y_{ik} is the test score for student i in subject k, with $k = r$ for reading and $k = m$ for math; g_i is a vector of indicators for each Hispanic immigrant generation; and x_{ik} is a vector of other covariates, including indicators for the non-Hispanic racial groups. The parameter vector α_k then contains the Hispanic-white test score gap for each immigrant generation in subject k. However, I must take a different approach to uncover these values since since I do not observe generation for the majority of Hispanic students.

I estimate the effect of generational status on test scores by maximum likelihood with a two-part likelihood function: The first component, which I use for observations where g_i is observed, comes from the standard likelihood for a linear model with a normally distributed error. The second component departs slightly from the standard likelihood to a finite mixture model since I must incorporate the fact that I do not observe immigrant generation for these students. In Appendix C, I show that estimates from my model are not substantially different from estimates obtained with ordinary least squares, as far as the overall Hispanic-white test score gap is concerned.

3.3.1 Likelihood function

In order to take full advantage of the information available, I jointly estimate achievement gaps for reading and math. This aspect of the setup is important for students with an unknown immigrant generation since it utilizes the correlation between the two scores; in other words, I use the fact that a high reading score may indicate that a student is of a higher immigrant generation in my estimation of the math coefficients, and vice versa. Specifically, I model the distribution of the shocks ε_{im} for math and ε_{ir} for reading as bivariate normal, where I parameterize the covariance matrix such that ρ is the correlation between the shocks and σ_k^2 is the variance of subject k's shock.

$$\begin{bmatrix} \varepsilon_{ir} \\ \varepsilon_{im} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_r^2 & \rho\sigma_r\sigma_m \\ \rho\sigma_r\sigma_m & \sigma_m^2 \end{bmatrix} \right) \tag{3.2}$$

For a non-Hispanic student or a Hispanic student with a birth certificate and thus a known immigrant generation, the contribution to the likelihood function is

$$L_{1i} = \Phi(y_{ir}, y_{im}|g_i, x_{ir}, x_{im}, \theta) \tag{3.3}$$

where $\Phi(\cdot)$ is the bivariate normal probability density function, and θ contains all the parameters.[10] For a Hispanic student with an unknown immigrant generation, the likelihood contribution is

$$L_{2i} = \sum_g \pi_{ig}\Phi(y_{ir}, y_{im}|g, x_{ir}, x_{im}, \theta) \tag{3.4}$$

where π_{ig} is the probability that student i belongs to generation g. In practice, this value is estimated in a first stage, as discussed below. Intuitively, this estimation method works because the objective function gives greater weight to observations that are likely to be of a particular immigrant generation when it determines the effect of being a member of that generation. The combined likelihood for all observations is

$$L = \prod_{i=1}^{n_1} L_{1i} \times \prod_{i=1}^{n_2} L_{2i} \tag{3.5}$$

Finally, taking logs and denoting the log-likelihood as ℓ, the function that I maximize

[10] That is,

$$\Phi(y_{ir}, y_{im}|g_i, x_{ir}, x_{im}, \theta) = \frac{1}{2\pi\sigma_r\sigma_m\sqrt{1-\rho^2}} \exp\left\{ -\frac{1}{2(1-\rho^2)} \left[\frac{(y_{ir} - g_i'\alpha_r - x_{ir}'\beta_r)^2}{\sigma_r^2} + \frac{(y_{im} - g_i'\alpha_m - x_{im}'\beta_m)^2}{\sigma_m^2} - \frac{2\rho(y_{ir} - g_i'\alpha_r - x_{ir}'\beta_r)(y_{im} - g_i'\alpha_m - x_{im}'\beta_m)}{\sigma_r\sigma_m} \right] \right\}$$

to estimate the model is

$$\ell = \sum_{i=1}^{n_1} \ln \Phi(y_{ir}, y_{im}|g_i, x_{ir}, x_{im}, \theta) + \sum_{i=1}^{n_2} \ln \left[\sum_g \pi_{ig} \Phi(y_{ir}, y_{im}|g, x_{ir}, x_{im}, \theta) \right] \tag{3.6}$$

In practice, some math and reading scores are missing, and this issue is more prevalent for Hispanic students' reading scores since students with LEP (limited English proficient) status are exempt from the reading tests. (They are still required to take the math tests.) By incorporating information about whether or not a student took the other type of exam, I gain an even better hold on the immigrant generation to which a student belongs. Accordingly, I include a dummy in the vector x_{ik} for whether the student took the other type of exam. When a student is missing one exam but not the other, I replace the bivariate normal density with a univariate normal density for the test the student did take.

3.3.2 Estimation with school fixed effects

Due to computing and data constraints, I cannot estimate a model with school fixed effects using traditional methods (i.e., direct estimation of the school effects or de-meaning all variables using the within transformation). Therefore, I adopt the following iterative procedure.

1. Given initial guesses $(\hat{\alpha}_k, \hat{\beta}_k)$ for $k = \{r, m\}$, calculate the average residual for each school s in the set of schools:

$$\hat{\mu}_{sk} = \frac{1}{n_s} \sum_{i \in s} \hat{\varepsilon}_{ik} \tag{3.7}$$

$$\hat{\varepsilon}_{ik} = y_{ik} - g_i' \hat{\alpha}_k - x_i' \hat{\beta}_k \tag{3.8}$$

   When g_i is not observed, I substitute the individual's immigrant generation probability distribution π_i. These average residuals by school represent the initial guesses for the school fixed effects.

2. Estimate the model with $(y_{ik} - \hat{\mu}_{sk})$ as the dependent variable to obtain new parameter guesses $(\hat{\alpha}_k, \hat{\beta}_k)$.

3. Repeat steps 1 and 2 until the estimated parameters converge.[11]

3.3.3 First-stage estimation of probability weights

I estimate the weights π_{ig} using data from the Census and the ACS. Since both the education data and the Census/ACS contain parent education, I construct weights based on this characteristic. Specifically, I take the population of school-age Hispanic students born outside of but living North Carolina and attending public school. Then, I calculate the probability that a student is of each immigrant generation given parent's educational attainment in five categories.[12] Finally, I adjust the weights to correct for undermatching between the birth and education data. I explain this adjustment procedure and the reasoning behind it in more detail in Appendix B. While I prefer the weights based on parent's education, unconditional weights yield the same qualitative results, which are also numerically very similar.[13]

3.4 Results

3.4.1 Descriptive statistics

Table 3.1 gives descriptive statistics for outcomes and covariates for the sample of students that are part of the cohorts in 3rd grade in 2000 and 2001. Pooling all grades together for each student gives over 1.2 million student-grade observations with an average of 4.7 grade observations per student, 3.7 for Hispanic students. In this sample, the average Hispanic-white reading gap is -0.74 standard deviations, and the

[11] In practice, I stop iterating when the percent change in the log-likelihood is below a very small threshold.

[12] For some years in my sample, I could also include location at the Super PUMA level in the calculation of the probability weights. This addition barely alters the results.

[13] Specifically, no generational coefficient moves by more than 0.06 standard deviations, and most are within 0.02 standard deviations in my adjusted gaps model.

average math gap is -0.61 standard deviations. The observation frequencies by grade in Table 3.2 justify normalization to the white mean and standard deviation for each cohort-grade combination—the number of white students in each grade is stable relative to the number of Hispanics. Due to this normalization, we would expect the year-to-year change in white test scores to be close to zero, though perhaps slightly greater than zero if there is positive selection into staying in the sample. For white math scores, this selection story appears to hold. However, the average Hispanic student gains about 0.06 standard deviations in each subject, conditional on having a score in the previous year. Given that there are fewer observations per student for Hispanics, the selection into having a previous year's score is likely higher. Nevertheless, this growth gives us a sense of the gains Hispanic students make with continuous enrollment.

This sample is 59% white, 31% black, and 5% Hispanic, while the remainder are another (single) race or multiracial.[14] Since Hispanic students move in and out of North Carolina public schools more often, the racial/ethnic composition looks slightly different without the grades pooled. This aspect of the data is also reflected in the proportion of student-grade observations that arrive late (defined as missing from the 3rd grade data) and exit early (missing from the 8th grade data). Most notably, 37% of Hispanic observations come from late arrivers, compared to 11% for whites, but the 13% probability that a Hispanic observation represents an early exiter is also significantly greater than the white probability. While white and Hispanic observations are equally likely to be missing a math score (0.1%), given that they have test score data, Hispanics are about 5 times as likely (0.4% versus 1.9%) not to have a reading score.

[14] For the race/ethnicity variable, the respondent could only mark one of the following: American Indian, Asian, Black, Hispanic, Multiracial, White, or Other. I interpret black as non-Hispanic black and white as non-Hispanic white, though I cannot know definitively what the student, parent, or teacher thought when marking the child's race/ethnicity.

Table 3.1: Descriptive statistics, NCERDC data

	All students	White	Hispanic
Reading score	-0.307	0.000	-0.743
	(0.001)	(0.001)	(0.004)
Change in reading score	0.001	0.000	0.062
	(0.001)	(0.001)	(0.003)
Math score	-0.301	0.000	-0.613
	(0.001)	(0.001)	(0.004)
Change in math score	0.011	0.004	0.063
	(0.001)	(0.001)	(0.003)
Race/ethnicity			
White	0.593		
Black	0.308		
Hispanic	0.050		
Other race	0.032		
Multiracial	0.018		
Parent's education			
Less than high school	0.124	0.081	0.467
High school graduate	0.462	0.423	0.412
Junior college or trade school	0.166	0.184	0.056
Four-year college	0.203	0.272	0.054
Graduate school	0.041	0.057	0.011
Free/reduced lunch	0.573	0.386	0.898
Female	0.492	0.490	0.495
Missing reading score	0.006	0.004	0.019
Missing math score	0.001	0.001	0.001
Late arrival	0.132	0.113	0.374
Early exit	0.092	0.091	0.127
Student-grade observations	1,257,894	745,796	63,365

Standard errors in parentheses. Sample: pooled grades 3-8 of the cohorts in 3rd grade in 2000 and 2001.

Table 3.2: Sample sizes by cohort and grade, NCERDC data

	All students		White		Hispanic	
	2000	2001	2000	2001	2000	2001
3rd	100,852	101,963	61,982	60,934	3,212	4,542
4th	101,303	102,481	61,821	60,954	3,668	4,964
5th	102,243	105,344	62,005	61,755	4,179	5,766
6th	105,046	106,953	62,628	62,111	4,932	6,330
7th	107,075	108,199	63,235	62,386	5,598	6,848
8th	107,897	108,538	63,540	62,445	6,072	7,245
Total	1,257,894		745,796		63,356	

Sample: cohorts in 3rd grade in 2000 and 2001.

The two socioeconomic status variables starkly reveal the relative average disadvantage of Hispanic students. Though about 40% of both whites and Hispanics have high school graduate for parent's education, the distributions skew in different directions: Forty-seven percent of Hispanic student-grade observations have parent educational attainment of less than high school compared to 8% for whites. A white student is much more likely to have a parent who has completed some level of higher education. This disparity in socioeconomic status shows up again in the proportion of students who receive free or reduced price lunch, which is only 39% for white students but 90% for Hispanics.

In Table 3.3, I report summary statistics for Hispanic students in the NCERDC data based on what I know about the students' immigrant generation. I separate the students not matched to a birth certificate by whether they arrived in North Carolina public schools by 3rd grade. From the Census and ACS (see Table 3.5), we know that about half the Hispanic students not born in North Carolina are first-generation immigrants. Between the late and early arrivers to North Carolina, there is remarkably little difference in socioeconomic status as represented by parent's education and subsidized lunch status. However, the reading test scores of the late arrivers are half a standard deviation lower and the math scores 0.38 standard deviations lower than the early arrivers. Even the early arrivers score about half a standard deviation

Table 3.3: Descriptive statistics, Hispanics in the NCERDC data

	Unknown generation		Known generation	
	Late arriver	Early arriver	2nd gen	3rd gen
Reading score	-1.012	-0.507	-0.621	-0.413
	(0.008)	(0.006)	(0.013)	(0.016)
Change in reading score	0.119	0.043	0.044	0.003
	(0.006)	(0.004)	(0.009)	(0.011)
Math score	-0.862	-0.482	-0.468	-0.400
	(0.006)	(0.005)	(0.012)	(0.015)
Change in math score	0.122	0.040	0.051	0.004
	(0.005)	(0.003)	(0.008)	(0.010)
Parent's education				
Less than high school	0.465	0.475	0.584	0.249
High school graduate	0.417	0.406	0.350	0.518
Junior college or trade school	0.050	0.056	0.038	0.117
Four-year college	0.057	0.052	0.024	0.099
Graduate school	0.011	0.012	0.005	0.018
Free/reduced lunch	0.899	0.908	0.926	0.769
Female	0.479	0.500	0.519	0.516
Missing reading score	0.039	0.008	0.007	0.005
Missing math score	0.001	0.001	0.000	0.001
Late arrival	1.000	0.000	0.082	0.097
Early exit	0.145	0.133	0.048	0.079
Student-grade observations	22,881	31,147	5,490	3,838

Standard errors in parentheses. Sample: pooled grades 3-8 of the cohorts in 3rd grade in 2000 and 2001.

below whites in both subjects. Late arrivers gain 0.12 standard deviations in math and reading from the previous year, though the caveat of having a previous year's test score applies more strongly for these students. Still, this change is three times the average gain of early arrivers. Finally, the late arrivers are missing a reading score in 3.9% of cases, compared to 0.8% for early arrivers.

Turning to students with a known immigrant generation, the third-generation Hispanic students lag behind whites by about 0.4 standard deviations in reading and math. The second generation's respective gaps are -0.62 and -0.47 standard deviations. While the students I know to be third-generation have the same average change in test scores as whites, known second-generation Hispanic students improve 0.05 standard deviations relative to whites. Not only are test score levels different by immigrant generation and time of arrival, so are the slopes. I also observe differences in the socioeconomic status of these two groups, as far as I can tell from the birth record matches. These third-generation Hispanic students on average have parents with less education than white students, and these second-generation students in turn have parents with less education than the third generation. More parents of second-generation Hispanics did not finish high school than did. For free/reduced price lunch, all groups of Hispanics in Table 3.3 have high levels of enrollment in the program, but the known third generation enrolls 77% compared to over 90% for the known second-generation and unknown generation students.

The Census and ACS provide a more complete classification of the immigrant generation of school-age Hispanic youth living in North Carolina as well as more background characteristics. I present descriptive statistics from this data for the years corresponding to the education sample (2000-2006) in Table 3.4.[15] Turning to average characteristics by immigrant generation, I generally find that a higher generational status means that a student comes from a more advantaged household. Average parent's education, household income, and mother's age at birth all suggest that third-generation Hispanic youth are positioned to do well in school. Exceptions are food stamp recipiency and mother's marital status, which appear to give first-generation children an edge. However, given household income, the low take-up rate of food stamps among families with first-generation Hispanic children likely reflects less familiarity with resources for low-income households.

[15] I observe some differences between the reported parent's education distributions in the NCERDC data for Hispanic youth (compare the last column of Table 3.1 to the last column of Table 3.4). This disparity could arise from conditioning on observing a test score in the NCERDC sample, teacher reporting of parent's education, or the fact that in the Census/ACS sample, I only have

Table 3.4: Descriptive statistics, Hispanics in the Census and ACS data

	1st gen	2nd gen	3+ gen	All Hispanics
Parent's education				
Less than high school	0.497	0.578	0.152	0.432
High school graduate	0.306	0.327	0.494	0.367
Junior college or trade school	0.054	0.038	0.148	0.074
Four-year college	0.079	0.041	0.143	0.082
Graduate school	0.065	0.017	0.063	0.045
Household income ($1000s)	40.9	37.6	53.9	42.9
	(1.8)	(1.6)	(2.5)	(1.1)
Food stamp recipiency	0.109	0.191	0.186	0.160
Mother's age at birth	24.2	25.3	26.1	25.2
	(0.4)	(0.2)	(0.3)	(0.2)
Mother married	0.810	0.716	0.718	0.748
Mother's employment status				
Employed	0.509	0.505	0.607	0.536
Unemployed	0.083	0.083	0.099	0.088
Not in labor force	0.408	0.412	0.294	0.377
Missing mother's characteristics	0.213	0.119	0.054	0.136
Migration status, one year				
Same house	0.701	0.759	0.776	0.742
Moved within state	0.179	0.164	0.156	0.167
Moved between states	0.034	0.063	0.060	0.051
Abroad one year ago	0.087	0.014	0.008	0.039
Place of origin				
Mexico	0.760	0.753	0.479	0.683
Central America	0.096	0.100	0.050	0.085
South America	0.065	0.022	0.049	0.045
Puerto Rico	0.036	0.040	0.285	0.104
Cuba	0.015	0.008	0.019	0.013
Other/not specified[a]	0.027	0.078	0.118	0.070
Maximum observations	1175	1225	1024	3424

Standard errors in parentheses. Sample: school-age Hispanic youth living in North Carolina and attending public school, 2000-2006. Used person-level weights provided by IPUMS.
[a] Other countries represented are Spain and the Dominican Republic, but most of these respondents did not specify a particular country or region.

Since first- and second-generation youth both have immigrant parents, one would expect their household characteristics to be similar but perhaps with higher socioeconomic status for households with parents that migrated before the birth of their child since they have resided in the U.S. for longer. Indeed, this pattern is close to what I find: Average household income is statistically no different, and the parent's education distributions appear to be similar. Mother's age at birth is understandably lower for first-generation children since some mothers migrate in the middle of their fertile years. Finally, over three quarters of first- and second-generation children come from Mexico with another 10% from Central America, whereas about half of third-generation children trace their origins to Mexico and a quarter to Puerto Rico.

3.4.2 Probability weights

Before estimating the test score model, I compute probability weights using the sample of school-age Hispanic youth attending public school and living in but not born in North Carolina. I then adjust the weights for undermatching using the procedure described in Appendix B. From Table 3.5, we see that there is substantial heterogeneity in generational status by parent's education for these unmatched records. To determine how parent's education helps predict immigrant generation, compare the probability of each parent education level to the overall probability in the last row of Table 3.5. For these students, higher educational attainment by a parent is generally associated with a higher generation.

3.4.3 Test score results

I lay out my test score results in a series of three sets of graphs, which correspond to different specifications of the test score model. Each graph plots the implied overall Hispanic-white test score gap by grade in addition to the gaps by immigrant generation. I present estimates for the implied gaps in Appendix C.

parent's education for youth living with at least one parent.

Table 3.5: Probability weights by parent's education

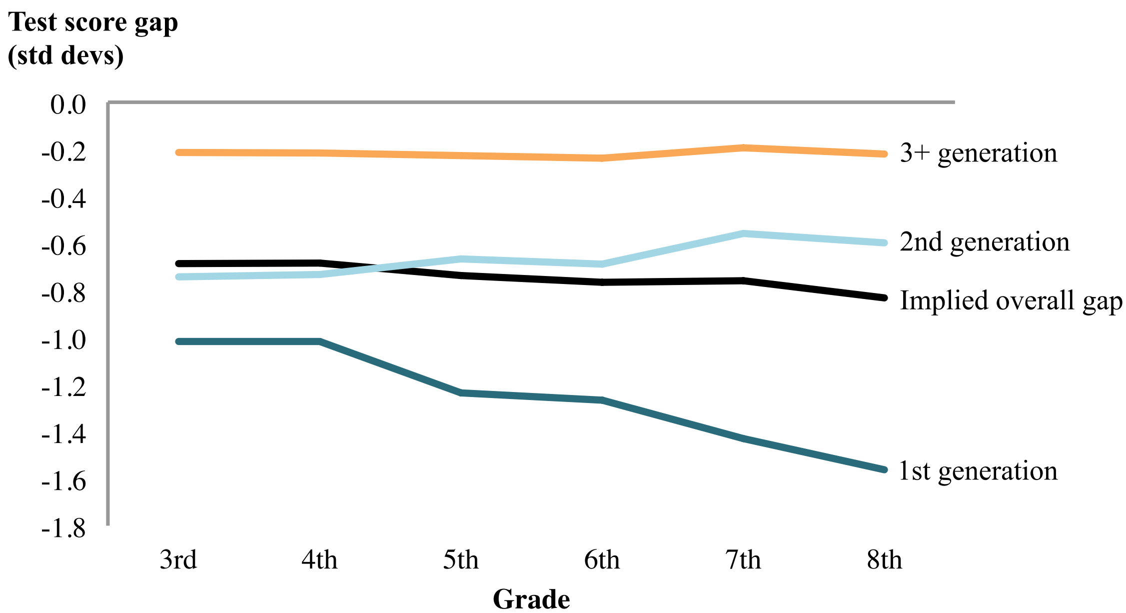
	1st gen	2nd gen	3+ gen
Less than high school	0.461	0.450	0.089
High school graduate	0.324	0.329	0.347
Junior college or trade school	0.269	0.285	0.446
Four-year college	0.367	0.175	0.458
Graduate school	0.513	0.146	0.341
All	0.452	0.332	0.216

Sample: school-age Hispanic youth attending public school and living in but not born in North Carolina, 2000-2006. The person-level weights provided by IPUMS are adjusted to account for unmatched education records (see text for details).

In Figure 3.2, I show the raw Hispanic-white test score gaps by grade for reading and math. The corresponding estimates and standard errors are in Tables 3.6 and 3.7. These unconditional differences reveal how far behind the average Hispanic student of some immigrant generation is compared to the average white student. A few features of these graphs stand out: First, each generation comes closer to eliminating the Hispanic-white test score gap in reading and math, though the difference between generations varies by grade. Second, the first generation drives the downward trend in the overall Hispanic-white gap, so much, in fact, that the overall trendline masks gains made by the second generation between 3rd and 8th grade. The third generation flatlines over this time period, showing no notable improvement or decline relative to whites. Clotfelter, Ladd, and Vigdor (2009) also find this modest decline in the overall Hispanic-white gap, but these graphs show more nuance.

Next, we turn to Figure 3.3, which plots raw gaps across grades for the population of students that enrolls continuously in school (estimates in Tables 3.8 and 3.9). To add these controls while still keeping the model tractable, I introduce a late arrival indicator for students without any 3rd grade test scores and an early exit indicator for students without any 8th grade test scores. Since I expect these effects to be different for first-generation Hispanics as late arrival more often than not means

(A) Reading



(B) Math

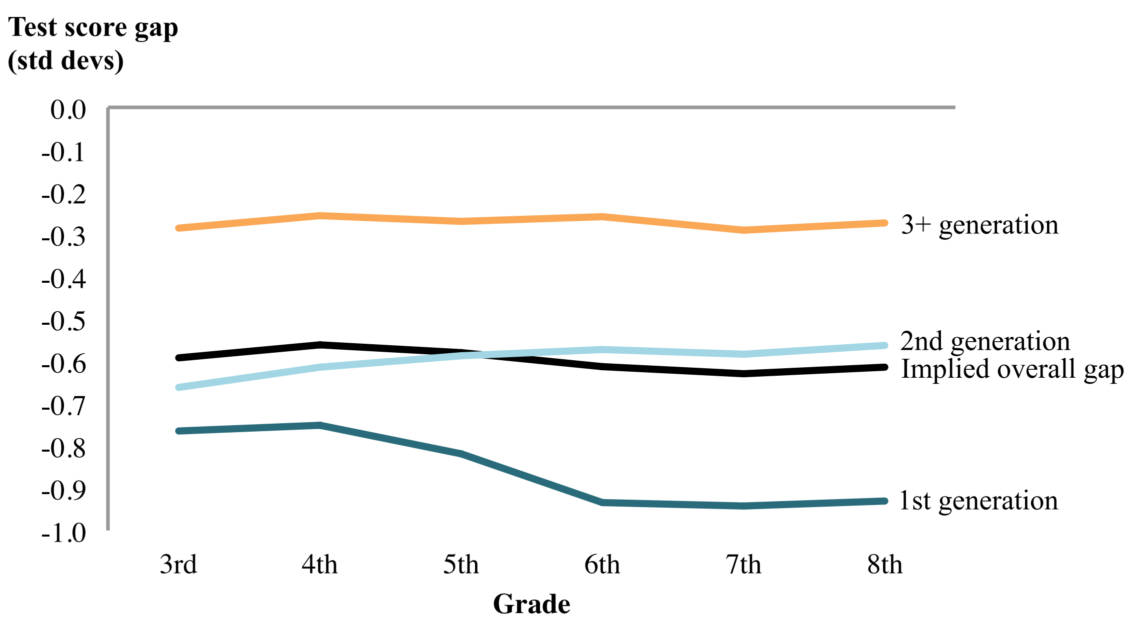


Figure 3.2: Raw Hispanic-white test score gaps by grade

Sample: cohorts in 3rd grade in 2000 and 2001. Regressions also include indicators for black, other race, multiracial, and missing other test score. See Tables 3.6 and 3.7 for point estimates and standard errors.

Table 3.6: Raw reading gaps by grade

Grade	3rd	4th	5th	6th	7th	8th
1st gen Hispanic	-1.017	-1.016	-1.235	-1.266	-1.429	-1.563
	(0.050)	(0.043)	(0.033)	(0.029)	(0.025)	(0.023)
2nd gen Hispanic	-0.742	-0.732	-0.666	-0.689	-0.557	-0.597
	(0.035)	(0.035)	(0.025)	(0.027)	(0.022)	(0.022)
3+ gen Hispanic	-0.213	-0.216	-0.227	-0.238	-0.194	-0.220
	(0.027)	(0.028)	(0.027)	(0.023)	(0.023)	(0.023)
Black	-0.780	-0.822	-0.830	-0.840	-0.837	-0.848
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Other race	-0.368	-0.369	-0.427	-0.378	-0.340	-0.332
	(0.013)	(0.013)	(0.013)	(0.012)	(0.013)	(0.012)
Multiracial	-0.240	-0.230	-0.230	-0.233	-0.241	-0.225
	(0.018)	(0.017)	(0.019)	(0.019)	(0.017)	(0.016)
Constant	0.000	0.000	0.000	0.001	0.000	0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Missing math score	-1.186	-1.037	-1.207	-1.463	-1.323	-1.388
	(0.100)	(0.082)	(0.069)	(0.084)	(0.060)	(0.056)

Standard errors in parentheses. Sample: cohorts in 3rd grade in 2000 and 2001.

Table 3.7: Raw math gaps by grade

Grade	3rd	4th	5th	6th	7th	8th
1st gen Hispanic	-0.765	-0.751	-0.819	-0.935	-0.942	-0.930
	(0.035)	(0.030)	(0.028)	(0.028)	(0.021)	(0.020)
2nd gen Hispanic	-0.662	-0.613	-0.587	-0.572	-0.583	-0.563
	(0.027)	(0.024)	(0.023)	(0.027)	(0.020)	(0.020)
3+ gen Hispanic	-0.286	-0.256	-0.270	-0.259	-0.291	-0.273
	(0.028)	(0.026)	(0.028)	(0.024)	(0.023)	(0.024)
Black	-0.876	-0.831	-0.841	-0.849	-0.816	-0.832
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Other race	-0.280	-0.249	-0.234	-0.175	-0.129	-0.094
	(0.013)	(0.013)	(0.012)	(0.011)	(0.012)	(0.012)
Multiracial	-0.293	-0.276	-0.279	-0.275	-0.313	-0.298
	(0.018)	(0.016)	(0.017)	(0.018)	(0.016)	(0.016)
Constant	0.005	0.006	0.005	0.002	0.002	0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Missing reading score	-0.851	-0.861	-1.006	-1.242	-1.072	-1.105
	(0.026)	(0.022)	(0.027)	(0.035)	(0.031)	(0.034)

Standard errors in parentheses. Sample: cohorts in 3rd grade in 2000 and 2001.

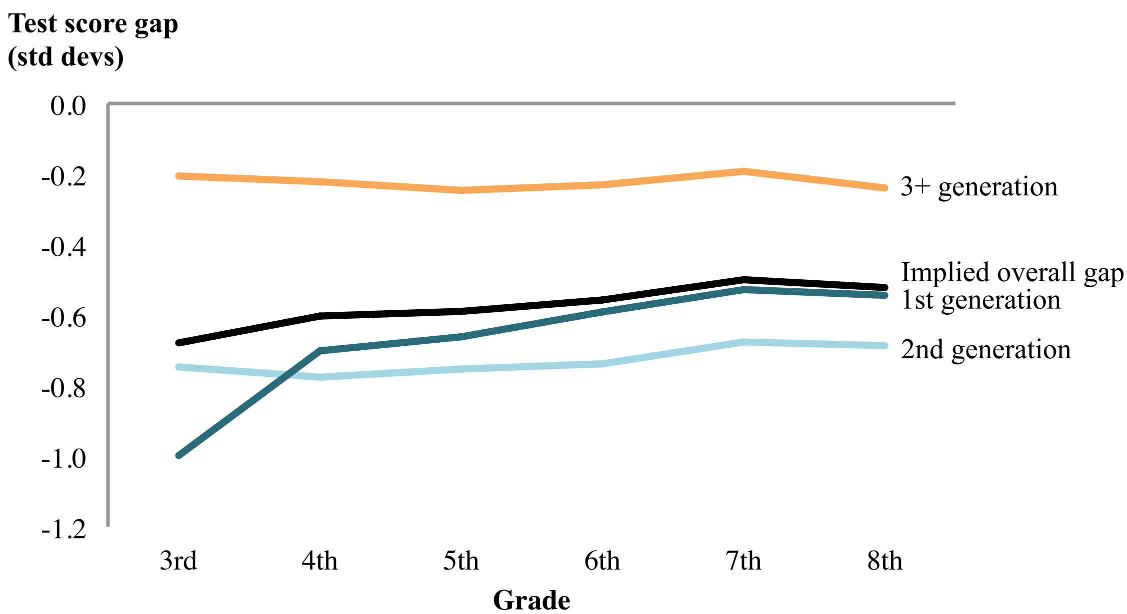
that the student just arrived in the U.S., I interact each of these indicators with first-generation Hispanic.[16] Indeed, Figure 3.3 shows a reversal in the trendlines for the first generation, which also flips the downward trend in the implied overall gap. This finding mimics the reversal in the overall Hispanic-white test score gaps that Clotfelter, Ladd, and Vigdor (2009) find after restricting their sample to continuous enrollees.[17] This paper shows that the immigrant students arriving in the schools drive the change, as opposed to all Hispanic students scoring lower after a move. In reading, first-generation Hispanics who arrive by 3rd grade and continuously enroll improve from -1.00 standard deviations to -0.54 standard deviations—an increase of 0.08 standard deviations per year. Second-generation students gain a modest 0.06 standard deviations to finish 8th grade with a gap of -0.69 standard deviations. The slopes of the math trendlines are similar for these earlier generations, but the average sizes of the gaps are about 0.1 standard deviations smaller. In contrast, third-generation Hispanic students hover around 0.2 standard deviations below whites in reading and 0.3 standard deviations below in math. As a point of comparison, the raw black-white gap in both subjects is consistently around -0.8 standard deviations, almost always below the first generation's gap with whites.

The coefficients on late arrival in Tables 3.8 and 3.9 show how this change takes place. We cannot place a purely causal interpretation on the late arrival and early exit coefficients; technically, these variables only represent whether a student was linked to an earlier or later test score record. Nevertheless, most of these students are moving in and out of the state, with some transfers in and out of private school

[16] When I allow full interactions of late arrival and early exit with each race/ethnicity and Hispanic immigrant generation, I do find that some of these interactions vary significantly by group. However, there is no clear pattern across grades outside of the main effects and first-generation interactions. Furthermore, only the interactions with first-generation Hispanic have a meaningful impact on the corresponding main effect (i.e. move the coefficient beyond the second decimal place). In other words, a complete set of interactions does not substantially change the second-generation and third-generation Hispanic coefficients.

[17] Comparable estimates for the years in my sample are in Appendix Table C.2.

(A) Reading



(B) Math

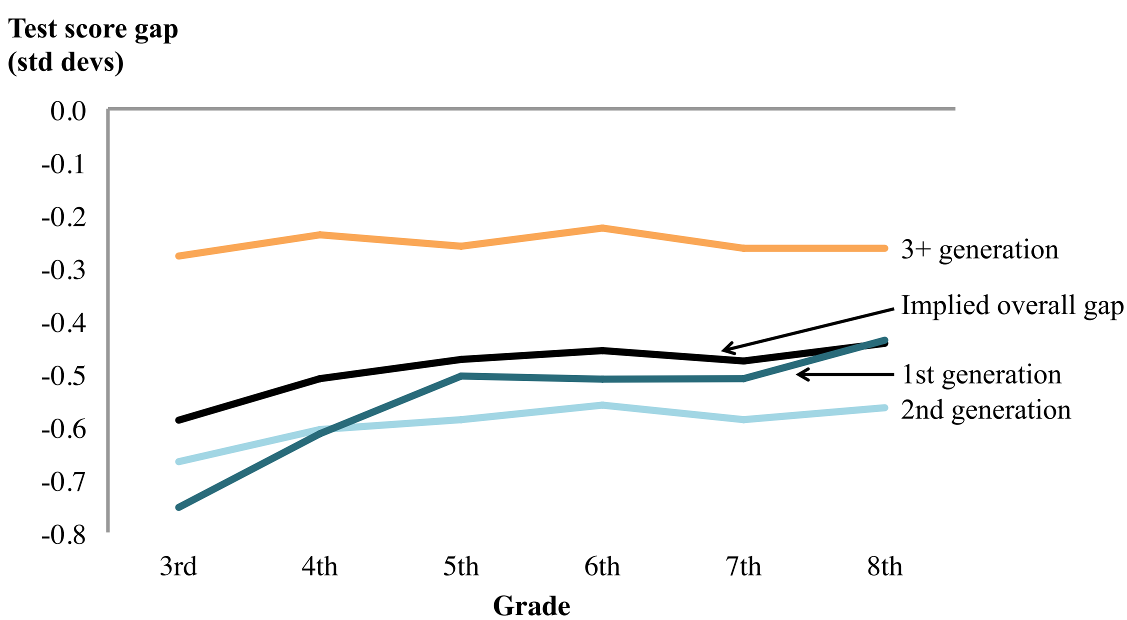


Figure 3.3: Raw Hispanic-white test score gaps by grade for continuous enrollees

Sample: cohorts in 3rd grade in 2000 and 2001. Test score gaps are adjusted for late arrival and early exit, main effects and interactions with first-generation Hispanic. Regressions also include indicators for black, other race, multiracial, and missing other test score. See Tables 3.8 and 3.9 for point estimates and standard errors.

Table 3.8: Raw reading gaps by grade with continuous enrollment

Grade	3rd	4th	5th	6th	7th	8th
1st gen Hispanic	-0.998	-0.701	-0.662	-0.591	-0.527	-0.544
	(0.033)	(0.043)	(0.120)	(0.045)	(0.112)	(0.038)
2nd gen Hispanic	-0.747	-0.776	-0.753	-0.738	-0.676	-0.686
	(0.029)	(0.029)	(0.057)	(0.027)	(0.028)	(0.022)
3+ gen Hispanic	-0.206	-0.221	-0.246	-0.231	-0.192	-0.239
	(0.026)	(0.027)	(0.043)	(0.024)	(0.023)	(0.026)
Late arrival		-0.094	-0.069	-0.111	-0.077	-0.081
		(0.009)	(0.009)	(0.006)	(0.007)	(0.006)
Early exit	-0.111	-0.103	-0.150	-0.286	-0.416	
	(0.007)	(0.007)	(0.008)	(0.008)	(0.009)	
1st gen Hispanic		-0.849	-1.051	-0.994	-1.140	-1.312
X Late arrival		(0.064)	(0.144)	(0.057)	(0.087)	(0.048)
1st gen Hispanic	0.014	-0.015	0.210	0.103	0.095	
X Early exit	(0.059)	(0.079)	(0.084)	(0.062)	(0.152)	
Black	-0.782	-0.823	-0.831	-0.837	-0.832	-0.847
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Other race	-0.366	-0.364	-0.419	-0.366	-0.329	-0.325
	(0.014)	(0.013)	(0.012)	(0.013)	(0.012)	(0.013)
Multiracial	-0.233	-0.219	-0.218	-0.213	-0.219	-0.215
	(0.018)	(0.018)	(0.017)	(0.017)	(0.017)	(0.017)
Constant	0.019	0.019	0.023	0.039	0.034	0.016
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Missing math score	-1.168	-0.987	-1.171	-1.351	-1.170	-1.370
	(0.147)	(0.062)	(0.065)	(0.074)	(0.069)	(0.051)

Standard errors in parentheses. Sample: cohorts in 3rd grade in 2000 and 2001.

or home school. From Table 3.4, we know that 72% of first-generation Hispanics that moved from out of state in the past year actually moved from another country. For reading, there is a relatively modest impact, usually around 0.1 standard deviations, for entering and leaving North Carolina public schools. However, the additional effect of late arrival for first-generation Hispanics is tremendous, ranging from -0.83 to -1.3 standard deviations. This group suffers no perceptible additional penalty for exiting early. In math, all students experience some drop in test scores for arriving late and exiting early, and again this drop is larger for first-generation Hispanics. However, their additional penalty is smaller than in reading, ranging between -0.35 and -0.58

Table 3.9: Raw math gaps by grade with continuous enrollment

Grade	3rd	4th	5th	6th	7th	8th
1st gen Hispanic	-0.753	-0.613	-0.504	-0.511	-0.509	-0.437
	(0.032)	(0.034)	(0.075)	(0.044)	(0.096)	(0.037)
2nd gen Hispanic	-0.666	-0.605	-0.587	-0.560	-0.587	-0.564
	(0.031)	(0.022)	(0.036)	(0.027)	(0.021)	(0.020)
3+ gen Hispanic	-0.278	-0.238	-0.259	-0.225	-0.263	-0.264
	(0.025)	(0.027)	(0.044)	(0.024)	(0.027)	(0.026)
Late arrival		-0.164	-0.150	-0.167	-0.130	-0.127
		(0.009)	(0.008)	(0.006)	(0.007)	(0.005)
Early exit	-0.128	-0.122	-0.163	-0.323	-0.400	
	(0.006)	(0.006)	(0.007)	(0.008)	(0.009)	
1st gen Hispanic		-0.346	-0.533	-0.571	-0.514	-0.583
X Late arrival		(0.059)	(0.088)	(0.051)	(0.067)	(0.044)
1st gen Hispanic	0.057	0.030	0.118	0.125	0.187	
X Early exit	(0.062)	(0.066)	(0.066)	(0.068)	(0.105)	
Black	-0.876	-0.837	-0.819	-0.846	-0.807	-0.838
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Other race	-0.277	-0.242	-0.221	-0.159	-0.114	-0.083
	(0.014)	(0.013)	(0.012)	(0.012)	(0.012)	(0.013)
Multiracial	-0.286	-0.0260	-0.261	-0.249	-0.286	-0.282
	(0.018)	(0.017)	(0.017)	(0.016)	(0.015)	(0.017)
Constant	0.027	0.032	0.039	0.052	0.043	0.025
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Missing reading score	-0.836	-0.782	-0.973	-1.123	-0.905	-1.043
	(0.025)	(0.022)	(0.026)	(0.037)	(0.033)	(0.039)

Standard errors in parentheses. Sample: cohorts in 3rd grade in 2000 and 2001.

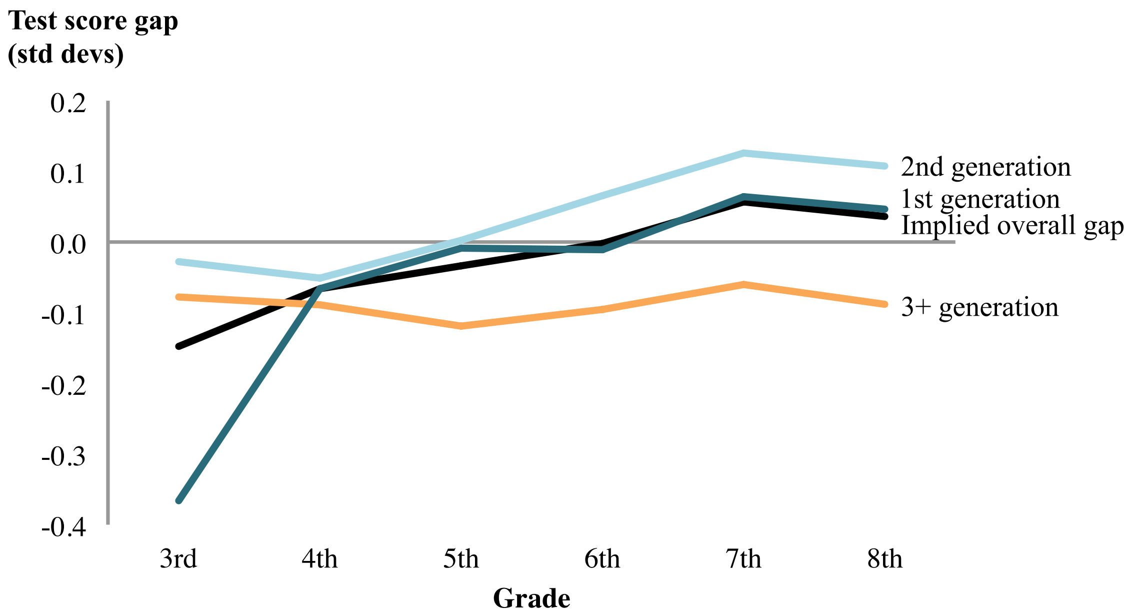
standard deviations. The more negative late arrival impacts in higher grades for the first generation drives the narrowing of their test score gap with whites.

These results give a more complete picture of why Clotfelter, Ladd, and Vigdor (2009) and others (Fryer and Levitt, 2004, 2006; Reardon and Galindo, 2009) find that a continuously followed sample of Hispanic students reduces the size of its test score gap with whites over time, while repeated cross sections, like in Phillips and Chin (2004), show a widening of the gap. Immigrant students that enter a longitudinal sample after the first wave, usually from out of the country, pull down the Hispanic average more each year. While others have speculated that this was the case, the evidence that I present here rules out the possibility that native-born Hispanics also make a significant contribution.

Last, I plot Hispanic-white test score gaps adjusted for parent's education in five categories, free/reduced lunch, gender, and school-by-year fixed effects as well as the late arrival and early exit indicators discussed above. We know from Table 3.1 that Hispanic and white students differ substantially in these measures of socioeconomic status. From Tables 3.3 and 3.4, we know that they also vary within the Hispanic population by immigrant generation. Finally, different generations of Hispanic students may attend different quality schools. Thus, Figure 3.4 (estimates in Tables 3.10 and 3.11) shows what differences remain between whites and Hispanic students of some immigrant generation after these additional controls are taken into account.

With this set of results, I find more support for the last scenario from Section 3.1.1, which I illustrate in Panel C of Figure 3.1. Third-generation Hispanics consistently perform about 0.1 standard deviations below whites in reading and math. If the Trejo critique holds, negative selection into Hispanic identification for later generations is responsible for this negative gap (Duncan and Trejo, 2011). The children of less successful Hispanics are on average more likely to be classified as Hispanic at school, which leads to the estimation of slightly lower scores relative to whites for this group. For the children of immigrants, I find astounding progress in both subjects with continuous enrollment from 3rd to 8th grade. After a gap of -0.37 standard deviations in reading in 3rd grade, the scores of first-generation Hispanics are statistically no different than whites from 4th grade onward. In math, these students score 0.21 standard deviations below whites in 3rd grade, surpass them in 5th grade, and exit 8th grade 0.12 standard deviations ahead of them. For the average immigrant student that arrives by 3rd grade, assimilation in test scores is complete in just a few years. This pattern of growth not only applies to Hispanic students who are immigrants themselves but also to the native-born children of immigrants.

(A) Reading



(B) Math

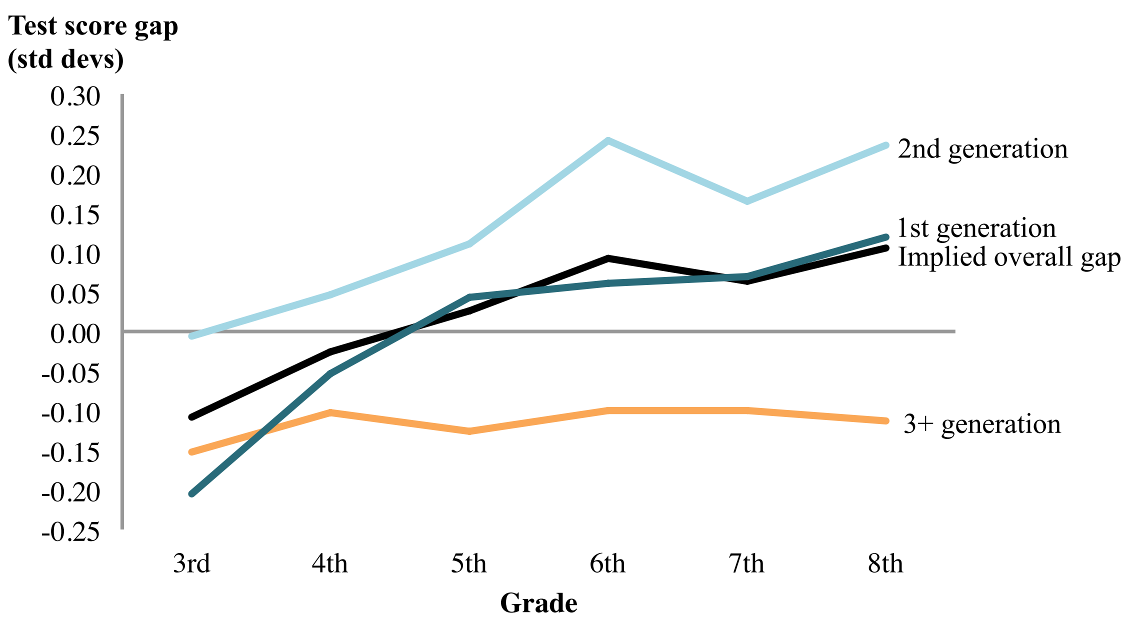


Figure 3.4: Adjusted Hispanic-white test score gaps by grade for continuous enrollees

Sample: cohorts in 3rd grade in 2000 and 2001. Test score gaps are adjusted for parent's education (5 categories), free/reduced price lunch, gender, school-by-year fixed effects, and late arrival and early exit, main effects and interactions with first-generation Hispanic. Regressions also include indicators for black, other race, multiracial, and missing other test score. See Tables 3.10 and 3.11 for point estimates and standard errors.

Table 3.10: Adjusted reading gaps by grade with continuous enrollment

Grade	3rd	4th	5th	6th	7th	8th
1st gen Hispanic	-0.366	-0.066	-0.009	-0.011	0.064	0.047
	(0.048)	(0.046)	(0.046)	(0.033)	(0.044)	(0.038)
2nd gen Hispanic	-0.028	-0.051	0.002	0.066	0.126	0.108
	(0.030)	(0.030)	(0.027)	(0.021)	(0.023)	(0.020)
3+ gen Hispanic	-0.078	-0.089	-0.119	-0.095	-0.060	-0.088
	(0.032)	(0.030)	(0.029)	(0.026)	(0.026)	(0.024)
Late arrival		-0.130	-0.118	-0.188	-0.149	-0.155
		(0.008)	(0.006)	(0.005)	(0.005)	(0.005)
Early exit	-0.118	-0.102	-0.131	-0.204	-0.287	
	(0.005)	(0.006)	(0.007)	(0.007)	(0.009)	
1st gen Hispanic		-0.809	-1.043	-0.974	-1.067	-1.271
X Late arrival		(0.069)	(0.057)	(0.038)	(0.051)	(0.045)
1st gen Hispanic	-0.106	-0.077	0.188	-0.029	-0.094	
X Early exit	(0.068)	(0.074)	(0.068)	(0.062)	(0.072)	
Black	-0.432	-0.456	-0.454	-0.444	-0.449	-0.458
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Other race	-0.137	-0.089	-0.113	-0.093	-0.056	-0.056
	(0.012)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Multiracial	-0.074	-0.060	-0.052	-0.044	-0.052	-0.042
	(0.015)	(0.015)	(0.015)	(0.014)	(0.015)	(0.015)
Constant	-0.568	-0.519	-0.536	-0.500	-0.523	-0.525
	(0.007)	(0.007)	(0.008)	(0.007)	(0.007)	(0.007)
Missing math score	-0.952	-0.853	-0.908	-1.073	-0.907	-1.023
	(0.108)	(0.073)	(0.058)	(0.063)	(0.059)	(0.050)
Parent's education						
High school graduate	0.414	0.371	0.390	0.347	0.373	0.372
	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)
Jr. college/trade	0.656	0.603	0.635	0.575	0.603	0.599
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Four-year college	0.956	0.915	0.920	0.868	0.896	0.886
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Graduate school	1.206	1.173	1.171	1.122	1.145	1.142
	(0.012)	(0.012)	(0.012)	(0.011)	(0.012)	(0.011)
Free/reduced lunch	-0.326	-0.340	-0.345	-0.335	-0.351	-0.356
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Female	0.188	0.167	0.174	0.205	0.211	0.190
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
School X year FEs	yes	yes	yes	yes	yes	yes

Standard errors in parentheses. Sample: cohorts in 3rd grade in 2000 and 2001.

Table 3.11: Adjusted math gaps by grade with continuous enrollment

Grade	3rd	4th	5th	6th	7th	8th
1st gen Hispanic	-0.205	-0.053	0.043	0.061	0.070	0.120
	(0.039)	(0.036)	(0.035)	(0.033)	(0.034)	(0.035)
2nd gen Hispanic	-0.006	0.046	0.111	0.241	0.165	0.235
	(0.026)	(0.024)	(0.022)	(0.021)	(0.019)	(0.019)
3+ gen Hispanic	-0.152	-0.102	-0.126	-0.100	-0.100	-0.113
	(0.030)	(0.027)	(0.026)	(0.025)	(0.023)	(0.023)
Late arrival		-0.199	-0.204	-0.247	-0.204	-0.205
		(0.008)	(0.006)	(0.005)	(0.005)	(0.005)
Early exit	-0.141	-0.122	-0.152	-0.239	-0.278	
	(0.005)	(0.006)	(0.006)	(0.007)	(0.008)	
1st gen Hispanic		-0.339	-0.533	-0.591	-0.486	-0.544
X Late arrival		(0.058)	(0.047)	(0.036)	(0.040)	(0.039)
1st gen Hispanic	-0.039	-0.029	0.113	0.036	0.020	
X Early exit	(0.065)	(0.060)	(0.058)	(0.053)	(0.056)	
Black	-0.543	-0.500	0.476	-0.439	-0.419	-0.424
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
Other race	-0.065	0.029	0.093	0.151	0.164	0.201
	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)	(0.010)
Multiracial	-0.135	-0.107	-0.099	-0.065	-0.099	-0.088
	(0.015)	(0.014)	(0.014)	(0.014)	(0.013)	(0.014)
Constant	-0.417	-0.374	-0.362	-0.387	-0.381	-0.412
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Missing reading score	-0.627	-0.599	-0.738	-0.822	-0.696	-0.843
	(0.022)	(0.019)	(0.022)	(0.030)	(0.029)	(0.030)
Parent's education						
High school graduate	0.358	0.309	0.296	0.302	0.294	0.301
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Jr. college/trade	0.579	0.518	0.516	0.523	0.501	0.517
	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Four-year college	0.894	0.852	0.845	0.845	0.847	0.847
	(0.008)	(0.007)	(0.008)	(0.007)	(0.007)	(0.007)
Graduate school	1.163	1.135	1.134	1.117	1.146	1.118
	(0.012)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Free/reduced lunch	-0.307	-0.316	-0.323	-0.336	-0.346	-0.346
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Female	0.003	0.013	0.019	0.063	0.066	0.070
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
School X year FEs	yes	yes	yes	yes	yes	yes

Standard errors in parentheses. Sample: cohorts in 3rd grade in 2000 and 2001.

Second-generation Hispanics start off in 3rd grade with scores on par with comparable whites in both subjects. In 8th grade, they outscore them by 0.11 standard deviations in reading and 0.24 standard deviations in math. Thus, these results show that the second generation more than assimilates; it outstrips whites as well as later-generation Hispanics.

Putting these three groups together, the trajectories of the implied overall gaps in math and reading appear very similar to those estimated in Clotfelter, Ladd, and Vigdor (2009). Between 3rd and 8th grade, the average Hispanic student that continuously enrolls reverses his position relative to the average white student from the same rough socioeconomic background, moving from 0.15 standard deviations below to 0.04 standard deviations above whites in reading, 0.11 below to 0.11 above whites in math. However, this overall trendline masks substantial heterogeneity in the growth and level of test scores by generational status.

The differences between the raw and adjusted test score plots come from adding controls for parent's education, free/reduced price lunch, gender, and school-by-year fixed effects. The magnitudes of the socioeconomic coefficients in Tables 3.10 and 3.11 reveal how important family background is to a child's performance in school. I estimate the effect of receiving a subsidized lunch as falling around -0.35 standard deviations, depending of the grade and subject. A student with a parent who graduated college scores almost a standard deviation higher on exams than a student with a parent that dropped out of high school. Besides parental human capital, this measure of educational attainment may also capture different meanings of educational attainment in the U.S. versus abroad. For example, having a parent that did not finish high school carries a more negative association in the U.S. than in a country like Mexico, where a higher proportion of the population does not finish secondary school.[18] The school-by-year fixed effects alone do not alter the estimated

[18] To investigate this issue further, I estimate my model on the sample of students with a high

gaps by much, a finding consistent with other papers that use this data, including Clotfelter, Ladd, and Vigdor (2009).

3.5 Conclusion

By decomposing the Hispanic-white test score gap by immigrant generation, this paper delves into why we see Hispanic test scores rise for intact cohorts and what it tells us about the academic progress of immigrant and native Hispanics. First, I show that the downward trend found with repeated cross-sections becomes an upward trend after conditioning on whether first-generation students arrive late to the public school system. Previous work could not distinguish between this cause and all Hispanics suffering disproportionately after a move between states (Clotfelter, Ladd, and Vigdor, 2009, 2012).

Second, I establish where the performance of later-generation Hispanics lies relative to whites. Just as is true for black students, Hispanic students with U.S.-born parents score below whites in early grades, and there is no evidence of improvement or deterioration in their relative position as they age through school. Furthermore, the small negative gap between later-generation Hispanics and whites is consistent with Duncan and Trejo's (2011) evidence that the later-generation Hispanics who identify as such are negatively selected. If this selection story is true, their test score gap is no real economic cause for concern; this difference would not exist with a true measure of ethnicity.

Lastly and most importantly, I show that the overall gains in Hispanic test scores are driven entirely by students with immigrant parents and that convergence in test scores is quite rapid. These students begin school with scores below their white

school educated parent since this level was common among whites and Hispanics. The pattern of test score gaps for the first generation is similar, but I find that the trajectory of the second generation looks closer to that of the third. However, I may merely be picking up differences in time of arrival for the immigrant parents themselves. If an immigrant parent completed high school, it is more likely that the parent migrated as a child and thus was educated in the U.S.

peers but by 8th grade often perform at substantially higher levels than them in both reading and math. The speed of assimilation in student test scores stands in contrast to the slower speed of wage assimilation for adults and is consistent with greater adaptability among younger immigrants (Borjas, 1985). The findings of Clotfelter, Ladd, and Vigdor (2009) therefore understate the impressive progress of children with immigrant parents with years in school. That immigrant students eventually outscore whites calls the importance of assimilation into question. Although factors like language acquisition surely contribute to test score growth for immigrants at early ages, assimilation does not explain gains after parity with native whites is reached. Moreover, native-born children with immigrant parents have contact with U.S. institutions from birth, which also goes against the notion that U.S. experience alone accounts for immigrant children's progress. One explanation that is consistent with all the evidence is that children with immigrant parents have home environments that value achievement, and this upbringing allows them to succeed over time. In any case, the progress made by Hispanic children of immigrants must come from a combination of noncognitive ability and home environment since I control for school environment and Cunha et al. (2006) rule out changes in cognitive ability in this age range.

The main takeaway is that Hispanic students from an immigrant background fare quite well under the current system, even better than suggested by Clotfelter, Ladd, and Vigdor (2009). Their main disadvantage is socioeconomic. While this hardship should not be downplayed—it still leads to the large raw achievement gaps between whites and immigrant Hispanics—the good news is that effective interventions targeted toward poor students should lift up Hispanic achievement as well.

4

Exploring the Racial Divide in Education and the Labor Market Through Evidence from Interracial Families

Differences in education, employment, and earnings between black and white Americans continue to be a social concern half a century after the passage of landmark civil rights legislation. While there has undoubtedly been some convergence, black-white gaps in outcomes remain stark. During the recent recession, black unemployment rates rose to more than double white unemployment rates. Among full-time workers, the median black male earns 80% less per week than his white counterpart.[1] Recent National Assessment of Educational Progress data reveal that black students score 0.7 standard deviations lower than whites on reading and mathematics tests. To the extent that test scores reflect underlying differences in human capital accumulation, we have seen little evidence of recent convergence (Neal, 2006).

Researchers across disciplines have advanced numerous theories on the sources of these gaps. In this paper, we provide new evidence that narrows the potential sources of these performance gaps using data from a longitudinal, school-based sur-

[1] See US Bureau of Labor Statistics (2011a,b).

vey, the National Longitudinal Study of Adolescent Health (Add Health). The data are rich in outcome variables, family background and parenting characteristics, and (crucially for this analysis) measures of maternal race. We exploit variation from interracial families generated by separate reporting of child and maternal race, along with interviewers' assessments of race and skin tone.[2] Using the linear decomposition method proposed by Gelbach (2009), we assess the relative importance of various sets of observed characteristics across a number of educational and early life labor market outcomes.

While controlling for observables typically explains only a fraction of the black-white gap on any particular outcome, our controls explain virtually the entire gap for all the outcomes we examine for black and Hispanic males, and they explain a larger part of the gap for females than what is typical in the literature. While there is a role for all factors, maternal race is the single largest factor for males, more important than the combined effects of all other characteristics of the mother, characteristics of the father, and school quality as measured by school fixed effects.

Using maternal race as part of the decomposition raises two econometric issues and an important issue of interpretation. Econometrically, both selection into interracial families and measurement error in race are a concern. We show that while white children come from families with significantly higher incomes, are more likely to come from two-parent families, and are more likely to attend high-quality schools than either black or Hispanic children with white mothers, controlling for these observables can account for virtually all of the differences in outcomes between interracial children with white mothers and white children.[3] Measurement error in either race of

[2] Throughout, we refer to Hispanic as a "race" although it is an ethnicity, simply for the sake of brevity in referring to multiracial, mixed-ancestry, and mixed-ethnicity families.

[3] While Hispanic children with white mothers have demographic characteristics that lie in between whites with white mothers and Hispanics with Hispanic mothers, we show that blacks with white mothers come from families that look observationally equivalent to the families of blacks with blacks mothers on a number of demographic measures.

the child or race of the mother could bias our estimates and this may be of particular concern given the low rates of interracial coupling. Our results are insensitive to using self-reported race, interviewer-reported race, and using interviewer-reported measures of child's skin tone instead of child race.

Our results suggest that discrimination is not occurring on the basis of the child's skin color alone, but that it must operate through characteristics associated with the race of the mother. The cultural environment for children raised by black mothers may produce characteristics that are later the source of discrimination. Grogger (2011) provides one example of this, showing that those who have distinctively black speech patterns suffer a wage penalty. The fact that maternal race better explains outcome gaps for males than for females is also consistent with Bertrand and Pan (2013), who show that noncognitive returns to parental inputs differ markedly by gender. Beyond differences in the home environment, black mothers may also be treated differently in the school system, resulting in worse classroom assignments and less teacher attention. Further, the legacy of discrimination may have resulted in black mothers not having access to the same information regarding prenatal care and parenting practices.[4] While we cannot distinguish between these different mechanisms (among many others correlated with mother's race) our results suggest channels correlated with race of the mother are likely to be the most fruitful in uncovering the sources of black-white inequities in education and labor market outcomes.

The rest of the paper proceeds as follows. In the following section, we discuss how our results fit into the long literature on racial inequality in education and the labor market. In Section 4.2, we describe the Add Health data and demographic characteristics of households with children and mothers of selected race combinations. Section 4.3 outlines our econometric methods. Section 4.4 reports differences in

[4] Currie and Grogger (2002) document differences in prenatal care between blacks and whites. Currie (2011) shows early life health disparities translate into inequality in school readiness.

educational and labor market outcomes in our sample and then goes on to examine how much maternal, paternal, family, and school characteristics can explain these differences. We conduct a series of robustness checks in Section 4.5 to confirm that race of the mother is indeed the driving factor in observed racial gaps in outcomes. Finally, we discuss the implications of our findings in Section 4.6.

4.1 Background

Over the course of the 20th century, black and white earnings converged substantially. While many factors contributed to this trend, two of the main drivers were improvements in the level and quality of education for blacks.[5] However, as differences in the amount of education between blacks and whites leveled off in the mid-1980s, so did differences in earnings, leaving a gap that has persisted for the past several decades.

This focus on educational attainment should not overshadow the importance of skills gained during (and before) formal schooling. Neal and Johnson (1996) argue that children with the same years of education can differ substantially in what they have learned. Using test scores from the Armed Services Qualifying Test (AFQT), rather than education, as a proxy for skills, they can explain most of the black-white gap in wages. Thus, much of the literature has focused on understanding why black children acquire less skill per year of schooling. The measure of "skill" most often used in empirical work is a standardized test score.

Most explanations for racial gaps in test scores fall under one of four categories: families, schools, discrimination, or genetics. We outline each in turn as our work has implications for all four channels.

The primary mechanism through which families foster growth in skills is parental

[5] See Smith and Welch (1989) for historical trends in wages and education. See Card and Krueger (1992) for black-white differences in educational quality and its connection to the wage gap.

human capital. Since black parents typically have lower levels of human capital relative to white parents, they may have a more limited capacity to aid in their children's skill accumulation. In addition to attaining less education on average, black parents may also have differential abilities to translate own education to child test scores (Currie and Thomas, 1999). The lower levels of wealth found among black parents limit the investments they can make in their children (Altonji, Doraszelski, and Segal, 2000). Home investments in children, particularly at a young age, are especially important for future test scores and thus explain some of the racial disparity (Todd and Wolpin, 2007; Cunha et al., 2006). All of these investment decisions are more complicated for children living outside of two-parent families, which is more common among black families.

Measures of school quality are known to vary considerably by race of the student (see Ferguson and Brown, 2000). Given the local public financing of schools, families sort into neighborhoods by school quality, leading students from low-income families to attend weaker schools. Ways in which school quality differs by race include average peer ability, average teacher quality, teacher turnover, and advanced course offerings. Even within schools, students have different experiences by race. Minority students are more likely to be put on a less rigorous academic track conditional on ability, and they more often have teachers with lower qualifications (see Hanushek and Rivkin, 2006).

The race of a child may also have a direct effect on outcomes through discrimination. For example, teachers may expect poorer performance from black students and act accordingly. Racist attitudes can influence the behavior of teachers, school administrators, and students in a variety of subtle and overt ways. Jencks and Phillips (1998) explore racial biases in the tests commonly used to assess student ability and learning. "Stereotype threat" may lead black students to perform poorly on a test when they believe the test is diagnostic of intellectual ability (Steele and Aronson, 1995).[6] If a student underperforms due to one of these latter two reasons, he might be sorted into a less challenging learning environment, which would in turn put him on a lower skill accumulation trajectory. Finally, blacks may face discrimination of another form if their peers view performing well in school negatively (see Austen-Smith and Fryer, 2005).

Finally, and most controversially, are genetic explanations. The study most related to ours directly deals with this issue. Willerman, Naylor, and Myrianthopoulos (1974) examine children of mixed-race families, both pairings of black mothers and white fathers as well as white mothers and black fathers. They show no cognitive differences at eight months—if anything, children of black mothers are at an advantage at this age. However, children of white mothers have significantly higher IQs at age four compared to children of black mothers, suggesting that the environment for children of white mothers was more conducive to cognitive development. Relatedly, Eyferth (1961) studies out-of-wedlock children of black and white U.S. soldiers during the post-World War II occupation of Germany. All of these children were raised by their German mothers. While there was considerable prejudice against blacks in Germany at the time when Eyferth gave these children a German version of the Wechsler IQ test, children with black fathers had almost identical scores to children with white fathers. Since this set of children did not live in segregated neighborhoods, did not attend segregated schools, and did not have mothers that were observably different, this study helped establish that race played no direct role in IQ score differences between black and white children.[7]

[6] Stereotype threat explanations, however, would seem unable to account for the very early emergence of black-white test score gaps (Carneiro, Heckman, and Masterov, 2008). Furthermore, the Steele and Aronson paper has been generally misinterpreted, actually providing evidence *against* standardized tests usually being given in a threatening environment. See Sackett, Hardison, and Cullen (2004).

[7] These results have also been called into question due to selection into the military. See Flynn (1999) and Jensen (1998) for the debate regarding selection and representation.

Like Eyferth (1961) and Willerman, Naylor, and Myrianthopoulos (1974), our study focuses on mothers. We argue that mothers are most often the parent primarily responsible for childrearing and show empirically that white mothers raising black children are remarkably similar to black mothers on a number of key traits. We compare the outcomes of black children raised by white mothers to two groups: white children raised by white mothers and black children raised by black mothers. While these divisions buy us separate identification of the effects of own race and maternal race on outcomes, the cost is that each of these measures of race is likely correlated with a host of unobserved factors. Of particular concern is whether white women who have children with black men are unobservably different from other white women in ways that are relevant for their children's outcomes. Nevertheless, the relative importance of child's race and mother's race helps us focus on the set of unobserved factors that most likely account for racial differences in outcomes.

While Eyferth (1961) and Willerman, Naylor, and Myrianthopoulos (1974) use parental race to narrow potential explanations for outcome gaps, there is a distinct but related literature on mixed-race children who self-identify as both black and white.[8] Fryer et al. (2012) find that mixed-race children are more likely to engage in risky behaviors compared to students who label themselves as only black or only white, but their self-reported grades are no different, a finding inconsistent with our results using transcript data. Also using Add Health data, Harris and Thomas (2002) show that test scores and grade point averages are higher among blacks from interracial families than blacks from single-race families. We explicitly attempt to separate the effect of race of the child from the effect of race of the mother. Note also that these studies focus on self-reported mixed-race individuals. Children with mixed heritage often only identify as one race or ethnicity, leading to selection issues

[8] Ruebeck, Averett, and Bodenhorn (2009) use Add Health to differentiate between mother's race from child's race, but they do not examine test score measures or labor market outcomes.

concerning which individuals report that they are mixed race. While the "one-drop rule" does not literally hold in U.S. survey data, several studies find that children with one white and one black parent often identify as black.[9]

4.2 Add Health data

We use data from the National Longitudinal Study of Adolescent Health (Add Health). The data is nationally representative; specifically, it is a school-based sample of seventh to twelfth grade students in 1995 within a randomly sampled set of 80 communities across the United States.[10] The first wave of the data, collected in the academic year 1994-95, attempted to survey all individuals at selected schools.

The data includes a subsample of students whose parents were also administered a survey. These in-home interviews provide information on race of the mother as well as assessments from the Add Health Picture Vocabulary Test (AHPVT).[11] Follow-up surveys were conducted in 1995-96, 2001-02 and 2008. Wave III (2001-02) includes transcript data, along with current education and labor market participation and wages. Wave IV (2008) provides information on completed education and labor market activity. The Add Health data also contains various non-representative oversamples, so throughout we use the cross-sectional probability weights provided to correct for the non-random sample design.[12] Wave-specific weights also correct for non-response between waves, and evaluation of these weights has been conducted in

[9] Roth (2005) discusses how changes in the U.S. Census between 1990 and 2000 reveal the prevalence of the one-drop rule. With 2000 Census data, Ruggles et al. (2010) show that 39% of black youth with white biological mothers identify only as black.

[10] A school pair, consisting of a high school and a randomly selected feeder school (middle school or junior high school from the same district) was taken from each community.

[11] The AHPVT is an abbreviated version of the Peabody Picture Vocabulary Test; a non-written test consisting of identifying pictures with verbal responses. It is designed to measure verbal scholastic aptitude.

[12] One of the groups oversampled was students from highly educated black families. However, this group is small and our results hold up with their exclusion. The sample weights correct for their overrepresentation as well.

a number of validation studies. Particularly relevant for our study, race distributions and the test scores show a bias of less than 0.5% after using the weights at Wave III. We discuss weighting further in Appendix D.[13]

4.2.1 Definition of race

We use a classification system that splits an individual's survey response into four distinct groups. If the respondent indicates that he is of Hispanic or Latino origin, then we classify him as Hispanic. If he marks that his race is black/African American but does not mark Hispanic, then we classify him as black. If he marks white but not Hispanic or black, then we classify him as white. If he marks a race category that does not fall into any of the above groups, then we classify him as other.[14] Table 4.1 shows a cross-tabulation of student and maternal race conditioning on maternal race being observed.[15]

4.2.2 Descriptive statistics for inputs

Table 4.2 shows how maternal characteristics vary by race of the mother and the child for Wave I. We focus on families with a white mother and those with either a black or Hispanic mother. We star (*) differences from the white mean that are significant at the 5% level, and † denotes differences from the own-group mean for

[13] We use the wave-specific weights all outcomes: Wave III weights are used for the results related to GPA; Wave IV weights are used for college completion and wages; and all other outcomes use the weights from Wave I. In Appendix D, we focus on attrition in test taking across race-groups in our estimation sample. The Carolina Population center has produced validation studies at: http://www.cpc.unc.edu/projects/addhealth/data/guides/.

[14] For children we use the Wave I In Home Questionnaire to define race. For mothers, we use the Parent Questionnaire. Mother's race is then the race of the surveyed parent when that parent is female, and the race of the surveyed parent's spouse (or resident mother) when the parental respondent is male.

[15] The parental survey response rate in Add Health is 85%. There is evidence that students whose parents did not respond had lower test scores. There is no evidence this gap is different across races with exception of the "other" race category, which we do not include in our regression analysis. We also drop those whose survey weights are zero. These students may be included in the data because they are twins, siblings, or unrelated co-residents of a sampled student.

Table 4.1: Student race and maternal race

	Maternal Race				
Race	White	Black	Hispanic	Other	Total
White	8151	14	75	78	8318
Black	132	2933	33	53	3151
Hispanic	319	35	2080	57	2491
Other	121	22	33	807	983
Total	8723	3004	2221	995	14943

Mother's race and student race are self-reported in separate survey instruments.

multiracial individuals (blacks and Hispanics).[16]

As is well known, white mothers are more educated, wealthier, are less likely to be a single parent, and they have fewer children than black mothers. What is less known is that white mothers with black children are demographically similar to black mothers in a number of respects. Average income, the probability of being on welfare, and the probability of having a single parent are similar for black students with white mothers and all black students. The one exception is years of schooling: white mothers with black children have more education than black mothers of black students. Although this difference is not significant, it is economically meaningful.[17] White mothers with Hispanic children have demographic characteristics that generally fall between Hispanics and whites.

The Add Health data contain a section designed to assess the relationship of adolescents to their biological fathers. We report a series of responses at Wave I by maternal race groups in Table 4.3.[18] White students generally see more involvement

[16] Means for the group black with white mother are tested against the means for blacks by excluding black children with white mothers when calculating the black mean. A similar test is used for Hispanic children.

[17] Because of this difference, we estimated the models below splitting the sample into two groups: mothers with a high school diploma or less, and mothers with at least some college. The patterns documented below were nearly identical in the two sets of results, which are available upon request. The fraction of black respondents completing high school in the Add Health data is very similar to NLYS97, Census, and Common Core data. See Heckman and LaFontaine (2010) for a comparison.

[18] Looking only at males, descriptive statistics show almost identical patterns.

Table 4.2: Mean mother characteristics

	White students	Black students	Black with white mom	Hispanic students	Hisp. with white mom
Income ($1000)	50.4	30.0*	33.0*	29.8*	48.5†
	(2.0)	(2.1)	(4.5)	(1.6)	(5.4)
	[7445]	[2558]	[120]	[1786]	[285]
On welfare	0.065	0.193*	0.211	0.184*	0.112†
	(0.008)	(0.016)	(0.093)	(0.023)	(0.031)
	[8286]	[2998]	[131]	[2158]	[318]
Single parent	0.206	0.564*	0.621*	0.291*	0.255
	(0.009)	(0.021)	(0.052)	(0.030)	(0.033)
	[8306]	[3004]	[132]	[2165]	[319]
Mother's age	41.3	41.5	39.6	40.7	40.7
	(0.2)	(0.4)	(1.1)	(0.3)	(0.5)
	[8032]	[2918]	[123]	[2096]	[312]
Mother's schooling	13.3	12.9*	13.5	11.1*	12.7*†
	(0.1)	(0.2)	(0.4)	(0.2)	(0.3)
	[7808]	[2759]	[117]	[1943]	[291]
Biological mother	0.914	0.871*	0.870	0.930	0.851†
	(0.005)	(0.010)	(0.040)	(0.008)	(0.033)
	[8265]	[2977]	[130]	[2137]	[315]
Household size[a]	3.33	3.63*	3.45	4.22*	3.69*†
	(0.04)	(0.08)	(0.18)	(0.11)	(0.15)
	[8304]	[3012]	[132]	[2164]	[319]
Siblings[b]	1.30	1.47*	1.42	1.86*	1.51*†
	(0.03)	(0.06)	(0.19)	(0.08)	(0.10)
	[8318]	[3019]	[132]	[2172]	[319]

Standard errors for means in parentheses, sample sizes in brackets. * significantly different from the white student mean at the 5% level. † significantly different from the own-minority group mean (black or Hispanic) at the 5% level. All variables from Parent Survey at Wave I.
[a] Household size is co-residents at Wave I.
[b] Siblings includes all non-biological siblings co-residing.

from their fathers, have more educated fathers, and receive more child support than black students regardless of the race of the black student's mother. One dimension on which black and white partners of black men may differ is in their bargaining power. Chiappori, Oreffice, and Quintana-Domeque (2011) show white women who intermarry are on average disadvantaged (heavier and less educated) relative to other white women. We see some evidence of this since most white mothers of black children are single and receive less child support than other white mothers. Looking down the columns for black students, no significant differences arise between black students and black students with white mothers. Although the differences are not significant, the fathers of black children with white mothers communicate less with their children are are less likely to co-reside with their children compared to the fathers of black students overall.

For Hispanics, the patterns are similar. Hispanics are less likely than whites to be living with their fathers and are less likely to know anything about their fathers, regardless of the race of their mother. Hispanic children with white mothers have less involvement with fathers than Hispanics overall; they are significantly less likely to live with their fathers and to speak with them weekly. Weighed against the lower involvement, Hispanic students with white mothers have more educated fathers than other Hispanic children, and they also receive higher child support payments.

Chiappori, Oreffice, and Quintana-Domeque (2011) address selection into interracial marriages. They show that on average the matchings between white women and black men involve lower "quality" partners using traditional metrics. White women who intermarry have higher BMI and lower education than other white women, while black men who intermarry are thinner and have lower wages than other black men.[19] In our own data, we can further investigate the role of selection into interracial couples by drawing on the Wave III relationship histories within Add Health, which also allows us to learn about selection into interracial Hispanic couplings. We calculate the mean test score within own- and cross-race relationships resulting in a pregnancy.[20] The results are presented in the bottom panel of Table 4.3. The highest

[19] One area in which we observe a small disparity is intergenerational differences in grandparent inputs. That white women who marry black men have lower education is suggestive that grandparent resources are negatively selected as well. Co-residence with grandparents is much higher among same-race black families than interracial families as well.

[20] Note that this sample contains matches from the next generation and selection into interracial relationships may have changed over time.

Table 4.3: Mean father characteristics

	White students	Black students	Black with white mom	Hispanic students	Hisp. with white mom
Know anything?	0.947	0.848*	0.927	0.874*	0.875*
	(0.004)	(0.012)	(0.038)	(0.015)	(0.027)
	[8308]	[3009]	[132]	[2163]	[319]
Currently live with?	0.651	0.325*	0.225*	0.590*	0.485*†
	(0.010)	(0.022)	(0.056)	(0.031)	(0.043)
	[8281]	[3009]	[131]	[2148]	[316]
Ever live with?	0.921	0.672*	0.731*	0.866*	0.828*
	(0.005)	(0.016)	(0.057)	(0.014)	(0.031)
	[8233]	[2961]	[130]	[2117]	[313]
Speak with weekly?	0.790	0.560*	0.384*	0.765	0.683*†
	(0.009)	(0.016)	(0.096)	(0.021)	(0.042)
	[7485]	[2413]	[110]	[1848]	[256]
Schooling	13.6	13.0*	13.2	12.0*	13.2†
	(0.1)	(0.1)	(0.5)	(0.2)	(0.3)
	[7584]	[2343]	[109]	[1757]	[264]
Child support ($/mo)	126.63	58.90*	68.57*	48.41*	101.97
	(5.57)	(4.88)	(23.15)	(6.73)	(29.17)
	[2902]	[1829]	[89]	[833]	[164]
Interracial matching[a]					
Female test score	0.080	-0.572*	-0.245*†	-0.486*	0.060†
	(0.039)	(0.081)	(0.141)	(0.091)	(0.104)
	[1044]	[826]	[81]	[417]	[96]
Male test score	0.298	-0.484*	-0.263*	-0.500*	-0.459*
	(0.046)	(0.090)	(0.118)	(0.107)	(0.227)
	[524]	[286]	[44]	[251]	[59]

Standard errors for means in parentheses, sample sizes in brackets. * significantly different from the white student mean at the 5% level. † significantly different from the own-minority group mean (black or Hispanic) at the 5% level. All variables measured at Wave I.

[a] The final two rows are drawn from the Add Health relationship histories gathered at Wave III. They present mean test scores where the unit of observation is a relationship that resulted in a pregnancy. Racial groups defined from male and female races: white when both partners were white, black when the mother was black, Hispanic when the mother was Hispanic. Interracial black and Hispanic groups are defined when the mother was white and father black or Hispanic, respectively.

test scores are found for white women with white men. White women who match with black men have lower test scores than white women who match with white men, but higher test scores than black women who match with black men.[21] Similarly, black men who match with white women have lower test scores than white men who match with white women, but higher test scores than black men who match with black women, though we cannot reject this latter difference is zero. For Hispanics couplings, we see that white women matched with Hispanic men have the same average ability as other whites, but there are no differences between Hispanic men who match with white women and those who match with Hispanic women. Below, we show that with observable measures, we can explain virtually all the outcome gaps between both black and Hispanic children with white mothers, as well as their respective same-race counterparts. This suggests that selection in the marriage market is not driving our results, or, at a minimum, that partner selection influences are mediated through other channels like parenting and schools.

Next, we examine neighborhood and school characteristics across groups. Table 4.4 shows that black children with white mothers live in neighborhoods and attend schools with characteristics in between those of white students and other black students. This holds true for both racial diversity as well as the percent of households below the poverty line. Looking at school-level characteristics reveals that the average test score at the schools blacks with white mothers attend also lies between the corresponding score for white students and other black students. The same patterns hold for Hispanic children of white mothers, with the demographics of their neighborhoods and schools lying in between those of white students and other Hispanic students. Overall, the patterns suggest that differences in choice of neighborhood and school may be important in explaining differences in schooling

[21] This finding lines up with education ordering in Chiappori, Oreffice, and Quintana-Domeque (2011).

Table 4.4: Mean location characteristics

	White students	Black students	Black with white mom	Hispanic students	Hisp. with white mom
Census block group[a]					
% white	0.921	0.371*	0.739*†	0.664*	0.835*†
	(0.008)	(0.029)	(0.035)	(0.032)	(0.021)
	[8238]	[2992]	[132]	[2164]	[318]
% black	0.048	0.592*	0.202*†	0.105*	0.065
	(0.007)	(0.031)	(0.032)	(0.015)	(0.017)
	[8238]	[2992]	[132]	[2164]	[318]
% in poverty[b]	0.108	0.267*	0.151†	0.200*	0.126†
	(0.008)	(0.016)	(0.028)	(0.015)	(0.013)
	[8238]	[2991]	[132]	[2164]	[318]
School					
% black	0.098	0.550*	0.264*†	0.169*	0.136
	(0.013)	(0.048)	(0.054)	(0.022)	(0.024)
	[8044]	[2981]	[131]	[2101]	[308]
Avg income	48.5	35.9*	43.0†	36.9	45.9†
	(1.8)	(2.2)	(4.4)	(2.2)	(2.6)
	[7445]	[2558]	[120]	[1786]	[285]
Avg test score	0.191	-0.324*	0.020†	-0.341*	-0.020*†
	(0.035)	(0.088)	(0.148)	(0.078)	(0.119)
	[7987]	[2876]	[127]	[2085]	[306]

Standard errors for means in parentheses, sample sizes in brackets. * significantly different from the white student mean at the 5% level. † significantly different from the own-minority group mean (black or Hispanic) at the 5% level. All variables measured at Wave I.
[a] Block group means taken from the 1990 Census.
[b] Poverty defined as below the 1989 poverty level.

and labor market outcomes.

4.2.3 Descriptive statistics for outcomes

We next examine outcome differences by the race of the student and the race of the mother, with the results reported in Table 4.5. Observed wage and education gaps between blacks and whites are similar to those found in the literature and administrative data. Black wages in the Add Health are around 80% lower than white wages, similar to Neal and Johnson (1996), and differences in college completion

Table 4.5: Mean outcomes

	White students	Black students	Black with white mom	Hispanic students	Hisp. with white mom
Test score	0.310	-0.540*	0.105†	-0.617*	0.138†
	(0.033)	(0.077)	(0.192)	(0.070)	(0.116)
	[7987]	[2876]	[127]	[2085]	[306]
Overall GPA	2.72	2.13*	2.54†	2.34*	2.54*†
	(0.01)	(0.07)	(0.12)	(0.04)	(0.08)
	[24953]	[8433]	[289]	[6254]	[690]
Math GPA	2.36	1.77*	2.15†	1.91*	2.18*†
	(0.01)	(0.06)	(0.12)	(0.04)	(0.09)
	[21455]	[7662]	[260]	[5434]	[590]
Science GPA	2.49	1.86*	2.15*	2.03*	2.26*
	(0.02)	(0.07)	(0.16)	(0.06)	(0.09)
	[19610]	[6879]	[231]	[4486]	[523]
College completion	0.341	0.219*	0.328	0.193*	0.247*†
	(0.020)	(0.029)	(0.070)	(0.018)	(0.050)
	[8170]	[3119]	[102]	[2107]	[249]
Wage	18.31	15.20*	16.92	17.62	17.59
	(0.327)	(0.543)	(1.783)	(0.624)	(1.101)
	[7122]	[2569]	[86]	[1835]	[212]
FT employment	0.909	0.886	0.870	0.900	0.801
	(0.009)	(0.019)	(0.049)	(0.012)	(0.112)
	[8170]	[3121]	[102]	[2108]	[250]

Standard errors for means in parentheses, sample sizes in brackets. * significantly different from the white student mean at the 5% level. † significantly different from the own-minority group mean (black or Hispanic) at the 5% level. Test scores measured at Wave I; GPA measured from transcripts at Wave III; completed education, wage, and full-time employment measured at Wave IV for males. The unit of observation for the GPA outcome is the individual-year; these standard errors are clustered at the individual level.

rates, 30% for whites versus 17% for blacks are similar to the overall gap among whites and blacks in the U.S. Census in 2008 (30% for whites and 18% for blacks).

Turning to differences by mother's race, white students and black students with white mothers have significantly higher test scores, math grades, and overall GPAs, and, for male students, are more likely to have finished college and have higher wages. For all these measures, there is no significant difference between white students and black students with white mothers. The only case where black students with white mothers are more similar to black students than white students is on grades in science classes. Hispanic students with white mothers also show significantly higher test scores than Hispanic students and have test scores that are not significantly different from white students. We view these differences as strong evidence against a genetic explanation: children from mixed-race couplings would have lower mean outcomes if they suffered from a genetic disadvantage. Given that white mothers of black children have low incomes and high rates of single parenthood, the notion that they could overcome a genetic disadvantage through increased child investment seems unrealistic. Hispanic students with white mothers have test scores and GPAs substantially higher than Hispanics as a whole, though they have slightly lower wages and employment levels.

Given work by Bond and Lang (2013) and Cascio and Staiger (2012) examining how test-score scaling influences mean comparisons, we also plot the distribution of test-scores for white students, black students, and black students with white mothers in Figure 4.1. The results are striking, and reveal not only similar means among children with white mothers, but that both distributions stochastically dominate the distribution for black students. Thus using a scale invariant measure, we observe similar patterns in the data.[22]

4.2.4 Comparison to a larger sample

Since the number of interracial children is small in the Add Health, we compare our descriptives to another data set where we observe the race of the mother, child, and a meaningful number of children from interracial couples. Namely, we examine vital statistics from the state of North Carolina. The means for background characteristics are presented in Table 4.6. On a number of measures, white mothers with black children look similar to black mothers. The percent receiving free or reduced lunch,

[22] The sets distributions for whites, Hispanics, and Hispanics with white mothers look nearly identical to those plotted in Figure 4.1, with the distributions for whites and Hispanics with white mothers overlapping, and both stochastically dominating the Hispanic distribution.

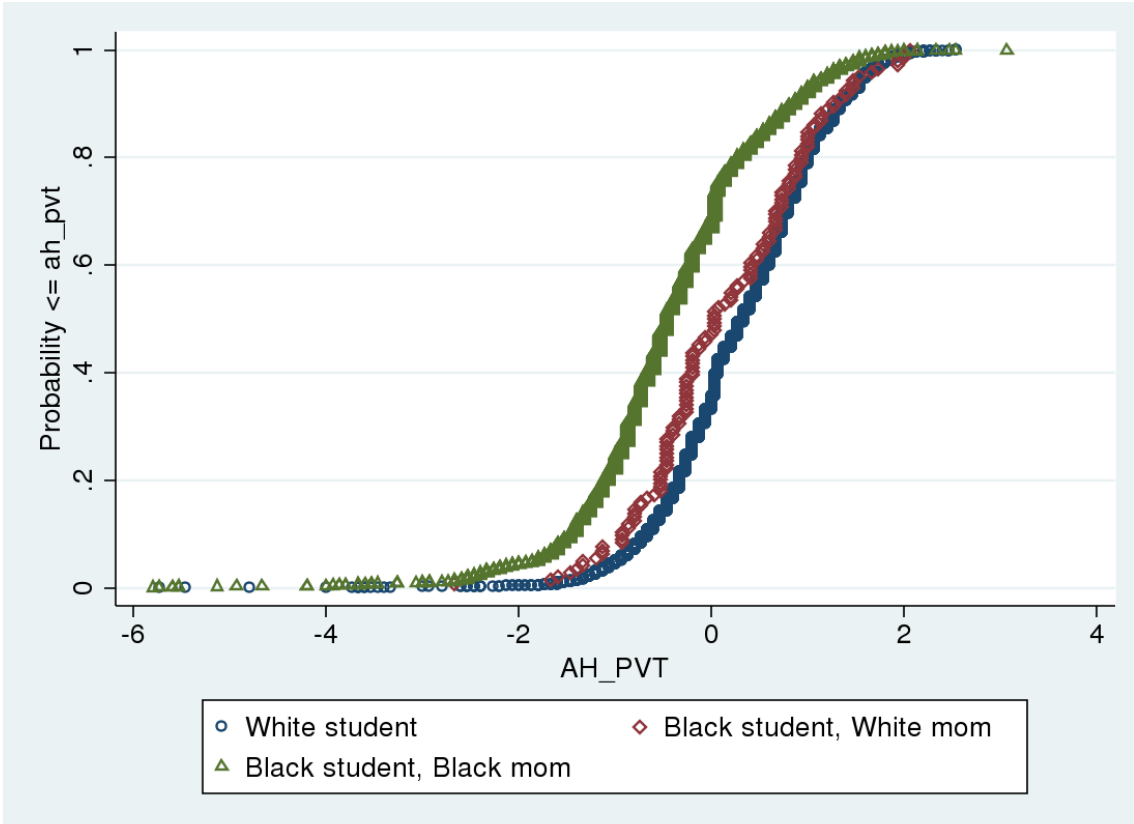


FIGURE 4.1: Cumulative distribution functions (CDFs) of test scores

our measure of economic need, is similar for both groups of black students and double the rate for white students. On outcomes like mother's education, married at the time of birth, and maternal age, white mothers with black children are disadvantaged relative to black mothers. Just as in the Add Health data, Hispanic children with white mothers have higher income and more educated mothers and fathers than Hispanic students as a group, but they still have lower income and less educated parents compared to those of white families. Overall, the demographic patterns look similare for Add Health and administrative data from a large U.S. state.

Table 4.6: Mean background characteristics in NCERDC data

	White students	Black students	Black with white mom	Hispanic students	Hisp. with white Mom
Free/reduced lunch	0.266	0.728*	0.639*†	0.702*	0.571*†
	(0.001)	(0.001)	(0.006)	(0.004)	(0.008)
	[306826]	[137094]	[6210]	[15806]	[3604]
Mother's education	13.0	12.3*	11.9*†	9.6*	11.6*†
	(0.003)	(0.004)	(0.02)	(0.02)	(0.03)
	[607471]	[266839]	[13072]	[49771]	[8279]
Mother married	0.860	0.389*	0.352*†	0.626*	0.672*†
	(0.0004)	(0.001)	(0.004)	(0.002)	(0.005)
	[607835]	[267077]	[14083]	[50133]	[8300]
Mother's age	26.9	24.7*	23.5*†	25.0*	24.9*
	(0.01)	(0.01)	(0.05)	(0.02)	(0.06)
	[607719]	[267035]	[14083]	[50131]	[8302]
Father's education	13.0	12.5*	12.3*†	9.5*	10.7*†
	(0.003)	(0.004)	(0.02)	(0.02)	(0.04)
	[577430]	[184505]	[11440]	[45441]	[7750]
First birth	0.457	0.399*	0.476*†	0.389*	0.389*
	(0.001)	(0.001)	(0.004)	(0.002)	(0.002)
	[607757]	[267050]	[14081]	[50120]	[8301]

Standard errors for means in parentheses, sample sizes in brackets. * denotes significantly different from the white-student mean at the 5% level. † denotes significantly different from the own-minority group mean (black or Hispanic) at the 5% level. All characteristics measured at birth except free/reduced lunch.

4.3 Methods

In a regression of the outcomes in Table 4.5 against an intercept, a black indicator, and a Hispanic indicator, the coefficient on the intercept would replicate column 1 of Table 4.5 and reflect the mean white outcomes; the coefficient on the black indicator would replicate the differences between mean white outcomes (column 1) and black outcomes (column 2); and similarly, the coefficient on the Hispanic indicator would replicate the difference between column 1 and mean Hispanic outcomes in column 4.

Because there are small differences in the gender and age distribution of whites, blacks and Hispanics in the sample, we estimate the model above separately by gender, and we control for age in each specification. Our baseline model that captures the difference in outcomes across groups is then:

$$Y_i = \alpha_0^B + \sum_r \alpha_{1r}^B I(Race_i = r) + \alpha_2^B X_{i1} + \varepsilon_i^B \tag{4.1}$$

where the superscript B denotes our baseline model. In a second regression, we then include a set of mother characteristics, including maternal race, and father characteristics. That is, we run:

$$Y_i = \alpha_0^I + \sum_r \alpha_{1r}^I I(Race_i = r) + \alpha_2^I X_{i1} + \alpha_3^I X_{i2} + \alpha_4^I X_{i3} + \varepsilon_i^I \tag{4.2}$$

where the superscript I denotes what we refer to as our intermediate model. In equation 4.2, mother's characteristics are denoted by X_{i2} and father's characteristics by X_{i3}. Our main interest here is how much do the additional factors in equation 4.2 help us explain the race effects found in equation 4.1. That is, α_{1r}^B describes the difference in outcomes between groups not conditioning on paternal and maternal factors, and α_{1r}^I does the same conditioning on those factors. Therefore, $(\alpha_{1r}^B - \alpha_{1r}^I)$ is the amount of the raw race effect accounted for by the two variable sets.

It is of interest to assess the relative importance of maternal and paternal factors in explaining the racial outcome gaps and, where relevant, to distinguish if any specific factor in those groupings of factors is especially important. To implement this, we use a method developed by Gelbach (2009), which nests the well-known Oaxaca-Blinder decomposition. Gelbach (2009) points out that if equation 4.2 is the true model, equation 4.1 is just a model with the variable sets X_{i2} and X_{i3} omitted, implying the well-known omitted variable bias formula applies. That is, the relationship between α_{1r}^B and α_{1r}^I is simply

$$\alpha_{1r}^B = \alpha_{1r}^I + \sum_{j=1}^{M} \delta_{3j}\alpha_{3j}^I + \sum_{j=1}^{P} \delta_{4j}\alpha_{4j}^I \tag{4.3}$$

where $(\alpha^I_{3j}, \alpha^I_{4j})$ are defined in equation 4.2 and there are M maternal characteristics and P paternal characteristics. The M δ_3s and P δ_4s are defined by the auxiliary regression:

$$I(Race_i = r) = \delta_0 + \delta_2 X_{i1} + \delta_3 X_{i2} + \delta_4 X_{i3} + \eta_i \tag{4.4}$$

A natural decomposition of how much each set of factors contribute to explaining the gap in outcomes is:

$$(\alpha^B_{1r} - \alpha^I_{1r}) = \sum_{j=1}^{M} \delta_{3j}\alpha^I_{3j} + \sum_{j=1}^{P} \delta_{4j}\alpha^I_{4j} \tag{4.5}$$

The first summation term is the part of the gap explained by maternal factors. Also notice that one can evaluate the contribution of a single factor within each group. For example, if the first element of X_{i2} is maternal race, then $\delta_{31}\alpha^I_{31}$ is the contribution of maternal race in accounting for the initial gap in outcomes by child's race.

Since the descriptive statistics reveal differences in the schools and neighborhoods across student and mother pairs, we also include specifications with school fixed effects to see how school quality contributes to the racial gaps. Our final regression is then:

$$\begin{aligned} Y_i = \alpha^F_0 + \sum_r \alpha^F_{1r} I(Race_i = r) + \alpha^F_2 X_{i1} + \alpha^F_3 X_{i2} + \alpha^F_4 X_{i3} + \sum_j^J \alpha^F_{5j} I(School_i = j) \\ + \varepsilon^F_i \end{aligned} \tag{4.6}$$

We can decompose the difference from equation 4.1 into:

$$(\alpha^B_{1r} - \alpha^F_{1r}) = \sum_{j=1}^{M} \theta_{3j}\alpha^F_{3j} + \sum_{j=1}^{P} \theta_{4j}\alpha^F_{4j} + \sum_{j=1}^{J} \theta_{5j}\alpha^F_{5j} \tag{4.7}$$

where the θ's are from an extended auxiliary regression that adds school fixed effects to equation 4.4 and uses θ for coefficients rather than δ. Now if we compare the contributions from maternal characteristics in equation 4.5 to those in equation 4.7, we see how much of the racial gap is still attributed to characteristics of the mother. If the attribution is reduced substantially by including school fixed effects, then one would conclude that the choice of school is an important mechanism through which characteristics of the mother operate.

4.4 Results

We present regression results in Table 4.7 for males where the outcome measures are test scores, math GPA, and log wages. We standardize test scores, so the interpretation of coefficients is as fractions of a standard deviation. Observations for math GPA are at the course-year level and are on a four-point scale.[23] Log wages come from the data in Wave IV and are conditional on working.[24] All the results reported in Table 4.7 use sampling weights. Unweighted results as well as results using different outcome measures are given in Appendix D.[25]

The first panel of Table 4.7 gives the results for test scores. Model 1 shows racial gaps when we only control for age of the child. To illustrate how quickly own-race effects disappear, the second column shows results where we condition only on characteristics of the mother. The final two columns show the results that are then used in the decomposition analysis, adding characteristics of the father and then adding school fixed effects. The baseline test score gaps for black and Hispanic males relative to white males are over 0.8 standard deviations. Adding characteristics of

[23] The course-specific GPA data come from each year an individual took a math course in school as recorded in the Wave III transcript file. The math GPA baseline regressions include an interaction between the level of math course (e.g., algebra or geometry) and the year of school it was taken (e.g., as a sophomore or junior). The standard errors for all GPA regressions are clustered at the individual level.

[24] We focus on Wave IV since the average age in this wave was around 28, beyond the period when most schooling is completed.

[25] The additional outcome measures we consider are science GPA, overall GPA, and whether the student obtained a four-year degree by Wave IV.

Table 4.7: Minority outcome gaps for males

Model:	(1)	(2)	(3)	(4)
Test scores				
Black	-0.855**	-0.037	-0.037	-0.155
	(0.074)	(0.139)	(0.124)	(0.108)
Hispanic	-0.819**	-0.133*	-0.086	-0.043
	(0.082)	(0.086)	(0.088)	(0.073)
Black mom		-0.740**	-0.655**	-0.420**
		(0.149)	(0.131)	(0.125)
Hispanic mom		-0.510**	-0.491**	-0.441**
		(0.102)	(0.114)	(0.104)
Math GPA[a]				
Black	-0.409**	0.067	0.136	0.187
	(0.051)	(0.103)	(0.099)	(0.103)
Hispanic	-0.228**	-0.090	-0.085	0.002
	(0.056)	(0.093)	(0.087)	(0.084)
Black mom		-0.415**	-0.481**	-0.397**
		(0.108)	(0.109)	(0.112)
Hispanic mom		-0.086	-0.164	-0.165*
		(0.099)	(0.095)	(0.091)
Log wages				
Black	-0.253**	0.018	0.042	0.054
	(0.041)	(0.095)	(0.095)	(0.093)
Hispanic	-0.071*	0.061	0.105	0.023
	(0.027)	(0.070)	(0.068)	(0.067)
Black mom		-0.207**	-0.211**	-0.212**
		(0.097)	(0.102)	(0.107)
Hispanic mom		-0.023	-0.019	-0.062
		(0.076)	(0.076)	(0.078)
Mother characteristics[b]	No	Yes	Yes	Yes
Father characteristics[c]	No	No	Yes	Yes
School FE	No	No	No	Yes

**,* significant at the 5%, 10% level. All regressions include child age. White and white mom are omitted. Missing indicators are used for all non-race variables.

[a] De-meaned course-level math GPA. Math GPA regressions include course-by-year fixed effects.

[b] Mother characteristics include income, on welfare, single parent, mother's age, mother's education, and biological mother.

[c] Biological father's characteristics are indicators for child knows anything about, child lives with, child ever lived with, child speaks to weekly, HS diploma, some college, college degree, no child support requirement, missing race, missing education, and monthly child support payment.

the mother alone reduces the effects of own race to be small and insignificant. Large gaps are present, however, when comparing male children of white mothers with male children of black and Hispanic mothers, with the sons of black (Hispanic) mothers scoring 0.74 (0.51) standard deviations worse than sons of white mothers. Only small drops in the estimates occur when we add father's characteristics. Including school fixed effects reduces the estimated gap between the male children of black mothers and white mothers to 0.42 standard deviations, suggesting that choice of schools is part of the reason for differences across these two groups.

For math grades and wages, similar patterns emerge for black males. Namely, large initial gaps exist between black and white students that disappear once we account for background characteristics, particularly mother's race. Children of black mothers have math grades that are almost a half point lower than children of white mothers, with the gap falling to 0.4 points once we add school fixed effects.[26] For wages, children of black mothers earn 20% lower wages than children of white mothers regardless of whether characteristics of the father or school fixed effects are included.[27] For Hispanics, the estimates are less precise but the same patterns emerge: virtually no effect of own race and negative estimates for children of Hispanic mothers relative to white mothers.

Table 4.8 shows results for females. For test scores, we see the same patterns as for males. Namely, large negative effects of own race are small and insignificant once we control for family background characteristics. We again see that children of black and Hispanic mothers have significantly lower test scores with the effects attenuated for blacks once we control for school fixed effects. The picture is more

[26] Given concerns about differential promotion and GPA scaling, we estimated an ordered probit of highest math level completed in school and found significant differences between black and white students as well as significant differences between black and white mothers, though these effects are very noisy and insignificant once we control for school fixed effects.

[27] Focusing instead on earnings (where the sample is then all those who have positive earnings) yielded identical findings.

muddled for math GPA and log wages, actually showing positive and significant effects for children of black mothers in the labor market once we account for school fixed effects. However, this latter result should be interpreted with caution given that selection into the labor market is a much bigger concern for females.

4.4.1 Decomposition

Given the large fraction of the raw gaps that can be explained with observables, we turn to decomposing the changes among the various sets of controls using equation 4.5. We report results for decompositions both with and without school fixed effects for each of the three outcome measures in Tables 4.9 and 4.10 for males and females, respectively. Three numbers are reported in each cell. First is the amount of the gap explained by the particular set of characteristics, with asterisks denoting the significance of the joint test that the variables in the group explain variation in the outcome equation. Second, in parentheses, is the standard error of this estimate. Finally, in brackets, is the fraction of the raw gap that is accounted for by this set, which is the variable-group coefficient divided by the raw gap. The bottom row in each panel then shows the total explained gap as well as the baseline gap.

The first column in Table 4.9 shows that absent school fixed effects, maternal race accounts for over 71% of the black and 55% of the Hispanic test score gap for males. Including school fixed effects (moving from column 1 to 4) drops this fraction to 46% for both blacks and Hispanics, suggesting school quality plays a much larger role for children of black mothers.[28] Other material characteristics are also important, accounting for 12.4% and 8.5% of the raw gap for blacks without and with school fixed effects respectively. The corresponding numbers for Hispanics are larger at

[28] Given the importance of school quality, one may be concerned that the effects of school quality are heterogeneous depending on student race. Specifications including school fixed effects interacted with race showed no significant changes in the maternal race coefficients, and in the case of blacks, we could reject their inclusion. Results are available on request.

Table 4.8: Minority outcome gaps for females

Model:	(1)	(2)	(3)	(4)
Test scores				
Black	-0.788**	-0.045	-0.040	-0.082
	(0.077)	(0.167)	(0.167)	(0.110)
Hispanic	-0.786**	-0.186**	-0.142	0.025
	(0.066)	(0.094)	(0.093)	(0.106)
Black mom		-0.620**	-0.553**	-0.375**
		(0.174)	(0.172)	(0.113)
Hispanic mom		-0.476**	-0.446**	-0.455**
		(0.107)	(0.104)	(0.101)
Math GPA				
Black	-0.491**	-0.153	-0.135	-0.136
	(0.040)	(0.171)	(0.169)	(0.153)
Hispanic	-0.390**	-0.202**	-0.195**	-0.067
	(0.052)	(0.094)	(0.091)	(0.087)
Black mom		-0.261	-0.255	-0.070
		(0.175)	(0.175)	(0.156)
Hispanic mom		-0.097	-0.036	-0.076
		(0.101)	(0.100)	(0.096)
Log wages				
Black	-0.177**	-0.181	-0.182	-0.219*
	(0.051)	(0.131)	(0.130)	(0.117)
Hispanic	0.007	0.013	0.041	-0.035
	(0.059)	(0.087)	(0.082)	(0.077)
Black mom		0.107	0.148	0.187*
		(0.116)	(0.114)	(0.107)
Hispanic mom		0.126	0.096	0.080
		(0.106)	(0.108)	(0.084)
Mother characteristics	No	Yes	Yes	Yes
Father characteristics	No	No	Yes	Yes
School FE	No	No	No	Yes

**,* significant at the 5%, 10% level. See Table 4.7 for a description of the dependent variables and controls.

Table 4.9: Decomposing minority outcome gaps for males

Model:	No school FE			With school FE		
Outcome:	Test	MGPA	Lwage	Test	MGPA	Lwage
Black gap						
Mother's char	-0.106**	-0.051*	-0.054**	-0.072**	-0.036	-0.044**
	(0.027)	(0.027)	(0.022)	(0.024)	(0.026)	(0.020)
	[0.124]	[0.126]	[0.215]	[0.085]	[0.088]	[0.174]
Mother's race	-0.612**	-0.443**	-0.197*	-0.393**	-0.366**	-0.197*
	(0.123)	(0.099)	(0.098)	(0.119)	(0.103)	(0.103)
	[0.716]	[1.082]	[0.778]	[0.460]	[0.894]	[0.777]
Father's char	-0.100*	-0.051	-0.043	-0.054	-0.009	-0.044
	(0.059)	(0.037)	(0.028)	(0.038)	(0.036)	(0.028)
	[0.117]	[0.125]	[0.171]	[0.063]	[0.022]	[0.175]
School FE				-0.181**	-0.184**	-0.022
				(0.065)	(0.047)	(0.034)
				[0.211]	[0.449]	[0.088]
Total	-0.818**	-0.545**	-0.294**	-0.700**	-0.594**	-0.307**
	(0.134)	(0.099)	(0.094)	(0.103)	(0.103)	(0.095)
	[0.957]	[1.333]	[1.164]	[0.819]	[1.453]	[1.215]
Baseline gap	-0.855**	-0.409**	-0.253**	-0.855**	-0.409**	-0.253**
	(0.073)	(0.051)	(0.041)	(0.073)	(0.051)	(0.041)
Hispanic gap						
Mother's char	-0.202**	-0.029	-0.082**	-0.155**	-0.014	-0.063**
	(0.028)	(0.020)	(0.017)	(0.023)	(0.019)	(0.016)
	[0.275]	[0.127]	[1.170]	[0.199]	[0.060]	[0.682]
Mother's race	-0.403**	-0.142*	-0.017	-0.360**	-0.141	-0.054
	(0.098)	(0.078)	(0.062)	(0.090)	(0.075)	(0.064)
	[0.550]	[0.620]	[0.242]	[0.464]	[0.614]	[0.580]
Father's char	-0.128**	0.027	-0.075**	-0.100**	0.030	-0.071**
	(0.048)	(0.034)	(0.025)	(0.042)	(0.032)	(0.025)
	[0.175]	-[0.119]	[1.058]	[0.129]	-[0.131]	[0.767]
School FE				-0.161**	-0.105	0.095**
				(0.056)	(0.051)	(0.042)
				[0.208]	[0.457]	-[1.029]
Total	-0.733**	-0.143*	-0.174**	-0.776**	-0.230**	-0.092
	(0.108)	(0.074)	(0.068)	(0.103)	(0.080)	(0.073)
	[0.895]	[0.628]	[2.470]	[0.948]	[1.009]	[1.309]
Baseline gap	-0.819**	-0.228**	-0.070*	-0.819**	-0.228**	-0.070*
	(0.082)	(0.056)	(0.037)	(0.082)	(0.056)	(0.037)

**,* group of characteristics significant at 5%, 10% level. Each cell contains the effect of each variable group on the white-minority outcome gap, the standard error in parentheses, and the fraction of the baseline outcome gap explained in brackets.

27.5% and 19.9%. Characteristics of the father are less important and are actually insignificant for blacks once school fixed effects are included.

Race of the mother is even more of a factor in accounting for gaps in math grades and log wages for black males, explaining over 77% of the raw gap regardless of whether school fixed effects are included. For Hispanic males, we see large effects of mother's race for math grades, which explain over 60% of the Hispanic-white gap, but the effects for wages are noisy. The latter is understandable given the base difference in log wages was small.

We report results for the same decomposition exercise for females in Table 4.10.[29] As with males, race of the mother is the dominant factor in explaining racial test score gaps. Mother's race accounts for 67% and 45% of the black-white gap without and with school fixed effects respectively, again suggesting school quality is one of the mechanisms through which race of the mother operates. Having a Hispanic mother is also the dominant factor for the Hispanic-white test score gap, accounting for around 44% of the raw gap regardless of controls for school fixed effects. However, we find no significant effects of mother's race for math grades.

4.4.2 Channels

Given the dominant role that race of the mother plays in accounting for racial gaps in school and in the labor market, particularly for black and Hispanic males, we next seek to understand whether parenting practices differ depending on race of the mother and whether these differences can begin to account for the larger mother race effects. Variables related to characteristics of the child's birth, behaviors of the mother, and how the parent and child interact are summarized in Table 4.11.

A number of differences across family types emerge. White mothers raising black

[29] We do not report decomposition results for female wages as the controls had no significant effects on racial wage gaps for this group.

Table 4.10: Decomposing minority outcome gaps for females

Model:	No school FE		With school FE	
Outcome:	Test	MGPA	Test	MGPA
Black gap				
Mother's characteristics	-0.089**	-0.028	-0.056**	-0.021
	(0.030)	(0.023)	(0.023)	(0.022)
	[0.113]	[0.057]	[0.071]	[0.043]
Mother's race	-0.530**	-0.244	-0.358**	-0.067
	(0.166)	(0.168)	(0.110)	(0.150)
	[0.673]	[0.498]	[0.455]	[0.136]
Father's characteristics	-0.131*	-0.084**	-0.091**	-0.076**
	(0.046)	(0.029)	(0.038)	(0.028)
	[0.166]	[0.171]	[0.116]	[0.154]
School FE			-0.200**	-0.191**
			(0.075)	(0.039)
			[0.254]	[0.389]
Total	-0.749**	-0.356**	-0.705**	-0.355**
	(0.178)	(0.165)	(0.135)	(0.151)
	[0.952]	[0.725]	[0.896]	[0.722]
Baseline gap	-0.788**	-0.491**	-0.788**	-0.491**
	(0.076)	(0.040)	(0.076)	(0.040)
Hispanic gap				
Mother's characteristics	-0.151**	-0.059**	-0.122**	-0.056**
	(0.025)	(0.018)	(0.023)	(0.018)
	[0.184]	[0.153]	[0.149]	[0.143]
Mother's race	-0.358**	-0.044	-0.365**	0.063
	(0.086)	(0.078)	(0.083)	(0.080)
	[0.438]	[0.114]	[0.445]	[-0.162]
Father's characteristics	-0.133**	-0.096**	-0.083**	-0.085**
	(0.029)	(0.025)	(0.030)	(0.024)
	[0.163]	[0.247]	[0.101]	[0.218]
School FE			-0.241**	-0.145**
			(0.093)	(0.043)
			[0.295]	[0.373]
Total	-0.643**	-0.643**	-0.810	-0.223**
	(0.097)	(0.097)	(0.129)	(0.083)
	[0.785]	[1.648]	[0.990]	[0.571]
Baseline gap	-0.819**	-0.390**	-0.819**	-0.390**
	(0.082)	(0.052)	(0.082)	(0.052)

, group of characteristics significant at 5%, 10% level. Each cell contains the effect of each variable group on the whtie-minority outcome gap, the standard error in parentheses, and the fraction of the baseline gap explained in brackets.

Table 4.11: Differences in parenting behaviors across families

	White students	Black students	Black with white mom	Hispanic students	Hisp. with white mom
AH HOME score					
Independence	0.823	0.892*	0.780†	0.795	0.841
	(0.005)	(0.008)	(0.039)	(0.017)	(0.028)
	[8055]	[2859]	[123]	[2009]	[309]
Hobby frequency	2.54	2.15*	2.39	2.21*	2.27
	(0.036)	(0.058)	(0.191)	(0.071)	(0.136)
	[8317]	[2931]	[132]	[2077]	[319]
No clubs	0.146	0.131	0.137	0.220*	0.162
	(0.008)	(0.014)	(0.045)	(0.019)	(0.032)
	[6293]	[2374]	[109]	[1522]	[242]
Mom talks about...					
Grades	0.624	0.615	0.683	0.606	0.584
	(0.009)	(0.013)	(0.082)	(0.021)	(0.036)
	[8066]	[2858]	[123]	[2014]	[310]
Behavior	0.340	0.283*	0.469*†	0.330	0.392
	(0.009)	(0.015)	(0.055)	(0.019)	(0.037)
	[8066]	[2858]	[123]	[2014]	[310]
School	0.145	0.134	0.168	0.118	0.176
	(0.007)	(0.013)	(0.050)	(0.012)	(0.038)
	[8066]	[2858]	[123]	[2014]	[310]
Birth characteristics					
Birth weight	7.47	6.97*	7.50†	7.34	7.11*
	(0.023)	(0.036)	(0.097)	(0.069)	(0.159)
	[7997]	[2678]	[119]	[1933]	[289]
Never breastfed	0.264	0.134*	0.192	0.302	0.305
	(0.008)	(0.013)	(0.054)	(0.022)	(0.037)
	[8055]	[2766]	[119]	[2006]	[294]
Mom home...					
Before school	0.608	0.579	0.477*	0.552*	0.560
	(0.011)	(0.016)	(0.061)	(0.022)	(0.044)
	[8054]	[2855]	[123]	[2014]	[209]
After school	0.249	0.349*	0.087*†	0.403*	0.270†
	(0.010)	(0.017)	(0.029)	(0.023)	(0.043)
	[8051]	[2856]	[123]	[2015]	[309]
At bedtime	0.746	0.802*	0.835	0.885*	0.782†
	(0.009)	(0.012)	(0.046)	(0.010)	(0.036)
	[8064]	[2863]	[123]	[2017]	[309]

Table 4.11 Differences in parenting behaviors across families

	White students	Black students	Black with white mom	Hispanic students	Hisp. with white mom
Mom employment					
Currently works	0.780	0.753	0.898*†	0.661	0.743
	(0.012)	(0.020)	(0.034)	(0.032)	(0.039)
	[8058]	[2843]	[122]	[2010]	[309]
Hours worked	36.53	38.08*	40.67*	36.32	35.68
	(0.282)	(0.489)	(1.130)	(0.400)	(1.000)
	[6146]	[2169]	[107]	[1325]	[215]
Child time use					
Sleep (hr/night)	7.931	7.651*	7.752	7.901	7.888
	(0.043)	(0.055)	(0.238)	(0.092)	(0.105)
	[8303]	[2918]	[132]	[2072]	[319]
Radio (hr/wk)	16.63	16.14	14.58	15.30	18.83†
	(0.441)	(0.891)	(3.080)	(0.699)	(1.670)
	[8926]	[2923]	[132]	[2074]	[319]
TV (hr/wk)	14.20	20.44*	17.30*†	16.74*	15.66
	(0.355)	(0.782)	(1.306)	(0.548)	(1.429)
	[8298]	[2914]	[132]	[2072]	[318]
Video games (hr/wk)	2.666	3.56*	3.456	2.713	2.863
	(0.121)	(0.265)	(0.668)	(0.229)	(0.702)
	[8311]	[2927]	[132]	[2077]	[319]
Hobbies (hr/wk)	1.543	1.306*	1.443	1.347*	1.383*
	-0.022	-0.034	-0.11	-0.042	-0.079
	[8317]	[2930]	[132]	[2077]	[319]

See Table 4.3 for notes.

children are less likely to agree that the mother fosters child independence but also encourage more participation in hobbies (e.g., reading, arts, and music).[30] These mothers are also more likely to discuss behavior problems than blacks or whites. Black children raised by white mothers have higher birth weights and lower rates of breastfeeding than black children raised by black mothers. Differences in time use are also present, with white mothers of black children more likely to work, working more hours, and correspondingly spending less time at home before and after school. Black children raised by white mothers also watch significantly less television than other black children, but more television than white children.

In Table 4.12, we add this large set of controls to equation 4.6 and present the accompanying decomposition for males.[31] Despite the differences in parenting behaviors shown in Table 4.11, the importance of maternal race for minority outcome gaps is unchanged.[32] Different sets of the additional variables have significant effects on the gaps depending on the outcome, but the effects are small. For example, birth characteristics (birth weight and breastfeeding) influence both test scores and wages for blacks, but their effect is about one-tenth of the size of the effect of mother's race. For Hispanics, the additional controls have even less explanatory power.

4.5 Robustness checks

There are at least four issues with the analysis conducted in the previous section. The first is selection into interracial relationships. Based on observables, as shown in Section 4.3, white mothers with black children appear to be negatively selected compared to other white mothers, yet their children have similar outcomes to white children given our controls. On observables white mothers with black children look very similar to black mothers, yet the outcomes for their children are very different. These two patterns suggest that selection is an unlikely explanation for our results.

Three additional issues remain. The first is measurement error in our race vari-

[30] Many authors have exploited the emotional support and cognitive stimulation HOME Scores from the National Longitudinal Survey of Youth 1979, to explain cognitive production and achievement (see, for example, Carneiro, Heckman, and Masterov, 2005; Cunha et al., 2006; Todd and Wolpin, 2007). Since a number of the questions overlap between these indices and the Add Health survey instrument, we examine the variables that overlap: the frequency of engaging in hobbies, arts, or playing music, whether the mother encourages independence, and whether the child is involved in no extra-curricular activities. These form our set of Add Health HOME Score variables.

[31] Following the discussion in Fryer and Levitt (2004), we also experimented with many school quality measures, none of which had significant impacts on the coefficients for most outcomes. This result is consistent with Fryer and Levitt (2013).

[32] Similar results are found for female test scores: maternal race is still the single most important factor (explaining roughly 50% of the outcome gap), even after conditioning on the large vector of controls.

Table 4.12: Decomposing minority outcome gaps for males using more channels

Model:	No school FE			With school FE		
Outcome:	Test	MGPA	Lwage	Test	MGPA	Lwage
Black gap						
Mother's char	-0.102**	-0.020	-0.027	-0.067**	-0.006	-0.022
	(0.030)	(0.028)	(0.022)	(0.024)	(0.027)	(0.022)
	[0.119]	[0.048]	[0.106]	[0.079]	[0.016]	[0.085]
Mother's race	-0.570**	-0.436	-0.174*	-0.385**	-0.375**	-0.175*
	(0.121)	(0.104)	(0.102)	(0.116)	(0.106)	(0.107)
	[0.667]	[1.065]	[0.686]	[0.450]	[0.917]	[0.693]
Father's char	-0.051	-0.015	-0.024	-0.029	0.028	-0.033
	(0.041)	(0.039)	(0.027)	(0.037)	(0.038)	(0.027)
	[0.060]	[0.038]	[0.096]	[0.034]	-[0.069]	[0.130]
AH HOME score	-0.002	0.007	0.005	-0.005	0.011	0.006
	(0.009)	(0.011)	(0.004)	(0.010)	(0.010)	(0.005)
	[0.002]	-[0.018]	-[0.021]	[0.006]	-[0.027]	-[0.022]
Mom's time use	-0.024	0.008	-0.013*	-0.024**	0.000	-0.011
	(0.015)	(0.015)	(0.007)	(0.011)	(0.011)	(0.008)
	[0.029]	-[0.020]	[0.051]	[0.028]	[0.000]	[0.043]
Child's time use	-0.020	-0.021	-0.016	-0.007	-0.014*	-0.025
	(0.018)	(0.009)	(0.011)	(0.009)	(0.009)	(0.023)
	[0.023]	[0.051]	[0.063]	[0.008]	[0.034]	[0.098]
Birth chars	-0.047**	-0.013	-0.018*	-0.042**	-0.007	-0.015
	(0.010)	(0.011)	(0.009)	(0.010)	(0.011)	(0.010)
	[0.055]	[0.031]	[0.070]	[0.049]	[0.017]	[0.058]
Parenting talks	-0.001	-0.042	-0.016	-0.003	-0.049**	-0.018
	(0.023)	(0.027)	(0.016)	(0.019)	(0.025)	(0.016)
	[0.001]	[0.104]	[0.064]	[0.003]	[0.120]	[0.071]
School FE				-0.160**	-0.166**	0.009
				(0.056)	(0.048)	(0.037)
				[0.187]	[0.405]	-[0.037]
Total	-0.816**	-0.531**	-0.283**	-0.721**	-0.579**	-0.283**
	(0.130)	(0.104)	(0.098)	(0.105)	(0.106)	(0.099)
	[0.955]	[1.298]	[1.116]	[0.843]	[1.414]	[1.118]
Baseline gap	-0.855**	-0.409**	-0.253**	-0.855**	-0.409**	-0.253**
	(0.073)	(0.051)	(0.041)	(0.073)	(0.051)	(0.041)

Table 4.12 Decomposing minority outcome gaps for males using more channels

Model:	No school FE			With school FE		
Outcome:	Test	MGPA	Lwage	Test	MGPA	Lwage
Hispanic gap						
Mother's char	-0.187**	-0.020	-0.067**	-0.147**	-0.016	-0.052**
	(0.029)	(0.028)	(0.018)	(0.023)	(0.019)	(0.017)
	[0.228]	[0.086]	[0.953]	[0.179]	[0.068]	[0.743]
Mother's race	-0.386**	-0.436	0.005	-0.353**	-0.129*	-0.042
	(0.096)	(0.104)	(0.062)	(0.088)	(0.078)	(0.065)
	[0.471]	[1.909]	-[0.068]	[0.431]	[0.567]	[0.591]
Father's char	-0.071**	-0.015	-0.042**	-0.062*	0.052	-0.040*
	(0.033)	(0.039)	(0.020)	(0.037)	(0.030)	(0.022)
	[0.087]	[0.068]	[0.591]	[0.076]	-[0.229]	[0.567]
AH HOME score	-0.017	0.007	-0.011	-0.030*	-0.018*	-0.014
	(0.018)	(0.011)	(0.009)	(0.019)	(0.011)	(0.010)
	[0.020]	-[0.032]	[0.155]	[0.037]	[0.077]	[0.199]
Mom's time use	-0.008	0.008	-0.008	-0.013	-0.004	-0.002
	(0.026)	(0.015)	(0.010)	(0.018)	(0.015)	(0.011)
	[0.009]	-[0.037]	[0.114]	[0.016]	[0.020]	[0.033]
Child's time use	-0.046	-0.021	-0.025	-0.028	-0.009*	-0.017
	(0.031)	(0.009)	(0.023)	(0.019)	(0.005)	(0.019)
	[0.056]	[0.091]	[0.355]	[0.034]	[0.041]	[0.242]
Birth chars	-0.007	-0.013	-0.005	-0.042**	0.006	-0.005
	(0.008)	(0.011)	(0.005)	(0.010)	(0.006)	(0.005)
	[0.009]	[0.056]	[0.066]	-[0.008]	-[0.027]	[0.065]
Parenting talks	-0.014	-0.042	-0.014**	-0.003	-0.004	-0.021*
	(0.019)	(0.027)	(0.007)	(0.014)	(0.012)	(0.011)
	[0.017]	[0.186]	[0.202]	[0.003]	[0.018]	[0.291]
School FE				-0.145**	-0.085*	0.106**
				(0.052)	(0.050)	(0.042)
				[0.177]	[0.372]	-[1.501]
Total	-0.735**	-0.109	-0.167**	-0.789**	-0.204**	-0.087
	(0.104)	(0.081)	(0.070)	(0.100)	(0.082)	(0.075)
	[0.897]	[0.478]	[2.370]	[0.964]	[0.893]	[1.231]
Baseline gap	-0.819	-0.228	-0.071	-0.819	-0.228	-0.071
	(0.082)	(0.056)	(0.037)	(0.082)	(0.056)	(0.037)

**,* group of characteristics significant at 5%, 10% level. See notes for Table 4.9.

ables. Below we use alternative measures of student race and show that our qualitative results hold. Second, differences could be driven by discrimination on the basis of skin tone, as children with white mothers are more likely to have lighter skin. We show that results for skin tone follow the same patterns as the results for own race: big effects for skin tone when mother's race is not in the set of controls, small or no effects when mother's race is in the set of controls. The final issue is small sample sizes for interracial families. While we obviously cannot increase the size of the Add Health data, we can improve precision by putting more structure on the model. In particular, rather than viewing each outcome in isolation, we estimate a joint model of all our outcome measures and impose some structure on the relationship between the covariates across the different outcomes.

4.5.1 Measurement error

Given that multiracial families are identified from measures of race that are self reported, measurement error may be a concern. Measurement error could manifest itself in at least two ways. First, individuals may choose to identify with a group in a way that does not match our standard definitions of race. If this is done by students, we would suspect that mixed-race students may identify with groups that hold values more like them. In this case, identifying as black could be correlated with identifying with a lower achieving group.[33] This would lead to negative effects of own race for blacks, effects we see little evidence of in our results. On the other hand, if it is the mother identifying with a particular racial group (e.g., if a mixed-race mother identifies herself as black), then mother's race is picking up the cultural environment with which the parent identifies, exactly the effects we hope to pick up by controlling for race of the mother.

The second way measurement error may manifest itself is through random misreporting. Suppose race of the mother and race of the child are noisy measures of the

[33] Add Health is one of the few data sets that allows mixed-race respondents to choose multiple races. The only way this form of measurement error could bias our results is if a sizable number of mixed-race adolescents marked only white. If they only mark black, or they mark white and black, they are classified as black.

Table 4.13: Interviewer-reported race cross-tabulation

	Interviewer-reported race		
Self-reported race	White	Black	Other
Full sample			
White	7854	11	49
Black	29	2901	17
Hispanic	1356	96	883
With a white mother			
White	7777	4	48
Black	22	93	8
Hispanic	219	4	83
With a black mother			
White	6	7	0
Black	6	2780	8
Hispanic	2	27	6
With a Hispanic mother			
White	71	0	1
Black	1	28	1
Hispanic	1135	65	794

Note: Mother's race is self reported.

same underlying factor. If children are more likely than their mothers to misreport, then more weight will be placed on the race of the mother. If this bias were large, we would expect changes in the maternal race coefficients when using a measure of child race that is less error laden.

To address this issue, we use interviewer reports of the child's race. Classification by interviewers included white, black/African American, and other races.[34] This classification misses Hispanics, who could be assigned to a number of these groups.

In Table 4.13, we report the cross-tabulation of adolescent self-reported race with the interviewer-reported race (white, black or other), first for the entire sample,

[34] The exact question from the survey is: "Interviewer: Please code the race of the respondent from your own observation alone: 1) White 2) Black/African American 3) American Indian/Native American 4) Asian/Pacific Islander 5) Other."

and by maternal race. The full-sample data show reports differ mainly for Hispanics. Very few individuals show up in the reverse categories for black and white, suggesting improper self-reporting is not driving the results above. In fact, for non-Hispanic respondents, 0.35% of interviewer-reported blacks and 0.36% of interviewer-reported whites self-report as the opposite race, suggesting this is the percentage of individuals making a reporting error. In the second panel, we see that 93 of the 123 self-reported black students from white mothers are identified by the interviewer as black.

To have as close to an error-free measure of child race as possible, we restrict our sample to cases where the interviewer report of the student's race matches the student's report, using observations along the diagonals of Table 4.13. We therefore do not include individuals who self-report as Hispanic in the results which follow.

Table 4.14 gives regression results for two subsamples. In the first column are results for black students with either black or white mothers as well as white students, conditioning on agreement of self-reported and interviewer-reported race. The same patterns emerge as in the previous tables. Namely, we see no significant negative effects of the student's race being black for male test scores, math grades, and log wages, nor do we see negative and significant own-race effects for female test scores. However, having a black mother is associated with worse outcomes along all these dimensions of around the same magnitudes as those presented in Table 4.7. In the second column, we further restrict the sample to only those who have white mothers. Again, we find no significantly negative effects of own race.

4.5.2 Skin tone

Another check on whether the self-reporting of race is driving the findings above is to examine skin tone. The difference between the own-race and maternal race coefficients in Tables 4.7 and 4.8 is identified from multiracial families. One potential channel for these effects is that there is less discrimination against children from

Table 4.14: Minority outcome gaps for alternative subsamples

	Subsample	
	Blacks and whites[a]	White mothers
Test scores: Males		
Black	-0.100	-0.109
	(0.127)	(0.129)
Black mom	-0.429**	
	(0.143)	
Wages: Males		
Black	0.059	0.037
	(0.114)	(0.111)
Black mom	-0.254*	
	(0.135)	
Math GPA: Males		
Black	0.240*	0.222*
	(0.122)	(0.125)
Black mom	-0.465**	
	(0.135)	
Test scores: Females		
Black	-0.074	-0.006
	(0.108)	(0.133)
Black mom	-0.387**	
	(0.124)	
Mother characteristics	Yes	Yes
Father characteristics	Yes	Yes
School FE	Yes	Yes

**,* significant at the 5%, 10% level. See Table 4.7 for a description of the dependent variables and controls.
[a] Subsample includes blacks with black mothers, whites with white mothers, and blacks with white mothers.

multiracial families, perhaps because they more frequently have a lighter skin tone.[35] The Add Health data collectors described the respondent's skin color in Wave III as "Black, Dark brown, Medium brown, Light brown or White."[36] Since this assessment may be subjective, we also include information on interviewer race.[37]

We provide a cross-tabulation of skin tone and maternal race in Table 4.15, which shows that the lighter-skinned categories contain a significant number of Hispanics and those of other races. In the lower panel, we present the distribution of skin tone for self-identified blacks, split by maternal race. There is substantial overlap in the distribution, with roughly 30% of each group having a "medium brown" skin tone, and "light brown" and "dark brown" groups also having nontrivial overlap. Nonetheless, the distribution is shifted toward lighter skin tones for blacks with white mothers.

Table 4.16 takes Table 4.7 and replaces self-identified race with interviewer-reported skin tone and adds controls as before. In the first set of columns that do not include mother's race, we see that darker individuals have lower test scores, math grades, and log wages. Adding controls, particularly mother's race, substantially reduces the estimated effects of skin tone. Darker skin is still associated with significantly lower test scores, but its magnitude is reduced by roughly 65% for each skin tone group.

In Table 4.17, we present the decomposition of the skin tone gaps. As in the previous decompositions, maternal race is the largest contributor to the test score gap, explaining 50-60% of the differences in test scores between each skin tone group

[35] Rangel (2014) examines this question in Brazil and finds differential investment among children within the same family but with different skin colors. Using data from the 1910 census, Mill and Stein (2012) find little difference in literacy rates between mulatto and black siblings, suggesting investment rates are fairly similar across skin color.

[36] The question interviewers answered was: "What is the respondent's skin color: 1) Black 2) Dark brown 3) Medium brown 4)Light brown 5) White?"

[37] Patterns look identical when using interviewer fixed effects as well, though the sample size shrinks.

Table 4.15: Student skin tone and maternal race cross-tabulation

	Student skin tone				
Maternal race	White	Light Brown	Medium Brown	Dark Brown/ Black	Total
White	6706	337	83	38	7164
Black	10	311	772	1383	2476
Hispanic	824	625	200	72	1721
	Skin tone of black students				
Maternal race	White	Light Brown	Medium Brown	Dark Brown/ Black	Total
Black	6	295	767	1356	2424
%	0.3	12.2	31.6	55.9	100.0
Non-black	14	48	32	26	120
%	11.7	40.0	26.7	21.7	100.0

Skin tone is reported by the interviewer at Wave III, race is self-reported from Wave I.

and those with white skin. Maternal race explains 40% of the wage gap for individuals with the darkest skin tone. Overall, the set of controls we introduce explain the vast majority of the outcome gaps for medium- and dark-skinned individuals in our sample (decompositions for light-skinned respondents are in Appendix Table D.6). With respect to the wage estimates, one plausible channel for the lack of wage penalties among those blacks raised by white mothers is less discrimination based on skin tone. However, our results do not support this hypothesis. Rather, the decomposition indicates that skin tone is important to wages through its correlation with maternal race.

4.5.3 Linking outcomes

The final issue with our analysis is small sample sizes, which sometimes result in large standard errors. While we cannot increase our sample, we can place some structure on the problem. So far we have analyzed each outcome independently. It is likely that all of these outcomes are influenced by common underlying factors

Table 4.16: Minority outcome gaps for males by skin tone

Model:	(1)	(2)	(3)	(4)
Test scores				
Light	-0.669**	-0.159**	-0.149**	-0.179**
	(0.083)	(0.066)	(0.060)	(0.055)
Medium	-0.862**	-0.213**	-0.240**	-0.162*
	(0.114)	(0.085)	(0.086)	(0.087)
Dark	-0.906**	-0.167*	-0.172*	-0.208**
	(0.090)	(0.097)	(0.093)	(0.091)
Black mom		-0.666**	-0.573**	-0.445**
		(0.079)	(0.075)	(0.079)
Hispanic mom		-0.565**	-0.506**	-0.411**
		(0.076)	(0.082)	(0.079)
Math GPA				
Light	-0.251**	-0.118*	-0.108	-0.081
	(0.062)	(0.067)	(0.064)	(0.066)
Medium	-0.202**	0.044	0.042	0.059
	(0.076)	(0.107)	(0.106)	(0.105)
Dark	-0.459**	-0.112	-0.109	0.003
	(0.063)	(0.121)	(0.123)	(0.120)
Black mom		-0.258**	-0.264**	-0.226*
		(0.110)	(0.117)	(0.122)
Hispanic mom		-0.126*	-0.203**	-0.146**
		(0.065)	(0.068)	(0.076)
Log wages				
Light	-0.079	0.014	0.028	0.025
	(0.045)	(0.051)	(0.047)	(0.049)
Medium	-0.096	0.069	0.072	0.076
	(0.050)	(0.061)	(0.060)	(0.062)
Dark	-0.355**	-0.116	-0.105	-0.110
	(0.069)	(0.092)	(0.092)	(0.095)
Black mom		-0.170**	-0.157**	-0.137*
		(0.060)	(0.071)	(0.083)
Hispanic mom		0.023	0.061	-0.052
		(0.045)	(0.042)	(0.052)
Mother characteristics	No	Yes	Yes	Yes
Father characteristics	No	No	Yes	Yes
School FE	No	No	No	Yes

**,* significant at the 5%, 10% level. See Table 4.7 for a description of the dependent variables and controls.

Table 4.17: Decomposition of outcome gaps by skin tone for males

Model:	No school FE			With school FE		
Outcome:	Test	MGPA	Lwage	Test	MGPA	Lwage
Medium tone						
Mother's char	-0.128**	-0.045**	-0.065**	-0.094**	-0.034	-0.051**
	(0.029)	(0.022)	(0.019)	(0.022)	(0.021)	(0.016)
	[0.147]	[0.223]	[0.674]	[0.108]	[0.170]	[0.532]
Mother's race	-0.424**	-0.184**	-0.064	-0.333**	-0.151**	-0.092**
	(0.057)	(0.067)	(0.039)	(0.055)	(0.070)	(0.044)
	[0.489]	[0.909]	[0.663]	[0.384]	[0.749]	[0.961]
Father's char	-0.074*	-0.015	-0.041**	-0.041	0.011	-0.044**
	(0.039)	(0.028)	(0.020)	(0.028)	(0.028)	(0.019)
	[0.086]	[0.074]	[0.431]	[0.048]	-[0.055]	[0.459]
School FE				-0.236**	-0.087*	0.004
				(0.093)	(0.047)	(0.026)
				[0.273]	[0.430]	-[0.043]
Total	-0.626**	-0.244**	-0.170**	-0.704**	-0.261**	-0.184**
	(0.080)	(0.071)	(0.040)	(0.105)	(0.078)	(0.044)
	[0.723]	[1.206]	[1.768]	[0.813]	[1.294]	[1.909]
Baseline gap	-0.866**	-0.202**	-0.096**	-0.866**	-0.202**	-0.096*
	(0.115)	(0.076)	(0.049)	(0.115)	(0.076)	(0.049)
Dark tone						
Mother's char	-0.116**	-0.061**	-0.061**	-0.082**	-0.041	-0.051**
	(0.025)	(0.029)	(0.023)	(0.024)	(0.028)	(0.022)
	[0.127]	[0.133]	[0.172]	[0.091]	[0.089]	[0.145]
Mother's race	-0.517**	-0.240**	-0.140**	-0.402**	-0.206*	-0.142**
	(0.070)	(0.107)	(0.062)	(0.073)	(0.112)	(0.073)
	[0.569]	[0.523]	[0.395]	[0.442]	[0.449]	[0.399]
Father's char	-0.103*	-0.049	-0.047*	-0.054	-0.007	-0.052**
	(0.065)	(0.038)	(0.028)	(0.041)	(0.038)	(0.027)
	[0.114]	[0.107]	[0.133]	[0.059]	[0.016]	[0.147]
School FE				-0.162**	-0.208**	-0.014
				(0.048)	(0.048)	(0.038)
				[0.179]	[0.453]	[0.039]
Total	-0.736**	-0.350**	-0.248**	-0.701**	-0.462	-0.259**
	(0.093)	(0.107)	(0.059)	(0.078)	(0.109)	(0.063)
	[0.810]	[0.763]	[0.700]	[0.771]	[1.008]	[0.730]
Baseline Gap	-0.909**	-0.459**	-0.355**	-0.909**	-0.459	-0.355**
	(0.073)	(0.063)	(0.069)	(0.073)	(0.063)	(0.069)

**,* group of characteristics significant at 5%, 10% level. Each cell contains the effect of each variable group on the whtie-minority outcome gap, the standard error in parentheses, and the fraction of the baseline gap explained in brackets.

related to human capital or discrimination. As a final specification, we consider an "achievement index" that allows each outcome to have a different intercept but constrains the effect of covariates to be proportional across outcomes. Here we include the outcomes in the main analysis (test scores, math GPA, and wages) as well as those in Appendix D (overall GPA, science GPA, and college completion).[38] To the extent that discrimination occurs on the basis of student race, the assumption is that it has the same relative effect as the other covariates across the different outcome measures. Specifically, we estimate a model of male outcomes of the form:

$$\begin{aligned} Y_{ik} = \lambda_{0k} + \lambda_{1k}\Bigg(&\sum_r \lambda_{2r} I(Race_i = r) + \lambda_3 X_{1i} + \lambda_4 X_{2i} + \lambda_5 X_{3i} \\ &+(1 - \lambda_{6k} I\{k = GPA\}) \sum_j \lambda_{7j} I(School_i = j)\Bigg) + \varepsilon_{ik} \end{aligned} \tag{4.8}$$

where ε_{ik} is distributed normally with mean zero and $Var(\varepsilon_{ik}) = \sigma_k^2$. Here, i denotes individual and k denotes outcome. The outcomes we consider are test scores, high school GPA, college attendance, and log wages. The scale parameter for test scores $\lambda_{1,test}$ is normalized to one, as is the variance of college completion σ_{coll}^2. We allow the school fixed effects to operate differently for grades than the other outcomes as grades are a relative measure: better schools may give lower grades conditional on observed characteristics but nonetheless produce more human capital.

Selected coefficients for the model outlined in equation 4.8 are presented in Table 4.18.[39] Across the three columns, the models vary by the sets of controls we include. In model 1, the only additional controls besides own race are course-by-year fixed effects for the GPA outcomes; in model 2, we add maternal and paternal characteristics; and in model 3, we add school fixed effects. The coefficients for maternal

[38] Our analysis includes all valid observations, implying we have an unbalanced panel. Using individuals who have valid observations for each outcome produces very similar results.

[39] The table shows results using the Add Health weights. Unweighted results produced similar patterns.

Table 4.18: Full factor model of outcomes

Model:	(1)	(2)	(3)
Black	-0.218*	0.022	0.044
	(0.145)	(0.084)	(0.047)
Hispanic	-0.213**	-0.004	0.049
	(0.079)	(0.052)	(0.030)
Black Mom	-0.748**	-0.562**	-0.470**
	(0.160)	(0.103)	(0.097)
Hispanic Mom	-0.464**	-0.248**	-0.290**
	(0.107)	(0.061)	(0.046)
Mother's education			
HS graduate		0.130**	0.136**
		(0.058)	(0.030)
Some college		0.258**	0.221**
		(0.061)	(0.052)
College graduate		0.332**	0.302**
		(0.066)	(0.067)
Graduate school		0.511**	0.515**
		(0.088)	(0.053)
Father's education			
HS graduate		0.029	0.037
		(0.058)	(0.062)
Some college		0.197**	0.222**
		(0.049)	(0.026)
College graduate		0.111**	0.111**
		(0.048)	(0.038)
Income ($1000)		0.008**	0.011**
		(0.003)	(0.003)
Single mom		-0.085*	-0.077*
		(0.044)	(0.041)
$\lambda_{1,wage}$	0.214**	0.347**	0.353**
	(0.050)	(0.036)	(0.048)
$\lambda_{1,coll}$	0.592**	1.510**	1.779**
	(0.071)	(0.131)	(0.225)
$\lambda_{1,GPA}$	0.458**	0.626**	0.632**
	(0.070)	(0.066)	(0.040)
$\lambda_{1,SciGPA}$	0.467**	0.648**	0.699**
	(0.063)	(0.069)	(0.075)
$\lambda_{1,MathGPA}$	0.667**	0.903**	0.957**
	(0.078)	(0.084)	(0.055)

Table 4.18 Full factor model of outcomes

Model:	(1)	(2)	(3)
$\lambda_{6,GPA}$			-9.959**
			(0.418)
$\lambda_{6,SciGPA}$			-8.777**
			(0.584)
$\lambda_{6,MathGPA}$			-9.900**
			(0.227)
Mother characteristics	No	Yes	Yes
Father characteristics	No	Yes	Yes
School FE	No	No	Yes

**,* significant at 5%, 10% level. Standard errors in parentheses.

race can be interpreted as the effects on test scores in standard deviation units, and multiplying them by the relevant λ_{1k} parameter gives the average effect for the other outcomes. The estimates from model 3 are directly comparable to the gaps presented in the last column of Table 4.7.

With the additional structure, the standard errors on the coefficients for own-race black or Hispanic fall when compared to Table 4.7. The coefficients on both own-race variables are small, positive, and insignificant. In contrast, the coefficients on maternal race variables are very large, negative, and statistically significant. The estimates of the own-race and maternal race parameters confirm the results of prior specifications but offer greater precision.[40]

4.6 Conclusion

Across a number of academic and early labor market outcomes, observables can fully account for differences in outcomes between black children with white mothers and white children with white mothers. Based on observables, white mothers with white children are negatively selected relative to white mothers with white children,

[40] We also estimated specifications on the subsample of blacks with white mothers and whites with white mothers, using the full model. Results are very similar: near zero own-race coefficients and large and significant maternal race coefficients.

suggesting that own race is not important for academic and labor market outcomes. Significant outcome gaps remain, however, between blacks with white mothers and blacks with black mothers.[41] Using the decomposition in Gelbach (2009), we assign almost half of the test score gap to unobserved factors correlated with mother's race for each race-gender combination analyzed.

This research has implications for how we formulate theories of human capital accumulation and discrimination. Our findings support the contention made by Heckman (2011) among others that the family environment is of primary importance in generating skill gaps observed later in life. While schools certainly play a role, we estimate that differences across schools only account for 20-30% of test score gaps. We also conclude that discrimination based on skin color is no longer the first-order concern. We argue instead that disparate outcomes must be operating through characteristics related to maternal race. Discrimination can still be important but must be operating through channels such as language (Grogger, 2011) that differ depending on race of the mother. The clear next step is to further isolate why it is that race of the mother correlates so strongly with education and labor market outcomes.

[41] While on demographics these two groups are similar, selection on unobservables may mean that these gaps are overstated.

Appendix A

Appendix to Chapter 2

This appendix describes in greater detail the original categorical responses for the time use data in Chapter 2. I also report the relative frequency of each response and the value I assign that response when I convert the categorical responses to a continuous measure.

Table A.1: Conversion of time use categories to continuous values

Response[a]	Relative frequency	Assigned value
Time on homework		*hrs/wk*
Has homework, but does not do it	0.013	0
Less than one hour each week	0.302	0.5
Between 1 and 3 hours	0.409	2
More than 3 but less than 5 hours	0.148	4
Between 5 and 10 hours	0.094	7.5
More than 10 hours	0.019	12
No homework is ever assigned	0.015	dropped
Amount of time spent free reading		*hrs/day*
None	0.116	0
About 30 minutes	0.481	0.5
About 1 hour	0.213	1
Between 1 and 2 hours	0.121	1.5
More than 2 hours	0.070	2.5
Student uses computer at home		*days/mo*
I use a computer at home for school work almost every day	0.067	24
Once or twice a week	0.167	6
Once or twice a month	0.182	1.5
Hardly ever	0.307	0.5
Never, even though there is a computer at home	0.154	0
There is no computer at home	0.123	dropped
TV watched at home each school day		*hrs/day*
None	0.045	0
1 hour or less each school day	0.258	0.5
2 hours	0.256	2
3 hours	0.201	3
4 to 5 hours	0.144	4.5
6 hours or more	0.097	7

[a] Variable titles and labels as they appear in the codebook.

Appendix B

Weight Adjustment Appendix

This appendix explains in detail how and why I adjust the probability weights π_{ig} for undermatching between the birth data and education data. Of the Hispanic 3rd graders, 19% are matched to a birth record. Though this match rate is low, we would expect it to be fairly low given that first-generation Hispanics will not be matched by definition. Furthermore, some second- and third-generation students living in North Carolina were not born there. According to the Census/ACS, 37% of school-aged Hispanics living in North Carolina are first generation. Census/ACS data also reveal that only half of later-generation Hispanics were born in the state. Thus, I would expect to match 32% of the school records to a birth record, based on this second data source. In contrast, the actual match rate for the pooled sample of Hispanics in 3rd through 8th grade is 15%. (The 19% above refers to the proportion of 3rd graders matched. Since each year Hispanics enter NC public schools for the first time, this match rate declines for each grade.) While there are enough differences between the two samples (e.g. the exact age ranges considered) that I would not expect the the two rates to line up exactly, the actual and expected match rates are different enough to merit some concern.

There are two main implications of this data set matching problem. One, I do not know as many immigrant generations with certainty. Two, matching fewer second- and third-generation Hispanics means that my weights might be off. These later generations, specifically the ones that were born in North Carolina, essentially need more representation in the weights. A third concern is that the second- and third-generation Hispanics that are not matched to a birth record might be selected.

I adjust the probability weights to correct for this under-matching problem. With the correction, the weights reflect the fact that more second- and third-generation students are not matched to a birth certificate than is indicated by the Census/ACS. To make the adjustment, I perform the following calculation. First, drawing on what I expect the match rates by generation to be from the Census/ACS, I calculate the proportion of observations that should have matched to birth record but did not. I then add these fractions to the proportion of students by generation that I would not expect to match anyway. The relative proportions for each of these generations are the unconditional (on parent's education) weights that I want, i.e. the weights that account for the under-matching. For each generation, I divide this number by the proportion of that generation that was acutally born out of state. Last, I multiply these weight adjusters by the sample weights provided by IPUMS (Ruggles et al., 2010). In my calculation of the probability weights by parent education (i.e. the weights used in estimation), I use these adjusted sample weights.

To get a sense of whether the later-generation Hispanics that are not matched are selected, I compare the parent education distributions for the Hispanic students in the NCERDC that are known to be second- or third-generation with what those distributions should be based on the Census/ACS. This comparison is in Table B.1. I also show the corresponding distributions for students that are not matched. However the generations mix together for the unmatched Hispanics, the distributions of parent

Table B.1: Actual vs. expected parent education distributions

	NCERDC known 2nd gen	Implied from ACS[a]	NCERDC known 3rd gen	Implied from ACS[b]	NCERDC unknown gen	Implied from ACS[c]
< HS	0.584	0.619	0.249	0.158	0.470	0.432
HS graduate	0.350	0.312	0.518	0.486	0.411	0.359
Jr. college	0.038	0.016	0.117	0.161	0.054	0.084
College	0.024	0.040	0.099	0.135	0.054	0.084
Grad school	0.005	0.014	0.018	0.061	0.011	0.052

NCERDC sample: pooled Hispanic students in grades 3-8 of the cohorts in 3rd grade in 2000 and 2001. Census/ACS sample: school-age Hispanic youth attending public school and living in North Carolina, 2000-2006. Used person-level weights provided by IPUMS.
[a] Second-generation Hispanics born in North Carolina.
[b] Third-generation Hispanics born in North Carolina.
[c] Hispanics of any generation not born in North Carolina.

education in the NCERDC are not far off from the expected distributions. For the second- and third-generation students in the NCERDC that are matched, the biggest inconsistency is that parents with the lowest education level are overrepresented for the third generation. This check suggests that selection into birth record matching is not a concern.

Appendix C

Model Fit Appendix

In this appendix, I compare estimates of the overall Hispanic-white test score gaps implied by my model (the "implied" gaps) to those obtained with ordinary least squares (the "actual" gaps). I calculate the implied gaps from a weighted average of the Hispanic-white gaps by generation. By definition, the raw gaps in Table C.1 should be identical—the estimates come from the same sample and the models use the same controls—and indeed, they are very close. To compare the models that focus on continuous enrollees, I restrict the sample to students present in 3rd and 8th grade to estimate the actual gaps. Table C.2 documents that these estimates are very similar. Thus, my limited controls for late arrival and early exit seem to adequately account for movement in and out of the sample. Finally, I compare actual and implied overall gaps in Table C.3. These gaps line up well, with the exception of 7th and 8th grade reading. Still, the differences are less than 0.1 standard deviations in these two cases. These tables offer a check on my econometric model, with its parsimonious controls for continuous enrollment and iterative estimation of school-by-year fixed effects.

Table C.1: Actual vs. implied raw Hispanic-white test score gaps

Grade	3rd	4th	5th	6th	7th	8th
A. Reading						
Actual	-0.681	-0.682	-0.731	-0.760	-0.745	-0.813
Implied	-0.685	-0.685	-0.738	-0.766	-0.760	-0.832
B. Math						
Actual	-0.588	-0.558	-0.575	-0.608	-0.622	-0.606
Implied	-0.592	-0.562	-0.580	-0.613	-0.630	-0.614

Sample: cohorts in 3rd grade in 2000 and 2001. Actual gaps are estimated by OLS from a model with race/ethnicity and missing other test as controls. Implied gaps are calculated from a weighted average of the test score gap for each Hispanic generation from Tables 3.6 and 3.7.

Table C.2: Actual vs. implied raw Hispanic-white test score gaps for continuous enrollees

Grade	3rd	4th	5th	6th	7th	8th
A. Reading						
Actual	-0.686	-0.603	-0.577	-0.540	-0.480	-0.487
Implied	-0.679	-0.603	-0.590	-0.557	-0.500	-0.522
B. Math						
Actual	-0.593	-0.511	-0.465	-0.438	-0.454	-0.406
Implied	-0.588	-0.509	-0.474	-0.457	-0.476	-0.443

Sample: cohorts in 3rd grade in 2000 and 2001. Actual gaps sample only uses students present in 3rd and 8th grades. Actual gaps are estimated by OLS from a model that includes race/ethnicity and missing other test as controls. Implied gaps are calculated from a weighted average of the test score gap for each Hispanic generation from Tables 3.8 and 3.9.

Table C.3: Actual vs. implied adjusted Hispanic-white test score gaps for continuous enrollees

Grade	3rd	4th	5th	6th	7th	8th
A. Reading						
Actual	-0.159	-0.077	-0.047	-0.030	-0.034	-0.032
Implied	-0.148	-0.066	-0.033	0.002	0.056	0.036
B. Math						
Actual	-0.114	-0.034	0.013	0.073	0.057	0.115
Implied	-0.108	-0.025	0.026	0.092	0.064	0.106

Sample: cohorts in 3rd grade in 2000 and 2001. Actual gaps sample only uses students present in 3rd and 8th grades. Test score gaps are adjusted for parent's education (5 categories), free/reduced lunch, gender, and school-by-year fixed effects. Actual gaps are estimated by OLS from a model that also includes race/ethnicity and missing other test as controls. Implied gaps are calculated from a weighted average of the test score gap for each Hispanic generation from Tables 3.10 and 3.11.

Appendix D

Appendix to Chapter 4

This appendix documents (1) selection into the sample, (2) results with different outcome measures (overall GPA, science GPA, and college completion), and (3) repeats the analysis of Table 4.7 without sample weights.

Table D.1 shows descriptive statistics in each wave, both conditional and unconditional on observing race of the mother. Those who persist to Wave IV have slightly higher test scores, but this is true for all racial groups. Validation studies show that the Wave I test score mean conditional on Wave III participation was 1.3% higher than the average among Wave I respondents.[1] Our Wave III sample differs because we require a transcript release in addition to participation in Wave III (and a valid maternal race observation). Comparing this measure of bias from attrition to our estimates shows that our selection criteria systematically increase test scores, but as shown in the final two rows of Table D.1, there are not large differences across races or time.

Tables D.2 and D.3 give estimation results for our alternative outcome measures

[1] Wave III and Wave IV validation studies are available at: http://www.cpc.unc.edu/projects/addhealth/data/guides/

Table D.1: Unweighted means across waves

	Wave I	Wave III[a]	Wave IV
Race			
Black	0.217	0.199	0.218
	(0.003)	(0.004)	(0.003)
Hispanic	0.171	0.157	0.159
	(0.003)	(0.003)	(0.003)
Other	0.081	0.076	0.070
	(0.002)	(0.002)	(0.002)
Race observations	18,906	11,540	14,788
Mom race			
Black mom	0.201	0.183	0.200
	(0.003)	(0.004)	(0.004)
Hispanic mom	0.149	0.135	0.139
	(0.003)	(0.004)	(0.003)
Other race mom	0.067	0.066	0.060
	(0.002)	(0.003)	(0.002)
Mom race observations	14,943	9,295	11,907
Test score			
Full sample	0.028	0.024	0.053
	(0.008)	(0.009)	(0.008)
Black	-0.389	-0.410	-0.345
	(0.016)	(0.021)	(0.017)
Mom race observed	0.028	0.057	0.082
	(0.008)	(0.010)	(0.009)
Black and mom race observed	-0.377	-0.390	-0.332
	(0.018)	(0.023)	(0.019)
Test score (weighted)			
Full sample	0.048	0.074	0.087
	(0.046)	(0.040)	(0.042)
Black	-0.524	-0.536	-0.504
	(0.071)	(0.060)	(0.070)
Mom race observed	0.073	0.101	0.109
	(0.047)	(0.040)	(0.044)
Black and mom race observed	-0.512	-0.525	-0.503
	(0.077)	(0.062)	(0.078)

[a] Wave III data are from the transcript file in Add Health.

for males and females, respectively. Overall GPA and science GPA are measured each year for each student enrolled in school during that year. The overall GPA regression controls for indicators of the math level, science level and year of schooling.[2] For males, we see the same patterns as we do with our other outcome measures. Namely, absent controls for mother's race, both black and Hispanic students have lower overall and science grades and are less likely to finish college. Once we control for mother's race, regardless of whether we account for school fixed effects, the coefficients on mother's race become small and insignificant. Having a black mother is negatively associated with each outcome. The same is true for Hispanic mothers but the estimates are imprecise. For females, the estimates are noisy, though adding controls does shrink the negative own-race effects.

The results in Table D.4 and D.5 shows that the estimates in Tables 4.7 and 4.8 are generally insensitive to whether weights are used. Gaps in test scores, math grades, and log wages are associated with the race of the mother, not the race of the child. For females, we observe similar patterns for test scores with no effects negative effects associated with race of the mother for wages.

Estimates in Table D.6 show decompositions for light-skinned individuals, complementing the analysis in Table 4.17 for this group. Maternal race is important in explaining differences in both test scores and grades, but it remains only a sizable predictor of wage gaps for light-skinned individuals when we condition on school fixed effects.

[2] From the Add Health codebook, "For each year of course taking, students are assigned to the category that reflects the highest level class they took for one semester or more, regardless of whether or not they received credit for the course. If a student took two different math courses in one year for example (such as Algebra II and Geometry), they are placed in the higher category (i.e., Algebra II)."

Table D.2: Minority outcome gaps for males, additional outcomes

Model:	(1)	(2)	(3)	(4)
Overall GPA[a]				
Black	-0.421**	-0.074	-0.011	0.056
	(0.042)	(0.085)	(0.082)	(0.083)
Hispanic	-0.216**	-0.038	-0.019	-0.049
	(0.045)	(0.073)	(0.073)	(0.073)
Black mom		-0.279**	-0.324**	-0.230**
		(0.090)	(0.090)	(0.092)
Hispanic mom		-0.085	-0.125	-0.139
		(0.077)	(0.078)	(0.075)
Science GPA[b]				
Black	-0.614**	-0.148	-0.067	-0.038
	(0.059)	(0.126)	(0.122)	(0.111)
Hispanic	-0.326**	-0.005	0.032	0.041
	(0.060)	(0.093)	(0.090)	(0.087)
Black mom		-0.340**	-0.426**	-0.311**
		(0.132)	(0.135)	(0.125)
Hispanic mom		-0.177*	-0.214**	-0.190**
		(0.100)	(0.099)	(0.092)
College completion[c]				
Black	-0.358**	0.131	0.405	0.249
	(0.107)	(0.209)	(0.254)	(0.286)
Hispanic	-0.424**	-0.093	0.285	0.288
	(0.096)	(0.144)	(0.181)	(0.205)
Black mom		-0.532**	-0.675**	-0.563**
		(0.221)	(0.272)	(0.319)
Hispanic mom		-0.411**	-0.271	-0.255
		(0.156)	(0.188)	(0.209)
Mother characteristics	No	Yes	Yes	Yes
Father characteristics	No	No	Yes	Yes
School FE	No	No	No	Yes

**,* significant at the 5%, 10% level. See Table 4.7 for a description of the controls.

[a] De-meaned overall GPA. Regressions include indicators for the math level, science level, and year of schooling.

[b] De-meaned course-level science GPA. Science GPA regressions include course-by-year fixed effects.

[c] Indicator of college completion by Wave IV.

Table D.3: Minority outcome gaps for females, additional outcomes

Model:	(1)	(2)	(3)	(4)
Overall GPA				
Black	-0.459**	-0.147	-0.040	-0.120
	(0.033)	(0.136)	(0.167)	(0.126)
Hispanic	-0.323**	-0.174**	-0.142	-0.151**
	(0.042)	(0.068)	(0.093)	(0.063)
Black mom		-0.231**	-0.553**	-0.036
		(0.139)	(0.172)	(0.127)
Hispanic mom		-0.041	-0.446**	0.051
		(0.076)	(0.104)	(0.073)
Science GPA				
Black	-0.528**	-0.200	-0.213	-0.263
	(0.044)	(0.213)	(0.214)	(0.199)
Hispanic	-0.424**	-0.250**	-0.244**	-0.270**
	(0.055)	(0.115)	(0.113)	(0.108)
Black mom		-0.208	-0.151	0.043
		(0.216)	(0.218)	(0.200)
Hispanic mom		-0.008	-0.003	0.054
		(0.125)	(0.126)	(0.127)
College completion				
Black	-0.252**	-0.042	-0.057	-0.107
	(0.109)	(0.193)	(0.255)	(0.302)
Hispanic	-0.377**	-0.334**	-0.137	-0.290
	(0.091)	(0.142)	(0.135)	(0.173)
Black mom		-0.219	-0.203	-0.121
		(0.199)	(0.260)	(0.307)
Hispanic mom		-0.051	-0.312**	-0.341**
		(0.159)	(0.145)	(0.171)
Mother characteristics	No	Yes	Yes	Yes
Father characteristics	No	No	Yes	Yes
School FE	No	No	No	Yes

**,* significant at the 5%, 10% level. See Table D.2 for a description of the outcomes and Table 4.7 for a description of the controls.

Table D.4: Minority outcome gaps for males, unweighted results

Model:	(1)	(2)	(3)	(4)
Test scores				
Black	-0.775**	-0.101	-0.096	-0.173**
	(0.027)	(0.081)	(0.080)	(0.079)
Hispanic	-0.741**	-0.142**	-0.115**	-0.088
	(0.029)	(0.055)	(0.055)	(0.055)
Black mom		-0.645**	-0.603**	-0.423**
		(0.082)	(0.083)	(0.082)
Hispanic mom		-0.422**	-0.402**	-0.354**
		(0.057)	(0.059)	(0.059)
Math GPA				
Black	-0.386**	-0.014	0.042	0.103
	(0.033)	(0.097)	(0.093)	(0.095)
Hispanic	-0.255**	-0.009	0.008	0.077
	(0.034)	(0.062)	(0.062)	(0.064)
Black mom		-0.352**	-0.416**	-0.380**
		(0.099)	(0.097)	(0.099)
Hispanic mom		-0.217**	-0.242**	-0.182**
		(0.064)	(0.066)	(0.069)
Log wages				
Black	-0.210**	0.015	0.037	0.022
	(0.024)	(0.077)	(0.077)	(0.078)
Hispanic	-0.021	0.015	0.042	-0.031
	(0.026)	(0.052)	(0.052)	(0.054)
Black mom		0.197**	0.220**	0.219*
		(0.078)	(0.079)	(0.080)
Hispanic mom		0.067	0.051	0.008
		(0.053)	(0.055)	(0.058)
Mother characteristics	No	Yes	Yes	Yes
Father characteristics	No	No	Yes	Yes
School FE	No	No	No	Yes

**,* significant at the 5%, 10% level. See Table 4.7 for a description of the dependent variables and controls.

Table D.5: Minority outcome gaps for females, unweighted results

Model:	(1)	(2)	(3)	(4)
Test scores				
Black	-0.704**	-0.159*	-0.132*	-0.209**
	(0.025)	(0.085)	(0.084)	(0.082)
Hispanic	-0.751**	-0.127**	-0.089	-0.063
	(0.029)	(0.059)	(0.058)	(0.057)
Black mom		-0.467**	-0.465**	-0.309**
		(0.086)	(0.085)	(0.083)
Hispanic mom		-0.471**	-0.462**	-0.397**
		(0.062)	(0.064)	(0.064)
Math GPA				
Black	-0.444**	-0.312**	-0.319**	-0.304**
	(0.028)	(0.108)	(0.108)	(0.112)
Hispanic	-0.374**	-0.143**	-0.130*	-0.066
	(0.033)	(0.071)	(0.070)	(0.068)
Black mom		-0.094	-0.064	0.035
		(0.099)	(0.097)	(0.099)
Hispanic mom		-0.191**	-0.171**	-0.073
		(0.074)	(0.074)	(0.073)
Log wages				
Black	-0.080**	-0.063	-0.056	-0.106
	(0.022)	(0.079)	(0.078)	(0.078)
Hispanic	0.087**	0.073	0.093	-0.008
	(0.026)	(0.053)	(0.052)	(0.053)
Black mom		0.039	0.060	0.104
		(0.079)	(0.079)	(0.079)
Hispanic mom		0.143**	0.136**	0.081
		(0.055)	(0.057)	(0.059)
Mother characteristics	No	Yes	Yes	Yes
Father characteristics	No	No	Yes	Yes
School FE	No	No	No	Yes

**,* significant at the 5%, 10% level. See Table 4.7 for a description of the dependent variables and controls.

Table D.6: Decomposition of outcome gaps for males with light skin tone

Model:	No school FE			With school FE		
Outcome:	Test	MGPA	Lwage	Test	MGPA	Lwage
Light tone						
Mother's char	-0.128**	-0.022	-0.057**	-0.102**	-0.014	-0.046**
	(0.026)	(0.019)	(0.016)	(0.023)	(0.018)	(0.014)
	[0.193]	[0.087]	[0.721]	[0.153]	[0.055]	[0.581]
Mother's race	-0.299**	-0.118**	0.004*	-0.240**	-0.089**	-0.045*
	(0.050)	(0.037)	(0.024)	(0.047)	(0.039)	(0.027)
	[0.451]	[0.470]	-[0.047]	[0.362]	[0.354]	[0.568]
Father's char	-0.087**	-0.003	-0.053**	-0.065**	0.006	-0.052**
	(0.026)	(0.024)	(0.013)	(0.024)	(0.023)	(0.014)
	[0.131]	[0.013]	[0.671]	[0.098]	-[0.024]	[0.663]
School FE				-0.077**	-0.074*	0.035
				(0.032)	(0.040)	(0.025)
				[0.116]	[0.293]	-[0.439]
Total	-0.514**	-0.143**	-0.106	-0.484**	-0.170	-0.108**
	(0.063)	(0.040)	(0.027)	(0.075)	(0.054)	(0.033)
	[0.775]	[0.570]	[1.345]	[0.729]	[0.677]	[1.373]
Baseline Gap	-0.663**	-0.251**	-0.079	-0.663**	-0.251	-0.079
	(0.083)	(0.062)	(0.045)	(0.083)	(0.062)	(0.045)

, group of characteristics significant at 5%, 10% level. Each cell contains the effect of each variable group on the white-minority outcome gap, the standard error in parentheses, and the fraction of the baseline outcome gap explained in brackets.

Bibliography

Aaronson, Daniel, Lisa Barrow, and William Sander. 2007. "Teachers and Student Achievement in the Chicago Public High Schools." *Journal of Labor Economics* 25 (1):95–135.

Altonji, Joseph G., Ulrich Doraszelski, and Lewis Segal. 2000. "Black/white differences in wealth." *Economic Perspectives* 24 (1):38–50.

Amato, Paul R. 2001. "Children of Divorce in the 1990s: An Update of the Amato & Keith (1991) Meta-Analysis." *Journal of Family Psychology* 15 (3):355–70.

Amato, Paul R. and Bruce Keith. 1991. "Parental Divorce and the Well-Being of Chidren: A Meta-Analysis." *Psychological Bulletin* 110 (3):26–46.

Austen-Smith, David and Roland G. Fryer, Jr. 2005. "An Economic Analysis of 'Acting White'." *Quarterly Journal of Economics* 120 (2):551–83.

Bertrand, Marianne and Jessica Pan. 2013. "The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior." *American Economic Journal: Applied Economics* 5 (1):32–64.

Björklund, Anders and Kjell G. Salvanes. 2011. "Education and Family Background: Mechanisms and Policies." In *Handbook of the Economics of Education*, vol. 3, edited by Eric A. Hanushek, Stephen Machin, and Ludger Woessman. Amsterdam: Elsevier, 201–47.

Bond, Timothy N. and Kevin Lang. 2013. "The Evolution of the Black-White Test Score Gap in Grades K-3: The Fragility of Results." *Review of Economics and Statistics* 95 (5):1468–79.

Borjas, George J. 1985. "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants." *Journal of Labor Economics* 3 (4):463–89.

Card, David and Alan B. Krueger. 1992. "School Quality and Black-White Relative Earnings: A Direct Assessment." *Quarterly Journal of Economics* 107 (1):151–200.

Carneiro, Pedro, James J. Heckman, and Dimitriy V. Masterov. 2005. "Labor Market Discrimination and Racial Differences in Premarket Factors." *Journal of Law and Economics* 48 (1):1–39.

———. 2008. "Understanding the Sources of Ethnic and Racial Wage Gaps and Their Implications for Policy." In *Handbook of Employment Discrimination Research*, edited by Laura Beth Nielsen and Robert L. Nelson. New York: Springer, 99–136.

Cascio, Elizabeth U. and Douglas O. Staiger. 2012. "Knowledge, Tests, and Fadeout in Educational Interventions." NBER Working Paper 18038.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014a. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review* 104 (9):2593–2632.

———. 2014b. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." *American Economic Review* 104 (9):2633–79.

Chiappori, Pierre-André, Sonia Oreffice, and Climent Quintana-Domeque. 2011. "Black-White Marital Matching: Race, Anthropometrics, and Socioeconomics." IZA Discussion Paper 6196.

Chiswick, Barry R. and Noyna DebBurman. 2004. "Educational attainment: analysis by immigrant generation." *Economics of Education Review* 23:361–79.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2006. "Teacher-Student Matching and the Assessment of Teacher Effectiveness." *Journal of Human Resources* 41 (4):778–820.

———. 2009. "The Academic Achievement Gap in Grades 3 to 8." *Review of Economics and Statistics* 91 (2):398–419.

———. 2012. "New Destinations, New Trajectories? The Educational Progress of Hispanic Youth in North Carolina." *Child Development* 83 (5):1608–22.

Cortes, Kalena. 2006. "The effects of age at arrival and enclave schools on the academic performance of immigrant children." *Economics of Education Review* 25:121–35.

Cunha, Flavio, James J. Heckman, Lance Lochner, and Dimitriy V. Masterov. 2006. "Interpreting the Evidence on Life Cycle Skill Formation." In *Handbook of the Economics of Education*, vol. 1, edited by Eric Hanushek and Finis Welch. Amsterdam: Elsevier, 697–812.

Cunha, Flavio, James J. Heckman, and Susanne M. Schennach. 2010. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica* 78 (3):883–931.

Currie, Janet and Jeffrey Grogger. 2002. “Medicaid expansions and welfare contractions: offsetting effects on prenatal care and infant health?” *Journal of Health Economics* 21 (2):313–35.

Currie, Janet and Duncan Thomas. 1999. “Intergenerational Transmission of ‘Intelligence’: Down The Slippery Slopes of *The Bell Curve*.” *Industrial Relations* 38 (3):297–330.

Currie, Janet M. 2011. “Health Disparities and Gaps in School Readiness.” *Future of Children* 15 (1):117–38.

Duncan, Brian and Stephen J. Trejo. 2011. “Intermarriage and the Intergenerational Transmission of Ethnic Identity and Human Capital for Mexican Americans.” *Journal of Labor Economics* 29 (2):195–227.

Eyferth, Klaus. 1961. “Leistungen verschiedener Gruppen von Besatzungskindern in Hamburg-Wechsler Intelligenztest für Kinder (HAWIK) [Performance of children on the Hamburg-Wechsler Intelligence Test for Children].” *Archhiv für die gesamte Psychologie* 113:222–41.

Felitti, Vincent J., Robert F. Anda, Dale Nordenberg, David F. Williamson, Alison M. Spitz, Valerie Edwards, Mary P. Koss, and James S. Marks. 1998. “Relationship of Childhood Abuse and Household Dysfunction to Many of the Leading Causes of Death in Adults: The Adverse Childhood Experiences (ACE) Study.” *American Journal of Preventative Medicine* 14 (4):245–58.

Ferguson, Ronald F. and Jordana Brown. 2000. “Certification Test Scores, Teacher Quality and Student Achievement.” In *Analytic Issues in the Assessment of Student Achievement*, edited by David W. Grissmer and J. Michael Ross. Washington, DC: National Center for Education Statistics, 133–56.

Figlio, David, Jonathan Guryan, Krzysztof Karbownik, and Jeffrey Roth. 2014. “The Effects of Poor Neonatal Health on Children’s Cognitive Development.” *American Economic Review* 104 (12):3921–55.

Fiorini, Mario and Michael P. Keane. 2014. “How the Allocation of Children’s Time Affects Cognitive and Noncognitive Development.” *Journal of Labor Economics* 32 (4):787–836.

Flynn, James R. 1999. “Searching for Justice: The Discovery of IQ Gains Over Time.” *American Psychologist* 54 (1):5–20.

Fryer, Roland G., Jr., Lisa Kahn, Steven D. Levitt, and Jörg L. Spenkuch. 2012. “The Plight of Mixed-Race Adolescents.” *Review of Economics and Statistics* 94 (3):621–34.

Fryer, Roland G., Jr. and Steven D. Levitt. 2004. "Understanding the Black-White Test Score Gap in the First Two Years of School." *Review of Economics and Statistics* 86 (2):447–64.

———. 2006. "The Black-White Test Score Gap Through Third Grade." *American Law and Economics Review* 8 (2):249–81.

———. 2013. "Testing for Racial Differences in the Mental Ability of Young Children." *American Economic Review* 103 (2):981–1005.

Gelbach, Jonah B. 2009. "When Do Covariates Matter? And Which Ones, and How Much?" Working Paper, University of Arizona.

Gill, Hannah. 2010. *The Latino Migration Experience in North Carolina: New Roots in the Old North State*. Chapel Hill, NC: The University of North Carolina State Press.

Grogger, Jeffrey. 2011. "Speech Patterns and Racial Wage Inequality." *Journal of Human Resources* 46 (1):1–25.

Hanushek, Eric. 1971. "Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro Data." *American Economic Review* 61 (2):280–88.

Hanushek, Eric A. and Steven G. Rivkin. 2006. "School Quality and the Black-White Achievement Gap." NBER Working Paper 12651.

Harris, David R. and Justin L. Thomas. 2002. "The Educational Costs of Being Multiracial: Evidence from a National Survey of Adolescents." Research Report 02-521, Population Studies Center, Institute for Social Research, University of Michigan.

Heckman, James J. 2011. "The American Family in Black and White: A Post-Racial Strategy for Improving Skills to Promote Equality." *Daedalus* 140 (2):70–89.

Heckman, James J. and Paul A. LaFontaine. 2010. "The American High School Graduation Rate: Trends and Levels." *Review of Economics and Statistics* 92 (2):244–62.

Jackson, C. Kirabo, Rucker C. Johnson, and Claudia Persico. 2015. "The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms." NBER Working Paper 20847.

Jencks, Christopher and Meredith Phillips. 1998. *The Black-White Test Score Gap*. Washington, DC: Brookings Institution.

Jensen, Arthur R. 1998. *The g Factor: The Science of Mental Ability*. Westport, CT: Praeger.

Kalil, Ariel and Kathleen M. Ziol-Guest. 2008. “Parental employment circumstances and children’s academic progress.” *Social Science Research* 37:500–15.

Kane, Thomas J., Jonah E. Rockoff, and Douglas O. Staiger. 2008. “What does certification tell us about teacher effectiveness? Evidence from New York City.” *Economics of Education Review* 27:615–31.

Kane, Thomas J. and Douglas O. Staiger. 2008. “Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation.” NBER Working Paper 14607.

Kao, Grace and Marta Tienda. 1995. “Optimism and Achievement: The Educational Performance of Immigrant Youth.” *Social Science Quarterly* 76 (1):1–19.

Krueger, Alan B. 1999. “Experimental Estimates of Education Production Functions.” *Quarterly Journal of Economics* 114 (2):497–532.

Mill, Roy and Luke C.D. Stein. 2012. “Race, Skin Color, and Economic Outcomes in Early Twentieth-Century America.” Working Paper, Stanford University.

Morgan, Michael and Larry Gross. 1980. “Television Viewing, IQ and Academic Achievement.” *Journal of Broadcasting* 24 (2):117–33.

Murnane, Richard J. 1975. *The Impact of School Resources on the Learning of Inner City Children*. Cambridge, MA: Ballinger Publishing.

Neal, Derek. 1997. “The Effects of Catholic Secondary Schooling on Educational Achievement.” *Journal of Labor Economics* 15 (1):98–123.

———. 2006. “Why Has Black-White Skill Convergence Stopped?” In *Handbook of the Economics of Education*, vol. 1, edited by Erik Hanushek and Finis Welch. Amsterdam: Elsevier, 511–76.

Neal, Derek A. and William R. Johnson. 1996. “The Role of Premarket Factors in Black-White Wage Differences.” *Journal of Political Economy* 104 (5):869–95.

Phillips, Meredith and Tiffani Chin. 2004. “School Inequality: What Do We Know?” In *Social Inequality*, edited by Kathryn Neckerman. New York: Russel Sage Foundation.

Rangel, Marcos A. 2014. “Is Parental Love Colorblind? Allocation of Resources within Mixed Families.” *Review of Black Political Economy* (Forthcoming).

Reardon, Sean F. and Claudia Galindo. 2009. “The Hispanic-White Achievement Gap in Math and Reading in the Elementary Grades.” *American Educational Research Journal* 46 (3):853–91.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. “Teachers, Schools, and Academic Achievement.” *Econometrica* 73 (2):417–58.

Rockoff, Jonah E. 2004. "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data." *American Economic Review* 94 (2):247–52.

Rong, Xue Lan and Linda Grant. 1992. "Ethnicity, Generation, and School Attainment of Asians, Hispanics, and Non-Hispanic Whites." *Sociological Quarterly* 33 (4):625–36.

Roth, Wendy D. 2005. "The End of the One-Drop Rule? Labeling of Multiracial Children in Black Intermarriages." *Sociological Forum* 20 (1):35–67.

Rothstein, Jesse. 2010. "Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement." *Quarterly Journal of Economics* 125 (1):175–214.

———. 2014. "Revisiting the Impacts of Teachers." Working paper.

Ruebeck, Christopher S., Susan L. Averett, and Howard N. Bodenhorn. 2009. "Acting White or Acting Black: Mixed-Race Adolescents' Identity and Behavior." *B.E. Journal of Economic Analysis and Policy* 9 (1):1–42.

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. 2010. *Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]*. Minneapolis: University of Minnesota.

Sackett, Paul R., Chaitra M. Hardison, and Michael J. Cullen. 2004. "On Interpreting Stereotype Threat as Accounting for African American-White Differences on Cognitive Tests." *American Psychologist* 59 (1):7–13.

Smith, James P and Finis R Welch. 1989. "Black Economic Progress after Myrdal." *Journal of Economic Literature* 27 (2):519–64.

Steele, Claude M. and Joshua Aronson. 1995. "Stereotype threat and the intellectual test performance of African Americans." *Journal of Personality and Social Psychology* 69 (5):797–811.

Stevens, Ann Huff and Jessamyn Schaller. 2011. "Short-run effects of parental job loss on children's academic achievement." *Economics of Education Review* 30:289–99.

Tillman, Kathryn Harker. 2007. "Family Structure Pathways and Academic Disadvantage among Adolescents in Stepfamilies." *Sociological Inquiry* 77 (3):383–424.

Todd, Petra E. and Kenneth I. Wolpin. 2003. "On the Specification and Estimation of the Production Function for Cognitive Achievement." *Economic Journal* 113 (485):F3–F33.

———. 2007. "The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps." *Journal of Human Capital* 1 (1):91–136.

US Bureau of Labor Statistics. 2011a. *The Employment Situation–August 2011*.

———. 2011b. *Usual Weekly Earnings of Wage and Salary Workers Second Quarter 2011*.

Willerman, Lee, Alfred F. Naylor, and Ntinos C. Myrianthopoulos. 1974. "Intellectual development of children from interracial matings: Performance in infancy and at 4 years." *Behavior Genetics* 4 (1):83–90.

Biography

Marie Claire Hull was born on February 27, 1987 in Lawrence, Kansas. She earned a B.A. in Economics, Mathematics, and Spanish from the University of Kansas in 2009. She also earned an M.A. in Economics in 2010 and a Ph.D. in Economics in 2015 from Duke University. Next fall, she will join the faculty in the Department of Economics at the University of North Carolina at Greensboro.