January 2017

The gender gap in math performance, self-concept, and anxiety: rural and urban China in an international context

Meichen Lu, Yaojiang Shi, Fang Chang, Kaleigh Kenny, Scott Rozelle

Abstract

Evidence from developed countries shows that there is a significant gender gap in STEM occupations. Girls may begin to underperform in math early as primary school. One possible explanation is the negative stereotype threat towards girls. However, this has been understudied in rural China. In this paper, we describe the math performance gender gap in rural China, compare the gender gap between rural and urban China, and finally compare the Chinese situation with other countries. We further examine possible explanations for the math performance gender gap from comparative perspectives. Using first hand datasets of 3,789 primary students and 12,702 junior high students in northwest China, combing with OECD 2012 Program for International Student Assessment (PISA) survey data, we find that in both rural and urban China, boys outperform girls in math. As students grow older, the gap widens. The size of the gender gap in rural China is larger than that in urban China, and larger than in many other countries. We further find that both the gender gaps in math self-concept and math anxiety and discriminatory family investment towards girls are not sufficient to explain the wide math performance gaps. Our study suggests the inequality of education in rural China still merits concern and calls for further work to explain the observed gender gap in math performance.

Working Paper 312

January 2017

reap.fsi.stanford.edu

The gender gap in math performance, self-concept, and anxiety: rural and urban China in an international context

Meichen Lu[a], Yaojaing Shi[a], Fang Chang[a*], Kaleigh Kenny[b], Scott Rozelle[b]

[a] Center for Experimental Economics in Education, Shaanxi Normal University, 710119 Xi'an, Shaanxi, China
[b] Rural Education Action Program, Freeman Spogli Institute for International Studies, Stanford University, Stanford, CA 94305-6055, United States

Abstract:
Evidence from developed countries shows that there is a significant gender gap in STEM occupations. Girls may begin to underperform in math early as primary school. One possible explanation is the negative stereotype threat towards girls. However, this has been understudied in rural China. In this paper, we describe the math performance gender gap in rural China, compare the gender gap between rural and urban China, and finally compare the Chinese situation with other countries. We further examine possible explanations for the math performance gender gap from comparative perspectives. Using first hand datasets of 3,789 primary students and 12,702 junior high students in northwest China, combing with OECD 2012 Program for International Student Assessment (PISA) survey data, we find that in both rural and urban China, boys outperform girls in math. As students grow older, the gap widens. The size of the gender gap in rural China is larger than that in urban China, and larger than in many other countries. We further find that both the gender gaps in math self-concept and math anxiety and discriminatory family investment towards girls are not sufficient to explain the wide math performance gaps. Our study suggests the inequality of education in rural China still merits concern and calls for further work to explain the observed gender gap in math performance.

Keywords: Gender gap; Math performance; Math self-concept; Math anxiety; Rural; China;

***Corresponding author**
Fang Chang
P.O. box 9, Shaanxi Normal University
620 Chang'an Rd West, Chang'an District, Xi'an City
Shaanxi Province, P.R.China, 710127.
E-mail: changfang4421@163.com
Tel.: +86 134 8495 0704; Fax: +86 029 8153 0872;

Acknowledgments: The authors would like to acknowledge the funding of the 111 project (高等学校学科创新引智计划资助, B16031), the International Initiative for Impact Evaluation (3ie) and the Ford Foundation.

The gender gap in math performance, self-concept, and anxiety: rural and urban China in an international context

Introduction

Many high-paying jobs in developed countries are in the fields of science, technology, engineering and math (STEM). For example, in the United States the average annual wages of STEM workers were 26 percent higher than wages in other fields in 2010 (Langdon et al., 2011). In Taiwan, eight of the top ten highest paying industries in 2011 were STEM-related fields (Marginson et al., 2013). High wages are found in these fields because they promote economic growth by encouraging scientific advances and technological innovation (Marginson et al., 2013; Carnevale et al., 2011; Atkinson and Mayo, 2010). However, compared to men, women are underrepresented in STEM fields in many countries (Beede et al., 2011; Hill et al., 2010). Although women occupied almost half of all jobs in the United States in the last decade, they held less than 25 percent of STEM occupations (Beede et al., 2011). This gender gap in STEM occupations may contribute to the overall gender gap in wages as well as social inequality (O'Neill, 2003; Beede et al., 2011; Barres, 2006; Xie and Shauman, 2003).

The STEM gender gap begins before men and women even enter the labor force. The gateway to many high-paying STEM occupations is a STEM degree or STEM skills-based education (Beede et al., 2011; Langdon et al., 2011; Atkinson and Mayo, 2010; Carnevale et al., 2011). However, women are underrepresented in STEM fields at colleges and universities (Hill et al., 2010; OECD, 2012). Evidence suggests the gap may begin even earlier in the course of schooling, potentially in junior high school or primary school. In many countries, girls perform worse than boys on standardized math examinations at these educational stages

(Else-Quest et al. 2010; Nosek et al. 2009; Osborne and Dillon, 2008; Guiso et al. 2008). For example, the 2012 Program for International Student Assessment (PISA) found that boys outperform girls in math by 11 points (about 0.1 standard deviations) on average across OECD countries (OECD, 2014); this gap is observed in 38 countries that participated in the test (e.g. Austria, Korea, Japan, Italy, Spain and New Zealand; OECD, 2014).

There are many possible structural or institutional explanations that might contribute to this persistent gender gap in math performance. One possible explanation is that there are inherent biological differences between men and women in cognitive and problem-solving abilities (Kimura, 2000; Byrnes, 2005; McClure, 2003; Halpern et al., 2007). However, research on sex differences in brain structure and hormones is inconclusive in explaining women's underrepresentation in math-related fields (Ceci et al, 2009). Other researchers state that the math performance gender gap could be attributed, in part, to discriminatory resource allocation (Fryer and Levitt, 2010; Hannum et al., 2009). That is, poor families may decide to invest scare resources in their sons rather than their daughters. It is also possible that, due to perceptions of differential labor market outcomes, boys are more likely to receive educational investments from parents and attention from teachers than girls (Hannum et al., 2009; Leinhardt et al., 1979).

Beyond these explanations, some researchers point to the existence of a negative stereotype threat that may affect the math performance of female students (Farenga and Joyce, 1999; Ambady et al., 2001; Osborne, 2001; Brown and Pinel, 2003). In other words, since girls are often told that their math skills are worse than those of boys, it undermines their confidence and, ultimately, their math performance (Hill et al., 2010; Nguyen and Ryan 2008;

Keller and Dauenheimer, 2003; Brown and Pinel, 2003; Steele and Aronson 1995). The threat of negative stereotyping is concerning because there is empirical evidence that one of the factors that shapes the behavior of students is their self-confidence, especially in competitive educational environments (Bandura, 1977).

In recent years researchers have focused on how math self-confidence affects student performance (Pajares and Miller, 1994; Osborne, 2001; Pietsch, Walker and Chapman, 2003; Furner and Berman, 2003; Marsh and Hau, 2004; OECD, 2014; OECD, 2015). Math self-confidence is related to two ideas: math self-concept (how a student views his/her own math ability) and math anxiety (stress that interferes with a student's performance in math). Evidence from developed countries shows that these two concepts are related to student performance in math and related subjects (Hembree, 1990; Ma, 1999; Ashcraft and Kirk, 2001; Marsh et al., 2006; Marsh and O'Mara, 2008). For example, Marsh and O'Mara (2008) found that academic self-concept and performance mutually reinforce each other among 10th graders in the United States. From a study conducted in 25 countries, Marsh et al. (2006) demonstrated that math self-concept is correlated with math performance (r=0.33). In a meta-analysis of 26 correlative studies, Ma (1999) found that math anxiety and math performance are significantly and negatively related.

Findings suggest that the apparent gender gap in math performance in many countries may be due, in part, to differences in math self-confidence. Using data from the 2012 PISA, the OECD reported that in most participating countries there was a considerable gap in math self-confidence between the genders, with boys generally being more confident in their math abilities. The same report suggests that even when math performance is the same for both genders, girls display lower levels of math self-concept and higher levels of math anxiety (OECD, 2014).

While international findings are relatively consistent, findings on the relative math performance of boys and girls in China have been more mixed (Turner, 1994; Lai, 2010). On the one hand, some studies find that male students outperform female students on math examinations (Turner, 1994; Gong et al., 2014). Using data from 235 junior high school students in the city of Wuhan, Turner (1994) found that the math test scores of males are significantly higher than those of females. On the other hand, studies have also found that female students perform better in math than male students (Lai, 2010; Wang et al., 2016). Lai (2010) found that girls performed better than boys in math on a Junior High School Graduation Exam among a sample of 7,235 students in Beijing. In addition, a 2014 study using nationally representative data found that the performance of 10 to 15 year old girls on a standardized math test indicated that their math skills were nearly one year ahead of those of boys of the same age range (Wang et al., 2016).

Due to these inconsistent findings, the gender differences in math performance in China still requires further study. And, if a gender gap exists, more research is necessary to determine the causes of the gap. To date, little nationally-representative data is available on gender differences in math performance, self-concept and anxiety in China. The PISA survey only provides results on the gender difference in math self-concept and math anxiety in Shanghai, China's richest city (OECD, 2014). Furthermore, to our knowledge, no large-scale, empirical study has been published that measures the gender gap in math performance between male and female students from rural areas of China.

The overall goal of this study is to determine whether there is a gender gap in math performance, self-concept, and anxiety among primary and junior high school students in rural China. To reach this goal, we first describe and compare the gender gap in math performance in primary school and junior high school in rural China. Next, we compare the gender gap in math performance between rural and urban China, and internationally with other countries to examine whether the size of the gap is relatively large or small. In the next part of the paper we begin to seek an explanation for why there is a gender gap in rural China. To do so, the paper describes and compares the gender gaps in math self-concept and anxiety in primary school and junior high school in rural China to determine whether these are drivers of a gender gap in math performance, if it exists. We then compare the gender gaps in math self-concept and anxiety in urban China with those from other countries.

The remainder of the paper is organized as follows. In Section 2, we describe the sampling process and data collection effort. Sections 3 reports and discusses the results. Section 4 presents our conclusions.

Data

In this paper we draw on three datasets. The *rural China dataset 1* is comprised of 3,789 grade 4 students in 103 primary schools in Shaanxi and Gansu provinces in 2014. The *rural China dataset 2* is comprised of 12,702 grade 7 students in 200 junior high schools in Shaanxi and Gansu provinces in 2015. We also use data from the OECD 2012 PISA survey (henceforth *PISA 2012 dataset*) as the source of data on our outcomes measures for students from urban China (specifically, Shanghai) and from other countries. In total, the PISA test was taken by 510,000 15-year-old students in 65 countries, with 6,374 students participating in the exam from Shanghai (OECD, 2012).[1]

Sampling

Both of our datasets from rural China were collected in Shaanxi and Gansu Provinces in Northwest China. Shaanxi and Gansu Provinces are located in northwest China. The provincial GDP per capita of Shaanxi is 46,928 yuan (approximately 7640 dollars); and Gansu is 26,427 yuan (approximately 4303 dollars). Shaanxi ranks 14th and Gansu ranks 31st among the 31 mainland provinces of China in terms of GDP per capita (CNBS, 2014).

Our sampling strategy for rural China dataset 1 consisted of three steps. First, we restricted our sampling frame to 16 nationally-designated "poverty" counties in Yulin Prefecture (Shaanxi Province) and Tianshui Prefecture (Gansu Province). Second, we obtained a list of all primary schools with grades 1 to 6 in the sample counties and randomly selected 103 schools. Finally, we randomly chose one 4th grade class from each school as a sample class. During the data collection process, we surveyed all students in sample classes.

The sampling strategy for rural China Dataset 2 was similar to that used for rural China Dataset 1. In this case, we restricted our sampling frame to 23 nationally-designated "poverty" counties in Yulin Prefecture and Shangluo Prefecture (both in Shaanxi Province) and Tianshui Prefecture (Gansu Province). Second, we obtained a list of all junior high schools in these counties and randomly chose 200 junior high schools as our sample schools. Third, we randomly chose one 7th grade class as our sample class in each sample school, and surveyed all students in sample classes.

[1] Due to the different education systems in different countries, the PISA survey uses an age-based sample (OECD, 2012).

According to reports by the OECD, the PISA data employed a two-step sampling procedure in most countries. The first step drew a (usually stratified) random sample of at least 150 schools per country that enrolled 15-year-old students (OECD, 2012). The second step was to randomly sample 15-year-old students in each school, with each 15-year-old student in a school having equal selection probability (OECD, 2012). The number of students sampled per school ranged from 20 to 35 students. This sampling procedure typically restricts a sample to between 4,500 and 100,000 tested students in each country (OECD, 2012).[2]

Data Collection

For rural China datasets 1 and 2, we visited each sample school and surveyed students. The survey for primary school students and junior high students were similar and both consisted of three blocks. In the first block, we asked each student to take a standardized math test. Items on the standardized math test were drawn from the Chinese mathematics curriculum; the validity of these items had been examined by multiple experts. Primary school students and junior high school students were given different tests according their grades. All math test items were drawn from the Chinese National Curriculum Framework for each grade surveyed (MOE, 2011).The psychometric properties of the test were then validated using data from extensive pilot testing.[3] The students were required to finish the math test in 30 minutes. We prepared and administered the test ourselves to ensure that students and teachers could not prepare for the test. The enumeration team closely proctored

2 Why are results from Shanghai the only city reported in our study from the PISA survey? At present, only data from Shanghai are available to the public. According to the explanation of OECD (2015), the national Ministry of China has carried out a piloting of the PISA test in several provinces in previous cycles as preparation for fuller participation of China. Only Shanghai participated fully in PISA 2009 and PISA 2012.

3 It should be noted that the standardized math test that we used in the rural China sample was different from that used in the PISA test (see section below). However, as we describe in the subsequent paragraphs the scales that we use for math self-concept and math anxiety are identical to those given in the PISA test.

the exams in order to strictly enforce time limits and minimize cheating. For the analysis, the scores were standardized by scaling them into z-scores, which was done by subtracting the mean score and dividing by the standard deviation (SD) of the math score distribution of all students in each dataset. These *standardized math test scores* are used as our measure for student math performance.

In the second block, students were asked to fill out a survey instrument that was designed to measure their math self-concept and math anxiety. The math self-concept and math anxiety scales were originally designed for (and were identical to those used in) the PISA survey (see below). In both surveys, students responded to items that were given with a 4-point Likert-type response of "*strongly agree, agree, disagree,* and *strongly disagree*." The measurement items we used in our surveys were consistent with those used in PISA survey (OECD, 2012). Specially, five items were used to measure the level of math self-concept: *"I am just not good at math," "I get good grades in math," "I learn math quickly," "I have always believed that math is one of my best subjects,"* and *"In my math class, I understand even the most difficult work.*" In addition, another five items were used to measure the math anxiety of students: *"I often worry that it will be difficult for me in math classes," "I get very tense when I have to do math homework," "I get very nervous doing math problems," "I feel helpless when doing a math problem,"* and *"I worry that I will get poor grades in math"* (OECD, 2012).

After collecting data on math self-concept and math anxiety, we constructed the variables to be used in our analysis. The two concept scales were independently standardized to a mean of 0 and a standard deviation of 1 across samples. Positive values on the math self-concept index suggest that a student reported higher math self-concept than the average student, meaning he/she has higher levels of confidence in his/her math ability. Meanwhile, positive math anxiety scores indicate that a student's level of math anxiety is higher than that of the average student. In other words, a high math anxiety score means that a student suffers from higher levels of stress when completing math problems.

In the third block, students were asked to provide information on basic demographic and family background characteristics. Questions asked for information on student age, gender, boarding status, whether his/her father or mother completed at least junior high school, and family assets. A summary of the characteristics of sampled students is provided in Appendix I.

The PISA 2012 survey was administered by the PISA National Center in each country. For interested readers, more information on the PISA 2012 survey is available through the OECD's website (http://www.oecd.org/pisa/keyfindings/pisa-2012-results.htm). In this study we utilize the PISA standardized math test scores and results from the math self-concept and math anxiety scales.

Results

Gender Gap in Math Performance

According to our data, there is indeed a gender gap in math performance in favor of boys among our sampled schools in rural China (Table 1). Specifically, in primary school, we calculated that the average math test scores of female students are 0.08 standard deviations below those of male students (significant at the 5% level – Table 1, column 4, row 1). This gap appears to increase as students age. Among students in our junior high school sample, the average standardized math test scores of girls are 0.18 standard deviations lower than those of boys and this difference is significant at the 1% level (Table 1, column 4, row 3). These findings suggest that the gender gap in math performance begins in primary school and becomes wider as students move through junior high school.

In context, we can see that the gender gap in math performance in rural China is much wider than the gap observed in most countries. While the standardized math test scores of junior high school girls in 49 participating countries were lower than boys' (Figure 2), the Rural China's math performance gender gap junior high students ranks twelfth among 65 participating countries and economies. The size of this gap is larger than both the OECD average (rank 28) and urban China's gap (rank 41; Figure 2).[4] Comparing rural China with other participating countries highlights the fact that the gender gap in math performance in rural China merits particular concern.

Gender gaps in math self-concept and math anxiety

We find that girls, in general, display lower levels of math self-concept and higher levels of math anxiety. In primary school, the average math self-concept measure for girls is 0.21 standard deviations lower than that for boys (significant at the 1% level – Table 2, column 4, row 1). In the same schools, the average math anxiety measure for girls is 0.09 standard deviations higher relative to that for boys (significant at 1% level – Table 2, column 4, row 4). The results of students in junior high school are consistent with those in primary school. In junior high schools, the average math self-concept measure for girls is 0.32 standard deviations lower than that for boys (significant at the 1% level – Table 2, column 4,

[4] All gender gap values reported in this paper are calculated in the same way: we subtracted girl outcomes from boy outcomes (for math scores, math self-concept and math anxiety).

row 2). Similarly, the math anxiety index of girls is 0.12 standard deviations higher than that for boys (significant at the 1% level –Table 2, column 4, row 5).

Importantly, our data suggest that the gender gaps in both math self-concept and math anxiety in junior high school are larger than the gender gaps observed in primary school. We find that, on average, junior high school students display measures of math self-concept 0.25 standard deviations lower than that of primary school students (significant at the 1% level – Table 2, column 1, row 3), while also suffering from levels of math anxiety 0.29 standard deviations higher than primary school students (Table 2, column 1, row 6).

When we examine how outcomes differ between primary and junior high school students by gender, we find that the math self-confidence of students decreases as they enter junior high schools from primary school. Junior high school girls displayed math self-concept measures 0.30 standard deviations lower that girls in primary school (significant at the 1% level – Table 2, column 2, row 3). In addition, girls in junior high school display measures of math anxiety 0.31 standard deviations higher than girls in primary school (significant at the 1% level – Table 2, column 2, row 6). Our results also show that boys in junior high school have lower levels of math self-concept and higher levels of anxiety than boys in primary school. However, these differences in both math self-concept (-0.19 standard deviations) and math anxiety (0.28 standard deviations) for male students are smaller than the gaps observed for female students (Table 2, column 3, rows 3 and 6).

How do our results in rural China compare with that of urban China? We find that the gender gap in math self-concept in urban students is actually 0.32 standard deviations wider than that of their rural counterparts (Table 3, column 4, rows 1 and 2). This holds true for math anxiety as well, with the urban gender gap 0.31 standard deviations wider than the rural gender gap (Table 3, column 4, rows 4 and 5). We also find that relative to their rural counterparts, urban boys suffer from less anxiety than urban girls, who have much higher levels of math anxiety (Table 3, columns 2 and 3, row 6). In other words, even though urban girls perform just as well in math as urban boys (Figure 1), our results show that girls in urban areas experience even more anxiety in math learning than girls in rural areas.

Surprisingly, on a global scale, our results indicate that the gender gaps in math self-concept and math anxiety in rural China are smaller than those in urban China and many other countries. Comparisons with countries participating in the PISA test show that the gender gap in math self-concept in rural China is smaller than the gap in urban China and the average gap in OECD countries, but larger than the non-OECD average (Figure 2). We also find that the gender gap in math anxiety in rural China is smaller than the gap in urban China, the OECD average, and the non-OECD average (Figure 3). These findings suggest that although there is a large gender gap in math performance among students in rural China, it is not accompanied by similarly large gender gaps in math self-concept and anxiety.

Discussion and Conclusion

In summary, we found that a gender gap exists in the math performance of rural Chinese students and this performance gap widens as students age. Moreover, we show that the size of the gap is wide by national and international standards. When comparing the size of the rural Chinese math performance gap to the gap of Shanghai students measured by the PISA test, we found that the gender gap in math performance among students in rural China is much wider than that in urban China and in most other countries participating in the PISA test. From these findings, we conclude that the gender gap in math performance in rural China merits attention and additional analysis.

Although there is clearly a gap in math performance between male and female students in rural China, to develop interventions that can address this gender gap requires determining how these differences arise. Therefore, we sought to test whether there was any obvious stereotype threat that was being directed at girls. If this were the case, we would expect to see significant gaps in math self-concept and anxiety between boys and girls (Marsh and O'Mara, 2008; Marsh et al., 2006). According to our analysis, we did indeed find that girls experience lower levels of math self-concept and higher levels of anxiety than boys in rural China. However, we do not believe that the gender gaps in math self-concept and math anxiety are wide enough to explain the math performance gap. Compared with the situations in urban China and in other countries, the sizes of the math self-concept and math anxiety gender gaps in rural China are relatively small. For this reason, we believe that stereotype threat cannot fully explain the gender gap in math performance in rural China.

Another possible factor that could drive the math performance gap is discriminatory investments in children (Hannum et al., 2009; Tsui, 2005). The argument is that poor families are more likely to invest scarce resources in the nutrition, education and health of their sons rather than their daughters. To examine whether this drives the gender gap in math performance, we compare the relative math scores of girls and boys from different types of families. If discrimination were playing a role, we would expect to see a larger gender gap among families that display characteristics associated with poverty (such as lower levels of household assets, parents with lower levels of education and children that board at school) compared to families that do not display these characteristics. To examine whether this is the case, we conducted heterogeneous analysis using the model described in Appendix II. According to the findings presented in Appendix III, we find that there are no significant heterogeneous effects in the gender gap among the types of families examined. In other words, our findings show that when families invest in the sons and daughters equally regardless of family income status. Based on these results, it appears that discriminatory family investment practices are not the reason behind the gender gap in math performance.

This leaves an addition possible source. It is possible that it is the teaching practices of math teachers that may be leading to the gender math performance gap. If teachers are exercising their biases that boys are inherently better and girls are inherently worse in math and so teachers spend more time and effort in teaching boys relative to girls, this may explain the observed gap.

Unfortunately, we do not have data that can be used to examine this empirically. The available international literature (as there is no literature from China), however, suggests that rural teacher gender attitudes may be the best candidate to explain the gap. There are several studies in developed countries that document that teachers give more attention and instruction in math to boys than girls (Kelly and Elliott, 1982; Leinhardt et al., 1979; Levy, 1973). For example, in a study of 33 second-grade teachers, Leinhardt et al. (1979) noted that teachers systematically made more academic contact with boys in math than girls. The Gaea paper also found that teachers spent relatively more time on math-related cognitively stimulating activities with boys. Meanwhile, Levy (1973), suggests that, while boys receive more intense stimuli, girls are often ignored or rewarded just for following directions in math classes. Of course, these are studies from other parts of the world and from earlier time periods. We can only speculate that this might be a large source of the current math performance gap in China. Further study is needed to substantiate the causal link between teacher instruction and the gender gap in math performance.

Our results suggest that differential math outcomes between the genders still exist in rural China even though the Chinese government has made efforts to improve education quality and equality in recent years (State Council Information Office of China, 2015). Not only do we find that girls lag behind boys in math performance, but our results also suggest that the gap begins as early as in primary school and becomes more pronounced as students continue their schooling. These findings suggest that government policies hoping to address the math performance gap between the genders would be well served to start early on in the course of schooling. Also, if the government is interested in eliminating this disparity, the solution will likely need to extend beyond building the math self-confidence of girls. More work is likely necessary to determine what other factors cause the gender gap in math performance in rural China.

References:

Ambady, N., Shih, M., Kim, A., and Pittinsky, T. L. 2001. Stereotype Susceptibility in Children: Effects of Identity Activation on Quantitative Performance. Psychological Science, 12(5): 385-390.

Ashcraft, M. H., and Kirk, E. P. 2001. The Relationships Among Working Memory, Math Anxiety, and Performance. Journal of Experimental Psychology: General, 130(2): 224.

Atkinson, R., and Mayo, M. 2010. Refueling the US Innovation Economy: Fresh Approaches to Science, Technology, Engineering and Mathematics (STEM) Education. Executive Summary. Information Technology and Innovation Foundation.

Bandura, A. 1977. Self-efficacy: Toward a Unifying Theory of Behavioral Change. Psychological Review, 84(2): 191.

Barres, B. A. 2006. Does Gender Matter? Nature, 442(7099): 133-136.

Beede, D. N., Julian, T. A., Langdon, D., McKittrick, G., Khan, B., and Doms, M. E. 2011. Women in STEM: A Gender Gap to Innovation. Economics and Statistics Administration Issue Brief, No. 04-11.

Brown, R. P., and Pinel, E. C. 2003. Stigma on My Mind: Individual Differences in The Experience of Stereotype Threat. Journal of Experimental Social Psychology, 39(6): 626-633.

Byrnes, J. P. 2005. Gender Differences in Math. Gender Differences in Mathematics: An Integrative Psychological Approach: 73-98.

Carnevale, A. P., Smith, N., and Melton, M., 2011. STEM: Science Technology Engineering Mathematics. Georgetown University Center on Education and the Workforce. Washington, DC.

Ceci, S. J., Williams, W. M., and Barnett, S. M. 2009. Women's Underrepresentation in Science: Sociocultural and Biological Considerations. Psychological bulletin, 135(2): 218.

China National Bureau of Statistics (CNBS). 2014. China Statistical Yearbook, China Statistics Press, Beijing.

Else-Quest, N. M., Hyde, J. S., and Linn, M. C. 2010. Cross-national Patterns of Gender Differences in Mathematics: A Meta-analysis. Psychological Bulletin, 136(1): 103.

Farenga, S. J., and Joyce, B. A. 1999. Intentions of Young Students to Enroll in Science Courses in the Future: An Examination of Gender Differences. Science Education, 83(1): 55-75.

Fryer, R. G., and Levitt, S. D. 2010. An Empirical Analysis of the Gender Gap in Mathematics. American Economic Journal: Applied Economics, 2(2): 210-240.

Furner, J. M., and Berman, B. T. 2003. Review of Research: Math Anxiety: Overcoming a Major Obstacle to the Improvement of Student Math Performance. Childhood Education, 79(3): 170-174.

Gong, X., Ding, Y., and Tsang, M. C. 2014. Gender Differences of Academic Performance in Compulsory Education in Rural Southwestern China. International Journal of Educational Development, 39: 193-204.

Guiso, L., Monte, F., Sapienza, P., and Zingales, L. 2008. Culture, Gender, and Math. Science-New York then Washington, 320(5880): 1164.

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., and Gernsbacher, M. A. 2007. The Science of Sex Differences in Science and Mathematics. Psychological Science in the Public Interest, 8(1): 1-51.

Hannum, E., Kong, P., and Zhang, Y. 2009. Family Sources of Educational Gender Inequality in Rural China: A Critical Assessment. International Journal of Educational Development, 29(5): 474-486.

Hembree, R. 1990. The Nature, Effects, and Relief of Mathematics Anxiety. Journal for Research in Mathematics Education: 33-46.

Hill, C., Corbett, C., and St Rose, A. 2010. Why so few? Women in Science, Technology, Engineering, and Mathematics. American Association of University Women, Washington, DC.

Keller, J., and Dauenheimer, D. 2003. Stereotype Threat in the Classroom: Dejection Mediates the Disrupting Threat Effect on Women's Math Performance. Personality and Social Psychology Bulletin, 29(3): 371-381.

Kelly, G. P., and Elliott, C. M. 1982. Women's Education in the Third World: Comparative Perspectives. State University of New York Press, Albany, NY.

Kimura, D. 2000. Sex and Cognition. MIT Press, Cambridge, MA.

Lai, F. 2010. Are Boys Left Behind? The Evolution of the Gender Achievement Gap in Beijing's Middle Schools. Economics of Education Review, 29(3): 383-399.

Langdon, D., McKittrick, G., Beede, D., Khan, B., and Doms, M. 2011. STEM: Good Jobs Now and for the Future. ESA Issue Brief No. 03-11. US Department of Commerce.

Leinhardt, G., Seewald, A. M., and Engel, M. 1979. Learning What's Taught: Sex Differences in Instruction. Journal of Educational psychology, 71(4): 432.

Levy, B. 1973. Sex Role Socialization in Schools. In Sex Role Stereotyping in the Schools. Washington, D.C.: National Education Association, 1-7

Ma, X. 1999. A Meta-analysis of the Relationship Between Anxiety Toward Mathematics and Achievement in Mathematics. Journal for Research in Mathematics Education, 520-540.

Marginson, S., Tytler, R., Freeman, B., and Roberts, K. 2013. STEM: Country Comparisons: International Comparisons of Science, Technology, Engineering and Mathematics (STEM) Education. Final Report. Australian Council of Learned Academies.

Marsh, H. W., and Hau, K. T. 2004. Explaining Paradoxical Relations Between Academic Self-concepts and Achievements: Cross-cultural Generalizability of the Internal/external Frame of Reference Predictions Across 26 Countries. Journal of Educational Psychology, 96(1): 56.

Marsh, H. W., Hau, K. T., Artelt, C., Baumert, J., and Peschar, J. L. 2006. OECD's Brief Self-report Measure of Educational Psychology's Most Useful Affective Constructs: Cross-cultural, Psychometric Comparisons Across 25 Countries. International Journal of Testing, 6(4): 311-360.

Marsh, H. W., and O'Mara, A. 2008. Reciprocal Effects Between Academic Self-concept, Self-esteem, Achievement, and Attainment Over Seven Adolescent Years: Unidimensional and Multidimensional Perspectives of Self-concept. Personality and Social Psychology Bulletin, 34(4): 542-552.

McClure, I. 2003. The Essential Difference: Men, Women and the Extreme Male Brain. British Medical Journal, 327(7405): 57.

Ministry of Education (MOE). 2011. Chinese National Curriculum Framework Accessed September 2, 2014, Government of China: http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s8001/201404/xxgk_1673 4 0.html

Nguyen, H. H. D., and Ryan, A. M. 2008. Does Stereotype Threat Affect Test Performance of Minorities and Women? A Meta-analysis of Experimental Evidence. Journal of Applied Psychology, 93(6): 1314.

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., and Kesebir, S. 2009. National Differences in Gender–Science Stereotypes Predict National Sex Differences in Science and Math Achievement. Proceedings of the National Academy of Sciences, 106(26): 10593-10597.

OECD. 2012. Education at a Glance 2012: OECD Indicators, OECD Publishing. http://dx.doi.org/10.1787/eag-2012-en

OECD. 2014. PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science, PISA, OECD Publishing. http://dx.doi.org/10.1787/9789264201118-en

OECD. 2015. The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence, PISA, OECD Publishing. http://dx.doi.org/10.1787/9789264229945-en

O'Neill, J. 2003. The Gender Gap in Wages, Circa 2000. The American Economic Review, 93(2): 309-314.

Osborne, J. W. 2001. Testing Stereotype Threat: Does Anxiety Explain Race and Sex Differences in Achievement?. Contemporary Educational Psychology, 26(3): 291-310.

Osborne, J., and Dillon, J. 2008. Science Education in Europe: Critical reflections (Vol. 13). London: The Nuffield Foundation.

Pajares, F., and Miller, M. D. 1994. Role of Self-efficacy and Self-concept Beliefs in Mathematical Problem Solving: A Path Analysis. Journal of educational psychology, 86(2): 193.

Pietsch, J., Walker, R., and Chapman, E. 2003. The Relationship Among Self-concept, Self-efficacy, and Performance in Mathematics During Secondary School. Journal of Educational Psychology, 95(3): 589.

State Council Information Office of China, 2015, Gender Equality and Women's Development in China. URL: http://www.scio.gov.cn/zfbps/ndhf/2015/Document/1449896/1449896.htm

Steele, C. M., and Aronson, J. 1995. Stereotype Threat and the Intellectual Test Performance of African Americans. Journal of Personality and Social Psychology, 69(5): 797.

Tsui, M. 2005. Family Income, Home Environment, Parenting, and Mathematics Achievement of Children in China and the United States. Education and Urban Society, 37(3): 336-355.

Turner, R. 1994. Sex Differences in Mathematical Performance Among Chinese Middle School Students. In Proceedings of the 17th Annual Conference of the Mathematics Education Research Group of Australasia (MERGA): Challenges in Mathematics Education: Constraints on Construction (Vol. 1, No. 2).

Wang, L., Min, W. 2016, Cognitive Learning Outcomes in Rural China: Comparisons Across Provinces and Among Subgroups of Children, Working Paper, The Rural Education Action Program, Freeman Spogli Institute for Foreign Policy, University of Stanford, Stanford, CA.

Xie, Y., Shauman, K. A., and Shauman, K. A. 2003. Women in Science: Career Processes and Outcomes (Vol. 26, No. 73.4). Harvard University Press, Cambridge, MA.

Table 1. Math performance of students using standardized math test in rural China, by gender

	Full sample		Girls		Boys		Difference:(3)-(2)	
	(1)		(2)		(3)		(4)	
	Dataset A: primary school students							
	Mean	SD	Mean	SD	Mean	SD		
1. Standardized math test score (SD)	0.00	(1.00)	-0.04	(0.98)	0.04	(1.01)	0.08**	(0.03)
2. Number of observations	3789		1791		1998			
	Dataset B: junior high school students							
	Mean	SD	Mean	SD	Mean	SD		
3. Standardized math test score (SD)	-0.00	(1.00)	-0.09	(0.97)	0.09	(1.02)	0.18***	(0.02)
4. Number of observations	12702		6256		6446			

Source: Authors' survey. *** p<0.01, ** p<0.05, * p<0.1

Table 2. Math self-concept and math anxiety of students in rural China, by gender

	Full sample		Girls		Boys		Difference:(3)-(2)
	(1)		(2)		(3)		(4)
	Math self-concept index (SD)						
	Mean	SD	Mean	SD	Mean	SD	
1. Primary school	0.19	(0.95)	0.08	(0.98)	0.30	(0.92)	0.21*** (0.03)
2. Junior high school	-0.06	(1.01)	-0.22	(0.98)	0.10	(1.00)	0.32*** (0.02)
3. Difference: junior high school - primary school	-0.25*** (0.02)		-0.30*** (0.03)		-0.19*** (0.03)		
	Math anxiety index (SD)						
	Mean	SD	Mean	SD	Mean	SD	
4. Primary school	-0.23	(0.97)	-0.18	(0.98)	-0.27	(0.96)	-0.09 *** (0.03)
5. Junior high school	0.07	(1.00)	0.13	(0.99)	0.01	(1.00)	-0.12*** (0.02)
6. Difference: junior high school - primary school	0.29*** (0.02)		0.31*** (0.03)		0.28*** (0.03)		

Source: Authors' survey. *** p<0.01, ** p<0.05, * p<0.1

Note: When comparing math self-concept/math anxiety across primary school with junior high school data, we appended the rural China dataset 1 with 2, reconstructed an index of math self-concept/ math anxiety with that sample, and standardized results to have a mean of 0 and a standard deviation of 1. In this case, positive values on the math self-concept index indicate that the student reported higher levels of math self-concept than the average student, which means he/she has more confidence in his/her math ability. The higher the math self-concept index is, the more confidence in math the student has. On the other hand, positive math anxiety scores indicate that student was prone to having higher levels of math anxiety than the average student, which means that a student suffered from more stress when dealing with math. The lower the math anxiety index is, the less stress the student suffers.

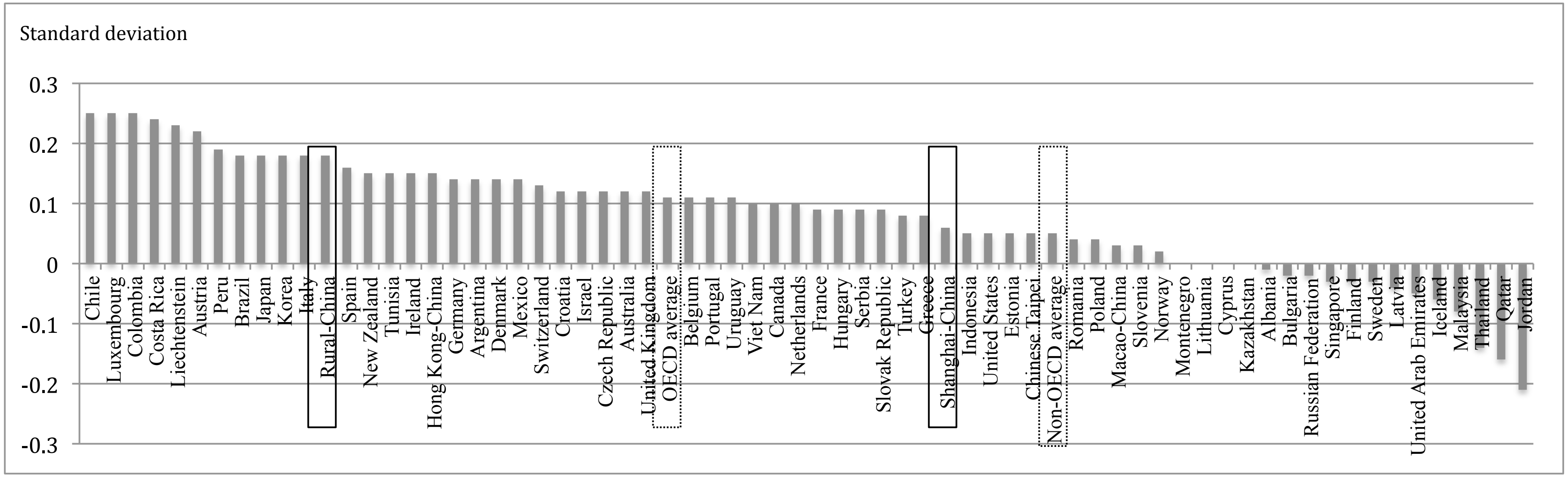
Table 3. Math self-concept and math anxiety of junior high school students in China, urban vs. rural, by gender

	Full sample		Girls		Boys		Difference:(3)-(2)	
	(1)		(2)		(3)		(4)	
	Math self-concept index (SD)							
	Mean	SD	Mean	SD	Mean	SD		
1. Rural	0.03	(0.96)	-0.12	(0.94)	0.18	(0.96)	0.30***	(0.02)
2. Urban	-0.13	(1.12)	-0.43	(1.03)	0.19	(1.11)	0.62***	(0.04)
3. Difference: Rural - Urban	0.16***	(0.02)	0.31***	(0.03)	-0.00	(0.03)		
	Math anxiety index (SD)							
	Mean	SD	Mean	SD	Mean	SD		
4. Rural	0.02	(0.97)	0.08	(0.96)	-0.03	(0.98)	-0.12 ***	(0.02)
5. Urban	-0.09	(1.09)	0.12	(1.07)	-0.31	(1.08)	-0.43***	(0.04)
6. Difference: Rural - Urban	0.11***	(0.02)	-0.04	(0.03)	0.27***	(0.03)		

Source: Urban China data is from OECD, 2012; rural China data is from author survey. *** p<0.01, ** p<0.05, * p<0.1

Note: When comparing math self-concept/math anxiety across rural China with urban China data, the indexes are different from those in table 2. We renormalized the indexes of math self-concept/math anxiety by appending the rural China dataset 2 with Shanghai data from PISA 2012 dataset, reconstructing an index of math self-concept/ math anxiety with that sample, and standardizing results to have a mean of 0 and a standard deviation of 1. In this case, positive values on the math self-concept index indicate that the student reported higher levels of math self-concept than the average student, which means he/she had more confidence in his/her math ability. The higher the math self-concept index is, the more confidence in math the student has. On the other hand, positive math anxiety scores indicate that student was prone to having higher levels of math anxiety than the average student, which means that a student suffered from more stress when dealing with math. The lower the math anxiety index is, the less stress the student suffers.

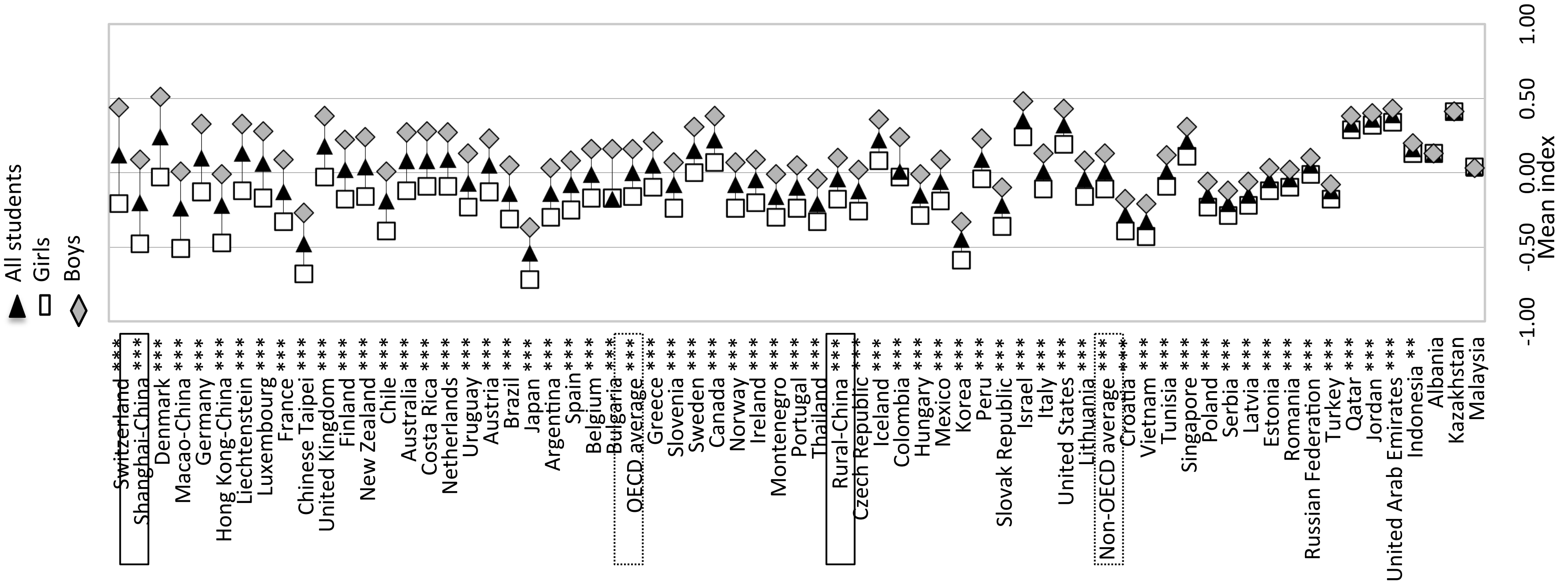
Figure 1. Gender gap in math performance using standardized math test, across countries/economies



Source: All data except for rural China is from OECD, 2012; the data used for rural China is from author survey.

Note: Countries and economies are ranked in ascending order of gender difference (boys – girls) in standardized math test score.

Figure 2. Math self-concept of students in junior high school, by countries/economies, by gender



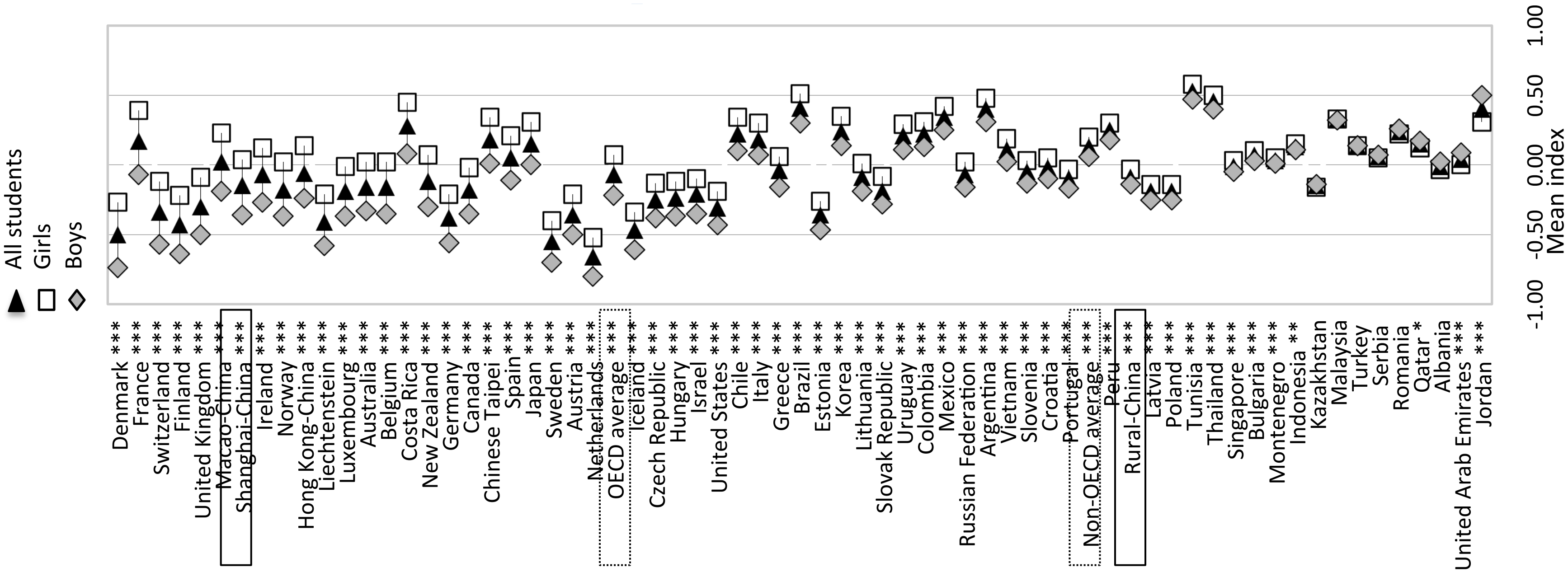
Source: All data except for rural China is from OECD, 2012; the data used for rural China is from author survey.

Note: a. Countries/economies where gender differences are significant are indicated with asterisk. *** p<0.01, ** p<0.05, * p<0.1

b. Countries and economies are ranked in descending order of gender differences on the index of math self-concept.

c. When comparing math self-concept across rural China with other countries and economies data, the index is different from those of table 2 and 3. We renormalized the index of math self-concept by appending the rural China dataset 2 with PISA 2012 dataset, reconstructing an index of math self-concept with that sample, and standardizing results to have a mean of 0 and a standard deviation of 1. In this case, positive values on the math self-concept index indicate that the student reported higher levels of math self-concept than the average student, which means he/she has more confidence in his/her math ability. The higher the math self-concept index is, the more confidence in math the student has.

Figure 3. Math anxiety of students in junior high school, by countries/economies, by gender



Source: All data except for rural China is from OECD, 2012; the data used for rural China is from author survey.

Note: a. Countries/economies where gender differences are significant are indicated with asterisk. *** p<0.01, ** p<0.05, * p<0.1

b. Countries and economies are ranked in ascending order of gender differences on the index of math anxiety.

c. When comparing math anxiety across rural China with other countries and economies data, the index is different from those of table 2 and 3. We renormalized the index of math anxiety by appending the rural China dataset 2 with PISA 2012 dataset, reconstructing an index of math anxiety with that sample, and standardizing results to have a mean of 0 and a standard deviation of 1. In this case, positive math anxiety scores indicate that student was prone to having higher levels of math anxiety than the average student, which means that a student suffered from more stress when dealing with math. The lower the math anxiety index is, the less stress the student suffers.

Appendix I. Description of student characteristics in rural China

Variable	Obs	Mean	SD	Min
Dataset A: primary school students				
1. Student age (year)	3789	10.8	0.96	9
2. Female student (1=yes, 0=no)	3789	0.47	0.50	0
3. Student lives at school (1=yes, 0=no)	3789	0.21	0.41	0
4. Mother completed junior high school (1=yes, 0=no)	3789	0.37	0.48	0
5. Father completed junior high school (1=yes, 0=no)	3789	0.56	0.50	0
6. Family asset (1=higher than median, 0=lower/equal to median)	3789	0.49	0.50	0
Dataset B: junior high school students				
7. Student age (year)	12702	13.2	0.96	9.92
8. Female student (1=yes, 0=no)	12702	0.49	0.50	0
9. Student lives at school (1=yes, 0=no)	12702	0.58	0.49	0
10. Mother completed junior high school (1=yes, 0=no)	12702	0.05	0.23	-0.0[illegible]
11. Father completed junior high school (1=yes, 0=no)	12702	0.09	0.28	-0.0[illegible]
12. Family asset (1=higher than median, 0=lower/equal to median)	12702	0.59	0.49	0

Source: Authors' survey

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

Note: The variable of family asset is based on the summed value of a set of assets, including e appliances, livestock, vehicles, etc. The variable equals 1 if the family asset value is higher th median value and it equals 0 if otherwise.

Appendix II. OLS model for examining heterogeneous effects of math performance on female students in junior high school

Our heterogeneous analysis is based on a basic model:

$$Y_i = \beta_0 + \beta_a X_i + \varphi_p + \varepsilon_i \qquad (1)$$

(1)

where Y_i represents the standardized math test score of student *i*; X_i represents a vector of student characteristics, including *female student (1=yes, 0=no), student lives at school (1=yes, 0=no), mother completed junior high school (1=yes, 0=no)*, *father completed junior high school (1=yes, 0=no),* and *family asset (1=higher than the median, 0=lower than or equal to the median*). It also includes a *school fixed effect*, φ_p . ε_i is a random error term.

To examine the heterogeneity in math performance on certain subgroups more than others, we added interaction terms between the *female student* and a set of key variables (*family asset, student lives at school, mother completed junior high school* and *father completed junior high school*) into the basic model. In all regressions, we included a school fixed effect.

Appendix III. OLS regression results showing the heterogeneous effects of math performance on female students in junior high school

Dependent variable: Student math score (SD)	(1)	(2)	(3)	(4)
1. Female (1=yes, 0=no)	-0.165***	-0.17***	-0.181***	-0.177***
	(0.025)	(0.024)	(0.016)	(0.016)
2. Female* Family asset	-0.024			
	(0.032)			
3.Female* Student lives at school		-0.017		
		(0.032)		
4. Female* Mother completed junior high school			0.031	
			(0.072)	
5. Female* Father completed junior high school				-0.029
				(0.057)
6. Family asset (1=higher than median, 0=lower/equal to median)	0.040*	0.028	0.028	0.028
	(0.024)	(0.018)	(0.018)	(0.018)
7. Student lives at school (1=yes, 0=no)	-0.030	-0.021	-0.030	-0.030
	(0.021)	(0.027)	(0.021)	(0.021)
8. Mother completed junior high school (1=yes, 0=no)	-0.002	-0.002	-0.017	-0.002
	(0.038)	(0.038)	(0.053)	(0.038)
9. Father completed junior high school (1=yes, 0=no) [a]	0.183***	0.184***	0.184***	0.197***
	(0.030)	(0.030)	(0.030)	(0.041)
School dummy	YES	YES	YES	YES
Constant	3.141***	3.144***	3.149***	3.146***
	(0.117)	(0.117)	(0.116)	(0.116)
Observations	12,702	12,702	12,702	12,702
R-squared	0.252	0.252	0.252	0.252

Source: Authors' survey

Notes:

[a] The variable of family asset is based on the summed value of a set of assets, including electric appliances, livestock, vehicles etc. The variable equals 1 if the family asset value is higher than the median value and it equals 0 if otherwise.

Cluster-robust standard errors adjusted for clustering at the school level in parentheses.

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