DIVERSITY

Culture, Gender, and Math

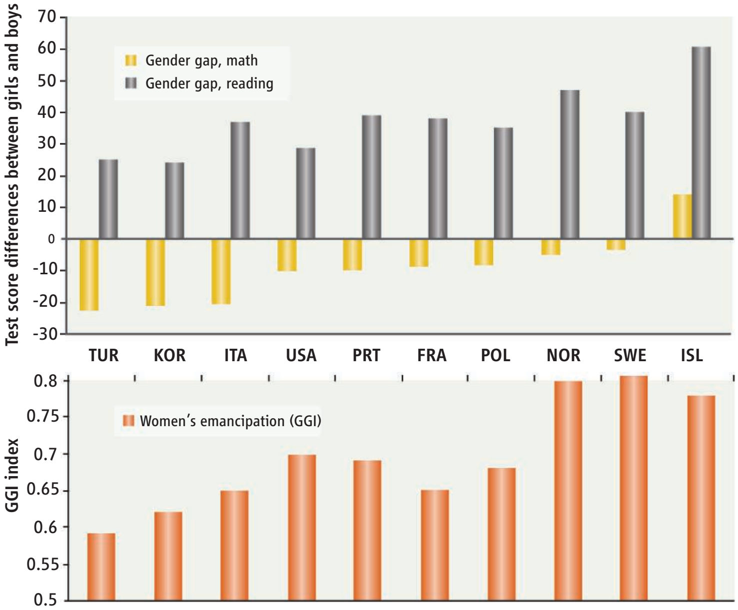
Analysis of PISA results suggests that the gender gap in math scores disappears in countries with a more gender-equal culture.

Luigi Guiso,[1]* Ferdinando Monte,[2]* Paola Sapienza,[3]*† Luigi Zingales[4]*

The existence (*1*), degree (*2*), and origin (*3*, *4*) of a gender gap (difference between girls' and boys' scores) in mathematics are highly debated. Biologically based explanations for the gap rely on evidence that men perform better in spatial tests, whereas women do better in verbal recall ones (*1*, *5*, *6*). However, the performance differences are small, and their link with math test performance is tenuous (*7*). By contrast, social conditioning and gender-biased environments can have very large effects on test performance (*8*).

To assess the relative importance of biological and cultural explanations, we studied gender differences in test performance across countries (*9*). Cultural inequalities range widely across countries (*10*), whereas results from cognitive tests do not (*6*). We used data from the 2003 Programme for International Student Assessment (PISA) that reports on 276,165 15-year-old students from 40 countries who took identical tests in mathematics and reading (*11*, *12*). The tests were designed by the Organisation for Economic Co-operation and Development (OECD) to be free of cultural biases. They are sufficiently challenging that only 0.6% of the U.S. students tested perform at the 99th percentile of the world distribution.

Girls' math scores average 10.5 lower than those of boys (2% less than the mean average score for boys), but the results vary by country (see chart, above): in Turkey, –22.6, whereas, in Iceland, 14.5. A similar variation exists in the proportion of girls over boys who score above 95%, or 99% of the country-level distribution (fig. S2A).



Math and reading gender gaps. In more gender-equal cultures, the math gender gap disappears and the reading gender gap becomes larger. **(Top)** Gender gaps in mathematics (yellow) and reading (gray) are calculated as the difference between the average girls' score and the average boys' score. A subset of countries is shown here (see SOM for complete data set and calculations). In many countries, on average, girls perform more poorly than boys in mathematics. In all countries, girls perform better than boys in reading. The gender gap in mathematics and reading correlates with country measures of gender status within the culture, one of which measures is the GGI **(bottom)**. Larger values of GGI point to a better average position of women in society. Besides USA, the countries are abbreviated as their first three letters, except for PRT, Portugal, and ISL, Iceland.

The gender gap is reversed in reading. On average, girls have reading scores that are 32.7 higher than those of boys (6.6% higher than the mean average score for boys), in Turkey, 25.1 higher and in Iceland, 61.0 higher (see chart). The effect is even stronger in the right tail of the distribution. In spite of the difference in levels, the gender gap in reading exhibits a variation across countries similar to the gender gap in math. Where girls enjoy the strongest advantage in reading with respect to boys, they exhibit the smallest disadvantage (sometime even an advantage) in math. [The correlation between the average gender gaps in mathematics and reading across countries is 0.59 (fig. S4)].

To explore the cultural inputs to these results, we classified countries according to several measures of gender equality. (i) The World Economic Forum's Gender Gap Index (GGI) (*10*) reflects economic and political opportunities, education, and well-being for women (see chart). (ii) From the World Values Surveys (WVSs) (*13*), we constructed an index of cultural attitudes toward women based on the average level of disagreement to such statements as: "When jobs are scarce, men should have more right to a job than women." (iii) The rate of female economic activity reflects the percentage of women age 15 and older who supply, or are available to supply, labor for the production of goods and services. (iv) The political empowerment index computed by the World Economic Forum (*8*) measures women's political participation, which is less dependent on math skills than labor force participation. These four measures are highly correlated (table S2).

We find a positive correlation between gender equality and gender gap in mathematics (fig. S5). If Turkey, a low gender-equality country (GGI = 0.59), were characterized by the degree of gender equality manifested in Sweden (GGI = 0.81), our statistical model suggests that the mean score performance in mathematics of girls relative to boys would increase by 23 points, which would eliminate the Turkish gender gap in math (see table, p. 1165). In more gender-equal countries, such as Norway and Sweden, the math gender gap disappears. Similar results are obtained when we use the other indicators of women's roles in society. These results are true not only at the mean level, but also in the tail of the distribution (table S3). In Iceland, the ratio of girls to boys who score above the 99th percentile of the country distribution in math scores is 1.17.

There are many unobserved reasons why countries may differ in a way that affects the

[1]European University Institute, Villa San Paolo, Via della Piazzuola 43, 50133 Florence, Italy. [2]Economics Department, University of Chicago,1126 East 59th Street, Chicago, IL 60637, USA. [3]Kellogg School of Management, Northwestern University, Evanston, IL 60208, USA. [4]Graduate School of Business, University of Chicago, 5807 South Woodlawn Avenue, Chicago, IL 60637, USA.

*These authors contributed equally to this work. †To whom correspondence should be addressed. E-mail: Paola-Sapienza@northwestern.edu

Published by AAAS

Differences in Test Scores Correlated with Indicators of Gender Equality ♀♂

	LHS: Gender difference in math				LHS: Gender difference in reading			
Women's emancipation (GGI)	105.49± 26.92**				83.56± 30.43**			
Avg. WVS indicators		13.21± 7.06				16.39± 8.46		
Female economic activity rate			0.45± 0.14**				0.34± 0.15*	
Women's political empowerment				29.10± 10.05**				24.35± 10.86*
Log GDP per capita, 2003	-6.56± 2.40**	1.09± 2.26	-3.12± 1.93	-4.95± 2.52	-2.23± 2.71	0.52± 2.71	-0.56± 2.15	-1.06± 2.73
Constant	-19.62± 20.01	-57.16± 23.27*	-2.75± 17.72	32.43± 23.72	-3.02± 22.62	-16.09± 27.90	21.49± 19.80	39.03± 25.63
Observations (no.)	37	32	39	36	37	32	39	36
R^2	0.32	0.15	0.23	0.21	0.20	0.14	0.12	0.15

Culture affects the gap. More gender-equal cultures are associated with reducing the negative gap in math and further enlarging the positive gap in reading in favor of women. Test scores are positively correlated with indicators of gender equality in society (GGI, WVSs, see text). Economic conditions are accounted for by per capita Gross Domestic Product (GDP). The correlation persists among high achievers on both tests (table S3). See SOM for details of statistical analysis. The constant is where the regression line intercepts the *y* axis, representing the amount the dependent *y* (gender gap) will be when all the independent variables are set to 0. LHS, left-hand side variable in the least-squares regression analysis. $^*P < 0.05$; $^{**}P < 0.01$.

math gender gap. Without appropriate controls, we run the risk of capturing a spurious correlation between the unobserved factors and our measures of gender equality. We reran our regression at the student level, inserting a dummy variable for each country, to control for unobserved heterogeneity (table S4). The interaction between gender and GGI index remains statistically significant at the 1% confidence level in a two-tailed *t* test, which suggests that the correlation between gender equality and girls' math scores is not driven by unobserved heterogeneity. This interaction between gender gap and GGI remains significant even when we insert an interaction between gender and log of GDP per capita, which suggests that the improvement in math scores is not just related to economic development, but to the improvement of the role of women in society.

To investigate whether the disappearance of the math gender gap in some countries translates into an overall improvement of girls or is simply limited to mathematics scores, we correlated reading performance differences with measures of women's equality (see table, above). In countries where women are more emancipated, girls' comparative advantage in reading widens. Comparing Turkey (GGI = 0.59) and Sweden (GGI = 0.81), we see an increase in the mean score performance of girls relative to boys in reading by 18 points, which almost doubles Turkey's reading gap in favor of girls.

To verify that these results are not driven by biological differences across countries, we analyzed whether they persist in populations that have a similar or identical evolutionary history. To assess history, we used a genetic distance measure (*14–17*) based on the frequency of each allele across DNA polymorphisms.

According to this measure, there are 13 European countries with genetic distance equal to zero and 26 European countries with genetic distance less than 100 (table S5). When we restrict the regression of the table (above) to either one of these two groups, our findings are substantially unchanged (table S6).

These results suggest that the gender gap in math, although it historically favors boys, disappears in more gender-equal societies. The same cannot be said for how boys score in mathematics compared with how boys score in readings. Boys' scores are always higher in mathematics than in reading, and although the difference between boys' math and boys' reading scores varies across countries, it is not correlated with the GGI index or with any of the other three measures of gender equality (table S7A). Hence, in countries with a higher GGI index, girls close the gender gap by becoming better in both math and reading, not by closing the math gap alone. The gender gap in reading, which favors girls and is apparent in all countries, thus expands in more gender-equal societies. Similarly, although the gender gaps in all math subfields decrease in societies with more gender equality, the difference between the gender gap in geometry (where the boys' advantage relative to the girls' is the biggest) and arithmetic (where the boys' advantage relative to the girls' is the smallest) does not (table S7B).

This evidence suggests that intra-gender performance differences in reading versus mathematics and in arithmetic versus geometry are not eliminated in a more gender-equal culture. By contrast, girls' underperformance in math relative to boys is eliminated in more gender-equal cultures. In more gender-equal societies, girls perform as well as boys in mathematics and much better than them in reading. These findings shed some light on recent trends in girls' educational achievements in the United States, where the math gender gap has been closing over time (*2*).

References and Notes

1. L. V. Hedges, A. Nowell, *Science* **269**, 41 (1995).
2. C. Goldin, L. F. Katz, I. Kuziemko, *J. Econ. Perspect.* **20**, 133 (2006).
3. C. P. Benbow, J. C. Stanley, *Science* **210**, 1262 (1980).
4. "The science of gender and science: Pinker vs. Spelke, a debate," *Edge*, 10 May 2005, no. 160; www.edge.org/documents/archive/edge160.html#d.
5. S. Baron-Cohen, *The Essential Difference: Men, Women, and the Extreme Male Brain* (Allen Lane, London, 2003).
6. D. Kimura, *Sex and Cognition* (MIT Press, Cambridge MA, 1999).
7. E. S. Spelke, *Am. Psychol.* **60**, 950 (2005).
8. D. Halpem, J. Wai, A. Saw, in *Gender Differences in Mathematics*, A. M. Gallagher and J. C. Kaufman, Eds. (Cambridge Univ. Press, New York, 2005), pp. 48–72.
9. Materials and methods are available as supporting material on *Science* Online.
10. R. Hausmann, L. D. Tyson, S. Zahidi, *The Global Gender Gap Report* (World Economic Forum, Geneva, Switzerland, 2006).
11. OECD, Programme for International Student Assessment (PISA), 2nd Assessment (OECD, Paris, 2003).
12. PISA includes originally 41 countries; we drop Liechtenstein because it contains only 165 observations, which makes problematic any calculation of the tail of the distribution. All other countries have at least 639 observations.
13. R. Inglehart *et al.*, "World Values Surveys and European Values Surveys, 1981–1984, 1990–1993, and 1995–1997" [Computer files; Interuniversity Consortium for Political and Social Research (ICPSR) version] (Institute for Social Research, Ann Arbor, MI, 2000), distributed by ICPSR.
14. This measure was originally computed at the population level by (*15*).
15. L. L. Cavalli-Sforza, P. Menozzi, A. Piazza, *The History and Geography of Human Genes* (Princeton Univ. Press, Princeton, NJ, 1996).
16. This measure has been mapped on modern countries by (*17*).
17. E. Spolaore, R. Wacziarg, "The diffusion of development," Centre for Economic Policy Research Discussion Paper 5630 (CEPR, London, 2006).

18. We thank S. Baliga; C Hoxby; seminar participants at Northwestern University, University of Chicago, and the National Bureau of Economic Research (NBER); and two anonymous referees for their helpful comments. The Initiative on Global Markets at the University of Chicago provided financial support.

10.1126/science.1154094

Supporting Online Material
www.sciencemag.org/cgi/content/full/320/5880/1164/DC1

www.sciencemag.org/cgi/content/full/320/5880/1164/DC1

Supporting Online Material for

Culture, Gender, and Math

Luigi Guiso, Ferdinando Monte, Paola Sapienza,* Luigi Zingales

*To whom correspondence should be addressed. E-mail: Paola-Sapienza@northwestern.edu

Published 30 May 2008, *Science* **320**, 1164 (2008)
DOI: 10.1126/science.1154094

This PDF file includes

Supporting Online Material

Overview

This Supporting Online Material provides details on the data used, the methods followed for their analysis, and robustness of the analysis for the results derived in the paper.

Materials and Methods

PISA data

The Programme for International Student Assessment (PISA) is an every-three-year international survey of 15-year-old students aimed at determining their knowledge and skills in different domains. We analyze data from the second cycle, which took place in 2003. In this survey, students' abilities were assessed in the three curricular domains: mathematics, reading, and science plus the cross-curricular domain of problem-solving. Students were also asked to provide individual information on their social and economic background. We choose to analyze the PISA second cycle since its main focus was on the domain of mathematics. In particular, 54% of the total student testing time was devoted to the assessment of mathematical skills (*1*).

The PISA target population is made up of all students in any educational institution between the ages of 15 years and 3 months and 16 years and 2 months at the time of the assessment. This specific age has been chosen because it is close to the end of compulsory education in most countries (*1*). Efforts have been made to insure the absence of cultural or national biases in the test items [see for example (*2*)] and in the evaluation of performance (*3*).

The student data set contains 276,165 observations, which roughly represent a population of 19,155,864 15-year-old students attending seventh grade or above in 41 countries, 30 of which belong to the Organisation for Economic Co-operation and Development (OECD). From this data set, we dropped Liechtenstein because the data contain only 165 observations, which make problematic any calculation of the tail of the distribution (all the other countries have at least 639 observations); moreover, Liechtenstein has no measure of gender equality.

To deal with possible differential drop-out rates between genders in different countries, we removed from the sample the students in the lower half of the country social-economic status (where dropping out is more likely).

The PISA data set contains one indicator of social status, called Economic, Social and Cultural Status (ESCS).This captures parental education, parental occupation, and home possessions, as reported by the student. This index has been normalized to have mean zero and variance one in the OECD student population. In each country, we computed the 50th percentile of ESCS (taking into account the students' final weights) and dropped all the observations below that threshold. We are then left with 138,305 observations representing more than 9,400,000 students. Nevertheless, our results are robust when we use the entire sample of students.

To verify whether PISA is a challenging test for students, we analyzed the U.S. students' results in details and compared them with other studies based on challenging tests in the U.S.A. In mathematics, U.S. students ranked 26th out of 40 countries, and this ranking was not only due to a large proportion of poor performers. The United States had a below-average number of top performers: Only 0.6% of the U.S. students tested perform at the 99th percentile of the world distribution. This result shows that PISA is not a minimal competency test.

Also, the gender differences for top performers confirm previous analyses done with U.S. data based on challenging tests. The ratio of U.S. girls to boys who are above the 75th percentile of the distribution in math is 0.85, drops to 0.59 for the 95th percentile, and to 0.30 for the 99th percentile. These ratios are not very different from the figures reported by Benbow and Stanley (*9*) who analyzed results from SAT-M given to selected samples of gifted seventh grade students in 1980, 1981, and 1982.

Our analysis is conducted using the students' final weight, so that when it is performed with student-level data, as in Table S4, each country will contribute in proportion to its share of the student population in the whole data set.

Questions in the mathematics test cover the subareas of "space and shape" (roughly geometry), "change and relationship" (algebra), "quantity" (arithmetic), and "uncertainty" (probability), in a range of difficulty from those that require simple mathematical operations to those that require complex thinking. The mathematics scores have been scaled in PISA to have a mean of 500 and a standard deviation of 100 in the OECD student population.

The reading scores have been scaled in PISA 2000, the first cycle, to have a mean of 500 and a standard deviation of 100 at OECD level. In the 2003 data, the average and the standard deviation reached 494 and 100, respectively, largely because of the arrival of new countries.

PISA asks students to report the number of hours of math classes they take. We construct the variable instructional time in mathematics dividing by 60 the number of minutes of weekly mathematical instruction. Homework in mathematics is already expressed in hours per week in the data set. Sample statistics of instructional and homework time by country, divided between boys and girls are reported in Table S1.

Measures of women's emancipation

In our discussion, we use variables on the position of women in society from four sources. The first is the "gender gap index," which we will refer to as women's emancipation (GGI), prepared by the World Economic Forum (we take the year 2006 series) (*7*). This indicator synthesizes the position of women in any given country by taking into accounts economic opportunities, economic participation, educational attainment, political achievements, and health and well-being. Larger values point to a better position of women in society.

The second series of data comes from our elaborations on the World Values Survey (WVS), all waves (*8*). The WVS asks, among other things, a series of questions on the role of women in society. The questions used are "When jobs are scarce, men should have more right to a job than women" (V61), "A working mother can establish just as warm and secure a relationship with her children as a mother who does not work" (V98), "Being a housewife is just as fulfilling as working for pay" (V99), "Both the husband and wife should contribute to household income" (V100), "On the whole, men make better political leaders than women do" (V101), "If a woman earns more money than her husband, it's almost certain to cause problems" (V102) and "A university education is more important for a boy than for a girl" (V103). For all but the first, levels of agreements had to be expressed on a scale from 1 to 4. In the first question the answers were "agree," "neither," and "disagree," to which we attributed the respective scores of 1.5, 2.5, and 3.5. ("Don't know" and missing answers are excluded.) We then inverted the answers to questions V98 and V100, so that higher values indicate a better position of women in society. The final index is the average by country of the answers to all these questions.

The "Female Economic Activity Rate (% ages 15 and older)" is the share of the female population of age 15 and older who supply, or are available to supply, labor for the production of goods and services (*10*).

The "Political Empowerment Index" is computed by the World Economic Forum and it is based on three components: (1) the ratio of females with seats in parliament over male value (International Parliamentary Union); (2) the ratio of females at ministerial level over male value (U.N. Human Development Report); (3) the ratio of number of years of a female head of state (last 50 years) over male value (World Economic Forum calculations) (*7*).

The raw data and sample statistics for these measures are presented in Table S2.

Other variables

The Gross Domestic Product per capita (GDP) is the 2003 real GDP per capita in real terms deflated with Laspeyres price index, taken from the Penn World Table (*11*).

Cavalli-Sforza *et al*. (*12*) computed genetic distance among different indigenous populations. The data we use is genetic distance among modern countries as mapped by Spolaore *et al*. (*13*).

Methods

The PISA program is designed to give estimates of population level parameters, rather than of individual level abilities. This makes the data set particularly apt to our purposes, but on the other hand, requires that specific procedures are followed for its analysis. OECD has produced a large body of documentation in this respect, and all the analysis have been carried out following the specified recommendations. Here we report the main issues we dealt with.

PISA assigns a probability distribution to each possible response pattern in each test, to describe the ability associated with that pattern. From this distribution, PISA draws a set of five values associated with each student. These values are called plausible values because they represent alternative estimates of the student ability that could have been obtained; we use plausible values in any analysis that involves test scores. In particular, any estimation procedure involves the calculation of the required statistic five times, one for each set of plausible values. The final estimate is the arithmetic average of the five estimates obtained. For example, the differences in math scores between girls and boys are calculated by running a regression of the math test scores on a constant and a gender dummy variable, defined as a variable that takes the value of one if the student is a girl and the value of zero if the student is a boy. In this case, the estimation has to be repeated five times, and the final gender difference is the average of the five coefficients on the gender dummy in the five regressions. Whenever we present an estimate involving test scores (e.g., gender difference in test or percentiles in test scores), we follow the procedure described above.

Standard errors are calculated with a replication method that takes into account the stratified, two-stage sample design for selection of schools and students within schools. In particular, OECD recommends the Balanced Repeated Replication (BRR) method: a set of 80 alternative weights are assigned to each student to form alternative samples at country level. When we do not use plausible values in the estimation, the standard error on any statistic is calculated as the square root of the average squared deviation of the estimates obtained from these alternative weights and the statistic obtained using the original students' weights. In the standard BRR method, schools are paired in pseudo-strata in the order of selection, and within each pseudo-stratum, one school at random is given zero weight and the other receives a double weight. By contrast, PISA adopts a particular Fay's variant, with a random school receiving a weight inflated by 1.4 and the other deflated by 0.6 in each pseudo-strata. This approach is used to avoid losing half of the sample, which would make it difficult to estimate parameters on sparse subgroups of the population. As a result, contrary to the standard BRR method, the sum of squared deviation is not divided by 80 but by $80(1 - 0.6)^2$. When plausible values are used, in addition to this sampling variance, the standard errors are corrected by a measurement error variance equal to 1.2 times the variance of the five estimates. All details can be found in OECD 2005 (*14*).

In any table, an asterisk indicates significance at 5%, two asterisks, significance at 1%.

For each test, we run a regression of the individual test score for each student on a constant and a gender dummy (defined as before), where each observation is weighted by final students' weights. The gender difference (in mathematics, or in reading, or in arithmetic, or in geometry) is the coefficient on the gender dummy in a country-by-country regression. When we control for hours of instruction in mathematics, the gender difference in mathematics is the coefficient on the gender dummy in a regression of test scores on a constant, a gender dummy, and the number of hours of instructional time in math, weighting each observation by final students' weights. These differences are calculated in the selected subsample, keeping only the upper half of the observations of ESCS.

For any given percentile, the ratio of girls to boys who score above that threshold (95th or 99th percentile) is calculated by finding the score that falls at the percentile of interest, summing the students' final weights for boys and girls with score at or above the cutoff score, and taking the ratio of the two (the sum of students' weight is an estimate of the target population in the country). These ratios have been calculated on the whole sample, not selecting away observations in the bottom 50% of ESCS.

To study whether boys score higher in mathematics than in reading and whether girls score higher in reading than in mathematics, we computed an indicator, the boys relative advantage in math, as the difference between boys' and girls' scores in mathematics less the difference between boys' and girls' scores in reading (which is, in turn, the negative of each respective gender dummy). Analogously, the boys' relative advantage in geometry is defined as the difference between boys' and girls' scores in geometry less the difference between boys' and girls' scores in arithmetic. The arithmetic benchmark has been chosen because it is the most gender-neutral subtest in mathematics among the four available.

Average instructional time and average homework times are calculated as the mean across all countries of the country level average of each variable, by gender, weighted by final student weight, using the selected sample.

SOM Text

The figure in the text shows the gender differences in mathematics and reading for a subset of countries. In Figure S1A, we show the gender difference in mathematics and reading for all the countries in our sample: The figure shows that in most countries boys have better scores in mathematics than girls. However, the differences between boys and girls are statistically significant in only 26 countries. In the remaining countries, there is no statistical difference between boys' and girls' performance in mathematics, with some cases in which girls outperform boys, though the difference is not always statistically significant.

Better average performance of boys in mathematics is present in spite of the fact that girls spend on average 19.5% more time doing math homework and is robust when we control for the fact that boys on average spend 2.3% more hours in math courses (Table S1). Figure S1B shows the gender gap in mathematics after controlling for differences in instructional time. The results are very similar to those presented in Figure S1A.

In Figure S2A, we plotted the ratio between girls and boys with a math score above the 95th percentile and the 99th percentile of the country-level distribution of scores. On average, there are 0.6 girls for every boy with a math score above the 95th percentile of the country-level distribution of scores, with a range from 0.4 in Korea to 1.1 in Indonesia.

Figure S2B shows that gender differences in mean scores and the ratio of girls to boys who score above 95th percentile are both appropriate ways to capture the same phenomenon in math scores.

In Figure S3 we plotted the ratio between girls and boys with a reading score above the 95th percentile and the 99th percentile of the country-level distribution of scores. Confirming the results of the figure in the text, girls outperform boys. On average, there are 1.83 girls for every boy with a reading score above the 95th percentile of the country-level distribution of scores, with a range from 1 in Turkey to 2.9 in Iceland (Fig. S3A). Once again, the high correlation between gender differences in reading mean scores and the ratio of girls to boys who score above the 95th percentile (Fig. S3B) suggests that both measures capture the same phenomenon in reading scores.

Figure S4 shows that gender differences in mathematics and reading are highly correlated across countries.

The table in the text shows the correlation between various measures of gender emancipation and gender differences in mathematics and reading scores. The results presented are the coefficient of an ordinary least squares regression (OLS) where the left-hand side variable (LHS) is either the gender gap in mathematics (first four columns) or the gender gap in reading (last four columns). Depending on the specification used, the two gender gaps are regressed on various measures of gender equality (GGI index in the first column and fifth column, average WVS indicators, in the second and sixth column, female economic activity rate in the fourth and seventh column, and women's political empowerment in the fourth and eight column) and the log of GDP per capita. A scatter plot (Fig. S5) illustrates the same point in a different way. To illustrate that our results apply to the right tail of the distribution, Table S3 reports the same regressions of Table 1 by using as dependent variables the ratio between girls and boys with scores respectively in mathematics and in reading above the 95th percentile or the 99th percentile of the country-level distribution of scores.

To control for unobserved heterogeneity, we reran our regression at the student level inserting a dummy variable for each country (Table S4).

References and Notes

1. OECD, *Learning for Tomorrow's World— First Results from PISA 2003* (OECD, Paris, 2004), p. 27.

2. (*1*), p. 333.

3. (*1*), p. 25.

4. (*1*), p. 23.

5. (*1*), p. 37.

6. (*1*), p. 272.

7. R. Hausmann, L. D. Tyson, S. Zahidi, *The Global Gender Gap Report* (World Economic Forum, Geneva, 2006).

8. R. Inglehart *et al.*, World Values Surveys and European Values Surveys, 1981–1984, 1990–1993, and 1995–1997 [Computer file, Interuniversity Consortium for Political and Social Research (ICPSR) version] (Institute for Social Research, Ann Arbor, MI, 2000), distributed by ICPSR.

9. C. P. Benbow, J. C. Stanley, *Science* **222**, 1029 (1983).

10. ILO (International Labor Organization), Estimates and Projections of the Economically Active Population, 1980–2020 [Database, ed. 5, rev. 2] (ILO, Geneva, 2005).

11. A. Heston, R. Summers, B. Aten, *Penn World Table*, Version 6.2, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, September 2006.

12. L. L. Cavalli-Sforza, P. Menozzi, A. Piazza, *The History and Geography of Human Genes* (Princeton Univ. Press, Princeton, NJ, 1996).

13. E. Spolaore, R. Wacziarg, “The diffusion of development” (Centre for Economic Policy Research Discussion Paper 5630, CEPR, London, 2006).

14. OECD, *PISA 2003 Data Analysis Manual* (OECD, Paris, 2005).

Legends

Table S1. World-level summary statistics in instructional time and homework time in mathematics, by gender, expressed in 60 min hours per week. Generally, girls report slightly less instructional time and substantially more homework time in math. These differences are significantly different from zero in the cross section of countries.

Table S2. (**A**) Different measures of gender equality in society, (**B**) their summary statistics, and (**C**) their correlation. As can be seen, all measures are significantly associated with each other, suggesting that even though these measures use different criteria they capture the same phenomenon.

Table S3. An OLS regression that replicates the results in the table in the text, using as left-hand side variable the ratio of girls to boys who score above (**A**) the 95th and (**B**) 99th percentile of scores in math and reading tests. The positive, significant correlation found between these test scores and indicators of gender equality in society is replicated, which suggests that our results also apply to the right tail of the distribution. Standard errors are reported below the estimated coefficients.

Table S4. An OLS regression that replicates the results in the table in the text using student-level data for the pooled sample of countries, without (column I) and then with (column II) interaction between gender and log of GDP per capita. In both cases, the interaction Gender*GGI is positive and strongly significant, which confirms that our findings are not simply spurious correlations between unobserved factors and measures of gender equality. Individual level controls include dummies for any students who are in a grade different from the modal one in the country. Standard errors are reported below the estimated coefficients.

Table S5. The countries in our sample with genetic distance equal to zero and with genetic distance between each other lower or equal to 72.

Table S6. The analysis in the table in the text is replicated in two subsamples of the data, (**A**) those with genetic distance lower or equal than 72 with each other and (**B**) those with genetic distance equal to zero with each other. The purpose of this is to check if, within genetically homogeneous countries, there is the same pattern of positive correlation between gender differences in test scores and gender-equality indicators. Although significances drop substantially due to the severe reduction in the sample size, the main evidence is confirmed. Standard errors are reported below the estimated coefficients.

Table S7. OLS regressions of relative advantage of boys in math vs. reading (**A**) and in geometry vs. arithmetic (**B**) with respect to girls. Generally, these measures are not correlated with gender-equality indicators, which suggests that, in countries with better gender equality, girls tend to close the gap in mathematics while improving in all the other test scores as well. Standard errors are reported below the estimated coefficients.

Figure S1. (**A**) The average gender difference in mathematics and reading scores (same graphic as the figure in the text) for the whole sample of countries. For mathematics, test scores have been normalized so that the average score across all OECD students in 2003 equals 500. For reading, the average score across all OECD students in 2003 is 494. To calculate the gender difference in these two tests, for each country, we ran a regression of individual test scores on a constant and a gender dummy. The coefficient on the gender dummy is the gender difference between girls and boys in every country. Since we are concerned about possible differential drop-out rates across genders, to compute the means we remove all the observations in the lower half of the distribution of social economic status, where probability of drop-out is more likely. In almost any country, on average, boys perform better than girls in mathematics, whereas in every country, girls perform better than boys in reading. The gender gap in mathematics and reading is significantly different across countries. (**B**) The average gender difference in math scores after controlling for individual differences in instructional time in mathematics. Although girls report slightly less instructional time in mathematics, average gender difference in math scores across countries do not change when controlling for instructional time.

Figure S2. (**A**) The ratio of girls to boys who score above the 95th and 99th percentile of math tests are again generally in favor of boys, but varies from country to country, which supports the findings presented in the figure in the text and the evidence in the paper. (**B**) Gender differences in mean scores and the ratio of girls to boys who score above the 95th percentile are both appropriate ways to capture the same phenomenon in math scores. Similar correlation is found between the gender differences in mean scores and the ratio of girls to boys who score above the 99th percentile.

Figure S3. (**A**) The ratios of girls to boys who score above the 95th and 99th percentile of reading tests are also almost always in favor of girls, thus supporting the findings presented in the figure in the text. (**B**) In parallel with the findings shown in Fig. S2B, gender differences in mean scores and the ratio of girls to boys who score above the 95th percentile are both good ways to capture the same phenomenon in reading scores.

Figure S4. The correlation between the average gender gap in reading and in mathematics. The figure shows that, in countries where girls have a bigger reading advantage over boys, girls also have relatively higher math scores vis-à-vis boys.

Figure S5. The correlation between the average gender gap in mathematics and the GGI. Consistent with the results presented in the table in the text, in more gender-equal societies, girls perform better in mathematics vis-à-vis boys.

Table S1

	Boys	Girls	Difference, girls-boys	N. obs
Weekly hours instructional time	3.49**	3.41**	-0.08**	40
Weekly hours homework time	2.51**	2.99**	0.49**	39

Table S2

Panel A

Country	Women emancipation (GGI)	Avg. WVS indicators	Female activity rate	Women political empowerment
Australia	0.72	2.76	56.1	0.16
Austria	0.7	2.41	49.3	0.28
Belgium	0.71	2.62	43.4	.
Brazil	0.65	2.78	56.3	0.06
Canada	0.72	2.71	60.2	0.16
Czech Republic	0.67	2.53	51.7	0.09
Denmark	0.75	2.97	59.4	0.31
Finland	0.8	3.03	56.9	0.47
France	0.65	2.71	48.2	0.1
Germany	0.75	2.87	50.4	0.37
Greece	0.65	.	42.7	0.06
Hong Kong - China	.	.	52.9	.
Hungary	0.67	2.5	42.1	0.07
Iceland	0.78	2.88	70.9	0.46
Indonesia	0.65		50.7	0.1
Ireland	0.73	2.55	51.9	0.32
Italy	0.65	2.61	37	0.09
Japan	0.64	2.55	48.5	0.07
Latvia	0.71	2.59	49.1	0.22
Luxembourg	0.67		44.1	0.14
Mexico	0.65	2.61	39.9	0.13
New Zealand	0.75	.	59.8	0.32
Norway	0.8	3.1	63.1	0.49
Poland	0.68	2.61	47.9	0.11
Portugal	0.69	2.82	55.2	0.14
Russian Federation	0.68	2.59	54.3	0.03
Serbia - Montenegro	.	2.74	48.6	.
Slovak Republic	0.68	2.49	51.9	0.08
Spain	0.73	2.8	44.2	0.42
Sweden	0.81	3.17	58.8	0.55
Switzerland	0.7	2.78	60.1	0.15
Thailand	0.68	.	65.4	0.06
Tunisia	0.63	.	27.9	0.11
Turkey	0.59	2.42	27.8	0.05
United Kingdom	0.74	2.68	55	0.31
United States	0.7	2.79	59.6	0.1
Uruguay	0.65	2.69	55.7	0.04
Average	0.7	2.71	51.36	0.19
Standard Deviation	0.05	0.19	9.04	0.15

Panel B

	Women emancipation (GGI)	Avg. all WVS	Female economic activity rate	Women Political Empowerment
Average	0.70	2.71	51.36	0.19
Std. Deviation	0.05	0.19	9.04	0.15
Observations	37	32	39	36

Panel C

	Women Emancipation (GGI)	Avg. all WVS	Female economic activity rate	Women Political Empowerment
Women Emancipation (GGI)	1			
Avg. all WVS	0.80**	1		
Female economic activity rate	0.65**	0.62**	1	
Women Political Empowerment	0.90**	0.72**	0.40*	1

Table S3

Panel A

	Ratio girls-to-boys above 95th percentile, math				Ratio girls-to-boys above 95th percentile, reading			
Women emancipation (GGI)	1.86± 0.55**				5.00± 1.32**			
Average World Values Survey (WVS) indicators		0.26± 0.12*				0.80± 0.39*		
Female economic activity rate			0.01± 0.00**				0.02± 0.01*	
Women Political Empowerment				0.54± 0.20*				1.47± 0.48**
Log GDP per capita, 2003	-0.17± 0.05**	0.01± 0.04	-0.10± 0.04*	-0.14± 0.05**	-0.12± 0.12	0.09± 0.12	-0.02± 0.10	-0.06± 0.12
Constant	0.96± 0.41*	-0.24± 0.39	1.06± 0.35**	1.88± 0.47**	-0.43± 0.98	-1.18± 1.28	1.06± 0.91	2.13± 1.14
Observations	37	32	39	36	37	32	39	36
R-squared	0.30	0.18	0.28	0.23	0.32	0.19	0.18	0.25

Panel B

	Ratio girls-to-boys above 99th percentile, math				Ratio girls-to-boys above 99th percentile, reading			
Women emancipation (GGI)	2.39± 0.81**				7.12± 2.59**			
Average World Values Survey (WVS) indicators		0.14± 0.19				1.08± 0.71		
Female economic activity rate			0.01± 0.00**				0.03± 0.01*	
Women Political Empowerment				0.69± 0.29*				2.18± 0.92*
Log GDP per capita, 2003	-0.17± 0.07*	0.08± 0.06	-0.08± 0.05	-0.13± 0.07	-0.15± 0.23	0.18± 0.23	0.02± 0.18	-0.07± 0.23
Constant	0.46± 0.60	-0.69± 0.62	0.63± 0.49	1.63± 0.68*	-1.25± 1.93	-2.51± 2.35	0.50± 1.66	2.44± 2.16
Observations	37	32	39	36	37	32	39	36
R-squared	0.22	0.11	0.25	0.16	0.21	0.13	0.15	0.17

Table S4

	Student individual score, mathematics	
Intercept	534.92± 2.99**	536.54± 2.97**
Gender	-69.87± 21.22**	-57.11± 22.43*
Gender * Women emancipation (GGI)	81.54± 30.93**	134.86± 38.19**
Gender * Log GDP per capita, 2003		-5.11± 2.38*
Country fixed effects	Yes	Yes
Individual level controls	Yes	Yes
Observations	132,124	132,124
R-squared	0.42	0.42

Table S5

Largest set of countries with genetic distance = 0	Australia Belgium Canada France Iceland Ireland Luxembourg New Zealand Poland Spain U.S.A. United Kingdom Uruguay	
Largest set of countries with genetic distance < 100	Australia Austria Belgium Brazil Canada Czech Republic Denmark France Germany Iceland Ireland Italy Luxembourg	Netherlands New Zealand Norway Poland Portugal Russian Federation Slovakia Spain Sweden Switzerland U.S.A. United Kingdom Uruguay

Table S6

Panel A

	Gender difference in test score, math				Gender difference in test score, reading			
Women emancipation (GGI)	91.89± 31.07**				57.39± 40.63			
Avg. WVS indicators		9.00± 8.67				6.15± 10.40		
Female economic activity rate			0.49± 0.18*				0.27± 0.23	
Women political empowerment				25.10± 10.04*				21.24± 12.37
log GDP per capita, 2003	-0.07± 3.00	4.45± 3.46	3.49± 2.70	0.19± 3.22	-0.26± 3.92	2.57± 4.15	2.01± 3.52	-0.89± 3.97
Constant	-75.20± 27.71*	-79.10± 34.36*	-71.93± 27.89*	-18.07± 31.13	-3.81± 36.24	-7.47± 41.23	0.00± 36.42	38.65± 38.32
Observations	26	24	26	25	26	24	26	25
R-squared	0.33	0.17	0.31	0.28	0.10	0.05	0.07	0.14

Panel B

	Gender difference in test score, math				Gender difference in test score, reading			
Women emancipation (GGI)	114.92± 56.13				79.18± 68.67			
Avg. WVS indicators		45.52± 25.53				58.48± 27.48		
Female economic activity rate			0.60± 0.24*				0.50± 0.30	
Women political empowerment				28.77± 17.65				27.45± 19.91
log GDP per capita, 2003	-2.11± 4.57	0.39± 5.36	-0.14± 4.15	-1.44± 4.97	-4.37± 5.59	-5.42± 5.76	-3.06± 5.09	-4.32± 5.61
Constant	-68.91± 52.79	-134.54± 73.97	-39.43± 43.39	-0.12± 49.31	22.93± 64.58	-67.78± 79.61	39.09± 53.11	72.97± 55.64
Observations	13	11	13	12	13	11	13	12
R-squared	0.30	0.31	0.38	0.23	0.14	0.37	0.24	0.19

Table S7

Panel A

	Relative advantage of boys in math vs. reading			
Women emancipation (GGI)	-21.93± 30.81			
Avg. WVS indicators		3.18± 7.45		
Female economic activity rate			-0.11± 0.15	
Women political empowerment				-4.76± 10.73
Log GDP per capita, 2003	4.33± 2.74	-0.57± 2.38	2.55± 2.14	3.89± 2.70
Constant	16.60± 22.90	41.07± 24.56	24.24± 19.70	6.60± 25.33
Observations	37	32	39	36
R-squared	0.07	0.01	0.04	0.06

Panel B

	Relative advantage of boys in geometry vs. arithmetic			
Women emancipation (GGI)	-10.71± 27.52			
Avg. WVS indicators		-15.77± 6.16*		
Female economic activity rate			0.10± 0.13	
Women political empowerment				-9.09± 9.34
Log GDP per capita, 2003	3.67± 2.45	5.14± 1.97*	2.07± 1.87	4.12± 2.34
Constant	-18.39± 20.46	2.38± 20.30	-15.37± 17.21	-28.66± 22.04
Observations	37	32	39	36
R-squared	0.07	0.25	0.07	0.09

Figure S1A

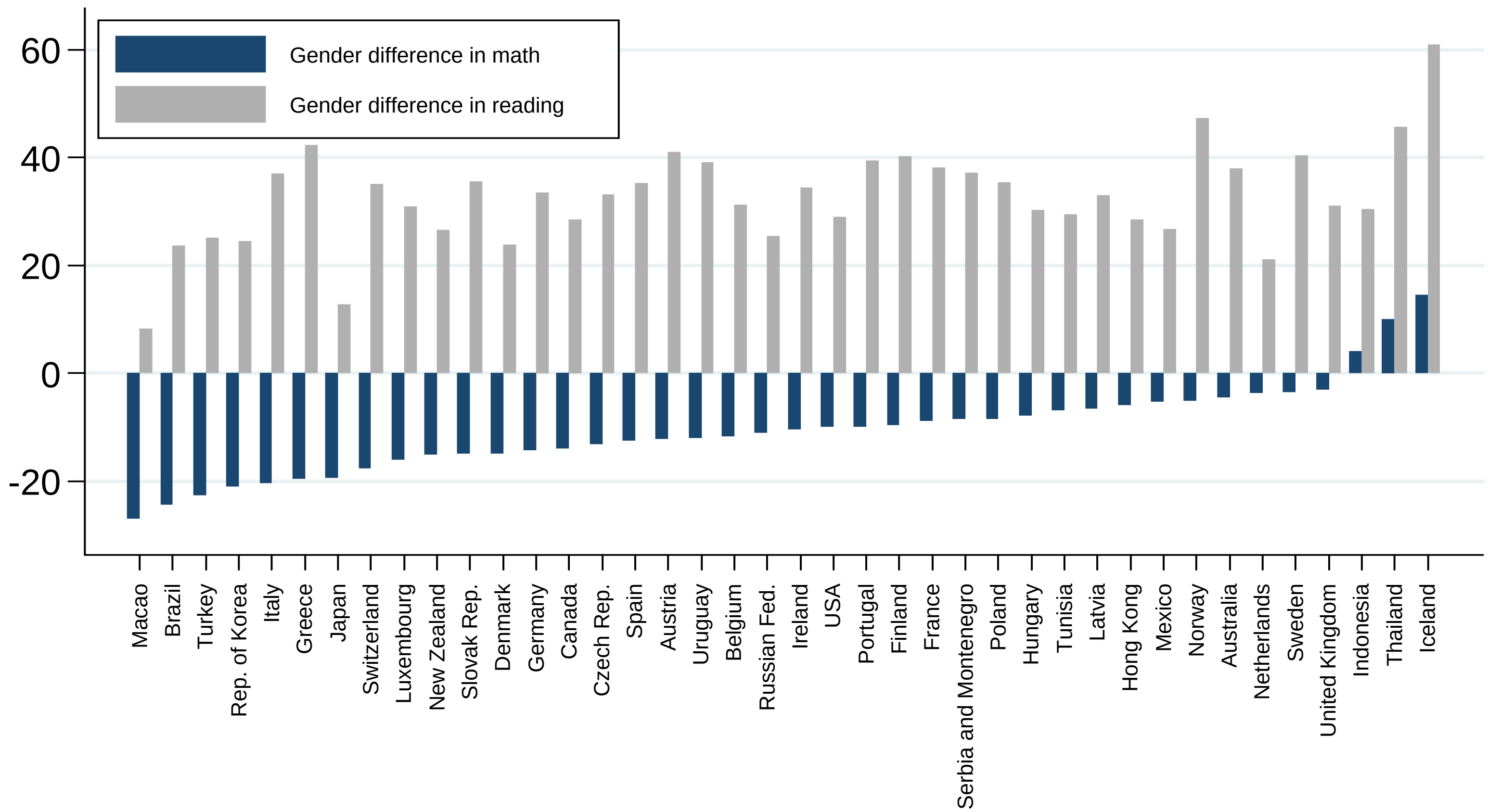

Gender difference in math
Gender difference in reading
60
40
20
0
-20
Macao
Brazil
Turkey
Rep. of Korea
Italy
Greece
Japan
Switzerland
Luxembourg
New Zealand
Slovak Rep.
Denmark
Germany
Canada
Czech Rep.
Spain
Austria
Uruguay
Belgium
Russian Fed.
Ireland
USA
Portugal
Finland
France
Serbia and Montenegro
Poland
Hungary
Tunisia
Latvia
Hong Kong
Mexico
Norway
Australia
Netherlands
Sweden
United Kingdom
Indonesia
Thailand
Iceland

Figure S1B

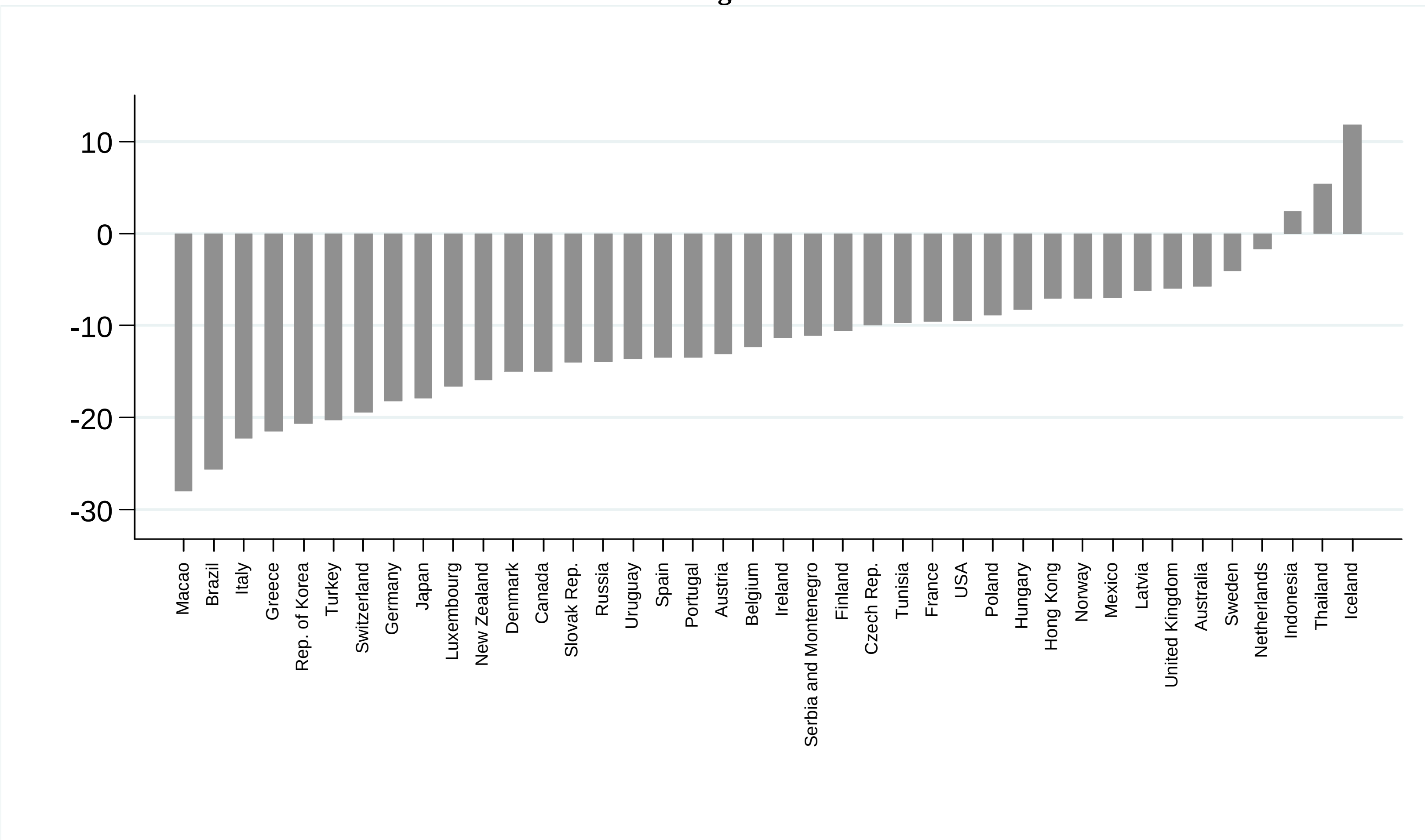
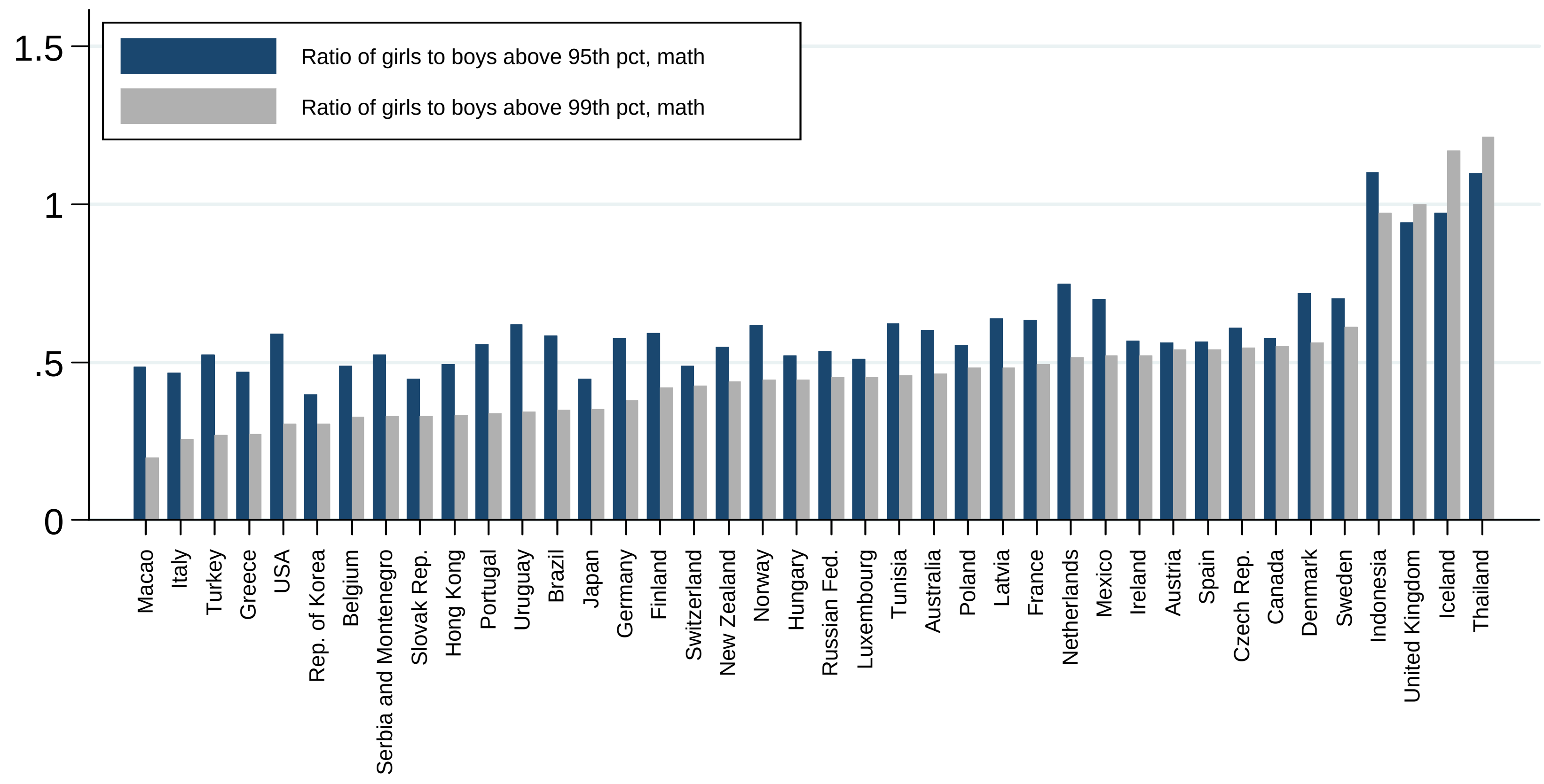

10
0
-10
-20
-30
Macao
Brazil
Italy
Greece
Rep. of Korea
Turkey
Switzerland
Germany
Japan
Luxembourg
New Zealand
Denmark
Canada
Slovak Rep.
Russia
Uruguay
Spain
Portugal
Austria
Belgium
Ireland
Serbia and Montenegro
Finland
Czech Rep.
Tunisia
France
USA
Poland
Hungary
Hong Kong
Norway
Mexico
Latvia
United Kingdom
Australia
Sweden
Netherlands
Indonesia
Thailand
Iceland

Figure S2A



Ratio of girls to boys above 95th pct, math
Ratio of girls to boys above 99th pct, math
1.5
1
.5
0
Macao
Italy
Turkey
Greece
USA
Rep. of Korea
Belgium
Serbia and Montenegro
Slovak Rep.
Hong Kong
Portugal
Uruguay
Brazil
Japan
Germany
Finland
Switzerland
New Zealand
Norway
Hungary
Russian Fed.
Luxembourg
Tunisia
Australia
Poland
Latvia
France
Netherlands
Mexico
Ireland
Austria
Spain
Czech Rep.
Canada
Denmark
Sweden
Indonesia
United Kingdom
Iceland
Thailand

Figure S2B

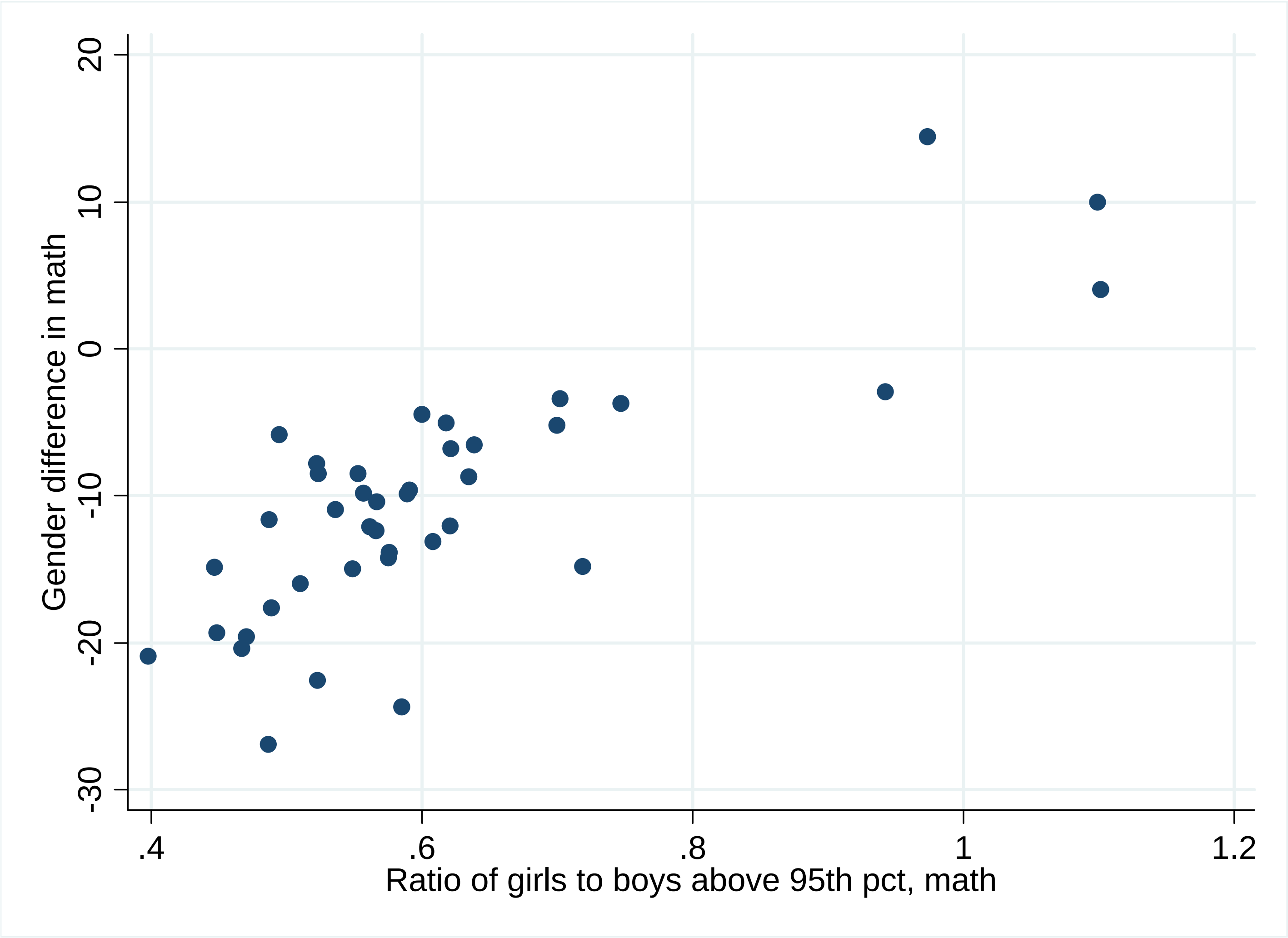

Gender difference in math
20
10
0
-10
-20
-30
.4
.6
.8
1
1.2
Ratio of girls to boys above 95th pct, math

Figure S3A

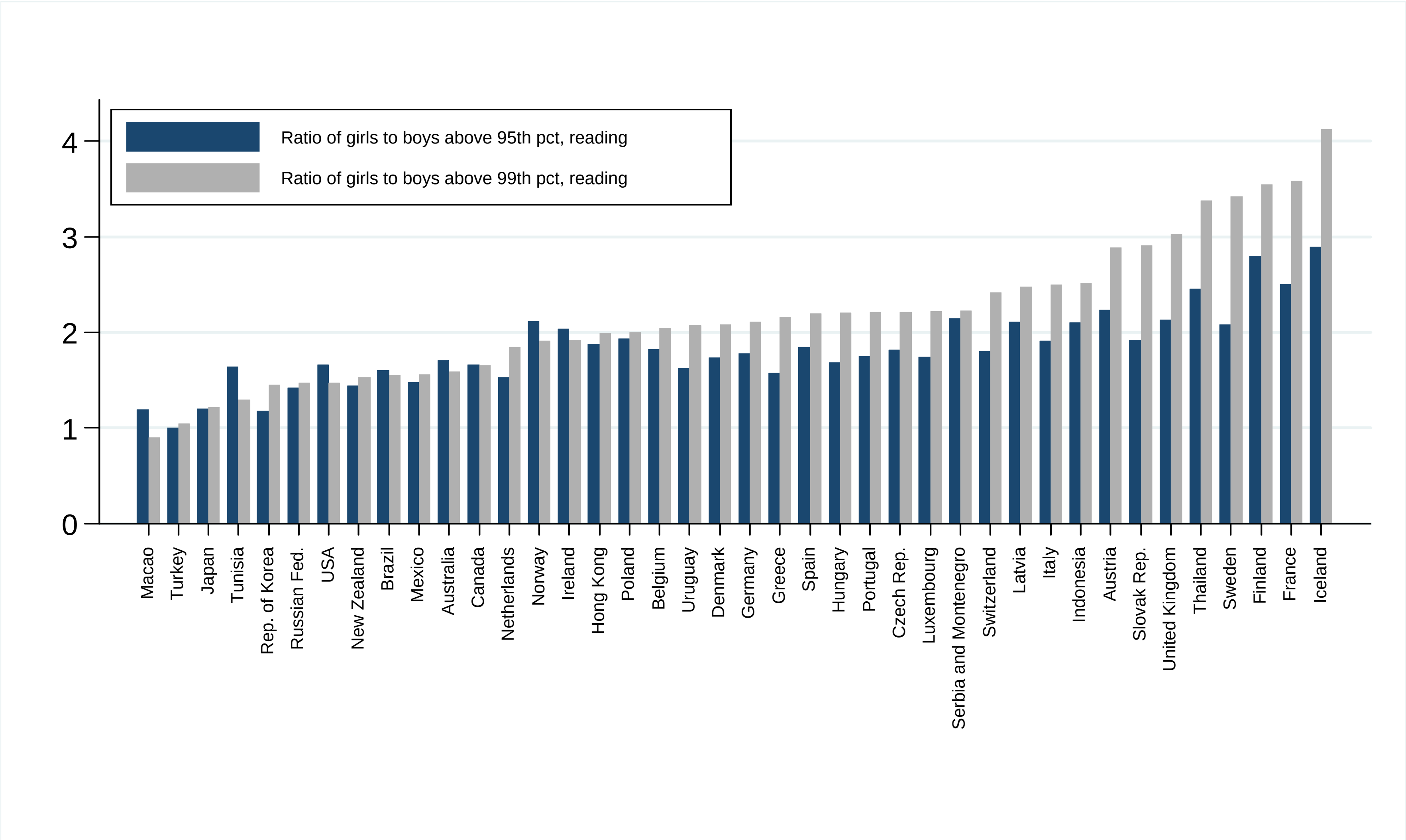

Ratio of girls to boys above 95th pct, reading
Ratio of girls to boys above 99th pct, reading
0
1
2
3
4
Macao
Turkey
Japan
Tunisia
Rep. of Korea
Russian Fed.
USA
New Zealand
Brazil
Mexico
Australia
Canada
Netherlands
Norway
Ireland
Hong Kong
Poland
Belgium
Uruguay
Denmark
Germany
Greece
Spain
Hungary
Portugal
Czech Rep.
Luxembourg
Serbia and Montenegro
Switzerland
Latvia
Italy
Indonesia
Austria
Slovak Rep.
United Kingdom
Thailand
Sweden
Finland
France
Iceland

Figure S3B

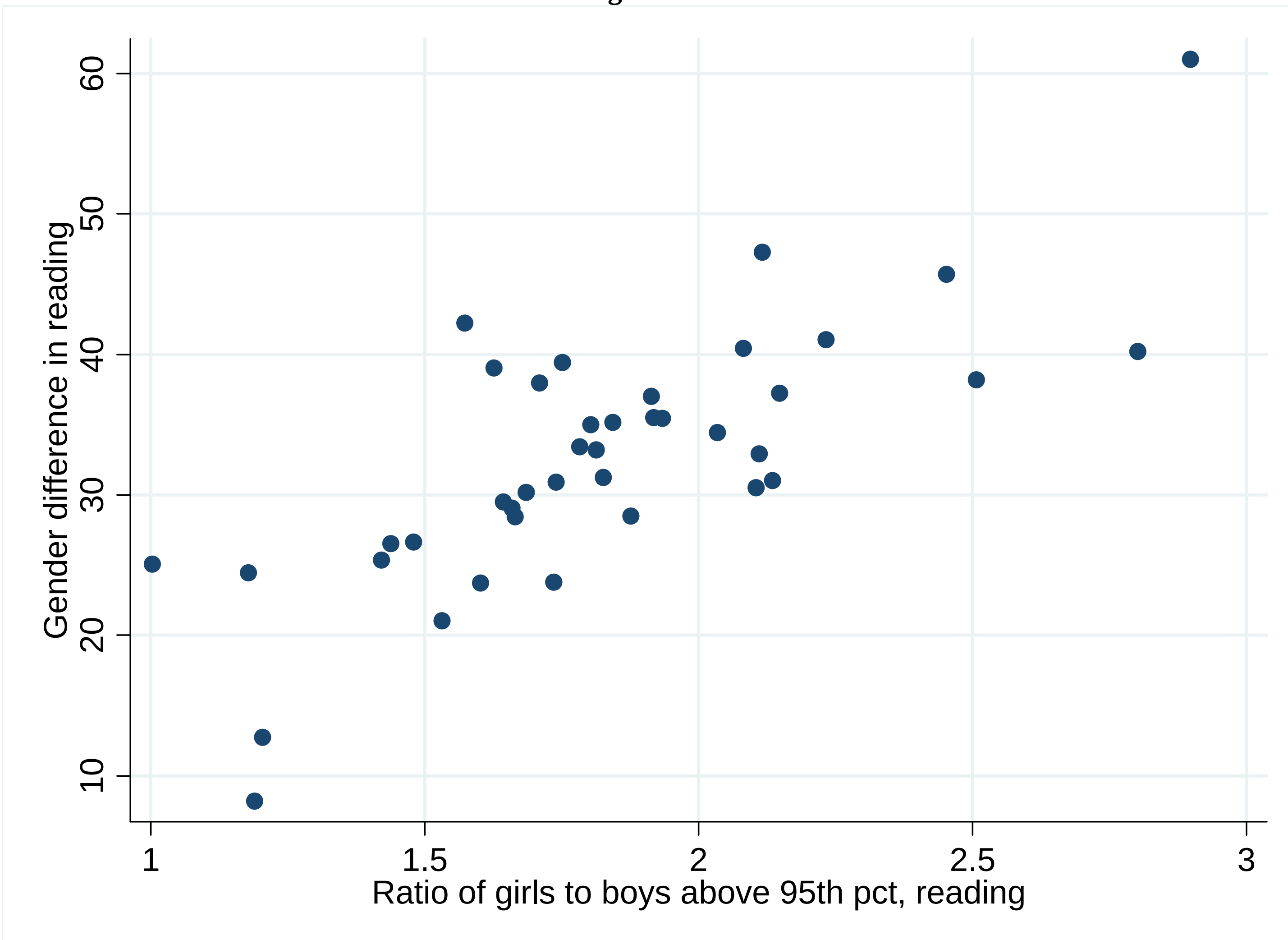

Gender difference in reading
10
20
30
40
50
60
1
1.5
2
2.5
3
Ratio of girls to boys above 95th pct, reading

Figure S4

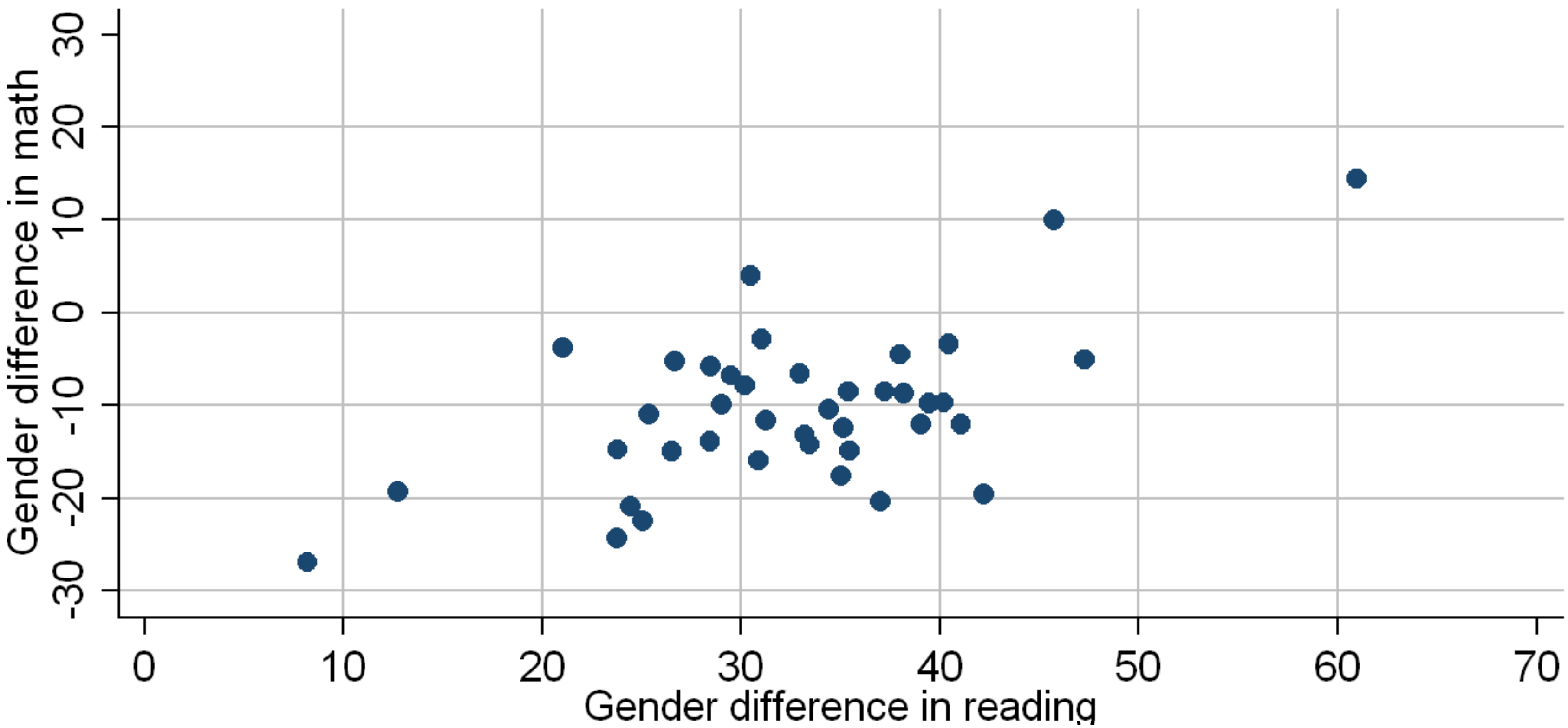
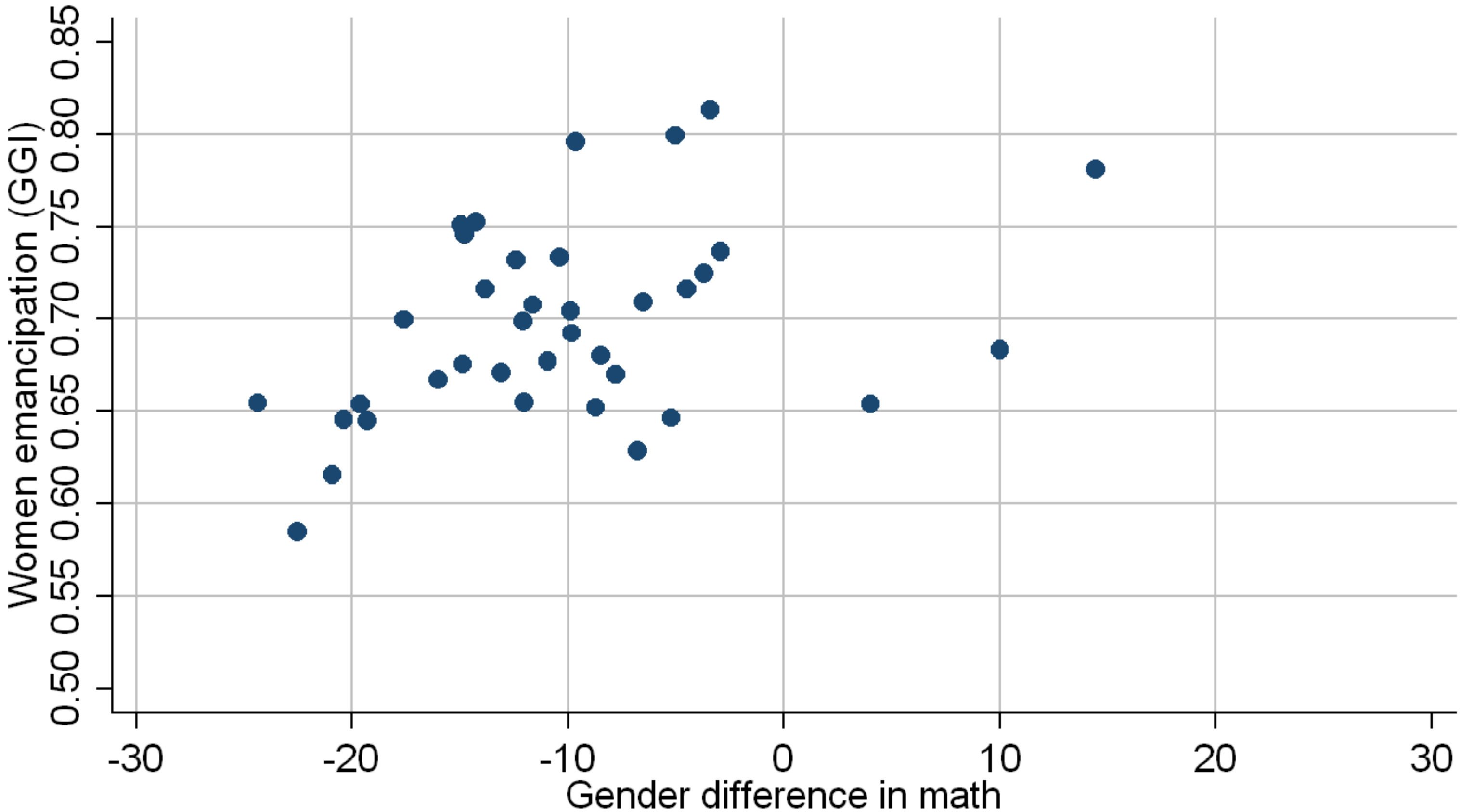

Gender difference in math
-30
-20
-10
0
10
20
30
0
10
20
30
40
50
60
70
Gender difference in reading

Figure S5


Women emancipation (GGI)
0.50 0.55 0.60 0.65 0.70 0.75 0.80 0.85
-30
-20
-10
0
10
20
30
Gender difference in math