Differences at the Extremes? Gender, National Contexts, and Math Performance in Latin America

Abstract

Studies of gender disparities in STEM performance have generally focused on average differences. However, the extremes could also be important because disparities at the top may shape stratification in access to STEM careers, while disparities at the bottom can shape stratification in dropout. This paper investigates determinants of gender disparities in math across the performance distribution in Latin American countries, where there is a persistent boys' advantage in STEM performance. Findings reveal disparate national patterns in gender gaps across the performance distribution. Further, while certain national characteristics are linked to gender gaps at the low- and middle-ranges of the performance distribution, female representation in education is the only characteristic associated with a reduced gender gap at the top level.

Keywords: STEM education, gender, Latin American

Introduction

The phenomenon of gender differences in STEM performance is a continuing concern, as it relates to the underrepresentation of women at the highest levels of science, technology, mathematics, and engineering (Else-Quest, Hyde, & Linn, 2010). Some scholars have argued that there may be greater variability in performance among boys than girls (Feingold, 1992; L. V. Hedges & Nowell, 1995; Hyde, Lindberg, Linn, Ellis, & Williams, 2008), which suggests the importance of considering patterns of gender disparity at the extremes of the performance distribution. The extremes of the distribution could also be important because disparities at high performance levels may shape stratification in access to high-level STEM education and careers (Fan, Chen, Matsumoto, & Fan, 1997; Xie & Shauman, 2003), while disparities at the bottom can shape stratification in grade repetition and dropout (Janosz, LeBlanc, Boulerice, & Tremblay, 1997; Jimerson, Egeland, Sroufe, & Carlson, 2000). However, cross-national studies of gender disparities in student STEM performance have generally focused on average differences.

There are other limitations in the comparative literature on gender gaps in STEM performance. Analyses of the national sources of variation in gender gaps in educational performance have tended to focus on national economic development and national gender equality indicators in various societal domains—education, economics, politics, and cultural norms (Else-Quest et al., 2010; Guiso, Monte, Sapienza, & others, 2008; Penner, 2008; Riegle-Crumb, 2005). National education system characteristics, such as degree of privatization, standardization, or stratification (branching or tracking), while studied extensively in the context of socioeconomic stratification (Park, 2008; Van de Werfhorst & Mijs, 2010), have not been considered as routinely in studies of gender disparities in math performance, with few exceptions

(Ayalon & Livneh, 2013). In addition, very little comparative research has analyzed factors shaping gender differences in math performance across Latin America.

This paper addresses these limitations. We investigate determinants of gender disparities in math performance across the performance distribution in Latin American countries, where there is a persistent boys' advantage in math performance in most countries despite a female-favoring gender gap in educational attainment (UNESCO, 2018). Using the Third Regional Comparative and Explanatory Study (TERCE) data from 15 Latin American countries, we address two questions: 1) Does the gender difference in math performance vary across the performance distribution? In particular, is there greater male variability? 2) Are gender differences across the math performance distribution associated with national-level factors, including economic development, gender equality regimes, and education system characteristics?

In addressing these questions, this paper begins to rectify a significant regional imbalance in empirical work on gender and STEM performance and will contribute a case to weigh in on whether there is a need to consider separately inequalities at high and low performance levels. The remainder of this paper reviews comparative and Latin America-specific literature in English and Spanish relevant to gender disparities in math performance, introduces the data and methods, presents results, and discusses implications.

Framework

Gender gaps at the extremes

Previous studies in the U.S. have shown that gender differences in mean mathematic performance are very small and sometimes favor girls, depending on the sample, measure, and educational stage. For example, a recent meta-analysis of U.S. studies found no significant gender difference in elementary and middle school, but small gender gaps in complex problem solving favoring male students in high school and college (Lindberg, Hyde, Petersen, & Linn, 2010). Similarly, using the National Assessment of Educational Progress data from 1990 to 2015, researchers find that male students have a negligible advantage in math in fourth grade and no advantage in eighth grade; a meaningful advantage only emerges in high school (Fahle & Reardon, 2018). However, a new study using the U.S. state accountability test data from third to eighth grade students found that although there is no overall gender achievement gap in math, there are considerable variations across school districts. Math gaps tend to favor male students more in socioeconomically advantaged school districts as well as in districts with larger gender differences in adult socioeconomic status (Reardon, Fahle, Kalogrides, Podolsky, & Zárate, 2019). This district variation suggests that gender gaps in math performance may be linked to broader social and structural contexts.

International studies also show considerable ambiguity: a gender gap in math performance persists in some countries while not in others. Using data from the 2003 Trends in International Mathematics and Science Study (TIMSS, which surveys 4th and 8th grade students) and the Programme for International Student Assessment (PISA, which surveys 15-year-old students), Else-Quest et al. (2010) find that although over 60 percent of the countries show a negligible to small gender difference in math performance, this gender gap varies greatly across countries. The gender effect size, measured by the difference between male and female means divided by the pooled within-gender standard deviation, varies from -0.42 in Bahrain to 0.40 in Tunisia.

One potentially significant limitation in the existing literature is its focus on mean differences in math performance, rather than differences at the extremes of performance.

Attention to the extremes is important, because gender differences at the extreme ends of performance are often more substantial than gender difference at the means (Baye & Monseur, 2016). Moreover, disparities at high performance levels may shape stratification in access to high-level STEM education and careers (Fan et al., 1997; Xie & Shauman, 2003), while disparities at the bottom can shape stratification in grade repetition and dropout (Janosz et al., 1997; Jimerson et al., 2000). For these reasons, an exclusive focus on average differences could elide socially significant disparities in math performance. An important exception to this characterization is Penner (2008), who examines extreme math performance through logistic regression and quantile regression models. Penner's sample of 22 countries, however, are mostly western developed countries.

Further, there is an ongoing debate about the "greater male variability hypothesis": the notion that independent of mean differences, male students have a greater variance than female students in math ability and therefore are more likely to be at both the top and the bottom of the distribution of math performance (Lindberg et al., 2010). This hypothesis is sometimes used to explain the underrepresentation of women in scientific research fields, given that if women had smaller variability in math performance, they would be underrepresented in the top of the distribution. Thus, a test of the greater male variability hypothesis could provide insights into the origin of the excess of male students at the top levels of math performance and math-intensive careers (Larry V. Hedges & Friedman, 1993; Lindberg et al., 2010).

Certain evidence is consistent with the greater male variability hypothesis. The variance of male students' math performance is larger compared to female students' in various datasets from the U.S. and other countries (Feingold, 1992; L. V. Hedges & Nowell, 1995; Hyde et al., 2008). A meta-analysis of 242 studies show that the overall variance ratio (VR) between male and female students' variance in math performance is 1.08, indicting a slightly greater male variability (Lindberg et al., 2010). But importantly, there is also contrary evidence that suggests smaller male variability in some national and international datasets. A cross-national study using the PISA dataset, for example, shows that in Germany, Lithuania, and the Netherlands, women have greater variability than men in math performance scores (Penner, 2008). Lindberg et al. (2010) analyze large U.S. adolescent datasets in the past 20 years, finding that male-female VR ranges from 0.88 to 1.34. These findings cast doubt on the universal applicability of the male variability hypothesis and suggest that features of national context may be linked to whether there are gender gaps in performance at the extremes (Hyde & Mertz, 2009; Lindberg et al., 2010; Penner, 2008).

Comparative perspectives on gender differences in math performance

The significant national variation suggests that patterns of gender difference in math performance may be shaped by macro-level structures. What macro structures might be tied to gender disparities in math performance? One classic line of thinking sometimes referred to as the modernization hypothesis implies that gender disparities recede with national economic development, as modern competitive pressures increase and egalitarian values become institutionalized (Baker & LeTendre, 2005; Inglehart & Norris, 2003). However, some studies find little effect of national economic development on gender gaps in math performance (Guiso et al., 2008) or even reveal larger gender gaps in math attitudes in more affluent countries (Charles, Harr, Cech, & Hendley, 2014; Sikora & Pokropek, 2012). Scholars have thus suggested that in countries with existential security and culture favoring individual self-expression, students’ instrumental concerns with lucrative careers would decrease, and pursuit of more personally expressive and gendered careers would increase, leading to larger gender gaps in math attitudes and, by possible extension, math performance (Charles et al. 2014).

Beyond theories about the role of economic development in driving gender disparities in educational performance, scholars have turned to what might be called the national gender equity context. Structural factors associated with more and less gender egalitarian societies may shape gender differences in math performance through two mechanisms: by creating incentive structures through promoting female representation in education, the labor market, and politics, and by attaching gendered values to different academic subjects and careers through gender norms and stereotypes (Penner, 2008). Using the Program for International Student Assessment (PISA) data and the Global Gender Gap Index (GGGI) developed by the World Economic Forum, Guiso et al. (2008) examined 40 countries and reported a smaller gender gap in math performance in countries with higher overall gender equity. Narrower gender gaps in math were also found in countries with higher gender equality in politics (Else-Quest et al., 2010; Guiso et al., 2008; Penner, 2008; Riegle-Crumb, 2005), school enrollment (Else-Quest et al., 2010), labor participation (Baker & Jones, 1993; Guiso et al., 2008; OECD, 2015), and research jobs (Else-Quest et al., 2010).

On the other hand, some studies have found counter-intuitive results that gender equality at the national level has no effect on or even exacerbates gender inequality in math performance. For example, using the same measurement as Guiso et al. (2008), Fryer and Levitt (2010) find no link between the Global Gender Gap Index and the gender gap in math performance. They argue that their different finding is due to the inclusion of countries in the Middle East, where, despite high levels of gender inequality, there is little or no gender gap in math performance. In addition, Riegle-Crumb (2005) found no association between gender equality in the labor force participation rate and gender gaps in math performance. Penner (2008) even found a negative association between gender equality in labor force participation and the gender gap in math performance with a sample of western developed countries. Penner suggests a possible explanation: in his sample, countries with greater female labor force participation also tend to have higher degrees of occupational gender segregation. When it comes to gendered cultural values, Penner (2008) and Riegle-Crumb both find that gender ideologies concerning the importance of home and children for women at the national level were not associated with gender gaps in math performance.

National education system characteristics

An important limitation in previous research is the scant attention to national education system characteristics that may shape gender disparities (Ayalon & Livneh, 2013). One such feature is the level of standardization, and especially the autonomy of schools on what and how they teach (Ayalon & Livneh, 2013; Park, 2008; Van de Werfhorst & Mijs, 2010). Among the few scholars who have examined the interaction between gender and standardization of education systems, Ayalon and Livneh (2013) find that standardization of curriculum helps reduce advantages of boys over girls in math performance. Tsui (2007) also argues that Chinese students achieve higher gender parity in math performance compared to their U.S. counterparts because of the rigorous and standardized national mathematics curriculum.

A second feature of education systems is stratification, which is sometimes called differentiation, branching, streaming, or tracking. Stratification usually refers to the extent to which education systems have differential curricula and tracks based on students' performance and aspirations at the secondary level (Han, 2016). Sikora and Pokropek (2012) find that higher levels of stratification are associated with a lower chance of expecting a career in computer science or engineering for girls but not for boys. Han (2016) also finds a positive association between the level of stratification and gender gaps in STEM occupational expectations. Current evidence suggests a greater gender gap in STEM aspirations and expectations in more stratified education systems, but implications of stratification for math performance itself are not established.

A third feature that might be significant is scope of privatization. Ceron (2016) found that in Latin American countries, achievement inequality by family background is greater in countries with higher levels of privatization of the education system. Consistent with this insight, Torche (2005) found that inequality increased for cohorts who received education during and after the privatization of education system in Chile. Accordingly, in Chile, the association between schools' aggregate family socioeconomic status and students' test scores is much greater for private-voucher schools than for public schools, which results in pronounced socioeconomic stratification (Mizala & Torche, 2012). The effect of privatization on gender differences remains underexplored.

We have noted limitations in the existing literature: a lack of attention to extreme performance and a dearth of attention toward national characteristics that might shape gender gaps across the distribution. A further point that might be viewed as a limitation is the geographic coverage of existing evidence: very few studies have analyzed factors shaping gender differences in math performance in Latin America in comparative perspective. We next provide a brief overview of the context of gender and education in Latin America.

Gender and education in Latin America

Latin America has achieved significant expansions of education coverage, access, and progression in most countries in recent decades, such that by 2012 the region's literacy rate reached an average of 93.3 percent, compared to 88.9 percent in 2000 (UNESCO, 2014). However, there is considerable heterogeneity in advances within the region and there are some equity issues within countries associated with class and location of residence. There are also remaining concerns regarding quality of education, as indicated by performance inequalities in the Third Regional Comparative and Explanatory Study (Tercer Estudio Regional Comparativo y Explicativo, hereafter TERCE). TERCE results show that 61 percent of third-graders and 70 percent of sixth-graders are in reading performance levels I and II (the two lowest out of four levels), while 71 percent of third-graders and 83 percent of sixth-graders are in the lowest two levels for math performance (UNESCO, 2015, pp. 7-8).

Overall, gender stratification in education in many countries in Latin America encompasses a girls' advantage in general educational attainment but a persistent boys' advantage in STEM performance. A policy paper for the UNESCO Global Education Monitoring Report indicates a significant access advantage for women in Latin America and the Caribbean: for every 100 women, 96 men completed primary, 94 completed lower secondary and 91 completed upper secondary education, while only 83 were attending some form of post-secondary education (UNESCO, 2018). A study of the impact of the 1980s economic crisis on inequality of educational opportunity in four Latin American countries for birth cohorts 1940 to 1975 finds a growing female advantage in educational attainment across cohorts, in line with trends in industrialized countries (Torche, 2010). Another study of educational stratification of adolescents growing up during the 1980s, 1990s and 2000s in Latin America shows that girls have higher probabilities of school enrollment in all years and countries studied (Marteleto, Gelber, Hubert, & Salinas, 2012). Gender parity indices for 2010 and 2013 show gender parity in access to primary education and indicate a slight advantage for women in secondary education (UNESCO, 2014).

However, evidence also suggests that in the majority of countries in the region, boys show a fairly consistent advantage in math and science. A UNESCO report on Latin America and the Caribbean concludes, "it is clear that girls in the region (with the exception of Cuba and the Dominican Republic) consistently achieve on average lower results in scientific subjects than the male students" (UNESCO, 2014, p. 98). Other studies have highlighted the unevenness in gender patterns across Latin American countries. Analyses of the 2007 TIMSS data show significant gender disparities favoring men in math scores in El Salvador and Colombia, while the 2006 PISA results show similar findings for Argentina, Brazil, Chile, Colombia, and Uruguay (Valverde & Näslund-Hadley, 2011). On the other hand, while seconding UNESCO's observation that women outperform men in reading and men outperform women in math and science, a report by the Inter-American Development Bank (IDB) contrasts evidence from different studies and finds that in some countries, this difference is not significant (Valverde & Näslund-Hadley, 2011). The 2009 Caribbean Certificate of Secondary Education (Certificado Caribeño de Educación Secundaria) entry exam even shows that women fared better than men in math and science in some English-speaking Caribbean countries.

Gender stratification patterns in education also vary across the performance distribution in Latin American countries. Abadía and Bernal (2017) analyze math, sciences, reading, and global performance by gender for Colombian 11th graders as reported by the 2014 entry exam to secondary education (SABER 11). The authors find a significant gender gap in math and science favoring boys that widens towards the top of the distribution. Another study analyzes the entry exam of the Universidade Federal de Pernambuco (UFPE), the major university in the Northeast of Brazil. Results indicate a male advantage in all three subjects and greater variation among boys than girls. Quantile regression results further indicate that the math gender gap varies across the distribution, and the male advantage is smaller at the tail (Guimaraes & Sampaio, 2008).

Lastly, most studies conclude that observable individual, family, and school characteristics only partially explain gender gaps in test scores across the region (Abadía, 2017, p. 15; Abadía & Bernal, 2017, p. 27). Unobserved factors contributing to the gap may include the broader national context of economic development, norms about women's roles in society, and educational system features. However, few studies in Latin American contexts have directly assessed the role of national contexts. Utilizing TERCE data and focusing on country-level characteristics, this study fills in the gap in the previous literature.

Hypotheses

Drawing on the comparative literature on gender and math performance and on evidence drawn from prior studies in Latin America, we first test whether there is greater male variability in students' math performance among Latin American countries. We then investigate whether national-level factors, including economic development, gender equality regime, and education system characteristics, are associated with gender differences across the distribution.

Hypothesis 1: There is a greater variability in math performance among boys than among girls.

The greater male variability hypothesis states that boys tend to have a greater variance in math performance. This means that the distribution of math performance among boys is flatter compared to that among girls, thus boys are not only more likely to be at the top, but also more likely to be at the bottom at the distribution. In addition, boys should be more advantaged at the top and more disadvantaged at the bottom of the distribution; this means that top-performing boys should have better performance compared to top-performing girls, while boys at the bottom should have worse performance compared to girls at the bottom.

Next, to evaluate national characteristics associated with modernization theories about gender gaps, we test the following hypotheses:

Hypothesis 2a: A higher level of national economic development is associated with smaller gender gaps in math performance across the distribution.

Since narrower gender gaps in math have been found in countries with higher gender equality (Guiso et al., 2008; Penner, 2008b; Riegle-Crumb, 2005), we also test the following hypothesis:

Hypothesis 2b: A higher level of national gender equality and lower level of gender segregation in occupations are associated with smaller gender gaps in math performance across the distribution.

Finally, drawing on the studies about education system characteristics, especially Ayalon and Livneh (2013) and Tsui (2007) on standardization, Sikora and Pokropek (2012) and Han (2016) on stratification, and Ceron (2016) on privatization, we test the following hypotheses:

Hypothesis 3a: A higher level of standardization of the curriculum is associated with smaller gender gaps in math performance across the distribution.

Hypothesis 3b: A higher level of stratification of the education system is associated with larger gender gaps in math performance across the distribution.

Hypothesis 3c: A higher level of privatization of the education system is associated with larger gender gaps in math performance across the distribution.

Methodology

The TERCE dataset

TERCE is a cross-national study of learning and achievement in Latin American countries administered in 2013 by the United Nations Educational Scientific and Cultural Organization (UNESCO); 15 countries participated in the study (Argentina, Brazil, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru and Uruguay). TERCE evaluated third and sixth grader performance in reading, science, writing and mathematics. In this study, we focus on gender differences in math performance among students in the sixth grade. The final combined sample includes 57,476 students in 15 countries.

TERCE presents student performance results in two different ways: the *test score* and the *performance level*. First, TERCE provides five scores called plausible values from a distribution with a regional mean of 700 and standard deviation of 100 points. Second, TERCE classifies students into four achievement levels based on their test scores. The fourth level represents the highest achievement (UNESCO, 2016). In this study, we use all five plausible values of test scores in quantile regression models and performance levels in logistic regression models as the dependent variables.

Country-level variables

Based on findings from previous studies, we include measures of three dimensions of country-level factors into our analysis: economic development, gender equality regimes, and education system characteristics.

First, to measure national level of economic development, we use *GNI per capita in 2013* (The World Bank, 2017).

Second, to measure national level of gender equality, we follow Guiso et al. (2008) and Fryer and Levitt (2010) to include four indices from the *Global Gender Gap Report 2013* (The World Economic Forum, 2013):

Index of educational attainment (EDU): derived from the female-to-male ratios of literacy rate, net primary enrollment rate, net secondary enrollment rate, and gross tertiary enrollment rate.

Index of economic participation and opportunity (ECON): derived from the female-to-male ratios of labor force participation, wage for similar work, total earned income, number of legislators, senior officials and managers, and number of professional and technical workers.

Index of health and survival (HS): derived from sex ratio at birth and female-to-male ratio of healthy life expectancy.

Index of political empowerment (PE): derived from the female-to-male ratios of current seats in parliament, positions at ministerial level, and number of years of a female head of state over the male value in the last 50 years.

As Penner (2008) points out, gender equality in economic participation might be correlated with gender segregation in the labor market, which may confound the findings. Therefore, we include gender segregation into the analysis. To measure national level of gender segregation in the labor market, we develop two Duncan Segregation Indices (Duncan & Duncan, 1955) based on data provided by the International Labor Organization (ILO): *Segregation by skills (SEGS)* is calculated from numbers of male and female employees working at different occupational skill levels (low, medium, and high). *Segregation by industry (SEGI)* is calculated from numbers of male and female employees working in different industries (agriculture; manufacturing; mining, quarrying, and electricity, gas and water supply; construction; market services; public administration or community, social and other services).

Else-Quest and Hamilton (2017) point out that composite gender equality measures may mask important factors and processes within each domain, and individual domain-specific gender-equality measures can be utilized to reveal specific mechanisms. Therefore, we also tested for the effects of selected domain-specific gender equality variables that are used to construct the composite measures. In the domain of economic participation, we include female-to-male ratio of labor participation rate, female-to-male ratio in professional and technical jobs, and an indicator of gender wage equality; in the domain of education attainment, we include female-to-male ratios in literacy rate, primary education enrollment rate, secondary education enrollment rate, and tertiary education enrollment rate; in the domain of health and survival, we include female-to-male ratio at birth and female-to-male ratio of life expectancy; in the domain of political empowerment, we include female-to-male ratio in parliament. These variables are also extracted from the *Global Gender Gap Report 2013* (The World Economic Forum, 2013). Due to data limitations, several variables are missing in certain countries: female-to-male ratio in professional and technical positions is missing in Guatemala; female-to-male ratio in primary enrollment rate is missing in Costa Rica; female-to-male ratio in secondary enrollment rate is missing in Brazil, Costa Rica, and Honduras. We use the composite measures in our main analysis to maximize the sample size of countries. We present results from models using the domain-specific indicators in Appendix B and Appendix C and discuss them in the notes.

Finally, to measure national education system characteristics, we develop the following three variables:

To measure *Standardization (STA),* we follow the definition and operationalization of Bol and Van de Werfhorst (2013) and construct a scale based on three questions from the principals' questionnaire in TERCE. These questions describe school autonomy in deciding textbooks, course contents, and which courses to offer. We perform a factor analysis on these three variables to create a standardized scale, and then aggregate this scale to the country level with student weights to indicate the level of standardization of the national education system.

To measure the level of stratification of education system, we use *Vocational Enrollment (VOP)*, or the proportion of students in vocational secondary education (number of students in vocational secondary education divided by total number of students in secondary education, regardless of age). This variable is derived from the World Development Indicators.

To measure *Privatization (PRIV),* we follow Ceron (2016) and use the weighted proportion of students in urban private schools in each country based on the TERCE data.

All national-level variables are standardized when included in the models. Table 1 shows all national-level variable values for the 15 countries included in the sample. For a correlation matrix between all country-level variables, please see Appendix A.

[TABLE 1 ABOUT HERE]

Models

Following Penner (2008), we use both logistic and quantile regression models to examine how gender differences vary across the distribution of math performance. The logistic models examine the likelihood of being at or above various performance levels. The quantile regression models examine the size of gender differences at different percentiles across the distribution. For example, results from logistic models using "being at level III or above" as the dependent variable report gender differences in the likelihood of being at level III or above; quantile regression models at the 90^{th} percentile, on the other hand, report the gender differences in test scores between the 90^{th} percentile of boys' and girls' distribution. In addition, ordinary least squares (OLS) regression models and ordered logistic models are also used for comparison. OLS models report difference in conditional means, while quantile regression models provide more information on conditional differences at specified percentiles. Ordered logistic models treat the dependent variable as ordinal and assume proportional odds across different levels, meaning that the relationship between each pair of outcome levels should be the same. Our analysis, however, shows that this is not the case.

Within each country, we apply logistic regression models in the standard form:

$$ln\left(\frac{p_i}{1-p_i}\right) = Female_i\beta + \epsilon_i,$$

where p_i is the probability of student i achieving a certain level or above; $\mathrm{Female_i}$ is a dummy variable that equals 1 if student i is a girl, 0 if student i is a boy; ϵ_i is the error term. Similarly, we apply the quantile regression models in the standard form:

$$\mathrm{Y_i} = \mathrm{X_i}\,\beta + \epsilon_i\,,$$

where $\mathrm{Y_i}$ is the math score for student i, $\mathrm{X_i}$ is a vector of independent variables, and ϵ_i is the error term. Standard errors were calculated using a Huber-White sandwich estimator adapted for quantile regression.

Next, to measure how national-level factors are associated with gender differences across the distribution of math performance, we estimate logistic and quantile regression models with gender at the individual level and include the cross-level interaction term between gender and national-level indices. The interaction terms estimate the effect of each national-level factor on gender differences at the individual level. To facilitate interpretation, each country-level factor and its interaction with gender is included in a separate model. In addition, we include country fixed effects in all models to control for unobserved heterogeneities across countries. The logistic regression models take the following form:

$$\ln\left(\frac{p_{ij}}{1-p_{ij}}\right) = Female_i\beta_1 + Female_i \times CountryVariable_j\,\beta_2 + \sum_{i=1}^{15}\sigma_j Country_j + \epsilon_{ij},$$

where p_{ij} is the probability for student i in country j to achieve a certain level or above, $CountryVariable_j$ is the value of the national-level variable at focus in country j, and $\sum_{i=1}^{15}\sigma_j Country_j$ is the country fixed effect. Similarly, the quantile regression models take the following form:

$$Y_i = Female_i\beta_1 + Female_i \times CountryVariable_j\,\beta_2 + \sum_{i=1}^{15}\sigma_j Country_j + \epsilon_{ij}$$

where the outcome Y_i is the math score for student i, and all the other terms are the same.[1] We describe our findings in the next section.

Analysis

Gender differences in mean and variability of math performance

[TABLE 2 ABOUT HERE]

Basic descriptive statistics show a gender gap favoring male students in most countries, but the size of this gender gap varies greatly across countries. Table 2 presents these descriptive statistics for each country in the TERCE dataset. Column 5 reports the mean differences between male and female students' math scores in each country. In most countries, except for Panama, there is a significant gender gap favoring male students. The size of this gender difference, however, varies across countries. Among the countries with a significant gender gap, the mean difference in column 5 ranges from 5.876 (Uruguay) to 22.795 (Peru), and the effect size in column 7 (mean difference divided by standard deviation) ranges from 0.060 (Uruguay) to 0.303 (Guatemala). This variation suggests the importance of national and social contexts when it comes to mean differences.

Results also partly contradict the male variability hypothesis. Column 6 in Table 2 reports the variance ratio (VR), or the male variance divided by the female variance (Hyde & Mertz, 2009). In most countries, male students indeed have greater variance in math performance (VR >1.0), but the size of the gender difference in variances is small and varies by country, ranging from VR=1.009 in Argentina to VR=1.092 in Colombia. Moreover, in Brazil and Peru, female students actually have greater variances than male students (VR < 1.0), which shows that the greater male variability hypothesis does not hold across the board and offers evidence of cross-country differences.

Gender representation across the distribution

[TABLE 3 ABOUT HERE]

In addition to positing that boys tend to have greater variance in math performance, the greater male variability hypothesis (Hypothesis 1) also implies that boys are more represented at both the top and bottom tails of the distribution. Results from our analysis, however, show contradicting evidence. While boys are indeed better represented at the top of the distribution, we find that girls are more represented at the bottom in many countries.

Table 3 shows results of an ordered logistic model predicting the level of math performance (column 1) and a series of logistic models predicting the probability of achieving level I (column 2), level II and above (column 3, which is a flipped version of column 2 and is presented for comparison with column 4 and 5), level III and above (column 4), and level IV (column 5). Findings show that in about half of the countries (Argentina, Brazil, Colombia, Guatemala, Honduras, Nicaragua, and Peru), girls are more represented at the bottom level; in other countries, there is no significant gender difference in the odds of being at the bottom level. In Guatemala, for example, the results show that girls are 1.368 times more likely to be at level I compared to boys (column 2); girls only have 73.1 percent of the chance of boys to be at or above level II (column 3), 54.3 percent of the chance of boys to be at or above level III (column 4), and 22.8 percent of the chance of boys to be at level IV (column 5). We can also roughly interpret odds ratio here as counts: for every boy at level I, there are 1.368 girls; for every boy at level II and above, there are 0.731 girls; for every boy at level III and above, there are 0.543 girls; for every boy at level IV, there are only 0.228 girls.

Columns 6 to 8 present the p-values from adjusted Wald tests that the coefficients at different levels are equivalent.[2] In Guatemala, all three pairs of coefficients are significantly different from each other, showing heterogeneous gender effects at different levels of math performance: boys tend to be better represented at higher levels. This pattern is also seen in other countries and regions: in Costa Rica, Dominican Republic, Ecuador, Guatemala and Nicaragua, boys are better represented at or above level III compared to level II; in Chile and Guatemala, boys are better represented at or above level IV compared to level II; in Guatemala, boys are better represented at level IV compared to at or above level III.

Gender gaps across the distribution

[FIGURE 1 ABOUT HERE]

We next present weighted quantile-quantile plots by country to compare male and female students' math score distributions within selected percentiles (Figure 1). Results show considerable variation across the performance distribution and distinct patterns across countries. The quantile-quantile plot is a plot of the quantiles of one data set against the quantiles of another dataset; here we show the plots of the quantiles of male students' math scores against the quantiles of female students' math scores. The red line is the reference line defined as y=x. If the two datasets come from a population with the same distribution, the points should fall approximately along this reference line. The greater the departure from this reference line, the greater the possibility that the two datasets have come from populations with different distributions. The points above the reference line indicate that boys' math scores are higher than girls' math scores at a certain percentile. The greater the departure from the reference line, the larger the gender difference is at the given percentile.

The plots in Figure 1 show distinct patterns. In Paraguay, Uruguay, Panama, and Chile, there seem to be no significant gender differences across the distribution. In Colombia, Ecuador, Dominican Republic, Guatemala, and Honduras, there are larger gender differences in favor of boys at the higher extreme. In Argentina and Brazil, there are larger gender differences at the lower extreme. To further test the significance and magnitude of gender differences across the distribution, we employ quantile regression models.

[TABLE 4 ABOUT HERE]

Table 4 further shows results from quantile regression models at different cutoffs of the math score. These results further confirm that the greater male variability hypothesis does not universally hold. Column 1 in table 4 reports results from an OLS regression as a reference; by comparing results from the OLS and quantile regression models, we can better examine whether there are heterogeneous gender effects across different percentiles of math performance.

Results in Table 4 show that across selected percentiles in all countries, whenever there is a significant gender gap, it is almost always in favor of boys. It is worth noticing that while the OLS coefficient in Chile shows no significant overall gender difference, the quantile regression results show that girls' scores are 13.517 points lower than boys at the 90^{th} percentile. Even in countries with consistent male advantages across different percentiles, the size of the gender gap varies at different positions of the distribution. For example, in Guatemala, boys score 14.052 points higher than girls at the 5^{th} percentile, but this advantage increases to 17.418 points at the median, 27.518 points at the 90^{th} percentile, and 32.015 points at the 95^{th} percentile; similarly, in Colombia, boys score 13.111 points higher than girls at the 5^{th} percentile, while this advantage increases to 19.813 points at the 75^{th} percentile and 35.433 points at the 95^{th} percentile.

Results in Table 4 extend the various patterns we see in Figure 1, suggesting distinct patterns across countries. For example, in Chile, there is a significant gender gap in favor of male students at a certain percentile, but no overall gender difference based on the OLS regression. In such cases, quantile regression helps to capture the nuances that the OLS regression alone would miss. In Brazil, the gender difference is larger at the lower extreme of the distribution. Conversely, in Guatemala, Ecuador, Mexico and Colombia, the gender differences are larger at the higher extreme of the distribution. These different patterns of distribution suggest that national context could be crucial in influencing gender gaps in math performance.

Estimating the effect of national-level characteristics

[TABLE 5 ABOUT HERE]

Table 5 reports the results of logistic regression models predicting the likelihood of being at or above certain levels. To model how country-level factors are associated with gender differences in math performance at different levels, we include country-level variables into the logistic regression models and interact them with the individual-level dummy variable for being female. Country fixed effects are also included to control for unobserved country-level characteristics. The coefficients of the cross-level interaction terms show how national-level variables moderate gender gaps across different levels. A significantly positive coefficient of the interaction term means a higher value of the country-level variable is associated with a smaller gender gap in math performance.

The interaction effects in Table 5 show that, consistent with hypothesis 2a, GNI per capita is generally positively associated with girls' likelihood of scoring at higher levels relative to boys. Other country-level characteristics, on the other hand, are mainly associated with gender differences in the likelihood of being at or above level II (column 2). For example, partly consistent with hypothesis 2b, in countries with higher gender equality in health and survival (HS) and political empowerment (PE), the gender difference in the probability of scoring at or above level II is smaller. Surprisingly, in countries with higher equality in economic participation and opportunities, this gender difference is larger. This is consistent with some of the previous studies (Penner 2008). Penner suggests that when countries provide female students with more opportunities of economic participation, it is plausible that these opportunities are more likely to be in female-dominated nontechnical sectors. To test this, we also included two gender segregation indices into the model. Results show that a higher level of gender segregation in skills is associated with a lower likelihood for girls to be at or above level II, which is partly consistent with Penner's assumption. On the other hand, the level of gender segregation in industries does not matter much to gender differences at any level.[3]

When it comes to education system characteristics, only the proportion of vocational secondary students (VOP) is found positively associated with girls' likelihood of scoring at or above level II, which is inconsistent with hypothesis 3b.

[TABLE 6 ABOUT HERE]

Quantile regression models present similar patterns. Table 6 reports results from quantile regression models of math scores estimated at different percentiles across the distribution. Similar to Table 5, the interaction terms show how national-level variables affect gender differences at selected percentiles. Results indicate that, partly consistent with hypothesis 2a, GNI per capita is associated with decreases in gender gaps in general (column 1) and at the lower end of the distribution (columns 3 to 6). The effects of national gender equality regimes, on the other hand, vary across the distribution. Female representation in education is associated with reduced gender gaps in math performance at the lower-middle (columns 5 and 6) and the top (columns 9 and 10) of the distribution; gender equality in economic participation is associated with reductions in the gender gap only at the 75th percentile (column 7); national gender equality in health and survival is associated with reductions in the gender gap in general but exhibits no particular effect at each percentile; national gender equality in political empowerment exhibits no significant effect. Further, gender segregation in skills and industries does not exhibit a significant effect on the gender difference in general or across the percentiles.[4]

When it comes to the education system characteristics, no significant effect emerges for the level of standardization or privatization of education systems. However, a higher proportion of vocational students is associated with a smaller gender gap at the lower to middle percentiles but not the top end of the distribution; this finding is inconsistent with hypotheses 3b. We will discuss implications in the next section.

Discussion and conclusion

Using cross-national data from Latin American countries, this paper examines whether the greater male variability hypothesis holds for math performance across different Latin American countries, and whether national-level factors are correlated with gender gaps across the performance distribution at the micro-level. Three findings are particularly important. First, the greater male variability hypothesis does not hold across the board. Second, a higher level of stratification of the education system is associated with a smaller gender gap in math performance at the lower end, which contradicts our hypotheses. Third, although many national-level factors are associated with gender differences from the lower end to the middle of the distribution, the same is not true at the top. Below, we discuss the implications of each of these findings.

Regarding the first finding, in some Latin American countries, girls are more likely to fall in the bottom of the distribution. This finding shows that analysis of mean differences alone obfuscates critical nuances in the gender gap across the distribution. Further, patterns of gender differences in mean and variance vary across Latin American countries. Although in general, there is a significant gender gap favoring male students, the size of the gender effect varies greatly across countries. Similarly, although male students in most countries tend to have a greater variability in math performance, in Brazil and Peru, female students have greater variances than male students. We further show that gender effects are not necessarily more pronounced at the extremes of the performance distribution: for example, in Costa Rica, the gender differences are actually smaller at both extremes of the distribution. Contrary to the greater male variability hypothesis, these complex findings show that there is not always greater representation of boys at both ends of the distribution.

Our findings do confirm a persistent gender gap favoring boys among the top performers in several countries. However, they also point out a problem that has not attracted as much attention: the vulnerability of girls at the bottom level. In about half of the countries, girls are more likely to fall into the bottom level of performance; even at the bottom percentiles of the distribution, boys still have a performance advantage in most countries. This finding is surprising. Low-performance is often associated with elevated risk of subsequent grade repetition and school dropout (Janosz et al., 1997; Jimerson et al., 2000), yet patterns of educational attainment across many Latin American countries are female-favorable. One possible explanation for this apparent contradiction points to expectations about girls' performance: if it is culturally assumed that girls "are not good at math", then the association between low performance in math and dropout for girls might be weaker than for boys, more so than in other regions. This speculation points to the need for further research on gender differences in the consequences of low performance.

Regarding the second main finding, our analysis shows that the association between education system characteristics and the gender gap in math performance is partly inconsistent with our hypotheses. We found no significant association between standardization or privatization and the gender gap. On the other hand, a higher proportion of vocational students is associated with a smaller gender gap at the lower performance range. One explanation of the stratification effect is that girls may be more likely to choose academic schools rather than vocational tracks than boys; as a quasi-experimental study using Finnish school data shows, in a comprehensive system where students are tracked into vocational and academic schools at age 15-16, girls are more likely to choose the academic track than boys (Pekkarinen, 2008). Considering that the TERCE data in our study focus on sixth-grade students in primary education, while the stratification measure is derived from secondary education data in each country, it is possible that girls in countries with a high proportion of vocational secondary students feel more motivated to achieve higher scores in order to successfully enter the academic track. Further research on stratification of education systems is needed to understand the actual mechanisms that affect gender gaps in math. More broadly, given reports that educational segregation has increased in Latin America during the last two decades, with lower-income students concentrating in often-under-resourced public schools (Arcidiacono et al., 2014), attention to measures of system stratification may be particularly important to monitor in the future.

Finally, results show that country-level factors are more consistently linked to gender gaps at the low- and middle- parts of the performance distribution, and less so among the top performers. For example, we find that higher GNI per capita is associated with smaller gender gaps in math performance at the lower to middle percentiles of the distribution. This finding is partly consistent with modernization theory, which predicts smaller gender differences in more developed countries. However, higher GNI per capita is not associated with a smaller gender gap at the top of the distribution. Similarly, findings also indicate that a higher level of stratification of the education system is associated with a smaller gender gap at the lower to middle percentiles, but not the top end of the distribution.

In fact, among country-level factors, only national context of female representation in education is associated with reductions of the gender gap in performance at the top of the distribution. This finding suggests that a gender-egalitarian national education context may be important for creating an incentive structure for top-performing girls. If we consider that a national context of gender equality might be more directly linked to cultural phenomena than measures of economic development, the fact that it is the only statistically significant national-level factor in our models might be an indicator of how relevant the cultural dimension of gender inequality in education is in Latin America. This interpretation is consistent with previous findings that only about half of the gender performance gap can be explained by observable individual, family, and school characteristics in Latin American countries, and the larger social-cultural context may play a role in affecting gender differences (Abadía, 2017; Abadía & Bernal, 2017). Considering that students from the upper tail of the math performance distribution are more likely to enter STEM fields (Fan et al., 1997), more work is needed to illuminate sources of the gender gap among top performers, as one part of the process that generates female underrepresentation in STEM careers.

There are two caveats to these analyses. First, this study focuses on selected macro-level measures and does not fully explore the complexity in national-level contexts. Although we use country fixed effects to control for country-level characteristics, additional national-level forces such as migration and urbanization and within-country variation across socioeconomic and socio-cultural groups may also be important for understanding patterns of gender difference.

Second, with cross-sectional data, we are only able to investigate associations and cannot make causal claims. Future studies should consider using longitudinal datasets with lagged outcome variables to further identify causal influences.

Despite these limitations, this study provides important insights into the problem of gender disparity in STEM education in Latin America. It also provides implications for future policies and initiatives. First, while it is important to study gender differences at the mean and the top, it is equally crucial to identify and provide assistance to disadvantaged girls at the bottom of the performance distribution. Second, while national gender equality in various domains may reflect structural opportunities for women, the measure most closely tied to gender parity among the highest performers is female representation in education. The policy implications of this relationship are complicated in a region where girls' secondary and tertiary enrollment outstrip boys' in many countries. Last, a greater share of students in vocational education does not necessarily result in a larger gender gap in STEM performance. Conversely, boys may be more vulnerable than girls under such circumstances. It is thus important to identify potential consequences of the characteristics and changes in the education system and provide targeted assistance to disadvantaged groups.

Notes

[1.] Using dummy variables for fixed effects in non-linear models with maximum likelihood estimation can produce bias when the number of clusters is large and number of cases within each cluster is small; however, when the number of clusters is small and the number of cases within each cluster is large, the bias will be minimized. Please refer to Allison (2009) for details.

[2.] The test is done using a general ordered logistic (gologit) model, which helps to relax the proportional odds assumption of the conventional ordered logistic model. After running the general ordered logistic model, we test whether coefficients across different levels are equivalent. For details, please refer to Williams (2016).

[3.] We further tested the effects of selected domain-specific gender equality measures using the same set of logistic regression models. Results are presented in Appendix B and are largely consistent with results in Table 5. The effect of female representation in education at the top level is reflected in the effects of female-to-male ratio in literacy rate and female-to-male ratio in secondary education enrollment rate; although the effects of these two specific indicators are only marginally significant, the magnitudes of effects are similar to that of the composite measure. Further, the gendering effect of gender equality in economic participation is mainly reflected in the effect of female-to-male ratio in labor participation rates. The effect of gender equality in health and survival is reflected in the effect of female-to-male sex ratio at birth. The effect of gender equality in political empowerment is reflected in the effect of female-to-male ratio in the parliament.

[4.] Using the same set of quantile regression models, we further tested the effects of selected domain-specific gender equality measures in education attainment, economic participation, and health and survival. Results are presented in Appendix C and are largely consistent with results in Table 6. The effect of female representation in education is reflected in the effects of female-to-male ratio in literacy rate and female-to-male ratio in secondary education enrollment rate. The effect of gender equality in economic participation is reflected in the effects of female-to-male ratio in labor participation rate and female-to-male ratio in professional and technical positions. The effect of gender equality in health and survival is reflected in the effect of female-to-male sex ratio at birth.

References

Abadía, L. K. (2017). Gender score gaps of Colombian students in the PISA test. *Vniversitas Economica*, *17*(8).

Abadía, L. K., & Bernal, G. (2017). A widening gap? A gender-based analysis of performance on the Colombian High School Exit Examination. *Revista de Economia Del Rosario, 20*(1), 5-31. https://doi.org/dx.doi.org/10.12804/revistas.urosario.edu.co/economia/a.6144

Allison, P. D. (2009). *Fixed Effects Regression Models*. Thousand Oaks, CA: SAGE Publications.

Arcidiacono, M., Cruces, G., Gasparini, L., Jaume, D., Serio, M., & Vazquez, E. (2014). *La Segregacion Escolar Publico-privada en America Latina* [Public-private School Segregation in Latin America] (No. 195). Santiago, Chile: Publicación de las Naciones Unidas.

Ayalon, H., & Livneh, I. (2013). Educational standardization and gender differences in mathematics achievement: A comparative study. *Social Science Research*, *42*(2), 432-445. https://doi.org/10.1016/j.ssresearch.2012.10.001

Baker, D. P., & LeTendre, G. K. (2005). *National Differences, Global Similarities: World Culture and the Future of Schooling*. Palo Alto, CA: Stanford University Press.

Baker, D. P., & Jones, D. P. (1993). Creating gender equality: Cross-national gender stratification and mathematical performance. *Sociology of Education*, *66*(2), 91-103. https://doi.org/10.2307/2112795

Baye, A., & Monseur, C. (2016). Gender differences in variability and extreme scores in an international context. *Large-Scale Assessments in Education*, *4*(1). https://doi.org/10.1186/s40536-015-0015-x

Bol, T., & Van de Werfhorst, H. G. (2013). *The measurement of tracking, vocational orientation, and standardization of educational systems: A comparative approach* (GINI Discussion Paper No. 81; p. 63). Retrieved from Amsterdam Institute for Advanced Labour Studies (AIAS) website: http://archive.uva-aias.net/uploaded_files/publications/81-3-3-1.pdf

Charles, M., Harr, B., Cech, E., & Hendley, A. (2014). Who likes math where? Gender differences in eighth-graders' attitudes around the world. *International Studies in Sociology of Education*, *24*(1), 85-112. https://doi.org/10.1080/09620214.2014.895140

Duncan, O. D., & Duncan, B. (1955). A methodological analysis of segregation indexes. *American Sociological Review*, *20*(2), 210-217. https://doi.org/10.2307/2088328

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, *136*(1), 103-127. https://doi.org/10.1037/a0018053

Fahle, E. M., & Reardon, S. F. (2018). Education. In Stanford Center on Poverty and Inequality (Series Ed.), *Special Issue, Pathways Magazine*. *State of the Union: The Poverty and Inequality Report*. Palo Alto, CA: Stanford.

Fan, X., Chen, M., Matsumoto, A. R., & Fan, X. (1997). Gender differences in mathematics achievement: Findings from the "National Education Longitudinal Study of 1988." *The Journal of Experimental Education*, *65*(3), 229-242.

Feingold, A. (1992). Sex differences in variability in intellectual abilities: A new look at an old controversy. *Review of Educational Research*, *62*(1), 61-84. https://doi.org/10.2307/1170716

Fryer, R. G., & Levitt, S. D. (2010). An empirical analysis of the gender gap in Mathematics. *American Economic Journal: Applied Economics*, *2*(2), 210-240.

Fu, V. K. (1999). Estimating generalized ordered logit models. *Stata Technical Bulletin*, *8*(44), 27-30.

Guimaraes, J., & Sampaio, B. (2008). Mind the gap: Evidence from gender differences in scores in Brazil. *Anais Do XXXVI Encontro Nacional de Economia*. Presented at the Associação Nacional dos Centros de Pósgraduação em Economia.

Guiso, L., Monte, F., Sapienza, P., & others. (2008). Culture, gender, and math. *Science*, *320*(5880), 1164.

Han, S. W. (2016). National education systems and gender gaps in STEM occupational expectations. *International Journal of Educational Development*, *49*, 175-187. https://doi.org/10.1016/j.ijedudev.2016.03.004

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, *269*(5220), 41-45.

Hedges, Larry V., & Friedman, L. (1993). Gender differences in variability in intellectual abilities: A reanalysis of Feingold’s results. *Review of Educational Research*, *63*(1), 94-105. https://doi.org/10.2307/1170561

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, *321*(5888), 494-495. https://doi.org/10.1126/science.1160364

Hyde, J. S., & Mertz, J. E. (2009). Gender, culture, and mathematics performance. *Proceedings of the National Academy of Sciences*, *106*(22), 8801-8807. https://doi.org/10.1073/pnas.0901265106

Inglehart, R., & Norris, P. (2003). *Rising Tide: Gender Equality and Cultural Change Around the World*. Cambridge, UK: Cambridge University Press.

Janosz, M., LeBlanc, M., Boulerice, B., & Tremblay, R. E. (1997). Disentangling the weight of school dropout predictors: A test on two longitudinal samples. *Journal of Youth and Adolescence*, *26*(6), 733-762. https://doi.org/10.1023/A:1022300826371

Jimerson, S., Egeland, B., Sroufe, L. A., & Carlson, B. (2000). A prospective longitudinal study of high school dropouts examining multiple predictors across development. *Journal of School Psychology*, *38*(6), 525-549. https://doi.org/10.1016/S0022-4405(00)00051-0

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin*, *136*(6), 1123-1135. https://doi.org/10.1037/a0021276

Marteleto, L., Gelber, D., Hubert, C., & Salinas, V. (2012). Educational inequalities among Latin American adolescents: Continuities and changes over the 1980s, 1990s and 2000s. *Research in Social Stratification and Mobility*, *30*(3), 352-375.

Mizala, A., & Torche, F. (2012). Bringing the schools back in: the stratification of educational achievement in the Chilean voucher system. *International Journal of Educational Development*, *32*(1), 132-144.

Morales, R., & Sifontes, D. (2014). Desigualdad de genero en ciencia y tecnologia: un estudio para America Latina [Gender Inequality in Science and Technology: a study for Latin America]. *Observatorio Laboral Revista Venezolana*, *7*(13), 95-110.

OECD. (2015). *The ABC of Gender Equality in Education: Aptitude, Behavior, Confidence*. PISA, OECD Publishing. https://doi.org/10.1787/9789264229945-en

Park, H. (2008). The varied educational effects of parent-child communication: A comparative study of fourteen countries. *Comparative Education Review*, *52*(2), 219-243. https://doi.org/10.1086/528763

Pekkarinen, T. (2008). Gender differences in educational attainment: Evidence on the role of tracking from a Finnish quasi-experiment. *The Scandinavian Journal of Economics*, *110*(4), 807-825. https://doi.org/10.1111/j.1467-9442.2008.00562.x

Penner, A. M. (2008). Gender differences in extreme mathematical achievement: An international perspective on biological and social factors. *American Journal of Sociology, 114 Suppl*, S138-170.

Reardon, S. F., Fahle, E. M., Kalogrides, D., Podolsky, A., & Zárate, R. C. (2019). Gender achievement gaps in U.S. school districts. *American Educational Research Journal*, *forthcoming*, 1–35. https://doi.org/10.3102/0002831219843824

Riegle-Crumb, C. (2005). The cross-national context of the gender gap in math and science. In Larry V. Hedges & B. Schneider (Eds.), *The Social Organization of Schooling* (pp. 227-243). New York, NY: Russell Sage Foundation.

Sikora, J., & Pokropek, A. (2012). Gender segregation of adolescent science career plans in 50 countries. *Science Education*, *96*(2), 234-264. https://doi.org/10.1002/sce.20479

The World Bank. (2017). The World Bank Databank. Retrieved March 19, 2017, from http://databank.worldbank.org/data/home.aspx

The World Economic Forum. (2013). *The Global Gender Gap Report 2013*. Geneva, Switzerland: The World Economic Forum.

Torche, F. (2005). Privatization reform and inequality of educational opportunity: The case of Chile. *Sociology of Education*, *78*(4), 316-343. https://doi.org/10.1177/003804070507800403

Torche, F. (2010). Economic crisis and inequality of educational opportunity in Latin America. *Sociology of Education*, *83*(2), 85-110.

Tsui, M. (2007). Gender and mathematics achievement in China and the United States. *Gender Issues*, *24*(3), 1-11. https://doi.org/10.1007/s12147-007-9044-2

UNESCO. (2014). *Latin America and the Caribbean Education for All 2015 Regional Review*. Retrieved from United Nations Educational, Scientific and Cultural Organization website: https://unesdoc.unesco.org/ark:/48223/pf0000232701

UNESCO. (2015). *Executive Summary TERCE Third Regional Comparative and Explanatory Study - Learning Achievements*. Retrieved from United Nations Educational, Scientific and Cultural Organization website: https://unesdoc.unesco.org/ark:/48223/pf0000243983_eng

UNESCO. (2016). *Gender Inequality in Learning Achievement in Primary Education: What Can TERCE Tell Us*? Retrieved from United Nations Educational, Scientific and Cultural Organization website: https://unesdoc.unesco.org/ark:/48223/pf0000244349

UNESCO. (2018). *Achieving Gender Equality in Education: Don't Forget the Boys*. Retrieved from United Nations Educational, Scientific and Cultural Organization website: https://en.unesco.org/gem-report/node/2426

Valverde, G., & Näslund-Hadley, E. (2011). *La Condición de la Educación en Matemáticas y Ciencias Naturales en América Latina y el Caribe* [The Condition of Education in Mathematics and Natural Sciences in Latin America and the Caribbean]. Retrieved from Division de Educacion, Banco Interamericano de Desarrollo website: https://publications.iadb.org/handle/11319/2757

Van de Werfhorst, H. G., & Mijs, J. J. B. (2010). Achievement inequality and the institutional structure of educational systems: A comparative perspective. *Annual Review of Sociology*, *36*(1), 407-428. https://doi.org/10.1146/annurev.soc.012809.102538

Williams, R. (2005). "GOLOGIT2: Stata module to estimate generalized logistic regression models for ordinal dependent variables," *Statistical Software Components S453401*, Boston College Department of Economics, revised 20 May 2019.

Xie, Y., & Shauman, K. A. (2003). *Women in Science: Career Processes and Outcomes*. Cambridge, MA: Harvard University Press.