Available online at www.sciencedirect.com

ScienceDirect

Social Science Research 37 (2008) 239–253

Social Science RESEARCH

www.elsevier.com/locate/ssresearch

Gender differences in mathematics achievement: Exploring the early grades and the extremes ☆

Andrew M. Penner *, Marcel Paret

Department of Sociology, University of California, Berkeley, CA 94720-1980, USA

Available online 10 August 2007

Abstract

Gender differences in mathematics achievement have important implications for the underrepresentation of women in science. Typically, gender differences in mathematics achievement are thought to emerge at the end of middle school and beginning of high school, yet some studies find differences among younger children. This paper utilizes data from the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 to analyze differences in a nationally representative sample of kindergartners as they progress from kindergarten to fifth grade. Using quantile regression models to examine gender differences across the distribution, differences are found among students as early as kindergarten. Initially boys are found to do better at the top of the distribution and worse at the bottom, but by third grade boys do as well or better than girls throughout the distribution. The male advantage at the top of the distribution among entering kindergartners is largest among families with high parental education, suggesting that gender dynamics in middle and upper class families have important implications for continuing gender segregation in science occupations. Gender differences for entering kindergartners also vary across race, with Asians exhibiting the largest male advantage at the top of the distribution. In contrast to the overall pattern, among Latino kindergartners girls have an advantage over boys at the top of the distribution.

Keywords: Gender differences; Mathematics achievement; Kindergarten; Quantile regression; Education

1. Introduction

The continued underrepresentation of women in science is noteworthy as it leads to a lack of qualified workers in science-related occupations and provides a particularly salient example of occupational gender segregation. Xie and Shauman (2003) go so far as to question whether science is "the 'final frontier' for occupational gender equality" (p.1). While many factors contribute to the underrepresentation of women in science, recently the role of the gender gap at the upper extreme of the mathematical achievement distribution has

☆ We are grateful to Doug Downey, Mike Hout, Andrew Noymer, Trond Petersen, Meredith Phillips and two anonymous reviewers for useful comments and discussions. A previous version of this paper was presented at the annual meeting of the American Sociological Association in Montreal, August 2006.

* Corresponding author.
E-mail address: penner@berkeley.edu (A.M. Penner).

0049-089X/$ - see front matter
doi:10.1016/j.ssresearch.2007.06.012

received much attention (e.g., Summers, 2005; National Academy of Sciences, 2006). Wise et al. (1979) find that science professionals overwhelmingly score at or above the 90th percentile in mathematics in high school, and Hedges and Nowell (1995) show that there are a disproportionately small number of females scoring at this level. Thus, even while females have been closing achievement gaps and surpassed males in college attendance and graduation, their continued underperformance in mathematics achievement has significant consequences (Bae et al., 2000).

The consensus in the existing literature is that gender differences in mathematics achievement typically arise at a fairly late stage in students' careers. Hyde et al. (1990), for example, find in a meta-analysis that differences in mathematical problem solving do not exist in elementary or middle school, but do exist in high school and college.[1] Similarly, Muller (1998) and Leahey and Guo (2001) find that there are no differences among middle school students, and that differences emerge as students progress through high school. Gender differences in mathematics are generally assumed to result from differences in curricular choice, so that the male advantage in mathematics is largely a function of their greater preparation (e.g., Pallas and Alexander, 1983). Together, the idea that gender differences in mathematics achievement originate later in students' educational careers, and the notion that they are created by differences in curricular choice serve as the foundation to many of the policies addressing the underrepresentation of women in mathematics and the sciences (see, e.g., United States Congress, 2000).

Recent findings, however, provide evidence that gender differences might emerge at an earlier age. For example, studies of gender differences in spatial abilities show a male advantage on mental rotation tasks at the age of 4 years 6 months (Levine et al., 1999). Similarly, Levine et al. (2005) show that second and third grade middle and high SES males outperform their female counterparts on two spatial abilities tasks. Given the link between spatial ability and mathematical performance (Halpern, 2000), it seems plausible to expect that differences in mathematics achievement follow similar patterns.[2] Descriptive results from the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) confirm this, showing that gender differences in mathematics achievement emerge as early as first grade (Rathbun et al., 2004). These findings are noteworthy as they suggest that gender differences exist before curricular choice, and that policy interventions aimed at curricular choices are likely insufficient.

Even if early gender differences are small relative to those found in later stages of students' educational careers, the mere presence of such differences is important because early achievement gaps lead to larger disparities later in the education system. Early academic achievement is associated with cognitive skills and social-psychological factors, both of which are likely to have lasting effects on educational success (e.g., Hauser et al., 1983). In terms of the former, the early grades are crucial for the development of basic cognitive skills which serve as a foundation for continued academic progress (Entwisle and Alexander, 1998). Due to the cumulative nature of academic curricula, low mathematics achievement early on is likely to lead to a pattern of low achievement in mathematics. Indeed, evidence suggests that failure to obtain mathematics skills during the early grades substantially limits later opportunities for learning and cognitive development (Kilpatrick et al., 2001). From a social-psychological standpoint, not only do beginning students develop an academic reputation among their peers, teachers, and parents, but they also develop images of themselves as students (Entwisle and Alexander, 1993). This "educational foundation" affects important factors such as track placement, teacher expectations, and student motivation, all of which place students on a trajectory of academic achievement (Farkas, 2003a).[3]

These findings suggest that if early gender differences in mathematics achievement do exist, they need to be considered seriously. If gender differences occur before curricular choices are made, then other explanations are needed. Given the salience of familial influences in young children (e.g., Downey and Condron, 2004; Far-

[1] It is worth noting that Maccoby and Jacklin (1974) find in an earlier review of the literature that quantitative differences between boys and girls emerge around the ages 9–13.

[2] As mathematics involves multiple mental tasks, we would expect that gender differences in mathematics would be some kind of a weighted average of the gender differences in the various mental tasks involved, and would vary somewhat from test to test depending on the composition. However, as mathematics is an important social construction, and is related to these underlying tasks, it is worth examining as a coherent entity.

[3] It is important to recognize that a wide variety of factors originating both inside and outside of educational institutions operate and interact to cause the expansion of academic achievement gaps. For example, research suggests that summer learning is crucial to understanding the growth of academic achievement gaps between students from different social class backgrounds (e.g., Burkam et al., 2004; Chin and Phillips, 2004; Downey et al., 2004).

kas, 2003), we draw from Levine et al. (2005) and examine how differences in family background, specifically parental education and race, affect gender differences. Over the past few decades, scholars, and feminist scholars in particular, have pointed to the intersecting nature of domination and inequality in society (e.g., Collins, 1986). Gender inequalities do not exist in a vacuum, but rather interact with other axes of inequality such as race and class. Previous research shows that gender gaps in academic achievement vary significantly across race and SES at the high school and postsecondary levels (e.g., Riegle-Crumb, 2006), and a substantial body of research confirms the importance of race and class for early academic achievement (e.g., Jencks and Phillips, 1998; Lee and Burkam, 2002). Thus, not only should any study of early gender achievement gaps take into account variation across these other axes of inequality, but doing so is likely to be extremely revealing.

This study makes three main contributions. First, we establish the extent to which early gender differences in mathematics exist using a nationally representative sample. Most research in this area has been based on small samples (e.g., Levine et al., 1999, 2005), and large sample research on early academic achievement that does exist does not focus heavily on gender (e.g., Rathbun et al., 2004; Fryer and Levitt, 2004).[4] Second, in addition to comparing differences in mean levels of mathematics achievement, we also employ quantile regression models to examine gender differences across the entire distribution. Focusing on differences throughout the distribution is particularly important in this case because evidence suggests that occupational gender segregation stems specifically from achievement gaps at the upper extreme. Third, the findings with respect to gender differences in the early grades and across the entire distribution have important implications for policies aimed at eliminating the gender gap in math-related fields. Our results suggest that policy strategies focused on students' curricular choices need to be refined to take into account the differences present among young children.

2. Data and variables

This paper uses data from the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), which provides detailed information about students and their educational experiences between kindergarten and fifth grade. The study began with a nationally representative sample of kindergarten students who were enrolled in a United States public or private school in 1998–99. Data were collected for this same group of students at six points in time: fall of kindergarten, spring of kindergarten, fall of first grade, spring of first grade, spring of third grade, and spring of fifth grade.[5] It is also important to note two sample modifications during the first grade data collections. First, the sample was freshened, meaning that students were added to the sample to make it nationally representative of the first grade population. Second, to conserve costs, a nationally representative subsample (approximately 30 percent) of students were included in the fall of first grade data collection. Although the ECLS-K was designed for longitudinal analyses of students over time, here we are mostly concerned with how math score distributions change as we move from grade to grade. Thus, while the samples do include the same students, we employ cross-sectional weights in order to treat each sample separately.[6] This has the added benefit of preserving as many observations as possible, as observations that fall out due to sample attrition are still included in the earlier analyses.

[4] Over the past decade or so scholars have paid increasing attention to academic achievement gaps present in early childhood. For example, Phillips et al. (1998) revealed that black–white test score differences emerge even before students enroll in school, and Lee and Burkam (2002) confirmed this finding with the ECLS-K data. As these studies suggest, however, the most influential research in this area has focused primarily on race and socioeconomic status (but see Riordan, 2002).

[5] The study followed students as they progressed through the early grades, though some students were retained in a grade. Thus, in the spring of 2000 (in most cases the spring of first grade) four percent of the students were enrolled in kindergarten; in the spring of 2002 (in most cases the spring of third grade), 11 percent of students were enrolled below third grade; and in spring of 2004 (in most cases the spring of fifth grade), 13 percent of students were enrolled below fifth grade. Less than one percent of the students were enrolled either above the modal grade or in an ungraded classroom. In addition to the waves mentioned above, data were also collected for a subsample of students (roughly 30 percent) in the fall of first grade. While data from all students in the cohort are used regardless of grade level, we refer to the different waves by the majority grade level.

[6] Due to the sample design, the kindergarten and first grade data are representative of the entire populations of 1998–1999 kindergartners and 1999–2000 first graders, respectively. Unfortunately, the sample was not freshened for the third grade and fifth grade data collections. Thus, in these grades the data are representative of students who were either kindergartners in 1998–1999 or first graders in 1999–2000, and are still in school.

Table 1
Cognitive skills evaluated in standardized mathematics assessments, by skill area

Level	Skill area	Cognitive skills evaluated
Math skills		
1	Number and shape	Identifying some one-digit numerals, recognizing geometric shapes, and one-to-one counting of up to ten objects.
2	Relative size	Reading all one-digit numerals, counting beyond ten, recognizing a sequence of patterns, and using nonstandard units of length to compare objects.
3	Ordinality and sequence	Reading two-digit numerals, recognizing the next number in a sequence, identifying the ordinal position of an object, and solving a simple word.
4	Addition and subtraction	Solving simple addition and subtraction problems.
5	Multiplication and division	Solving simple multiplication and division problems.
6	Place value	Understanding place value in integers to the hundreds place.
7	Rate and measurement	Using knowledge of measurement and rate to solve word problems.
8	Fractions	Understanding the concept of fractional parts.
9	Area and volume	Solving word problems involving area and volume, including change of units of measurement.

The measures of mathematics achievement are based on standardized tests of cognitive ability in mathematics. The tests were untimed and administered one-on-one, assessing student proficiency in nine mathematical skill areas: (1) number and shape, (2) relative size, (3) ordinality and sequence, (4) addition and subtraction, (5) multiplication and division, (6) place value, (7) rate and measurement, (8) fractions, and (9) area and volume. The specific skills associated with each area are provided in Table 1. The parental education and race-ethnicity measures are based on parent-reported questionnaire responses. Parental education (highest level of attainment among both parents) is broken into five categories: (1) less than a high school diploma, (2) high school diploma, (3) some college, (4) bachelor's degree, and (5) advanced degree. Race-ethnicity is coded white (non-Latino), black (non-Latino), Latino, Asian, and other.

3. Methods

On the subject of the golden mean English theologian Charles Simeon said, "The truth is *not in the middle*, and *not in one extreme; but in both extremes*" (Simeon, 1847, p. 600). The lesson is applicable here, as research on educational differences has focused almost exclusively on differences in means and variances. While this is understandable, given that means and variances are simple to compute and easy to understand, it is not clear that analyzing differences in means and variances always provides the relevant information. The mean is a measure of the central tendency of the distribution, and the variance provides information about distances from the mean. While thinking about differences in terms of the center of the distribution is useful, two other obvious reference points, the top and the bottom, have been neglected. Regarding issues such as women in the sciences, the top and bottom tails of the distribution are of more interest than the center. While processes occurring in the center of a distribution and average group differences are often of great interest, the assumption that the factors affecting central tendencies are the same as the factors affecting extreme tendencies is too often made without thought. Additionally, predictions about the extremes based on OLS regressions (which essentially provide conditional means) implicitly assume that the top and the bottom of the distribution are subject to identical processes. While this is not necessarily a bad assumption, it seems foolhardy to believe that: (1) all of the processes affecting the low-achievers affect the high-achievers, (2) only processes affecting the low-achievers affect the high-achievers, and (3) these processes have effects of the same magnitude for the low- and high-achievers.

There are many studies that, in addition to looking at mean differences, also examine differences in variances. This too is not necessarily inappropriate (see, for example, Archer and Mehdikhani, 2003). However, for our purposes variance-based comparisons are not sufficient, as they still privilege the center of the distribution. As Grissom and Kim argue, "It is far more informative to compare distributions at various quantiles, such as deciles, than to compare them only at their centers" (2001, p. 142). While it might not always be necessary to compare distributions at multiple points, it is certainly worth thinking about which points are worthy of comparison. Indeed, given the degree of discussion that surrounds questions of measurement and scale construction, it seems irresponsible not to at least consider non-mean-based comparisons.

There is some precedent for an extreme-oriented approach in the educational literature on gender differences. Feingold (1992) argues that differences in means and variances should be observed together to better capture the extremes, and Hedges and Friedman (1993) go a step further by arguing that it is preferable to observe gender differences in the extremes directly, proposing the ratio of boys to girls in the extremes as a measure. This approach is adopted in several studies that report not only gender differences in the means and variances, but also the ratios of girls to boys at the extremes of the distribution (Hedges and Nowell, 1995; Nowell and Hedges, 1998; Stumpf and Stanley, 1996; Xie and Shauman, 2003). Recent work in education by Konstantopoulos (2004) and Levin (2001) uses quantile regression to look at differences across the distribution. While quantile regression models are used widely in economics to observe how effects vary across the distribution, there is yet little work using these models in sociological analyses of education.[7]

Quantile regression uses least absolute value (LAV, also known as least absolute deviations) estimation to estimate conditional differences in the median and other quantiles in the distribution. This can be thought of as estimating the percentiles for boys and the percentiles for girls separately, and then observing the difference between the boys' P percentile and the girls' P percentile. Thus, where OLS reports conditional differences in means, quantile regression reports conditional differences in percentiles. More formally, the models estimated in this paper take the standard form:

$$y_i = X_i\beta + \varepsilon_i \tag{1}$$

where y_i is the mathematics score for student i, and X_i includes the independent variables, in this case the constant and a dummy variable for being female. As laid out in Koenker and Bassett (1978), this model can be estimated at the θth quantile by minimizing Eq. (2):

$$\min_{\beta}\left[\sum_{\{i|y_i \geqslant X_i\beta\}} \theta|y_i - X_i\beta| + \sum_{\{i|y_i < X_i\beta\}} (1-\theta)|y_i - X_i\beta|\right] \tag{2}$$

Intuitively, what we are doing here is estimating β at different quantiles by changing the weights (θ and $1-\theta$) on the positive and negative residuals. For example, at the median ($\theta = .5$) positive and negative residuals are given equal weight so that the sum of absolute deviations is minimized.

The figures presented in this paper graph the effect of being female across a series of quantiles. For any given quantile the female effect is obtained from Eq. (3):

$$y_i = \beta_0 + G_i\beta_1 + \varepsilon_i \tag{3}$$

Here β_0 is the constant and G is a dummy variable for being female. Specifically, the figures plot the size of β_1 on the y-axis with θ on the x-axis. [8]

If it is strange to think about gender effects at different percentiles in the distribution, this is likely because of a common usage of the concept of effects in reporting OLS regression results. The gender effect, as it is often called, indicates a difference in conditional means, or the average difference between males and females net of other variables. This has come to signify the difference between Everyman and Everywoman. Using quantile regression to look at the extremes is confusing then because while we (wrongly) think of Everyman and Everywoman apart from the population that produces them, it is harder to do this for Extreme-man and Extreme-woman. However, in the same way that it does not make sense to think of those at the 10th percentile apart

[7] Quantile regression models are often used in analyses of wage differences between genders (e.g., Buchinsky, 1994), sectors (e.g., Lucifora and Meurs, 2006), and education levels (e.g., Buchinsky, 1998). Other analyses have used quantile regression to examine birthweight (Koenker and Hallock, 2001), the demand for alcohol (Manning et al., 1995), and school effects on student performance (Eide and Showalter, 1998).

[8] Models are estimated using Koenker's quantreg package for R. The standard errors reported are Huber sandwich estimates using Hall and Sheather's bandwidth and kernel estimation (see Koenker and Hallock, 2000 for details). For an excellent non-technical introduction to quantile regression, see Koenker and Hallock (2001).

from the rest of the distribution they represent, we should not think of the mean apart from the distribution that produces it. Put differently, the mean, like the 10th percentile, is a property of the distribution, and not of any individual, mythical or not. In an attempt to minimize confusion around this here, in describing the results we will generally refer to the gender effect as the difference at that percentile.

4. Results

Table 2 reports basic descriptive statistics over the different grade levels, and also includes gender differences in standard deviation units at the mean and various percentiles. Overall, we see that the gender differences vary substantially across the distribution. For example, at kindergarten entry the difference in average mathematics scores for boys and girls is small and statistically insignificant. At the 95th percentile, however, boys have an advantage of 0.15 standard deviation units. To put this in perspective, this is roughly four times the size of the previously established mean difference reported for spring of first grade (Rathbun et al., 2004). Looking across time we see that while the gap at the 95th percentile does not seem to grow as students progress through school, the gender difference at the 75th percentile grows to be about the same size. The nature of the female disadvantage changes over time as well, so that in kindergarten and first grade the gap exists mostly at the top of the distribution, while in third and fifth grade the gap exists more or less throughout the distribution and is largest at the median. The quantile regression models provide similar results, which we describe in more detail. The quantile regression results are presented graphically in Fig. 1.

Fig. 1 graphs the results from a series of bivariate quantile regressions of mathematics score on gender for different grade levels. The *x*-axes on the graphs are the percentiles, and the *y*-axes are the effects, so that the graphs show the "gender effect" at different percentiles. The black line provides the estimate of the gender difference at that percentile, and the grey shading shows the 90 percent confidence interval. The grey horizontal line depicts the OLS regression estimate of the gender effect, and the faint dashed grey lines provide its 90 percent confidence interval. We see that in fall of kindergarten girls do better at the lower end of the distribution than boys. That is, if we compare the 1st through 40th percentiles for boys to the 1st through 40th percentiles for girls, we find that girls score higher. At the top of the distribution, on the other hand, boys score higher than girls. While there are differences at both ends of the distribution, there are no differences at the median and mean. This suggests that analyses of mean differences are unable to detect the important gender gaps that are already present before students begin school. Spring of kindergarten reveals a similar overall pattern, though both the female advantage at the bottom and the male advantage at the top are slightly less widespread and the middle area with no difference includes a broader range of percentiles. Both fall and spring of first grade show similar results, with a female advantage in the bottom half of the distribution and a male advan-

Table 2
Descriptive statistics and sex differences for mathematics scores, by data collection point

	Round of data collection					
	Fall kindergarten	Spring kindergarten	Fall first grade	Spring first grade	Spring third grade	Spring fifth grade
Mean	19.12	27.07	32.41	42.79	83.25	111.22
Standard deviation	7.23	8.81	9.61	9.49	18.28	22.41
Differences by sex (female–male), in standard deviation units						
Mean	0.00	−0.01	0.00	−0.04	−0.15	−0.16
Percentile						
5%	0.12	0.06	0.16	0.02	−0.01	−0.11
25%	0.06	0.04	0.05	0.04	−0.12	−0.15
Median	0.04	0.02	0.01	−0.04	−0.26	−0.22
75%	−0.05	−0.03	−0.10	−0.16	−0.17	−0.14
95%	−0.15	−0.21	−0.15	−0.16	−0.17	−0.11
N	18,635	19,647	5,223	16,635	14,374	11,274
Sum of weights	3,801,035	3,824,462	3,819,311	3,915,655	3,909,152	3,903,170

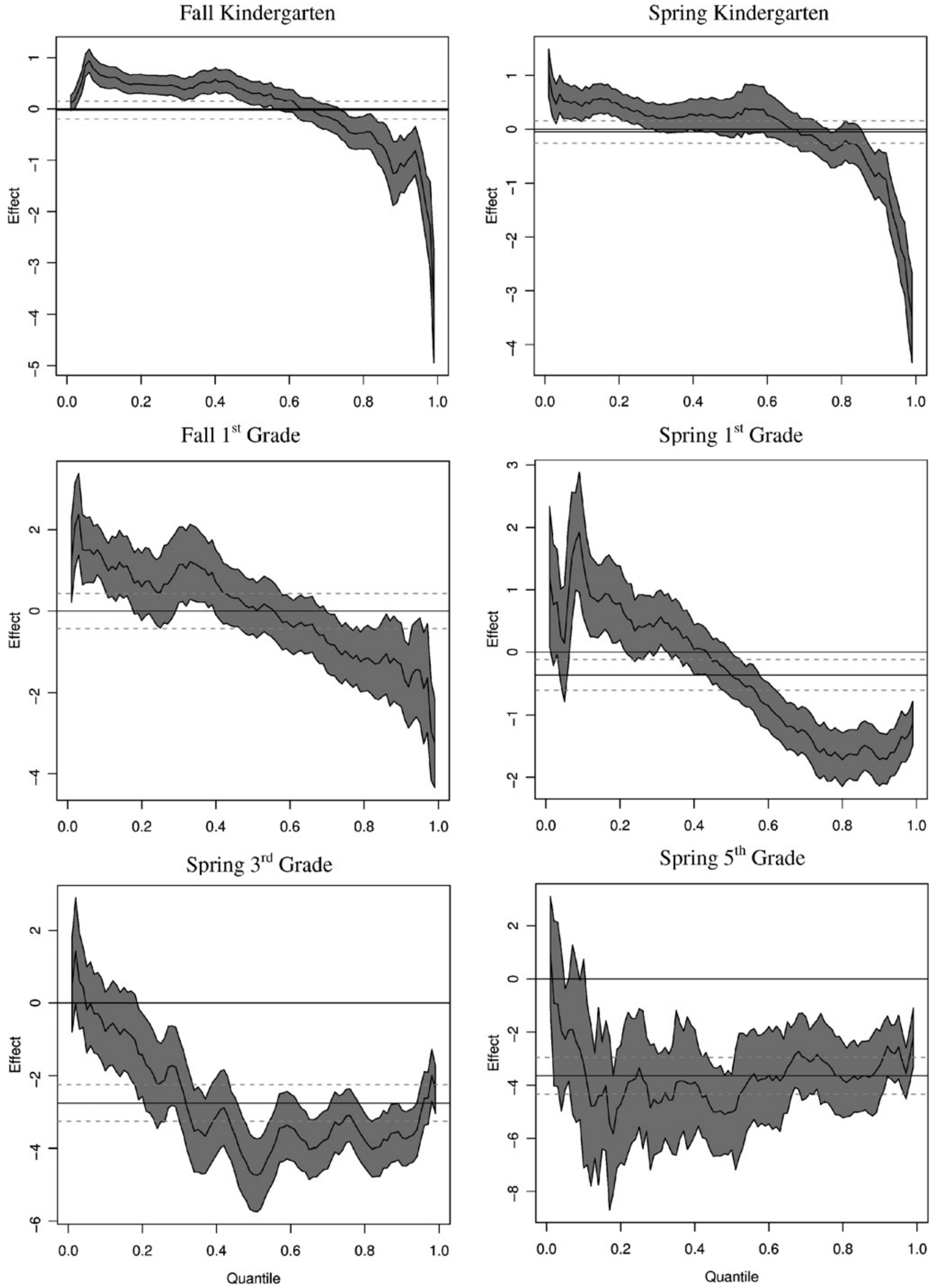

Fig. 1. The effect of gender (female) on mathematics achievement across the distribution for different ages. *Note:* Black, quantile regression estimate of gender effect for different percentiles. Grey shading, confidence interval of quantile regression estimate. Grey line, OLS estimate. Dashed grey line (faint), OLS confidence interval.

tage in the top half. The male advantage at the very top is smaller, but the portion of the distribution exhibiting a male advantage is larger.[9] It is also worth noting that in the spring of first grade the mean gender effect

[9] The decline of the male advantage at the very top of the distribution is interesting as it goes against the simple idea that the male advantage at the very top just keeps getting bigger.

is significant for the first time. By spring of third grade the female advantage at the bottom of the distribution is gone, and there is a relatively consistent male advantage from about the 35th percentile up. Fifth grade exhibits a strikingly stable gender difference across the distribution from the 10th percentile up. In both third grade and fifth grade the differences at the top of the distribution do not vary substantially from the differences across the rest of the distribution, so that the top of the distribution is no longer the part with the largest male advantage.

Four points from Fig. 1 are worth highlighting. First, as Rathbun et al. (2004) show, mean differences in mathematics achievement are visible as early as spring of first grade. This is significant because it contradicts the conventional wisdom that differences emerge later in students' academic careers, and does so with nationally representative data. Second, we find that differences in the tails of the distribution arise earlier than the mean differences, and are visible even in the fall of kindergarten. This indicates not only that differences are present at an early age, but also that they are not exclusively a function of the educational system. Indeed, the commonly cited male advantage at the upper extreme of the mathematics distribution is found here to be present at the beginning of kindergarten. Third, differences at various points across the distribution change over time. The amount of change visible in the effect of gender over time shows that differences are not static, suggesting that it may be possible to reduce the gaps through social interventions. Fourth, differences at the extremes are different than at the mean, sometimes quite substantially. These results should serve as a cautionary tale for OLS regression results, as they show that mean differences are not necessarily indicative of the distribution as a whole.

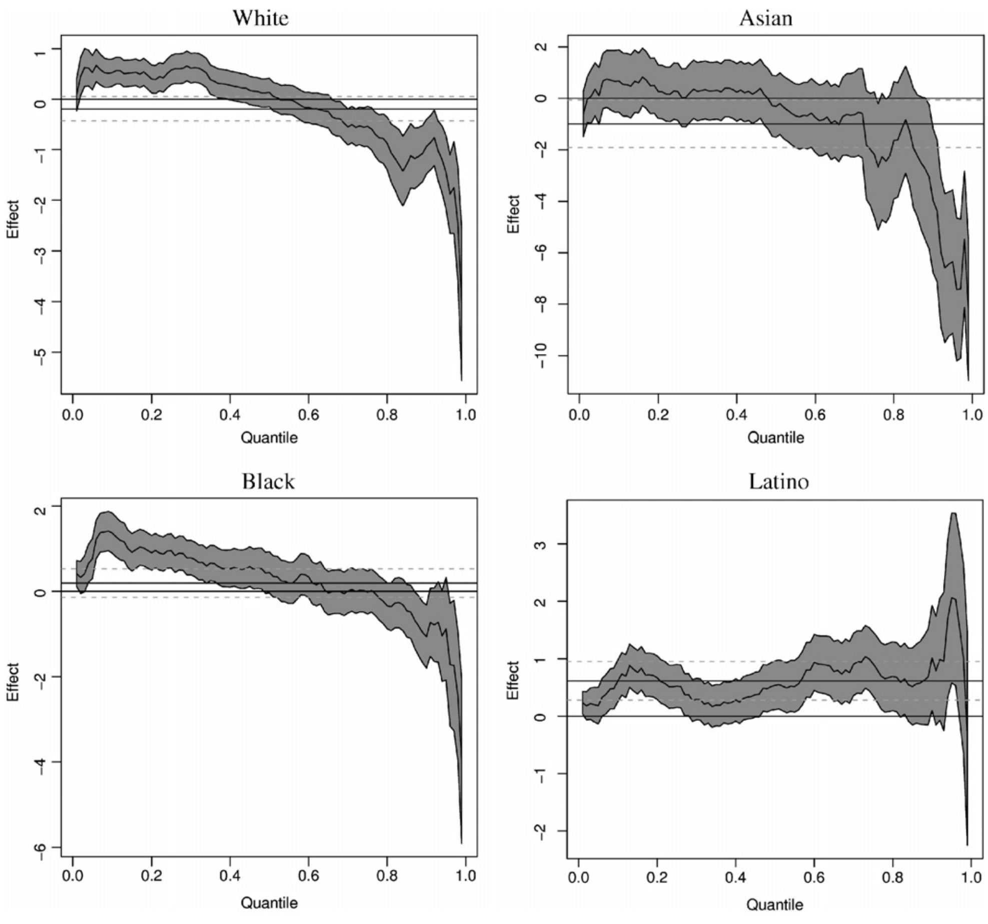

Fig. 2. The effect of gender (female) on mathematics achievement across the distribution in fall of kindergarten separately by race. *Note:* Black, quantile regression estimate of gender effect for different percentiles. Grey shading, confidence interval of quantile regression estimate. Grey line, OLS estimate. Dashed grey line (faint), OLS confidence interval.

Given the difference between the effect of gender at the mean and in other places throughout the distribution, it is worth examining how gender differences are impacted by introducing other independent variables into the equation. For this purpose we build on the fall of kindergarten model by estimating the gender effect separately by race and parental education. This is similar to adding controls and interactions into the model, but is preferable because the effects remain simple to interpret. This examination is useful because it reveals the extent to which the observed patterns are being driven by certain social groups. For example, if the male advantage at the top of the distribution is mainly reflected among certain socioeconomic or racial groups, then this provides a window into the possible mechanisms underlying the gender gap.

Table 3
Descriptives for fall of kindergarten mathematics scores, by race and parental education

Population	*N*	Sum of Weights	Mean	Percentiles				
				5%	25%	Median	75%	95%
Overall	18,635	3,801,035	19.12	9.03	13.88	18.16	22.68	32.83
Male	9,479	1,948,457	19.13	8.72	13.64	18.01	22.86	33.42
Female	9,156	1,852,578	19.11	9.57	14.10	18.31	22.51	32.33
Race								
White								
Overall	10,433	2,176,886	21.0	10.9	15.7	19.9	24.7	34.7
Male	5,351	1,124,527	21.1	10.7	15.4	19.8	24.9	35.5
Female	5,082	1,052,359	20.9	11.2	16.0	20.0	24.4	34.1
Black								
Overall	2,860	613,143	16.6	8.5	12.8	16.0	19.7	26.4
Male	1,431	309,970	16.5	8.2	12.2	15.7	19.7	27.0
Female	1,429	303,172	16.7	9.0	13.2	16.2	19.7	26.1
Latino								
Overall	3,390	737,744	15.8	8.1	11.7	14.6	19.0	27.6
Male	1,716	375,011	15.5	8.1	11.4	14.4	18.5	26.6
Female	1,674	362,733	16.1	8.3	11.9	14.9	19.4	28.7
Asian								
Overall	898	90,052	22.2	11.4	16.5	20.4	26.7	39.0
Male	444	44,989	22.7	11.0	16.4	20.6	27.1	41.9
Female	454	45,063	21.7	11.5	16.6	20.2	25.3	35.8
Other race								
Overall	1,023	176,246	17.4	8.3	12.5	16.2	20.8	31.4
Male	519	90,031	17.4	8.3	12.4	16.0	20.6	32.3
Female	504	86,215	17.4	8.5	12.6	16.7	20.8	30.7
Parental education								
Less than high school								
Overall	1,743	383,899	14.2	8.0	10.6	13.5	17.0	22.3
Male	907	200,936	14.1	7.9	10.5	13.2	16.8	22.7
Female	836	182,963	14.2	8.0	10.7	13.5	17.2	21.9
High school								
Overall	4,646	1,001,683	16.9	8.4	12.7	16.1	20.0	28.7
Male	2,356	513,436	16.8	8.2	12.4	15.9	20.1	29.5
Female	2,290	488,247	17.0	8.8	13.0	16.4	20.0	27.8
Some postsecondary								
Overall	5,783	1,190,363	19.1	10.2	14.5	18.5	22.3	31.8
Male	2,906	598,594	19.1	9.8	14.3	18.3	22.4	32.3
Female	2,877	591,769	19.1	10.4	14.7	18.7	22.2	31.1
Bachelor’s degree								
Overall	3,581	682,167	22.5	12.5	17.2	21.2	26.3	36.2
Male	1,816	350,328	22.6	12.3	17.0	21.2	26.6	37.0
Female	1,765	331,839	22.3	12.9	17.4	21.2	26.2	34.7
Advanced degree								
Overall	2,060	379,632	24.9	13.0	19.2	23.4	30.5	40.7
Male	1,061	196,968	25.2	13.0	18.8	23.5	30.8	42.4
Female	999	182,664	24.6	13.0	19.6	23.2	29.9	38.7

Fig. 2 estimates the gender effect for the fall of kindergarten separately by race. We find that there are substantial differences across racial groups. Whites and blacks both exhibit the same pattern as seen overall, with a female advantage at the bottom of the distribution and a male advantage in the upper tail. The results for Asians have a similar shape, but there is no female advantage at the bottom of the distribution, and the male

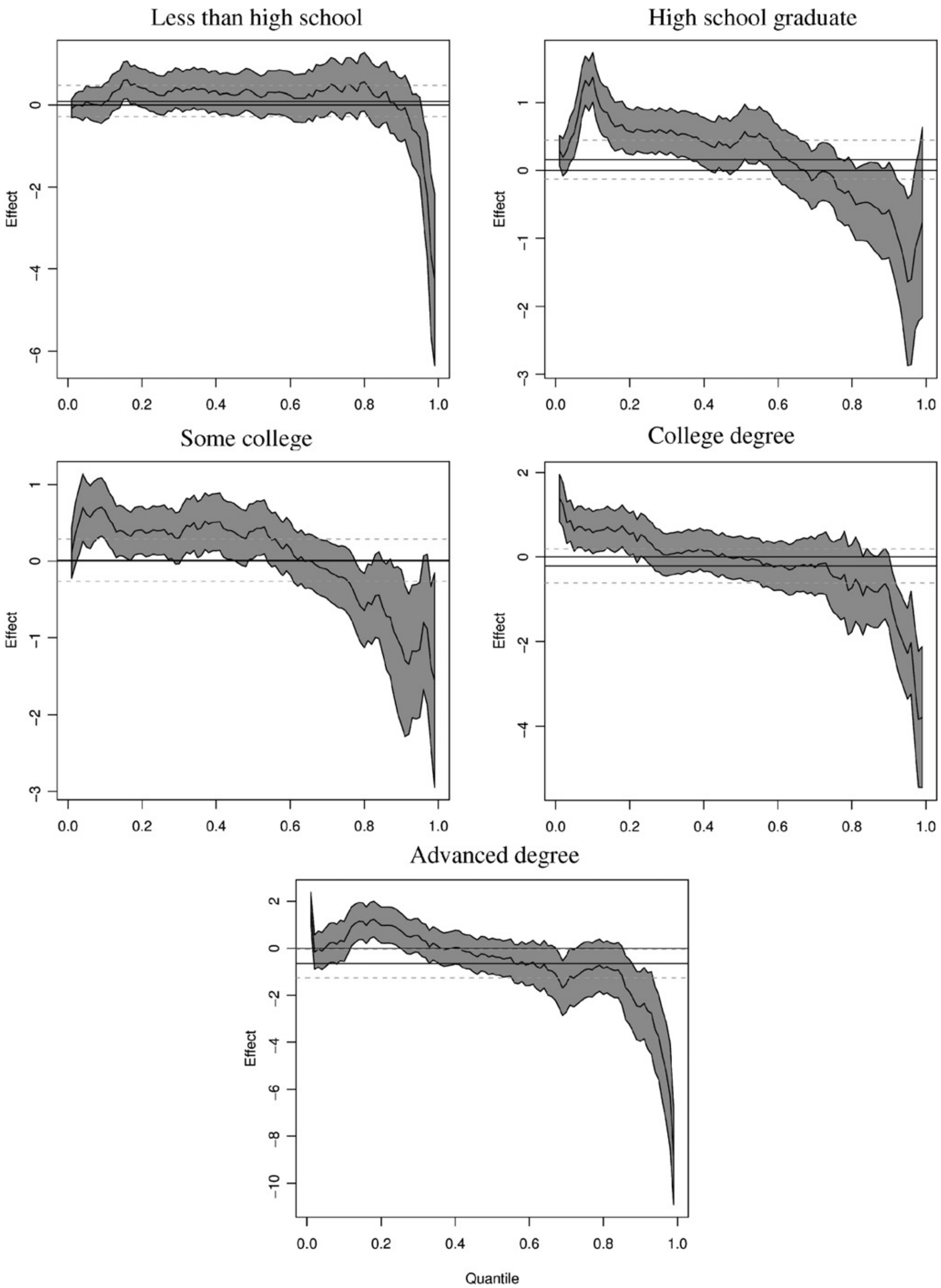

Fig. 3. The effect of gender (female) on mathematics achievement across the distribution in the fall of kindergarten separately by parental education. *Note:* Black, quantile regression estimate of gender effect for different percentiles. Grey shading, confidence interval of quantile regression estimate. Grey line, OLS estimate. Dashed grey line (faint), OLS confidence interval.

advantage at the top is considerably larger than in the general population. Among Latinos we find that not only is there a female advantage at the bottom of the distribution, but girls have an even larger advantage at the top of the distribution. The variation across racial groups shows rather definitively that the pattern of gender differences across the distribution is affected by family background.

In thinking about how these differences combine to form the overall gender difference there are two points that are important to bear in mind. First, as with any analysis broken down separately by racial groups, it is important to remember that the groups are not of equal size in the population. Whites are 57% of the sample, blacks 16%, Asians 2%, and Latinos 19%. Second, in thinking about differences across the distribution, it is important to remember that while the overall male and female differences are a composite of the race-specific male–female differences, the scores also vary across race. For example, the 95th percentiles for boys and girls are 33.4 and 32.2, respectively. But the 95th percentiles for Asian boys and Asian girls are 41.9 and 35.8, respectively, and the 95th percentiles for Latino boys and Latino girls are 26.6 and 28.7, respectively. Thus, while Fig. 2 shows that Asians females are scoring 6 points lower than Asian boys, and Latina girls are scoring 2 points higher than Latino boys, it is worth noting that these comparisons are being made at scores that differ from each other, as well as those in the overall comparisons. Table 3 provides information about the overall, male, and female distributions for different racial and parental education groups, providing a sense of where the various comparisons are being made along the range of scores. Due to the fact that blacks and Latinos have generally lower math scores, the gender differences at the tops of their distributions have less impact on gender differences at the top of the overall distribution than do the gender differences at the tops of the white and Asian distributions. Thus, the male advantage at the top is likely being driven by gaps among whites and Asians. Nonetheless, the female advantage at the top of the Latino distribution is important for thinking about the mechanisms facilitating gender gaps, and is something that future research should investigate further.

Fig. 3 examines how gender differences in the fall of kindergarten vary across parental education groups. The same overall pattern of a male advantage at the top and female advantage at the bottom holds for all of the parental education groups, but there are noteworthy differences among groups. In particular, the higher the parental education, the further the male advantage extends over the distribution. Thus, among the less than high school group, males only have an advantage above the 90th percentile, but among the advanced degree group the male advantage extends from the 40th percentile up. The magnitude of the male advantage at the very top of the distribution is relatively stable among the lowest four parental education groups, but is substantially larger among the advanced degree group. This latter finding is consistent with previous research suggesting that gender differences are driven by students from upper and middle class backgrounds (Levine et al., 2005). Taken together these results suggest that male advantages in math are mediated by parental education, and in particular that the male advantage is most pronounced among students with higher parental education. This suggests that future research on gender gaps in mathematics achievement should pay close attention to mechanisms and practices associated with the middle and upper classes.

5. Discussion

This study estimates quantile regression models using nationally representative data on kindergarten through fifth grade students to examine gender differences throughout the distribution of mathematics achievement. We find that gender differences emerge at an early age and vary across the distribution. These findings have important substantive, methodological, and policy implications.

Substantively, we are able to confirm that gender differences in mathematics achievement exist at the outset of students educational careers using nationally representative data. This confirms the findings of Levine et al. (1999), and extends them to mathematical achievement more generally. We find that in the fall of kindergarten boys do better than girls at the top of the distribution, and girls do better than boys at the bottom. By the spring of third grade, however, the female advantage at the bottom disappears and the male advantage extends to most of the distribution. These findings are particularly important because they demonstrate that the male advantage at the top is present at the beginning of the educational process, and shows that it only spreads to the rest of the distribution after children have been in school for a while. The achievement gaps present at kindergarten entry differentially position students in the education system, paving the way for the development of the familiar achievement gaps often found at the end of high school.

This suggests that the gender differences in mathematics achievement associated with occupational gender segregation are likely to be more than just the product of gendered educational processes, though our study does not provide any evidence that such processes are absent. Because girls are less likely to enter kindergarten at the very top of the distribution, they are also subject to the same institutionalized processes (inside or outside of school) that contribute to expanding achievement gaps among racial and class groups. Increasing our understanding of the mechanisms that facilitate expanding gaps between low achievers and high achievers will therefore benefit our understanding of the eventual gender gaps that in turn facilitate occupational gender segregation. At the same time, future researchers may choose to focus their attention on the early childhood years before kindergarten, as our evidence also suggests that gender differences are being created at a very young age.

In addition to documenting the existence of gender differences among young children, we also show that family background plays a substantial role. Gender differences are found to vary substantially across racial groups, and also somewhat by parental education. Where race is concerned it is interesting to note that gender differences vary across race not only in magnitude but also in pattern. This suggests that there is nothing predetermined about these differences, and that simple biological theories are unlikely to be as helpful as social and cultural factors in explaining these differences.[10] The overall gender gap at the top of the distribution, present at the beginning of kindergarten, is being driven by whites, blacks, and Asians—not Latinos—though the differences are most pronounced among Asian students. Thus, future research on young Asian students and their families might be useful for illuminating overall patterns of gender difference. Conversely, research comparing Latinos to other racial groups is also likely to be useful for this purpose, as Latinos represent the one population where females actually have an advantage at the top of the distribution at kindergarten entry.

We are also able to confirm Levine et al.'s (2005) finding that the development of gender gaps varies by socio-economic status. While the female advantage at the bottom of the distribution is similar for all students regardless of parental education, the male advantage at the top of the distribution is most pronounced among students whose parents have a college or advanced degree. This is interesting in that it indicates that the male advantage at the top of the distribution is mediated by socioeconomic factors in a way that the female advantage at the bottom does not appear to be. More specifically, gender segregation in science occupations may be related to gender dynamics in middle and upper class families.[11]

In thinking about the many factors that could contribute to race and SES variation in gender differences, two seem worth noting. The first follows from Correll's (2001, 2004) examination of societal stereotypes. Given the effect of broader cultural stereotypes on performance (Steele, 1997), it seems plausible that the variation in the gender differences in mathematics scores could be due to variation in the gender stereotypes prevalent in different groups. While most stereotype threat research focuses on adults, there is some evidence that suggests that stereotype threat exists even in young children (Ambady et al., 2001). Another potentially important factor involves the transmission of cultural resources between parents and children. For example, it could be that the childrearing practices of middle and upper class families that have been shown to benefit students (e.g., Lareau, 2003), are, when it comes to mathematics, more beneficial to boys than to girls. Parents may be more likely to pass on mathematics related "cultural resources" to boys than to girls, or boys may be more likely than girls to be able to "activate" these resources. This is congruent with Muller's (1998) finding that

[10] Simple biological arguments that men are better than women in mathematics are unsatisfactory given the variation in gender differences across the distribution, across racial categories, and across educational groups. Approaches combining biological and social factors, like Guo and Stearns' (2002) research showing that high-SES environments allow individuals to better achieve their "genetic potential," are more promising. However, even this relatively sophisticated theory of biological and social factors still needs considerable refinement to explain our results. In particular, the steady female advantage at the bottom of the distribution is difficult to explain using this line of reasoning, and even the SES mediated male advantage at the top would require the existence of undocumented genetic gender differences in mathematical potential.

[11] In separate analyses we found that in the fall of kindergarten the commonly noted female advantage in reading is smallest at the top of the distribution. In fact, among high-SES families boys actually have an advantage, albeit a statistically insignificant one, at the very top of the distribution. Although reading and mathematics are quite different in terms of the cultural norms surrounding them, these results suggest an interesting gender and SES interaction that is worth exploring further. Results are available upon request.

among high school students parental involvement matters more for boys than for girls.[12] It will be important to pursue these explanations for the SES-mediated gender gaps, as well as other potential explanations, in future research.

The methodological implications of this study are both straightforward and noteworthy. In plotting the OLS results alongside the quantile regression results, we are able to compare the findings obtained by the two methods. We find that while OLS results often provide an accurate summary of differences across the distribution, they miss important variation in differences across the distribution. Thus, while differences in the mean are important and tell us about how populations differ on average, differences at the extremes of the distribution often vary substantially from the mean differences. In cases such as gender differences in mathematics, where the extremes of the distribution are of more interest than the middle, quantile regression or other extreme-sensitive methods should be employed. This method also suggests a host of new research topics; for example, do factors such as school funding, parental encouragement, and experienced teachers matter more for the top of the distribution than the bottom?

Finally, our results have important implications for policy on gender inequality in mathematics and science. We show that achievement differences at the top of the distribution are visible as early as kindergarten, and previous research shows that early academic achievement is a good predictor of later educational success. Thus, if we aim to eliminate the gender gap in science occupations, we must begin to acknowledge the differentiation between boys and girls that occurs as early as kindergarten, and probably earlier. More specifically, focusing our attention on girls' decisions about course enrollment is not likely to solve the problem, as gender differences emerge before students are even allowed to make such decisions. The current study suggests that if we seek to develop policy that addresses the roots of the gender gap in mathematics and science, we must first develop an understanding of the social factors that contribute to gender differences during early childhood.

Of course, this means that policies aimed at changing course-taking decisions are not likely to be the solution. Rather, effective strategies for reducing gaps in mathematics achievement will need to be based on a much broader view that situates differential levels of achievement within larger structures of inequality. This is not to say that attempts to change course taking are not useful, but by themselves they are unlikely to eliminate the gender gap. As previous research has shown, educational inequalities have a tendency to both reflect and reinforce structures of inequality, whether they are related to gender, race, or class divisions. In the case of math achievement these three axes of difference interact to produce gender inequalities among boys and girls, and impact the gender gaps in science occupations. Thus, rather than focusing narrowly on the course-taking decisions of middle and high school girls, it is important to focus on the ways in which processes of academic achievement are embedded in and shaped by hierarchical structures of difference and inequality.

References

Ambady, Nalini, Shih, Margaret, Kim, Amy, Pittinsky, Todd L., 2001. Stereotype threat in children: effects of identity on quantitative performance. Psychological Science 12, 385–390.

Archer, John, Mehdikhani, Mani, 2003. Variability among males in sexually selected attributes. Review of General Psychology 7, 219–236.

Bae, Yupin, Choy, Susan, Geddes, Claire, Sable, Jennifer, Snyder, Thomas, 2000. Trends in Educational Equity of Girls and Women. National Center for Education Statistics, Washington, DC.

Buchinsky, Moshe, 1994. Changes in the US wage structure 1963–1987: application of quantile regression. Econometrica 62, 405–458.

Buchinsky, Moshe, 1998. The dynamics of changes in the female wage distribution in the USA: a quantile regression approach. Journal of Applied Econometrics 13, 1–30.

Burkam, David T., Ready, Douglas D., Lee, Valerie E., LoGerfo, Laura F., 2004. Social-class differences in summer learning between kindergarten and first grade: model specification and estimation. Sociology of Education 77, 1–31.

Chin, Tiffani, Phillips, Meredith, 2004. Social reproduction and child-rearing practices: social class, children's agency, and the summer activity gap. Sociology of Education 77, 185–210.

Collins, Patricia Hill, 1986. Learning from the outsider within: the sociological significance of black feminist thought. Social Problems 33, S14–S32.

[12] Although these explanations could explain the role of SES at the top of the distribution, they do not explain why the female advantage at the bottom of the distribution does not vary by SES. This would require that the cultural norms or stereotypes regarding gender difference at the bottom of the distribution differ from those at the top.

Correll, Shelley J., 2001. Constraints into preferences: gender, status, and emerging career aspirations. American Sociological Review 69, 93–113.
Correll, Shelley J., 2004. Gender and the career choice process: the role of biased self-assessments. American Journal of Sociology 106, 1691–1730.
Downey, Douglas B., von Hippel, Paul T., Broh, Beckett A., 2004. Are schools the great equalizer? Cognitive inequality during the summer months. American Sociological Review 69, 613–635.
Downey, Douglas B., Condron, Dennis J., 2004. Playing well with others in kindergarten: the benefits of siblings at home. Journal of Marriage and Family 66, 333–350.
Eide, Eric, Showalter, Mark H., 1998. The effect of school quality on student performance: a quantile regression approach. Economics Letters 58, 345–350.
Entwisle, Doris R., Alexander, Karl A., 1993. Entry into school: the beginning school transition and educational stratification in the United States. Annual Review of Sociology 19, 401–423.
Entwisle, Doris R., Alexander, Karl A., 1998. Facilitating the transition to first grade. Elementary School Journal 98, 351–364.
Farkas, George, 2003. Cognitive skills and noncognitive traits and behaviors in stratification processes. Annual Review of Sociology 29, 541–562.
Farkas, George, 2003a. Racial disparities and discrimination in education. Teachers College Record 105, 1119–1146.
Feingold, Alan, 1992. Sex differences in variability in intellectual abilities. Review of Educational Research 62, 61–84.
Fryer, Roland G., Levitt, Steven D., 2004. Understanding the Black–White test score gap in the first two years of school. The Review of Economics and Statistics 86, 447–464.
Grissom, Robert J., Kim, John J., 2001. Review of assumptions and problems in the appropriate conceptualization of effect size. Psychological Methods 6, 135–146.
Guo, Guang, Stearns, Elizabeth, 2002. The social influences on the realization of genetic potential for intellectual development. Social Forces 80, 881–910.
Halpern, Diane F., 2000. Sex Differences in Cognitive Abilities. Erlbaum Associates, NJ.
Hauser, Robert M., Tsai, Shu-Ling, Sewell, William H., 1983. A model of stratification with response error in social and psychological variables. Sociology of Education 56, 20–46.
Hedges, Larry V., Friedman, Lynn, 1993. Gender differences in variability in intellectual abilities: a reanalysis of Feingold's results. Review of Educational Research 63, 94–105.
Hedges, Larry V., Nowell, Amy, 1995. Sex differences in mental test scores, variability, and numbers of high scoring individuals. Science 269, 41–45.
Hyde, Janet S., Fennema, Elizabeth, Lamon, Susan J., 1990. Gender differences in mathematics performance: a meta-analysis. Psychological Bulletin 107, 139–155.
Kilpatrick, Jeremy, Swafford, Jane, Findell, Bradford, 2001. Adding It Up: Helping Children Learn Mathematics. National Academy Press, Washington DC.
Jencks, Christopher, Phillips, Meredith, 1998. The Black–White Test Score Gap. Brookings Institution Press, Washington, DC.
Koenker, Roger, Bassett Jr., Gilbert, 1978. Regression quantiles. Econometrica 46, 33–50.
Koenker, Roger, Hallock, Kevin F., 2000. Quantile regression: an introduction. Retrieved, February 20, 2006. <http://www.econ.uiuc.edu/~roger/research/intro/rq.pdf>.
Koenker, Roger, Hallock, Kevin F., 2001. Quantile regression. Journal of Economic Perspectives 15, 143–156.
Konstantopoulos, Spyros, 2004. Trends in Asian-white differences in achievement. Presented at the Annual Meeting of the American Sociological Association, August 14, San Francisco, CA.
Lareau, Annette, 2003. Unequal Childhoods: Class, Race, and Family Life. University of California Press, Berkeley, CA.
Leahey, Erin, Guo, Guang, 2001. Gender differences in mathematical trajectories. Social Forces 80, 713–732.
Lee, Valerie E., Burkam, David T., 2002. Inequality at the Starting Gate: Social Background Differences in Achievement as Children Begin School. Economic Policy Institute, Washington, DC.
Levin, Jesse, 2001. For whom the reductions count: a quantile regression analysis of class size and peer effects on scholastic achievement. Empirical Economics 26, 221–246.
Levine, Susan C., Vasilyeva, Marina, Lourenco, Stella F., Newcombe, Nora S., Huttenlocher, Janellen, 2005. Socioeconomic status modifies the sex difference in spatial skill. Psychological Science 16, 841–845.
Levine, Susan C., Huttenlocher, Janellen, Taylor, Amy, Langrock, Adela, 1999. Early sex differences in spatial skill. Developmental Psychology 35, 940–949.
Lucifora, Claudio, Meurs, Dominique, 2006. The public sector pay gap in France, Great Britain and Italy. Review of Income and Wealth 52, 43–59.
Maccoby, Eleanor E., Jacklin, Carol N., 1974. The Psychology of Sex Differences. Stanford University Press, Palo Alto, CA.
Manning, William G., Blumberg, Linda, Moulton, Lawrence H., 1995. The demand for alcohol: the differential response to price. Journal of Health Economics 14, 123–148.
Muller, Chandra, 1998. Gender differences in parental involvement and adolescents' mathematics achievement. Sociology of Education 71, 336–356.
National Academy of Sciences, 2006. Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering. National Academies Press, Washington, DC.
Nowell, Amy, Hedges, Larry V., 1998. Trends in gender differences in academic achievement from 1960 to 1994. Sex Roles 39, 21–43.

Pallas, Aaron M., Alexander, Karl L., 1983. Sex-differences in quantitative SAT performance—new evidence on the differential coursework hypothesis. American Educational Research Journal 20, 165–182.
Phillips, Meredith, Crouse, James, Ralph, John, 1998. Does the black–white test score gap widen after children enter school? In: Jencks, C., Phillips, M. (Eds.), The Black–White Test Score Gap. Brookings Institution Press, Washington, DC, pp. 229–272.
Rathbun, Amy, West, Jerry, Germino-Hausken, Elvira, 2004. From Kindergarten Through Third Grade: Children's Beginning School Experiences. National Center of Education Statistics, Washington, DC.
Riegle-Crumb, Catherine, 2006. The path through math: course sequences and academic performance at the intersection of race-ethnicity and gender. American Journal of Education 113, 101–122.
Riordan, Cornelius, 2002. Male/female early achievement and development: gender gaps among first time kindergarten children in the public schools. Presented at the Annual Meeting of the American Educational Research Association, April 4, New Orleans, LA.
Simeon, Charles, 1847. In: Carus, W. (Ed.), Memoirs of the Life of the Rev. Charles Simeon: With a Selection from his Writings and Correspondence. Hatchard and Son, London.
Steele, Claude M., 1997. A threat is in the air: how stereotypes shape intellectual identity and performance. American Psychologist 52, 613–629.
Stumpf, Heinrich, Stanley, Julian C., 1996. Gender-related differences on the college board's advanced placement and achievement tests, 1982–1992. Journal of Educational Psychology 88, 353–364.
Summers, Lawerence H., 2005. Remarks at NBER conference on diversifying the science & engineering workforce. Retrieved, September 12, 2006. <http://www.president.harvard.edu/speeches/2005/nber.html>.
United States Congress, 2000. A Review of the Morella Commission Report. United States Government Printing Office, Washington, DC.
Wise, Lauress L., Steel, Lauri, MacDonald, Charlotte, 1979. Origins and Career Consequences of Sex Differences in High School Mathematics Achievement. American Institute for Research, Palo Alto, CA.
Xie, Yu, Shauman, Kimberlee A., 2003. Women in Science. Harvard University Press, Cambridge, MA.