SOEP The German Socio-Economic Panel study

993 2018

SOEPpapers
on Multidisciplinary Panel Data Research

SOEP — The German Socio-Economic Panel study at DIW Berlin 993-2018

Math, Girls and Socialism

Quentin Lippmann and Claudia Senik

DIW SOEP

SOEPpapers on Multidisciplinary Panel Data Research at DIW Berlin

This series presents research findings based either directly on data from the German Socio-Economic Panel study (SOEP) or using SOEP data as part of an internationally comparable data set (e.g. CNEF, ECHP, LIS, LWS, CHER/PACO). SOEP is a truly multidisciplinary household panel study covering a wide range of social and behavioral sciences: economics, sociology, psychology, survey methodology, econometrics and applied statistics, educational science, political science, public health, behavioral genetics, demography, geography, and sport science.

The decision to publish a submission in SOEPpapers is made by a board of editors chosen by the DIW Berlin to represent the wide range of disciplines covered by SOEP. There is no external referee process and papers are either accepted or rejected without revision. Papers appear in this series as works in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be requested from the author directly.

Any opinions expressed in this series are those of the author(s) and not those of DIW Berlin. Research disseminated by DIW Berlin may include views on public policy issues, but the institute itself takes no institutional policy positions.

The SOEPpapers are available at
http://www.diw.de/soeppapers

Editors:
Jan **Goebel** (Spatial Economics)
Stefan **Liebig** (Sociology)
David **Richter** (Psychology)
Carsten **Schröder** (Public Economics)
Jürgen **Schupp** (Sociology)

Conchita **D'Ambrosio** (Public Economics, DIW Research Fellow)
Denis **Gerstorf** (Psychology, DIW Research Fellow)
Elke **Holst** (Gender Studies, DIW Research Director)
Martin **Kroh** (Political Science, Survey Methodology)
Jörg-Peter **Schräpler** (Survey Methodology, DIW Research Fellow)
Thomas **Siedler** (Empirical Economics, DIW Research Fellow)
C. Katharina **Spieß** (Education and Family Economics)
Gert G. **Wagner** (Social Sciences)

ISSN: 1864-6689 (online)

German Socio-Economic Panel (SOEP)
DIW Berlin
Mohrenstrasse 58
10117 Berlin, Germany

Contact: soeppapers@diw.de

Math, Girls and Socialism

Quentin Lippmann* and Claudia Senik†

June 11th, 2018

Abstract

This paper argues that the socialist episode in East Germany, which constituted a radical experiment in gender equality in the labor market and other instances, has left persistent tracks on gender norms. We focus on one of the most resilient and pervasive gender gaps in modern societies: mathematics. Using the German division as a natural experiment, we show that the underperformance of girls in math is sharply reduced in the regions of the former GDR, in contrast with those of the former FRG. We show that this East-West difference is due to girls' attitudes, confidence and competitiveness in math, and not to other confounding factors, such as the difference in economic conditions or teaching styles across the former political border. We also provide illustrative evidence that the gender gap in math is smaller in European countries that used to be part of the Soviet bloc, as opposed to the rest of Europe. The lesson is twofold: (1) a large part of the pervasive gender gap in math is due to social stereotypes; (2) institutions can durably modify these stereotypes.

JEL codes: I2, J16, J24, P36, Z13

Keywords: Gender Gap in Math, Institutions, German Division, Gender Stereotypes.

*Paris School of Economics, qlippmann@gmail.com

†Sorbonne University and Paris School of Economics, senik@pse.ens.fr

‡We are grateful to participants at the Milano MILLS workshop, the Seventh ECINEQ Meeting, the Kiev « People Matter » conference, and seminars at the Paris School of Economics, Berlin Frei Universität, Magdeburg University, and bar Ilan University for useful comments. We thank CEPREMAP and the French National Research Agency, through the program Investissements d'Avenir, ANR-10-LABX-93-01, for financial support.

1 Introduction

Since the 1980's, girls have started to reverse their initial disadvantage in educational investment (Goldin et al. 2006, Goldin 2014, Kane & Mertz 2012, Autor & Wasserman 2013, Fortin et al. 2015), and they now account for a disproportionate share of "the worldwide boom in higher education" (Becker et al. 2010). However, they still accumulate a specific mix of human capital, neglecting mathematics-intensive fields (Ceci et al. 2014, Blau & Kahn 2017). The broad picture is that girls have closed the gap and conquered most of the avenues to professional success, such as business, medicine, law and biology, not talking about their traditional and intact advantage in reading and literature (Fryer & Levitt 2010), but they stall at the door of math-based curricula and occupations.

Beyond being intriguing, this resilient male advantage in math bears important implications in terms of well-being and quality of life, as mathematics are generally associated with higher earnings (Altonji 1995, Altonji et al. 2012) and more prestigious occupations (Blau & Kahn 2017). This could simply be because math training enhances cognitive and non-cognitive skills, such as clarity in expressions, logical reasoning and inference (Joensen & Nielsen 2009, Arcidiacono 2004). It could also be due to the increasing value of math skills in a period of rapid math-intensive technological change. As a consequence, the number of math-skilled people in the labor force is a positive ingredient for growth, as illustrated by Kimko & Hanushek (2000). Hence, both equity and efficiency motives plead for understanding and reducing the gender gap in math.

A large literature has examined whether girls' lower appetence and performance in math is the outcome of natural brain-based skills or of social stereotypes (on behalf of pupils, parents and teachers). This paper argues that a large part of the gap is indeed due to social norms and stereotypes, which can be undone by institutions. We use the German division and reunification as a natural experiment that offers the possibility to study the lasting effect of highly gender-equal institutions. In view of its ambitious growth objectives (and maybe its egalitarian ideology), the former GDR, as alike other socialist countries, made employment a universal right, but also a duty, for women as well as men, and adopted a host of accompanying measures ensuring the compatibility between fertility and employment. These gender-equal policies reflected on the work values of women, and, more generally, modified the conception of gender roles (Campa & Serafinelli 2016, Lippmann et al. 2016).

We first show that the gender gap in mathematics is smaller in East Germany as compared to West Germany. This is true both in self-declared math grades but also in standardized tests. Then, we establish that this smaller gap is accompanied by different attitudes towards mathematics. In particular, girls in the East feel less anxious and more confident about their aptitude in math than their counterparts from West Germany. They also are more competitive, especially at intermediate levels of performance. Importantly, this higher performance of girls in math does not come at the price of a lower performance of girls in reading (their traditional advantage). Finally, we generalize our results to all European countries by contrasting former socialist "Eastern" countries to capitalist "Western" countries; we uncover a similar picture: the gender gap in math is much smaller, and even sometimes inexistent, in Eastern countries.

Most of our findings are based on the PISA, the OECD Program for International Student Assessment. To assess the robustness of our results, we also use the German Socio-Economic Panel. We also provide substantive evidence that the reduction in the gender gap is not due to differences in teaching methods, organization of classes or math workloads in Eastern versus Western schools. Finally, we check that other structural differences, such as poorer economic conditions in the Eastern regions, do not affect the gender gap.

Our paper is related to a large literature, which has tried to determine whether gender differences in mathematics are innate, e.g. brain-based, or acquired. Halpern et al. (2007) provides an exhaustive review of the large scientific literature dedicated to the brain-based conjecture, but finds it to be inconclusive, as experience can alter brain structure and functioning. One particular nature-based explanation of the preponderant presence of boys at high achievement levels in math is the so-called "male greater variability hypothesis". This hypothesis has been tested many times but no consensus has been reached: some studies found it consistent with data on American students (Benbow & Stanley 1980), while other found it implausible, based on PISA and TIMSS (Kane & Mertz 2012) or IMO data (Hyde & Mertz 2009).

In a totally different spirit, the gender gap in math can be explained as a rational response to the unequal opportunities offered to boys and girls. In countries where women are bound to spend shorter years on the labor market, with interruptions due to pregnancy and childcare, leading to lower-profile careers, they naturally opt for less ambitious, less competitive, and thus less rewarding, education tracks. They also chose fields that are

related to the type of activities they expect to be exerting in the future, i.e. relational, caregiving or educational, hence their taste for language, psychology, healthcare, etc. As suggested by the economics of identity and culture, rational expectations can also be reinforced by social roles, which become sources of utility per se (Akerlof & Kranton 2000, 2013). Norms may be self-sustained as they create expectations that influence educational choices, which, in turn, contribute to the dynamic persistence of stereotypes (see Altonji et al. 2012 for instance). Cultural economics have shown how such beliefs and attitudes could persist over time and across generations (Fernandez & Fogli 2009, Fernandez 2010, Blau et al. 2013, Bisin & Verdier 2001, 2010). Other behavioral motives such as girls' self-confidence, biased priors about their chances of success and lower appetence for competition, in general or against boys (Gneezy et al. 2003, Niederle & Vesterlund 2010), lie somewhere between the expectations and the cultural hypotheses.

Beyond consequences on educational choices, such stereotypes have important implications in terms of well-being as they provide strong ground for statistical discrimination against women. Employer's beliefs about women's mathematical abilities are likely to hinder female careers (Reuben et al. 2014 and Moss-Racusin et al. 2012) and directly impact their mental and physical health (Pascoe & Richman 2009 and Schmitt et al. 2014).

This paper belongs in the second line of interpretation that posits that the gender gap in math is sustained by stereotypes. It argues that institutions can durably shape expectations, social norms and gender roles, that eventually result in a gender gap in math. We directly contribute to the the empirical evidence supporting this conjecture.

Various measures of economic and social gender inequality have been shown to correlate with the size of the gender gap in math and science, as measured by PISA scores (Fortin et al. 2015, Guiso et al. 2008, Nollenberger et al. 2016), TIMSS scores (Baker & Jones 1993), IMO data (Hyde & Mertz 2009), or all of these measures together (Kane & Mertz 2012, Ellison & Swanson 2010), as well as American data (Pope & Sydnor 2010). Some studies have also documented the association between stereotypes about gender roles and the width of the gender gap in math across American states (Pope & Sydnor 2010, Else-Quest et al. 2010) or across Spanish regions (Gonzalez de San Roman & de la Rica 2012). Our paper is closer to that of Schnepf (2007) and Amini & Commander (2012) who noted the smaller gender gap in education that prevails in Central and Eastern European countries. We add to this literature by providing evidence of the causal influence of socialism. We also analyze the gender gap in some types of competitions that have not been explored

before.

This paper is far from being the first attempt to use the German division as a "natural experiment". Before us, some articles have illustrated the smaller gender gap in East Germany, as compared with West Germany, in terms of household behavior (Kuenzler et al. 2001, Cooke 2004, 2007), self-reported work preferences and beliefs about gender role (Breen & Cooke 2005, Bauernschuster & Rainer 2012, Gorges & Beblo 2015, Campa & Serafinelli 2016, Lippmann et al. 2016). Other papers have documented the lasting (and sometimes progressively withering) effect of East German institutions on mentalities (Alesina & Fuchs-Schündeln 2007, Rainer & Siedler 2009). This paper adds to this literature, by focusing on the gender gap in mathematics.

The rest of this work is organized as follows. Section 2 describes the socialists episode in East Germany. Section 3 describes the data and the empirical strategy. Section 4 presents the results. Section 5 demonstrates the robustness of our results. Section 6 enlarges the results to the rest of Europe and to two supplementary international competitions. Finally, section 7 concludes.

2 The Socialist Episode in East Germany

In 1949, Germany was divided in two parts based on the occupation zones of Western forces for the Federal Republic of Germany (FRG) and the Soviet army for the German Democratic Republic (GDR). Once separated, the economic and political regimes of the two countries rapidly departed. The GDR immediately settled institutions in favor of gender equality in education and occupations. This was not only for ideological reasons, but also, and perhaps mostly, in view of the industrialization objective, in particular, in the East German context of wartime destruction, postwar labor shortage and mass exodus from the country. The policy enacted by the socialist party explicitly aimed at three objectives: (1) the realization of legal equality between men and women, (2) the promotion of women's work, and (3) special protection of mothers and children. The 1965 Family code stated that *"relationships between spouses must be organized in a way that allows the wife to conciliate maternity with her professional and social activities"* (Richter 2014). Work-family compatible policies were put in place, included kindergartens and other childcare facilities (often within firms), generous maternity leaves, paid days-off for doing housework, exclusion from dangerous or strenuous jobs, and promotion of women's

enrollment in factories and other productive units (Cooke 2007). In the 1970s and 1980s, the "mothers' policy" (*Mutti-Politik*) made women's work hours more flexible, gave them more day-offs, holidays and paid leaves, preserved some jobs and some places for them in universities, and introduced the free "desired child pill" (*Wunschkindpill*) in order to allow them to plan the time of childbirth.

In the meantime, the FRG strengthened the traditional male-breadwinner model. Irregular school schedules and scarce childcare facilities inhibited female employment (Cooke 2007). The tax system favored single earner families as unemployed spouses and children could get public health insurance at no extra cost. Until 1977, the Marriage and Family law stated that: "*The wife is responsible for running the household. She has the right to be employed as far as this is compatible with her marriage and family duties*" (Rheinstein & Glendon 1978). Subsequent policies then alternated more or less conservative incentives for female participation in the labor market.

The FRG and GDR also radically diverged as concerns the representations of the ideal womanly models, as illustrated by newspapers, movies, television programs, posters and advertisements (Richter 2014). In the GDR, Clara Zetkine, whose portrait was represented on 10 DDM (Ostmark) bills, incarnated the ideal model of a feminist, politically engaged and influent woman. Female stereotypes represented in the *Neue Berliner Illustrierte* between 1949 and 1989 were professionally active and "emancipated" women, working as journalists, professors, brigadiers or factory workers, participating in the construction of socialism. By contrast, in West Germany, the female stereotype was a perfect and modern housewife and mother, inspired by the American standard of the 1950s (Richter 2014).

As a result of these very different ideologies and policies, the rates of female labor market participation rapidly diverged. At the end of the division episode, in 1990, women's labor market participation rate in East Germany was about 89% compared to 92% for men, whereas in West Germany, only 56% of women were in the labor force, as compared to 83% of men (German Statistical yearbooks 1933-1991, Rosenfeld et al. 2004). Ten years later (in 2000), the labor force participation was still approximately the same across gender in the former GDR (around 80%), whereas the gap remained wide in West Germany, with 65% of women in the labor force against 81% of men (Schenk 2003). These objective differences are matched by opinions regarding gender roles (Bauernschuster & Rainer 2012, Campa & Serafinelli 2016).

Our main argument is that equality on the labor market has changed East German

women's expectations, hence their educational attitudes, but we do not exclude the possibility that some differences in the school systems of East versus West Germany also played a role in shaping persistent gender attitudes towards math and science curricula. Traditionally, the German education system was managed at the Länder level. Several unification policies took place under the Prussian empire, the Weimar republic, and the third Reich. During the division, Western Länder recovered their autonomy over schools' curricula, whereas, in the GDR, education was standardized over the entire territory (Schnepf 2007). A unique education track was compulsory until Grade 10, without any formal differentiation between boys and girls (Campa & Serafinelli 2016). By contrast, in West Germany, boys and girls followed different curricula until the 1970s. Even after this was abolished, the allocation of pupils into different types of schools (*Hauptschule, Realschule, Gymnasium*) after grade 4 allowed for more differentiation between boys and girls. In terms of contents, Fuchs-Schundeln & Masella (2016) indicate that "*the teaching of mathematics was of similar importance in East and West, even though GDR schools devoted significantly more time to natural sciences, while FRG schools devoted more time to teaching of "softer" subjects like foreign languages, sports, arts and music, and religious education*".

After the German reunification, the Western system spread to the Eastern part. About 20 percent of teachers of the GDR were dismissed. Curricula became more flexible (Fuchs-Schundeln & Masella 2016). Regional differentiation may have appeared, given the greater autonomy given by the Federal system. In spite of these changes, we cannot exclude a priori that the socialist episode has left some persisting legacy on the gender gap in math. However, as we will show, the information collected by PISA about current school' s teaching styles, in particular in math, reveals no statistically significant difference between Eastern and Western schools for that matter.

3 Data and Empirical Strategy

3.1 Two Datasets

To investigate the consequences of the socialist episode in Germany, we use the German Socio-Economic Panel (GSOEP) and the PISA international evaluation scores of 15 years old pupils.

3.1.1 GSOEP

We first use self-stated math grades contained in the German Socio-Economic Panel, a longitudinal survey run by the German Institute for Economic Research (DIW, Berlin). This survey was started in 1984 in West Germany and was extended to East Germany after the reunification. We use 22 waves, from 1991 to 2012. A question about their last grade in math is asked to adults upon their entry in the panel, and to 17 years old teenagers: *Can you remember your last report card? What grade did you have in mathematics?* (1-6 scale with 1 being the highest grade). For adults, we restrict the sample to individuals who were born before 1971, hence completed their education before 1990; we define a dummy East that code 1 for individuals who, in 1991, declared that they used to live in East Germany before 1990 (and 0 otherwise). For teenagers, East is a dummy that codes 1 for individuals of German nationality who reside on the territory of the former GDR at the time they are surveyed. We drop people who were born outside Germany. We also drop Berlin from the sample, as the data does not allow distinguishing East Berlin from West Berlin.

Descriptive statistics (Table A3) show that the self-reported grades in mathematics are generally higher for adults as compared to teenagers but lower for girls as compared to boys. The gender gap is smaller in East Germany. For adults, 55% of East German men report a high grade in math (below 3 on a 1-6 scale), against 53% of East German women. In West Germany, these proportions are of 49% for men against 42% of women. Among teenagers, the pattern is similar. 34% of West German boys report a high grade against 30% of girls. But in East Germany, the gender gap in self-reported grade is reversed, as 35% of boys report a high grade, against 39% of girls.

3.1.2 PISA

To investigate objective tests of performance, we rely on the Programme for International Student Assessment (PISA) survey[1], launched in 2000, in order to measure (and monitor) the outcomes of OECD countries' education systems in terms of student achievement, within a common international framework. It assesses the Scientific, Reading and Mathematical Literacy of 15 years old students (more precisely from 15 years 3 months to 16 years 2 months). The survey takes place every three years.

Paper-and-pencil tests are used, with assessments lasting a total of two hours for each

[1] https://www.oecd.org/pisa/

student. Test items are a mixture of multiple-choice items and questions requiring students to construct their own responses. A total of about 390 minutes of test items is covered, with different students taking different combinations of test items. Students answer a background questionnaire, which takes 30 minutes to complete, providing information about themselves and their homes. School principals are given a 20-minute questionnaire about their school.

The international survey is typically administered to a range of 4 500 to 10 000 students in each country. In addition to the international survey, countries have the possibility to run PISA on a larger national sample (PISA-E). Because we need information about the localization of students, we use the national German PISA-E survey. Each of PISA cycles looks in depth at a major domain, to which two-thirds of the testing time is devoted. Major domains were: *reading literacy* in 2000, *mathematical literacy* in 2003 and *scientific literacy* in 2006. We thus focus on PISA-E 2003 Germany, which is dedicated to math.

The mathematics section of PISA includes four content domains: Quantity (which assesses the understanding of numeric phenomena, quantitative relationships, and patterns), Space/Shape (spatial and geometric phenomena and relationships), Change/Relationships (mathematical manifestations of change, functional relationships, and dependency among variables), and Uncertainty (probabilities and statistics).

We restrict the sample to teenagers who were born in Germany. We also drop Berlin from the sample as it is not possible to distinguish East Berlin from West Berlin. Simple descriptive statistics (Table A1) show that the scores are generally lower for girls than for boys, but higher in East Germany than in West Germany, and that the gender gap is smaller in East Germany.

In addition to objective tests of performance, PISA includes a student questionnaire and a school questionnaire. The school questionnaire asks the director a series of questions concerning teaching styles in mathematics. The student questionnaire elicits a host of attitudes to math, spanning motivation, self-confidence, stress, pleasure, as well as competitiveness in math (see Appendix B.2). It also asks pupils about the time spent on math in their class. Finally, it asks pupils to indicate the last grade they obtained in mathematics in their last school report. Since these grades are self-declared, we checked using the PISA dataset, that they are correlated with objective test scores. In Table A2, we show that objective test scores are indeed higher for people with higher self-reported grade, which suggests that the latter are reliable.

3.2 Empirical Strategy

3.2.1 Econometric Specification

We are interested in the gender gap in math scores (or math attitudes) and in the influence of the socialist episode on this gap. Hence, all of our estimates regress the considered math score of individuals (indexed by subscript i) on gender, a dummy for East Germany (as opposed to West), and the interaction of the two later terms, controlling for socio-demographic variables and the relevant contextual magnitudes. The typical form of our estimates is described by equation 1:

$$MathScore_i = \gamma_1 Female_i + \gamma_2 Female_i * East_i + \gamma_3 East_i + \beta X_i + \epsilon_i \quad (1)$$

Where i is the subscript for an individual. We expect the coefficient γ_1 to be negative, reflecting the usual disadvantage of girls in math; we have no prior on the coefficient γ_3 and our test is about the sign of γ_2, i.e. the specific gender gap in math in East Germany.

In estimates based on PISA, the controls X_i include a quadratic in age, the pupil's class level, a quadratic in wealth, education, employment status of both parents, size of the location area (1: village, hamlet or rural area (fewer than 3 000 people); 2: small town (3 000 to 15 000 people); 3: town (15 000 to 100 000 people); 4: city (100 000 to 1 000 000 people); 5: large city (over 1 000 000 people), the share of girls in the school and Länder fixed-effects. As education is under the responsibility of the Länder, these fixed-effects are crucial to control for potential differences in school system.

In regressions that use GSOEP, the controls include age, household income and Länder fixed-effects. We run linear probability models of declaring a good grade, i.e. strictly below 3 on a 1-6 scale (where 1 is the best grade and 6 the worst). We cluster the standard errors at the household level.

3.2.2 Validity Tests - Pre-Division Differences

The identification of the influence of the socialist period in East Germany rests on the assumption that Eastern and Western regions were identical before the division. It is now well established that the output and employment structure, as well as the rate of female labor market participation were similar in the regions of East and West Germany before the

division (Alesina & Fuchs-Schündeln 2007, Bauernschuster & Rainer 2012, Schenk 2003, Lippmann et al. 2016). We collected data from statistical yearbooks before the division in 1933 to provide substantive evidence on this matter. Table 1 describes the differences between Eastern and Western regions in 1933. As we can see from column 1, 44.64% of the worforce was in the industrial sector in East Germany, against 40.08% in West Germany. There was little difference in terms of industrial structure. Looking at indicators of gender gap, we see that the female share of employees and the female share of high school students were respectively of 35.92% and 31.61% in West Germany, against 33.12% and 34.87% in East Germany. We then tested whether these differences, between East and West Germany, are similar to the average regional differences in Germany at that time. To do so, we run a permutation test comparing the actual differences to the mean of all the possible regional differences. Column 4 displays the average regional differences and column 5 the p-value from a Fisher exact test. It shows that, in 1933, the structural dissimilarities between the two regions that would later become East Germany and West Germany (excluding Berlin) were not any different from what would stem form a random division of the 15 Länder into two groups of 5 + 10 Länder.

Table 1: East and West Germany in 1933. Descriptive Statistics

	$East$	$West$	$\lvert East - West \rvert$	Average Regional Differences	p-value
	(1)	(2)	(3)	(4)	(5)
Industry and Handcraft %	44.64	40.08	4.55	4.69 (2.92)	0.47
Retail and Transport %	16.32	16.65	0.33	2.05 (1.77)	0.89
Agriculture %	16.01	21.21	5.20	5.79 (3.66)	0.51
Services%	9.10	9.60	0.51	0.49 (0.46)	0.36
Free occ. / Self-employed %	13.95	12.98	0.97	0.66 (0.49)	0.25
Female share of employees** %	35.92	33.12	2.80	2.94 (1.91)	0.43
Female share of high school students%	31.61	34.87	3.26	2.32 (1.57)	0.28
Marriages per 1000 inhabitants	9.87	9.45	0.42	0.52 (0.34)	0.57
Births per 1000 inhabitants	13.02	14.97	1.95	1.02 (0.78)	0.15
Population (in Millions)	11.43	35.44			

Notes: Sources: Own calculations based on Statistisches Reichsamt (1936:27, 37, 1935:297) for 1933. We use the regions of the former German Empire that coincide with the later boundaries of GDR, excluding Berlin and FRG, excluding Berlin. East: Anhalt, Mecklenburg, Prov. Sachsen, Sachsen, Thüringen. West: Baden, Bayern, Bremen, Hamburg, Hessen, Hohenzollerische Lande, Lippe, Lübeck, Oldenburg, Prov. Hannover, Prov. Hessen-Nassau, Prov. Westfalen, Rheinprovinz, Schaumburg-Lippe, Württemberg. ** For these statistics, the divide is based on the State Employment Office Districts, i.e. for GDR: Mitteldeutschland, Sachsen, and for FRG: Bayern, Hessen, Niedersachsen, Rheinland, Sudwestdeutschland, Westfalen. Column (4) displays the average absolute value of differences between regions in all possible regional partitions of 20 regions into 5 + 15, as well as the standard error of these averages in parenthesis. Column (5) displays the probability that these differences are higher than the East/West difference displayed in column (3), which corresponds to the share of these differences that is higher than the East/West difference. In row 6 (female share of employees), the calculation was made on the basis on the 8 regions available in official statistics (instead of 20). In row 7 (female share of high school students), data are missing for Schaumburg-Lippe, and the resulting number of available Länder is 19.

4 Results

4.1 Self-Declared Math Grades

We start with self-declared math grades collected by the GSOEP. Concerning adults, Table 2 shows that, as expected, women score lower in terms of remembered grades, i.e. they are 7 percentage points less likely to report good grades in math. Eastern people report slightly higher grades. But, notwithstanding memory-biases, East German women totally close the subjective gender gap in math.

Table 2: Self-declared Math Grades by German Adults - GSOEP

Dependent Variable: Self-Declared Last Math Grade (0-1 scale)	
	(1)
Female	-0.07***
	(0.01)
East	0.04*
	(0.02)
East*Female	0.07***
	(0.02)
Observations	15245

* p < 0.1, ** p < 0.05, *** p < 0.01. Linear probability model. The data comes from the German Socio-Economic Panel. Standard errors clustered at the household level are given in parentheses. Sample restricted to individuals born before 1971. East=1 if the household head lived in the GDR before 1990. Controls: age, household Income and Länder fixed-effects. Question: Can you remember your last report card ? What grade did you have in mathematics (1-6 scale with 1 being the highest grade)? Answer originally on a 1-6 scale recoded on a 0-1 scale, 1 corresponding to a grade of 1 or 2.

Concerning teenagers, the picture is similar. The upper panel of Table 3 reveals a gender gap in math grades of the same magnitude, no statistically significant difference between East and West Germany, and again, the absence of gender gap in math in East Germany. The lower panel of Table 3 is based on the German PISA-E survey of 2003. As PISA also asks pupils about their remembered grades in math (*In your last school report, what was your mark in Mathematics?*), we use this information to double-check the findings based on GSOEP. The gender gap in self-declared math grades, collected by PISA, is of the same

order of magnitude as the one found using GSOEP (9%). Again, in PISA, the gender gap in self-reported math grades is entirely closed in East Germany. Hence, from ten to twenty years after the reunification, self-declared math grades are similar for men and women in Eastern Länder, whereas they are lower for women living in Western Länder.

Table 3: Self-declared Math Grades by Teenagers (0-1 Scale)

Panel A: GSOEP - 17 years old	
Female	-0.05***
	(0.02)
East	-0.04
	(0.05)
Female*East	0.09**
	(0.04)
Observations	3446
Panel B: PISA-E - 15 years old	
Female	-0.09***
	(0.01)
East	-0.01
	(0.02)
Female*East	0.08***
	(0.01)
Observations	23513

* p < 0.1, ** p < 0.05, *** p < 0.01. Linear probability model. *Panel A*: the data comes from the German Socio-Economic Panel. The sample is restricted to individuals who were born in Germany. Standard errors clustered at the household level are given in parentheses. East=1 if the Länder is part of the former GDR. Controls: age, household income and Länder fixed-effects. Question : What Grade did you get in your last report card in mathematics ? Answer originally on a 1-6 scale recoded on a 0-1 scale, 1 corresponding to a grade of 1 or 2. *Panel B*: the data comes from the German National Evaluation of Pisa. The sample is restricted to individuals who were born in Germany. Standard errors clustered at the school level are given in parentheses. Individual controls: a quadratic in age, class, a quadratic in wealth, education and employment status of both parents. School controls: Size of location area and share of girls in the school and Länder fixed-effects.

4.2 Objective Math Scores - PISA-E Germany 2003

4.2.1 The Gender Gap in Mathematics is Smaller in East Germany

We now calculate the average score obtained by each student in each of the four categories of math skills (Space and Shape, Relationship and Change, Uncertainty, and Quantity), as well as in an additional category: Problem Solving, assessed by PISA-E 2003.

The estimates displayed in Table 4 show that, in average, the scores of Eastern Länder pupils do not differ from those of Western Länder. Girls' scores are lower than boys', by 18 points (Problem Solving) to 39 points (Uncertainty), where the average score is about 500. However, the interaction terms show that this gap is reduced for girls from East Germany, by 5 points (Quantity) to 11 points (Uncertainty). In average, the gender gap in the global score amounts to 31 points, and is reduced by 4 points in East Germany, i.e. about one eighth. In general, the gender gap is math is thus reduced in East Germany, as compared to West Germany, by about one tenth to one third, depending on the subjects.

Table 4: Maths Scores - PISA 2003-E Germany

Dependent Variable: Math Score						
	(1) Global	(2) Space and Shape	(3) Relationships and Change	(4) Uncertainty	(5) Quantity	(6) Problem Solving
Female	-31.36***	-34.70***	-32.32***	-39.06***	-23.78***	-17.86***
	(1.31)	(1.59)	(1.61)	(1.47)	(1.51)	(1.40)
East	2.70	5.69	6.80	-5.15	-3.75	-11.62
	(7.60)	(8.43)	(8.97)	(7.69)	(7.54)	(7.94)
Female*East	4.35*	-2.72	6.06**	10.90***	4.86**	6.71***
	(2.26)	(2.91)	(2.58)	(2.36)	(2.35)	(2.35)
Observations	23619	23619	23619	23619	23619	23619

* p < 0.1, ** p < 0.05, *** p < 0.01. The data comes from the German National Evaluation of Pisa. The sample is restricted to individuals who were born in Germany. Standard errors clustered at the school level are given in parentheses. Individual controls: a quadratic in age, class, a quadratic in wealth, education and employment status of both parents. School controls: Size of location area and share of girls in the school and Länder fixed-effects. OLS estimation.

Because the gender gap in math is often found at higher levels of achievement, we run quantile regressions of the previous estimates. Table A4 in Appendix distinguishes math

scores from the 10th to the 90th decile, with an increment of 20 points. The gender gap increases at higher levels of performance for Relationship and Change and Uncertainty (by one third), in Problem Solving (where it doubles) and in Quantity (slightly), but not in Space and Shape. Again, this gap is attenuated in East Germany, sometimes particularly so at the top level of performance, i.e. for Relationship and Change and Uncertainty, but this is not systematic.

4.2.2 The Case of the Space and Shape Content Domain

It is remarkable that the only type of exercise for which the gender gap in math is not reduced in East Germany is the Space and Shape category. This goes back to the discussion about the brain foundation of the gender gap in math. Else-Quest et al. (2010) already noticed that the largest gender difference in math achievement concerns the Space and Shape content domain of PISA. A possible evolutionist explanation is that hunting has formatted males' brain, making them more apt to orientation in space. However, the authors notice that spatial skill instruction is often neglected in schools, and that small amounts of instruction can produce large increases in spatial skills. They also mention research showing that playing videogames can improve mental rotation skills and in some cases eliminate gender differences in this type of skill. Without further research, we can only acknowledge the particularly resilient gender gap in visuo-spatial skills.

4.2.3 Is It Detrimental to Reading Scores?

One could wonder whether the better performance of East German girls in mathematics comes at the price of lesser efforts in other domains, such as Reading. Table 5 shows that this is not the case. There is no average difference between East and West Germany in terms of Reading scores. As is generally the rule, girls outperform boys in Reading, in average, and this is even more so in East Germany. This is consistent with the findings by Guiso et al. (2008), who report that in more gender-equal countries, the gender gap in math is smaller, but girls outperform boys in Reading.

Table 5: Reading Score - PISA 2003-E Germany

Dependent Variable: Reading Score	
	(1)
Female	17.60***
	(1.48)
East	-6.04
	(7.50)
Female*East	6.56***
	(2.50)
Observations	23619

* p < 0.1, ** p < 0.05, *** p < 0.01. The data comes from the German National Evaluation of Pisa. The sample is restricted to individuals who were born in Germany. Standard errors clustered at the school level are given in parentheses. Individual controls: a quadratic in age, class, a quadratic in wealth, education and employment status of both parents. School controls: Size of location area, share of girls in the school and Länder fixed-effects. OLS estimation.

4.3 Subjective Attitudes to Math - PISA-E Germany 2003

If the gender gap in math is at least partly a social construct, it should reflect on students' perceptions of their skills. Bharadwaj et al. (2016) showed that boys and girls differ significantly in perceptions about their own ability in math. Else-Quest et al. (2010) also found that boys felt more confident and less anxious in their math abilities and were more extrinsically and intrinsically motivated to do well in math than were girls. Boys also scored higher than girls on math self-concept and self-efficacy.

We use PISA 2003-E questions about students' subjective attitudes towards math. (The questionnaire is presented in the Appendix, Section B.2). Table 6 shows that girls generally express a lower appetence for math, lower self-confidence, more stress and less pleasure in the practice of math. Their score is about 0.4 point lower than that of boys, on a 4 points scale, for these measures. Living in East Germany does not significantly influence these attitudes. But being a girl in East Germany reduces the subjective gender

gap. The magnitude of this effect varies from about one tenth to one third. This is quite impressing, as this effect is measured more than ten years after the dissolution of the GDR. The last column of the table shows that this East-West difference is not due to the learning methods practiced by students.

Table 6: Subjective Attitudes to Math - PISA 2003-E Germany

Dependent Variable: Aggregate Subjective Variables					
	(1) Attitude	(2) Confidence	(3) Stress	(4) Pleasure	(5) Methods
Female	-0.37***	-0.32***	0.35***	-0.48***	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
East	0.01	-0.03	0.02	-0.01	-0.01
	(0.03)	(0.03)	(0.04)	(0.04)	(0.02)
Female*East	0.11***	0.04***	-0.04**	0.07***	0.02
	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)
Observations	23598	23561	23553	23536	23550

* p < 0.1, ** p < 0.05, *** p < 0.01. The data comes from the German National Evaluation of Pisa. The sample is restricted to individuals who were born in Germany. Standard errors clustered at the school level are given in parentheses. Individual controls: a quadratic in age, class, a quadratic in wealth, education and employment status of both parents. School controls: Size of location area, share of girls in the school and Länder fixed-effects. OLS estimation. Full labelling of the questions in Appendix B.2.

4.4 Stereotype Threat. Are East German Girls More Competitive?

In line with our interpretation, we expect the stereotype threat concerning math that weights on girls to be smaller in East Germany than in the West. In particular, one of the main channels through which gender stereotypes are thought to influence girl's attitude to math is that they "shy away from competition" (Niederle & Vesterlund 2010), especially when competitors are boys. Part of the student questionnaire targets exactly this notion of competition, namely: Thinking about your mathematics classes: To what extent do you agree with the following statements?

- I would like to be the best in my class in mathematics.

- I try very hard in mathematics because I want to do better on the exams than the others.
- I make a real effort in mathematics because I want to be one of the best.
- In mathematics I always try to do better than the other students in my class.
- I do my best work in mathematics when I try to do better than others.

Table 7 shows that girls do express much less competitive attitudes than boys (a difference of 0.2 on a 1-4 scale). There is generally no difference in competitive attitudes of Eastern versus Western pupils (except for one question). But Eastern girls express much more competitive attitudes than Western ones.

Table 7: Competitive Attitudes in Math - PISA-E 2003 Germany

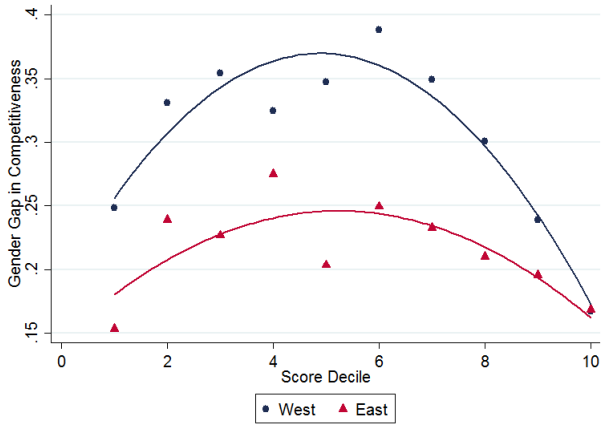
Subjective Variable: Competitiveness in Mathematics					
	(1) Be the Best	(2) Try hard	(3) Effort	(4) Outperform	(5) Challenge
Female	-0.20***	-0.22***	-0.27***	-0.39***	-0.32***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
East	-0.16***	0.03	-0.02	-0.06	0.03
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Female*East	0.08***	0.07***	0.10***	0.09***	0.04*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Observations	23474	23393	23370	23362	23340

* p < 0.1, ** p < 0.05, *** p < 0.01. The data comes from the German National Evaluation of Pisa. The sample is restricted to individuals who were born in Germany. Standard errors clustered at the school level are given in parentheses. Individual controls: a quadratic in age, class, a quadratic in wealth, education and employment status of both parents. School controls: Size of location area, share of girls in the school and Länder fixed-effects. The explained variables are, for each column: Column 1 : I would like to be the best in my class in mathematics; Column 2 : I try very hard in mathematics because I want to do better in the exams than the others; Column 3 : I make a real effort in mathematics because I want to be one of the best; Column 4 : In mathematics I always try to do better than the other students in my class; Column 5 : I do my best work in Mathematics when I try to do better than others. Answers are given on a 1-4 scale, estimation by OLS.

Additionally, Figure 1 plots a quadratic fit of the gender gap in competitive attitudes (based on the average score on the five questions mentioned above) in East (triangle, red

line) and West (dot, blue line) Germany, depending on the actual PISA average math score of pupils (by deciles, computed for the distribution of scores for East and West Germany separately). The gender gap in competitive spirit is generally greater in West Germany. But the difference between East and West Germany is particularly important at intermediate levels of performance. It grows from the first to the fifth decile, and then declines. This exactly corresponds to the predictions of a model à la Altonji et al. (2012): the greater the uncertainty about their skills, the more girls underestimate themselves and shy out from competition. At very low levels of math skills, it is likely that boys and girls have a precise notion of their (low) performance, and we do not expect any gender gap in competitiveness. The same reasoning goes for very high levels of performance. It is in-between the two extremes that there is room for under/over-confidence, and this is where the gender gap takes place. This finding is thus in line with an interpretation in terms of persisting cultural attitudes inherited from the time of the division.

Figure 1: Competitive Attitudes in Math by Deciles - PISA-E 2003 Germany



Notes: The data comes from the German National Evaluation of Pisa. The sample is restricted to individuals who were born in Germany. The red line (triangle) shows the quadratic fit of the indicator of competitive attitudes in mathematics by deciles for East Germany (computed on the basis of PISA global mathematics scores in East or West Germany). The blue line (circle) represents West Germany.

5 Robustness

Could the difference across the former Berlin wall be due to other structural differences? Two main confounding factors are obvious: different teaching practices, and structural

economic differences across the former political border between the FRG and the GDR. We also discuss the potential different religious orientations.

5.1 Are Eastern and Western Schools Different?

As discussed in Section 2, mathematics could be taught in a different way, which could happen to be more favorable to girls, in Eastern Länder. Testing this idea thoroughly would be the object of a separate study. However, we can make use of PISA's student and school questionnaires to enquire.

To start with, PISA student questionnaire contains a question about the time allotted to math in their class. We aggregate the answers at the school level and look at the difference between East and West schools for this matter. If Eastern schools devoted more time to math training, say, than Western schools, this could have an impact on the gender gap in math. It could, for instance, reduce the gender gap in math because of increasing returns, or on the contrary, discourage girls even more, if they are initially little attracted to this field. It turns out (Table 8) that the time allotted to math does not differ in East versus West German schools.

Table 8: Total Math Time in Class per Week - PISA 2003-E Germany

Dependent Variable: Minutes of Mathematics per Week - School level	
	(1)
East	-10.93
	(7.33)
Constant	251.57***
	(8.99)
Observations	1199

* p < 0.1, ** p < 0.05, *** p < 0.01. The data comes from the German National Evaluation of Pisa. School controls: Size of location area, share of girls in the school and Länder fixed-effects.

In addition, we analyze the school questionnaire, which interviews each school's director, and covers a lot of teaching aspects. The upper panel of Table A5 displays an aggregate score of answers to a series of questions about the organization of classes, in particular the allocation of students into ability groups. The second panel displays the average score of

questions pertaining to the degree of innovative teaching methods for mathematics adopted by teachers. The third panel uses questions related to the optimal level of difficulty to target. Finally, the bottom panel presents questions related to the relative weight of cognitive (math) versus non cognitive skills, i.e. whether mathematics teachers grant more importance to the social and emotional development of students or to their acquisition of Mathematical skills and knowledge in Mathematics classes.

Again, the result of our analysis shows that there is little statistical difference between East German versus West German school. To be sure, this does not rule out definitively the idea that schools may differ, but with the available data, we are unable to reject the hypothesis that they are similar.

5.2 Are the Results Driven by Different Economic Conditions?

The difference in girls' math scores across the Berlin wall could also be due to different economic conditions prevailing in East versus West Germany, instead of persisting gender norms. For instance, it could be the case that a higher rate of regional unemployment in the East increases the necessity for women to work, hence the motivation for girls to invest in math and other human capital skills.

To rule out this alternative explanation, we need to look at the possible relationship between economic conditions and the gender gap in math. The difficulty is that, for confidentiality reasons, we cannot match PISA scores to regional and macroeconomic statistics, because in the database the Länder have been "anonymized", so that we know which schools belong to the same Länder, and whether the latter was part of the former FRG or GDR, but we cannot precisely identify it. In order to go around this problem, we construct our own measures of economics conditions on the basis of the information available in PISA data.

We focus on West Germany, and, using the answers given by students about their parent's labor market status; we calculate the ratio of parents who are declared to be unemployed by their children to the total number of parents. These ratios based on PISA are of 6% in West Germany and 13% in the East. Although close, they do not match the official figures for the rate of unemployment in 2003 (8.4% in the West and 18.2% in the East), because the concepts are different, but the difference is of similar magnitude.

Concerning wealth, we use the household wealth indicator constructed by the OECD

on the basis of the answers given by students about the size and equipment of their house, and the number of books that their parents possess. We use this variable and aggregate it at the Länder level. We obtain a score of wealth varying from -6.17 to 3.15, with a higher score indicating a plausibly higher wealth.

We restrict our sample to West Germany and relate these ratios to the gender gap in mathematics. The upper panel of Table A6 shows that a higher unemployment rate is associated with a lower general score in math, but that there is no association with the gender gap in math. The lower panel shows a positive relationship between regional wealth and the score in mathematics, altough not significant for every category, but this relationship does not seem to differ for girls or boys. Extrapolating these results to the East-West difference, we conclude that the poorest economic conditions that prevail in East Germany, as opposed to West Germany, cannot explain the smaller gender gap in math in the Eastern part of the country.

5.3 Are the Results Driven by Different Religious Orientations?

Because Germany is divided into predominantly protestant and catholic regions, religion is one of the main candidates, in the list of potentially cofounding factors. Could the difference between East and West Germany be the simple reflection of the Catholic/Protestant Divide? Spenkuch (2017) helps elucidate this issue with a map reproduced in Figure A1. The latter is built using self-declared religious orientations of German citizens collected in the SOEP in 2000-2008. Spenkuch draws a map of religious affiliations in the 17th century and in the years 2000's. As is obvious from Figure A1, individuals who currently live in historically Protestant areas are considerably more likely to self-identify as Protestant than residents of historically Catholic regions. However, an important exception is the former GDR, where most people self-identify as non religious. In the 17th century, although the divide between Protestant and Catholic Länder is more of an North-East/South-West divide than a simple East/West divide, a large part of the territories of the former GDR used to be Protestant Länder.

To rule out the possibility that this different cultural background between East and West Germany affects our results, we replicated the same exercize as in the previous section using religious orientation as the dividing line instead of the former iron curtain. Given that there is no information on religion in PISA, we relied on SOEP data and computed

the shares of catholics, protestants, as well as the ratio of the two in each Western Länder.

Table A7 shows that the share of protestant people is negatively associated with the probability of reporting a good grade in mathematics. The opposite is true of the share of catholics in the local population. However, none of these coefficients are statistically significant and, more importantly, nor is their interaction with the variable *Female*, which is very close to zero. If anything, the higher the share of protestants, the wider the gender gap in math. Thus, there is no indication that the religious orientation affects the gender gap in math in Germany.

6 Extensions

The case of the German division is particularly adapted to the demonstration of the causal influence of institutions. Nonetheless, for illustration, we now extend the analysis to all European countries, and contrast former socialist "Eastern" countries to capitalist "Western" countries. We uncover a similar picture: the gender gap in math is much smaller, and even sometimes inexistent, in Eastern countries. This is also true for other math-related domains such as Chess competition ratings.

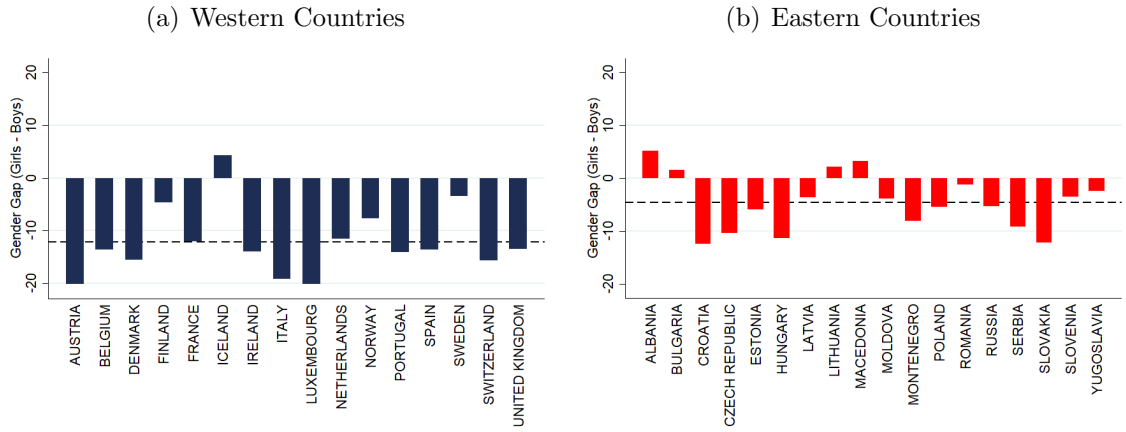
6.1 Empirical Evidence from PISA Europe

We first look at the sample of European countries participating in PISA international assessments from 2000 until 2012[2]. We consider this time span in order to include the largest possible number of countries. Given that the participation to PISA is not mandatory, countries may not be present at every wave and may join the survey after 2000. We begin by representing graphically the raw gender gap in mathematics test scores in each European country. We compute for each country the difference in the average math score of girls and boys. Figure 2 represents this gender gap for Western (a) and Eastern countries (b). Visually, the gender gap is generally smaller in Eastern countries than in Western ones. On the Western side, girls perform better than boys only in Iceland. Besides Iceland, the

[2]We selected all the countries who participated at least once in the PISA international assessments from 2000 until 2012. Countries included in the Western group are: Austria, Belgium, Denmark, Finland, France, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Switzerland, Sweden and United-Kingdom. Countries included in the Eastern group are: Albania, Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Macedonia, Moldova, Montenegro, Poland, Romania, Serbia, Slovakia, Slovenia, Russia and Yugoslavia.

three smallest gender gap are found in Scandinavian countries - namely Finland, Norway and Sweden. In the remaining other Western countries, the gender gap ranges from -10 to -20. On the Eastern side, there are four countries where girls perform better than boys (Albania, Bulgaria, Lithuania and Macedonia). In only four countries is the negative gender gap larger than -10 (Croatia, Czech Republic, Hungary and Slovakia). In the remaining Eastern countries, the gender gap is rather small, in the range of Scandinavian countries.

Figure 2: Gender Gap in Mathematics in Western and Eastern European Countries



Notes: The data comes from the 2000, 2003, 2006, 2009 and 2012 waves of PISA. The horizontal dashed line represents the average gender gap for each region.

Table 9 displays a more systematic analysis of this East-West difference. Controlling for country and year fixed-effects, as well as the usual socio-demographic controls, the average gender gap in math is of 16 points, where the sample's average score is of 500. Students from former socialist countries do not score differently, in average, than the rest of the sample. However, girls from these countries score 9 points above their Western counterpart, i.e. they close more than half of the gender gap (conditional on coming from a former socialist country).

Table 9: Math Scores - PISA Europe 2000 to 2012

Dependent Variable: Global Math Score	
	(1)
Female	-16.38***
	(0.38)
Former Socialist	-5.73
	(4.19)
Female*Former Socialist	9.23***
	(0.63)
Observations	724784

* p < 0.1, ** p < 0.05, *** p < 0.01. The data comes from the 2000, 2003, 2006, 2009 and 2012 waves of PISA. Controls: quadratic in age, class, quadratic in household wealth, education of both parents as well as country and year fixed effects. Sample restricted to national born individuals.

We also reproduced the same analysis on the gender gap in math to see whether the better performance of girls comes at the cost of their traditional advantage: Reading. Consistent with the results on Germany, we see in Table A8 that the gender gap in reading is accentuated in Eastern countries. Girls outperform boys by 33 points on this matter and this gap increases by an average of 10 points in former socialist countries. This pattern, as noted in Section 4.2.3, is regularly observed in more gender equal countries (see Guiso et al. 2008).

Furthermore, in line with Section 4.2, we ran quantile regressions to investigate wether the attenuation of the gender gap in Eastern European countries is present at every percentile. The results are displayed in Table A9. While the gender gap increases as we move towards the top of the distribution, the reduction of this gap in former socialist countries seems stable at every quantile and slightly lower at the 90th quantile, as compared to the 10th one.

6.2 International Mathematical Olympiad

In order to explore the gender gap amongst high achievers in math, we now focus on students who are sent to the International Mathematical Olympiad (IMO), an annual world championship mathematics competition for high school students under 20 years old[3]. The IMO involves an extremely difficult, proof-based, essay-style examination in mathematical problem solving (Kane & Mertz 2012).

The first IMO was held in 1959 in Romania, with 7 countries participating. Over the past 50 years, it has grown from a small contest among Soviet-bloc nations to a true worldwide contest among 100 countries. The number of high school students that each country may send was progressively reduced to six students. These students are often winners of the country's national Olympiad, but the manner in which teams are selected varies. The top scorers on the IMO have exceptional skills in mathematics, that is, at the 1-in-a-million level (Ellison & Swanson 2010, Andreescu et al. 2008). Several contestants have become famous mathematician, some have won the Fields Medal.

We use the IMO data spanning from 1959 until 2015. We employ three different variables in order to capture the gender gap in the participation in the IMO: a dummy that equals 1 if a country sends at least one girl, the share of girls in the team and the number of girls in the team. We regress each of these variables on a dummy that equals 1 if the country belongs or used to belong to the socialist bloc, controlling for year and country fixed-effects as well as for the size of the teams. Our sample includes only European countries[4] and distinguishes two periods: before and after 1991.

Table 10 displays the results. In panel A column 1, before 1991, we see that the probability that a socialist country sends at least one girl to the Olympiad was 27 percentage points higher than for a Western country. Consistently, the share of girls sent to the IMO, as well as the number of girls in the team, were higher. After 1991, former socialist countries kept sending more girls to the IMO; the probability that a former socialist country sends at least one girl was 38 percentage points higher than in the West (column 1).

[3] https://www.imo-official.org/?language=en

[4] This sample includes, on the socialist side: Albania, Bielorussia, Bosnia, Bulgaria, Croatia, Estonia, Hungary, Poland, GDR, Romania, Czechoslovakia, USSR and Yugoslavia. On the Western side: Spain, Denmark, UK, Finland, Norway, Italy, FRG, Sweden, Belgium, Austria and the Netherlands.

Table 10: Presence of Girls at the International Mathematical Olympiad

Dep. Var	(1) At Least One Girl	(2) Share Girls	(3) N Girls
Panel A: Before 1991			
Socialist	0.27**	0.05**	0.32**
	(0.11)	(0.02)	(0.15)
Observations	414	414	414
Panel B: After 1991			
Former Socialist	0.38**	0.14***	0.83***
	(0.19)	(0.05)	(0.27)
Observations	530	530	530

* p < 0.1, ** p < 0.05, *** p < 0.01. Linear probability model. The data comes from the International Mathematical Olympiad. Panel A is restricted to the years before 1991 and Panel B to the years after. Controls include country fixed-effects, year fixed-effects and the size of the team. Dependent variables are a dummy that equals 1 if a country sends at least one girl (column 1), the share of girls in the team (column 2) and the number of girls in the team (column 3). Observations are at the level of a country per year.

6.3 Chess

Chess championships are another type of top level competition for math-minded people. The FIDE, French acronym for the World Chess Competition (Federation Internationale des Echecs) publishes ratings for the very top of the distribution of chess players, i.e. the top 100 and top 50 players by sex, each year, for standard, rapid and blitz competitions. It also ranks the top 100 junior boys and girls players (separately) along the same categories. The ranking of participants is based on their Elo rating. The data indicates the origin country of each player, men and women, in the sample. We use data from the FIDE chess ratings[5].

We compare the proportion of men coming from former socialist/capitalist countries to the same ratio for women[6]. Among men, top level players coming from a former socialist

[5]https://ratings.fide.com/

[6]Formally we test whether : $Prop^{com}_{men} = Prop^{com}_{women}$ where $Prop^{com}_{men}$ corresponds to the proportion

country account for 72.9%. This proportion is 81% for women. The difference between the two ratios is statistically significant at the 1% level. When we reduce the sample to one observation per player (there are on average 5.26 observations per player), we obtain similar results: the ratio of top players coming from the former Soviet bloc is 73% for men and 83.8% for women. These proportions are statistically different at the 5% level. It thus seems that the socialist episode has exerted long lasting effects on girls' performance in math, as measured by international standardized PISA scores, as well as by the International Mathematical Olympiads, and Chess competitions.

7 Conclusions

The claim of this paper is that the socialist episode has exerted long lasting multi-directional effects on women's expectations, self-confidence and choices. Previous studies have shown that women's attachment to paid work was greater in East Germany. This paper extends the same conjecture to girls' appetence for mathematics and achievement therein. It is a general stylized fact that girls underperform in math; this constitutes one of the most resistant gender gaps of modern societies. We show that this specific handicap has been sharply attenuated in East Germany. Even in recent years, girls' performance in math, as measured by international standardized PISA scores, is closer to that of boys in the regions of the former GDR, as opposed to the former FRG. Evidence from the international standardized PISA scores, the International Mathematical Olympiads and International Chess competitions, suggest that the gender gap in math and math-minded competitions is generally smaller in countries of the former Soviet bloc, as opposed to other European countries.

We interpret this stylized fact as a legacy of socialist institutions and policies that enacted, in a particularly compelling way, the objective of female full-employment. Policies that facilitated labor market participation and maternity for women were accompanied by official propaganda sustaining stereotypes of professionally active women, whose work was praised as a political engagement in the construction of socialism. Overall, this normative pressure changed the conception of gender roles and identity in many dimensions, including girls' school curricula and performance.

of men from former socialist countries among the top 100 players and $Prop^{com}_{women}$ corresponds to the proportion of women from former socialist countries among the top 50 women.

The claim of this paper is not that the socialist ideology in general was favorable to women. Nor do we discuss whether the gender policy that was implemented in socialist countries was dictated by equality motives or rather by the imperious rapid growth objective that commanded the 5-year plans. On the normative side, needless to say that we do not advocate for the implementation of similarly authoritarian gender policy, nor, of course, for the host of other measures that came with it. Some of the aspects of the gender policy (child care, maternity leaves, etc.) have already been successfully replicated in other contexts, in particular in Nordic countries such as Sweden and Denmark, without the same degree of authoritarianism. A minima, this paper wishes to show that, to a large extent, gender gaps, even in domains that seem to be grounded in nature, can actually be dissolved by institutions.

References

Akerlof, G. A. & Kranton, R. E. (2000), 'Economics and identity', *The Quarterly Journal of Economics* 115(3), 715–753.

Akerlof, G. A. & Kranton, R. E. (2013), *Identity economics: How our identities shape our work, wages, and well-being*, Princeton University Press.

Alesina, A. & Fuchs-Schündeln, N. (2007), 'Good-bye lenin (or not?): The effect of communism on people's preferences', *The American Economic Review* 97(4), 1507–1528.

Altonji, J. G. (1995), 'The effects of high school curriculum on education and labor market outcomes', *The Journal of Human Resources* 30(3), 409–438.

Altonji, J. G., Blom, E. & Meghir, C. (2012), Heterogeneity in human capital investments: High school curriculum, college major, and careers, Working Paper 17985, National Bureau of Economic Research.

Amini, C. & Commander, S. (2012), 'Educational Scores: How does Russia fare?', *Journal of Comparative Economics* 40(3), 508–527.

Andreescu, T., Gallian, J. A., Kane, J. M. & Mertz, J. E. (2008), 'Cross-cultural analysis of students with exceptional talent in mathematical problem solving', *Notices of the American Mathematical Society* .

Arcidiacono, P. (2004), 'Ability sorting and the returns to college major', *Journal of Econometrics* 121(1), 343 – 375. Higher education (Annals issue).

Autor, D. & Wasserman, M. (2013), Wayward sons: The emerging gender gap in labor markets and education, Technical report, Third Way Report. Washington D.C. Higher education (Annals issue).

Baker, D. P. & Jones, D. P. (1993), 'Creating gender equality: Cross-national gender stratification and mathematical performance', *Sociology of Education* 66(2), 91–103.

Bauernschuster, S. & Rainer, H. (2012), 'Political regimes and the family: how sex-role attitudes continue to differ in reunified germany', *Journal of Population Economics* 25(1), 5–27.

Becker, G. S., Hubbard, W. H. J. & Murphy, K. M. (2010), 'The market for college graduates and the worldwide boom in higher education of women', *American Economic Review* 100(2), 229–33.

Benbow, C. & Stanley, J. (1980), 'Sex differences in mathematical ability: fact or artifact?', *Science* 210(4475), 1262–1264.

Bharadwaj, P., Giorgi, G. D., Hansen, D. & Neilson, C. A. (2016), 'The gender gap in mathematics: Evidence from chile', *Economic Development and Cultural Change* 65(1), 141–166.

Bisin, A. & Verdier, T. (2001), 'The Economics of Cultural Transmission and the Dynamics of Preferences', *Journal of Economic Theory* 97(2), 298–319.

Bisin, A. & Verdier, T. (2010), The economics of cultural transmission and socialization, NBER Working Papers 16512, National Bureau of Economic Research, Inc.

Blau, F. D. & Kahn, L. M. (2017), 'The Gender Wage Gap: Extent, Trends, and Explanations', *Journal of Economic Literature* 55(3), 789–865.

Blau, F., Kahn, L., Liu, A. & Papps, K. (2013), 'The transmission of women's fertility, human capital, and work orientation across immigrant generations', *Journal of Population Economics* .

Breen, R. & Cooke, L. P. (2005), 'The persistence of the gendered division of domestic labour', *European Sociological Review* 21(1), 43–57.

Campa, P. & Serafinelli, M. (2016), Politico-Economic Regimes And Attitudes: Female Workers Under State-Socialism, Working Papers 089, "Carlo F. Dondena" Centre for Research on Social Dynamics.

Ceci, S. J., Ginther, D. K., Kahn, S. & Williams, W. M. (2014), 'Women in academic science: A changing landscape', *Psychological Science in the Public Interest* 15(3), 75–141.

Cooke, L. P. (2004), 'The gendered division of labor and family outcomes in germany', *Journal of Marriage and Family* 66(5), 1246–1259.

Cooke, L. P. (2007), 'Persistent policy effects on the division of domestic tasks in reunified germany', *Journal of Marriage and Family* 69(4), 930–950.

Ellison, G. & Swanson, A. (2010), 'The gender gap in secondary school mathematics at high achievement levels: Evidence from the american mathematics competitions', *Journal of Economic Perspectives* 24(2), 109–28.

Else-Quest, N., Shibley Hyde, J. & Linn, M. (2010), 'Cross-national patterns of gender differences in mathematics: A meta-analysis', 136, 103–27.

Fernandez, R. (2010), Does Culture Matter?, NBER Working Papers 16277, National Bureau of Economic Research, Inc.

Fernandez, R. & Fogli, A. (2009), 'Culture: An empirical investigation of beliefs, work, and fertility', *American Economic Journal: Macroeconomics* 1(1), 146–77.

Fortin, N. M., Oreopoulos, P. & Phipps, S. (2015), 'Leaving boys behind: Gender disparities in high academic achievement', *Journal of Human Resources* 50(3), 549–579.

Fryer, R. G. & Levitt, S. (2010), 'An empirical analysis of the gender gap in mathematics', *American Economic Journal: Applied Economics* 2(2), 210–40.

Fuchs-Schundeln, N. & Masella, P. (2016), 'Long-lasting effects of socialist education', *The Review of Economics and Statistics* 98(3), 428–441.

Gneezy, U., Niederle, M. & Rustichini, A. (2003), 'Performance in competitive environments: Gender differences', *The Quarterly Journal of Economics* 118(3), 1049–1074.

Goldin, C. (2014), 'A Grand Gender Convergence: Its Last Chapter', *American Economic Review* 104(4), 1091–1119.

Goldin, C., Katz, L. F. & Kuziemko, I. (2006), 'The Homecoming of American College Women: The Reversal of the College Gender Gap', *Journal of Economic Perspectives* 20(4), 133–156.

Gonzalez de San Roman, A. & de la Rica, S. (2012), Gender Gaps in PISA Test Scores: The Impact of Social Norms and the Mother's Transmission of Role Attitudes, IZA Discussion Papers 6338, Institute for the Study of Labor (IZA).

Gorges, L. & Beblo, M. (2015), Breaking down the wall between nature and nurture: An exploration of gendered work preferences in East and West Germany, Technical Report 112825, Verein fur Socialpolitik / German Economic Association.

Guiso, L., Monte, F., Sapienza, P. & Zingales, L. (2008), 'Culture, gender, and math', *Science* 320(5880), 1164–1165.

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S. & Gernsbacher, M. A. (2007), 'The science of sex differences in science and mathematics', *Psychological Science in the Public Interest* 8(1), 1–51.

Hyde, J. S. & Mertz, J. E. (2009), 'Gender, culture, and mathematics performance', *Proceedings of the National Academy of Sciences* 106(22), 8801–8807.

Joensen, J. S. & Nielsen, H. S. (2009), 'Is there a causal effect of high school math on labor market outcomes?', *Journal of Human Resources* 44(1), 171–198.

Kane, J. & Mertz, J. (2012), 'Debunking myths about gender and mathematics performance', *Notices of the American Mathematical Society* 59.

Kimko, D. D. & Hanushek, E. A. (2000), 'Schooling, Labor-Force Quality, and the Growth of Nations', *American Economic Review* 90(5), 1184–1208.

Kuenzler, J., Walter, W., Reichart, E. & Pfister, G. (2001), Gender division of labor in unified germany, Worc report, Tilburg University Press.

Lippmann, Q., Georgieff, A. & Senik, C. (2016), Undoing Gender with Institutions. Lessons from the German Division and Reunification, Technical report.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J. & Handelsman, J. (2012), 'Science faculty's subtle gender biases favor male students', *Proceedings of the National Academy of Sciences* 109(41), 16474–46479.

Niederle, M. & Vesterlund, L. (2010), 'Explaining the gender gap in math test scores: The role of competition', *Journal of Economic Perspectives* 24(2), 129–44.

Nollenberger, N., Rodriguez-Planas, N. & Sevilla, A. (2016), 'The Math Gender Gap: The Role of Culture', *American Economic Review* 106(5), 257–261.

Pascoe, E. A. & Richman, L. S. (2009), 'Perceived discrimination and health: A meta-analytic review', *Psychological Bulletin* 135(4), 531–554.

Pope, D. G. & Sydnor, J. R. (2010), 'Geographic variation in the gender differences in test scores', *The Journal of Economic Perspectives* 24(2), 95–108.

Rainer, H. & Siedler, T. (2009), 'Does democracy foster trust?', *Journal of Comparative Economics* 37(2), 251–269.

Reuben, E., Sapienza, P. & Zingales, L. (2014), 'How stereotypes impaire women's careers in science', *Proceedings of the National Academy of Sciences* 111(12), 4403–4408.

Rheinstein, M. & Glendon, M. A. (1978), 'West german marriage and family law reform', *The University of Chicago Law Review* 45(3), 519–552.

Richter, T. J. (2014), Stéréotypes, représentations et identités en R.D.A. et en R.F.A. : une comparaison transnationale des discours journalistiques de Der Spiegel et de la Neue Berliner Illustrierte entre 1949 et 1989, PhD thesis. Thèse de doctorat dirigée par Humbert, Geneviève Etudes germaniques Strasbourg 2014.

Rosenfeld, R. A., Trappe, H. & Gornick, J. C. (2004), 'Gender and work in germany: Before and after reunification', *Annu. Rev. Sociol.* 30, 103–124.

Schenk, S. (2003), 'Employment opportunities and labour market exclusion: towards a new pattern of gender stratification', *See Kolinsky & Nickel* 2003, 53–77.

Schmitt, M., Branscombe, N., Postmes, T. & Garcia, A. (2014), 'The consequences of perceived discrimination for psychological well-being: A meta-analytic review', 140.

Schnepf, S. (2007), *Women in Central and Eastern Europe: Measuring Gender Inequality Differently*, VDM, Verlag Dr. Müller.

Spenkuch, J. L. (2017), 'Religion and work: Micro evidence from contemporary germany', *Journal of Economic Behavior And Organization* 135(Supplement C), 193–214.

Appendix. For Online Publication

A Descriptive statistics

A.1 PISA 2003-E Germany

Table A1: Descriptive statistics. PISA

	West Germany		*East Germany*	
	Mean	S.D.	Mean	S.D.
Male Score Global	524.33	94.92	531.34	86.65
Female Score Global	505.43	91.49	515.00	83.33
Grade 8	0.14	0.35	0.09	0.29
Grade 9	0.61	0.49	0.61	0.49
Grade 10	0.25	0.44	0.30	0.46
Female	0.50	0.50	0.49	0.50
Wealth	0.17	0.95	-0.06	0.86
Age	15.69	0.28	15.71	0.28
Mother Educ.	3.41	1.59	4.03	1.24
Father Educ.	3.82	1.67	4.07	1.29
Father's professionnal situation				
Working Full-Time	0.24	0.43	0.63	0.48
Working Part-Time	0.46	0.50	0.18	0.38
Looking for Work	0.04	0.19	0.11	0.32
Other	0.26	0.44	0.07	0.26
Mother's professionnal situation				
Working Full-Time	0.84	0.37	0.81	0.40
Working Part-Time	0.07	0.26	0.06	0.24
Looking for Work	0.04	0.19	0.09	0.29
Other	0.06	0.23	0.04	0.19
Share of Female in School	0.51	0.09	0.50	0.07
Size of the community				
Less than 3000 people	0.15	0.35	0.01	0.09
Between 3000 and 15000	0.02	0.14	0.14	0.35
Between 15000 and 100000	0.25	0.44	0.34	0.48
Between 100000 and 1000000	0.36	0.48	0.37	0.48
More than 1000000	0.22	0.41	0.14	0.35
Observations	16401		7922	

Notes: The data comes from the German National Evaluation of Pisa. The sample is restricted to individuals who were born in Germany.

In Table A2, we relate the self-declared mathematics grades to the test scores obtained for each of the mathematics categories. Looking at the first line, we see in column 1 that individuals who declared a grade of 1 or 2 in mathematics (highest grades) scored on average 555 points on the global score whereas individuals who declared a grade strictly higher than 2 scored 499 points on average. The difference between the two quantities is of 55 points and is significant at the 1% level.

Table A2: Relationship Between Self-Declared Grades and Objective Test Scores

Reported Grade	High Grade 1 or 2	Low Grade > 2	Difference
Objective Score	(1)	(2)	(1)-(2)
Global	555.804	499.841	55.963***
			(1.272)
Space and Shape	559.428	493.747	65.680***
			(1.501)
Relationships and Change	560.623	501.877	58.747***
			(1.483)
Uncertainty	537.516	484.614	52.902***
			(1.351)
Quantity	566.134	511.369	54.765***
			(1.346)
Problem Solving	556.167	508.261	47.905***
			(1.289)
Observations	6807	14162	21244

* p < 0.1, ** p < 0.05, *** p < 0.01. The data comes from the German National Evaluation of Pisa. The sample is restricted to individuals who were born in Germany. Column 1 is restricted to individuals who declared a grade in mathematics of 1 or 2 (Highest Grades). Column 2 is restricted to individuals who declared a grade strictly higher than 2 (low grades). Column 3 computes the difference between Column 1 and Column 2. Each line displays the objective test scores for one of the mathematics categories.

A.2 German Socio-Economic Panel

Table A3: Descriptive statistics. GSOEP

	West Germany		*East Germany*	
	Mean	S.D.	Mean	S.D.
		Panel A : Adults		
High Grade in Math Male	0.49	0.50	0.55	0.50
High Grade in Math Female	0.42	0.49	0.53	0.50
Age	49.57	16.66	48.59	17.03
Female	0.52	0.50	0.52	0.50
Household Income (EUR)	3080.08	2575.18	2223.61	1322.18
Observations	14906		4799	
		Panel B : Teenagers		
High Grade in Math Male	0.34	0.47	0.35	0.48
High Grade in Math Female	0.30	0.46	0.39	0.49
Age	17.10	0.39	17.11	0.41
Female	0.50	0.50	0.47	0.50
Household Income (EUR)	3238.77	1946.33	2479.61	1346.98
Observations	2618		827	

Notes: The data comes from the German Socio-Economic Panel. The sample is restricted to individuals who were born in Germany. High reported grades: 1 and 2 on a 1-6 scale.

B PISA-E 2003 Germany

B.1 Quantile Regressions

We run quantile regressions to investigate whether the gender gap increases at higher levels of performance and how this gap is attenuated in East Germany. To test whether coefficients between quantiles are different, we compute the p-value from a Wald Test between the 10th and 90th percentile. From the coefficients, it seems that the gender gap increases by the largest extent for Relationship and Change (p-value=0.000), followed by Uncertainty (p-value=0.003) and then Problem Solving (p-value = 0.005). For the Space

and Shape and Quantity categories, there is no significant increase in the gender gap in math between the 10th and 90th percentiles.

The attenuation of the gender gap exhibits a similar impact across all quantiles except for the Relationship and Change category (p-value = 0.074) where the gap is sharply attenuated in the 90th percentile as compared to the 10th one. While there seems to be an increase in the Uncertainty category, the difference is not significant (p-value= 0.23). The differences between the 10th and 90th percentiles in the impact of $Female * East$ is not significant for the other categories.

Table A4: Maths Scores by Quantiles - PISA 2003-E Germany

	(1)	(2)	(3)	(4)	(5)
Quantile	10	30	50	70	90
Dependent Variable: Space and Shape Score					
Female	-31.52***	-34.56***	-34.11***	-35.03***	-34.81***
	(2.76)	(1.76)	(1.23)	(1.80)	(2.26)
East	1.79	5.49	-0.39	2.04	-1.74
	(6.00)	(5.19)	(4.51)	(5.04)	(6.20)
Female*East	-6.73	-3.32	-4.09	-1.19	1.12
	(4.91)	(2.55)	(2.64)	(2.84)	(3.88)
Dependent Variable: Relationships and Change Score					
Female	-24.45***	-30.57***	-31.79***	-33.68***	-38.43***
	(2.51)	(1.59)	(1.43)	(2.04)	(2.35)
East	15.96**	8.39*	9.56***	4.60	-13.08**
	(7.58)	(4.46)	(3.40)	(4.10)	(6.27)
Female*East	-2.77	6.43**	6.17**	2.53	7.28**
	(5.53)	(2.85)	(3.01)	(4.35)	(3.00)
Dependent Variable: Uncertainty Score					
Female	-33.19***	-38.27***	-41.82***	-41.82***	-43.48***
	(2.24)	(1.73)	(1.59)	(1.41)	(2.43)
East	-2.99	-4.39	-6.90*	-9.74**	-22.32***
	(4.95)	(4.78)	(4.00)	(4.08)	(5.99)
Female*East	7.99**	10.19***	13.10***	11.83***	14.61***
	(3.43)	(3.03)	(1.94)	(2.97)	(4.24)
Dependent Variable: Quantity Score					
Female	-23.78***	-23.83***	-24.22***	-24.57***	-26.48***
	(1.99)	(1.81)	(1.49)	(1.89)	(1.54)
East	-3.05	-5.29	-3.13	-5.56	-11.46*
	(6.41)	(6.17)	(5.59)	(4.77)	(5.87)
Female*East	5.33*	4.15*	5.65	4.29	2.54
	(2.75)	(2.43)	(3.56)	(3.37)	(3.46)
Dependent Variable: Problem Solving Score					
Female	-12.92***	-14.67***	-20.58***	-21.51***	-23.27***
	(2.89)	(1.88)	(1.59)	(1.87)	(1.97)
East	-9.46	-6.94**	-14.71***	-16.46***	-16.85***
	(6.31)	(2.91)	(3.48)	(3.95)	(6.05)
Female*East	7.02	3.82	9.04***	9.49***	7.10**
	(4.78)	(2.73)	(2.82)	(3.16)	(3.39)

* p < 0.1, ** p < 0.05, *** p < 0.01. The data comes from the German National Evaluation of Pisa. The sample is restricted to individuals who were born in Germany. Standard errors are calculated with a bootstrap method. Individual controls: a quadratic in age, class, a quadratic in wealth, education and employment status of both parents. School controls: Size of location area, share of girls in the school and Länder fixed-effects.

B.2 Data Appendix - Student Questionnaire

- Thinking about your views on Mathematics: to what extent do you agree with the following statements:

- I enjoy reading about Mathematics
- Making an effort in mathematics is worth it because it will help me in the work I want to do later.
- I look forward to Mathematics lessons.
- I do Mathematics because I enjoy it.
- Learning Mathematics is worthwhile for me because it improves my career prospects, chances.
- I am interested in the things is learn in Mathematics.
- Mathematics is an important subject for me because I need it for what I want to study later on.
- I will learn many things in Mathematics that will help me get a job.

- How confident do you feel about having to do the following Mathematics tasks?
  - Using a train timetable to work out how long it would take to get from one place to another.
  - Calculating how much cheaper a TV would be after a 30% discount.
  - Calculating how many square feet of tile you need to cover a floor.
  - Understanding graphs presented in newspapers.
  - Solving an equation like 3x+5= 17.
  - Finding the actual distance between two places on a map with a 1:100 scale.
  - Solving an equation like 2(x+3)=(x + 3)(x - 3).
  - Calculating the gas mileage of a car.

- Thinking about studying mathematics: To what extent do you agree with the following statements?
  - I often worry that it will be difficult for me in Mathematics classes.
  - I am just not good at Mathematics.
  - I get very tense when I have to do Mathematics homework.
  - I get good grades in Mathematics.

  - I get very nervous doing Mathematics problems.
  - I learn Mathematics quickly.
  - I have always believed that Mathematics is one of my best subjects.
  - I feel helpless when doing a Mathematics problem.
  - In my Mathematics class, I understand even the most difficult work.
  - I worry that I will get poor grades in Mathematics.

- Thinking about your Mathematics classes: To what extent do you agree with the following statements?
  - I would like to be the best in my class in Mathematics.
  - In Mathematics I enjoy working with other students in groups.
  - I try very hard in Mathematics because I want to do better on the exams than the others.
  - When we work on a project in Mathematics, I think that it is a good idea to combine the ideas of all the students in a group.
  - I make a real effort in Mathematics because I want to be one of the best.
  - I do my best work in Mathematics when I work with other students.
  - In Mathematics I always try to do better than the other students in my class.
  - In Mathematics, I enjoy helping others to work well in a group.
  - In Mathematics I learn most when I work with other students in my class.
  - I do my best work in Mathematics when I try to do better than others

- There are different ways of studying Mathematics.To what extent do you agree with the following statements?
  - When I study for a Mathematics test, I try to work out what are the most important parts to learn
  - When I am solving Mathematics problems, I often think of new ways to get the answer.

- When I study Mathematics, I make myself check to see if I remember the work I have already one.
- When I study Mathematics, I try to figure out which concepts I still have not understood properly.
- I think about how the Mathematics I have learnt can be used in everyday life.
- I go over some problems in Mathematics so often that I feel I could solve them in my sleep.
- When I study for Mathematics, I learn as much as I can by heart.
- I try to understand new concepts in Mathematics by relating them to things I already know.
- In order to remember the method for solving a Mathematics problem, I go through examples again and again.
- When I cannot understand something in Mathematics, I always search for more information to clarify the problem.
- When I am solving a Mathematics problem, I often think about how the solution might be applied to other interesting questions.
- When I study Mathematics, I start by working out exactly what I need to learn.
- To learn Mathematics, I try to remember every step in a procedure.
- When learning Mathematics, I try to relate the work to things I have learnt in other subjects.

Respondents have to tick one answer on a 4 points scale, labeled as: 1 'strongly agree', 2 'agree', 3 'disagree' and 4 'strongly disagree'. We recoded the variables in ascending order of positive attitude.

C Robustness

C.1 Are Eastern and Western Schools Different?

Table A5: Teaching Methods per School - PISA-E 2003 Germany

	(1)	(2)	(3)	(4)
Panel A: Organization of Classes and Abilities				
East	0.13	-0.05	0.16	-0.05
	(0.14)	(0.12)	(0.12)	(0.15)
Observations	1175	1164	1171	1155
Panel B: Teachers and Innovation				
East	0.13	-0.12	0.05	
	(0.11)	(0.11)	(0.12)	
Observations	1191	1191	1191	
Panel C: Level of Difficulty				
East	0.30**	-0.09	0.03	
	(0.12)	(0.12)	(0.10)	
Observations	1187	1188	1188	
Panel D: Cognitive vs Non-Cognitive Skills				
East	0.13	-0.37***	0.18*	
	(0.11)	(0.12)	(0.11)	
Observations	1194	1189	1190	

* p < 0.1, ** p < 0.05, *** p < 0.01. The data comes from the German National Evaluation of Pisa. School level regressions. School controls: Size of location area, share of girls in the school and Länder fixed-effects.

Panel A: Which of the following options describe what your school does for 15-year-old students in Mathematics classes? 1-3 scale. Column 1: Mathematics classes study similar content, but at different levels of difficulty; Column 2: Different classes study different content or sets of Mathematics topics that have different levels of difficulty; Column 3: Students are grouped by ability within their Mathematics classes; Column 4: In mathematics classes, teachers use a pedagogy suitable for students with heterogeneous abilities (i.e. students are not grouped by ability). *Panel B*: How much do you agree with these statements about innovation in your school? 1-4 scale. Column 1: Mathematics teachers are interested in trying new methods and teaching practices; Column 2: There is a preference among Mathematics teachers to stay with well-known methods and practices; Column 3: There are frequent disagreements between innovative and traditional Mathematics teachers. *Panel C*: How much do you agree with these statements about teachers expectations in your school? 1-4 scale. Column 1: There is a consensus among Mathematics teachers that academic achievement must be kept as high as possible; Column 2: There is a consensus among Mathematics teachers that it is best to adapt academic standards to the students' level and needs; Column

3: There are frequent disagreements between Mathematics teachers who consider each other to be too demanding or too lax. *Panel D*: How much do you agree with these statements about teaching goals in your school? 1-4 scale. Column 1: There is consensus among Mathematics teachers that the social and emotional development of the student is as important as their acquisition of Mathematical skills and knowledge in Mathematics classes;Column 2: There is consensus among Mathematics teachers that the development of Mathematical skills and knowledge in students is the most important objective in Mathematics classes; Column 3: There are frequent disagreements between Mathematics teachers who consider each other as too focused on skill acquisition or too focused on the affective development of the student.

C.2 Are the Results Driven by Different Economic Conditions?

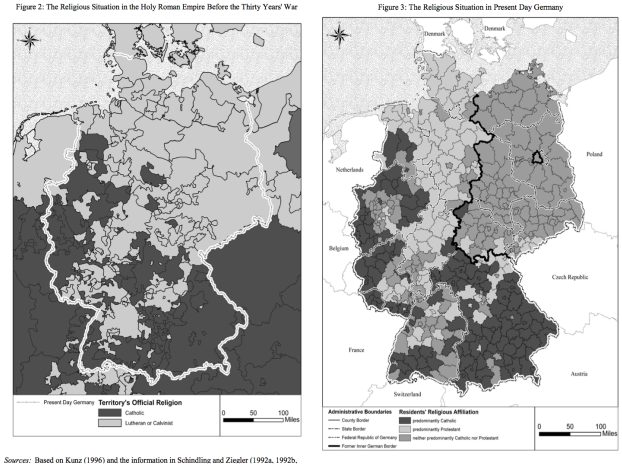
Table A6: Economic Conditions and the Gender Gap in Mathematics - PISA-E 2003. Within West Germany

	(1) Global	(2) Space and Shape	(3) Relationships and Change	(4) Uncertainty	(5) Quantity	(6) Problem Solving
Panel A	*Indicator of economic conditions : Rate of unemployment*					
Female	-31.09***	-33.03***	-33.25***	-40.29***	-22.07***	-19.92***
	(1.85)	(2.33)	(2.37)	(2.14)	(2.15)	(2.13)
Unemployment Rate	-3.98***	-3.67***	-4.65***	-4.11***	-3.88***	-4.10***
	(0.36)	(0.38)	(0.42)	(0.38)	(0.38)	(0.38)
Unemployment Rate*Female	0.04	-0.11	0.29	0.34	-0.16	0.50
	(0.26)	(0.32)	(0.32)	(0.29)	(0.31)	(0.32)
Observations	15886	15886	15886	15886	15886	15886
Panel B	*Indicator of economic conditions : Wealth*					
Female	-30.73***	-33.07***	-31.75***	-38.20***	-23.04***	-16.07***
	(1.29)	(1.64)	(1.65)	(1.52)	(1.56)	(1.44)
Wealth Lander	99.18	111.90	157.67*	109.01	62.33	79.68
	(67.02)	(69.82)	(82.00)	(69.70)	(66.60)	(67.68)
Wealth Lander*Female	-2.01	6.08	-8.86	-2.44	-6.55	12.65
	(8.39)	(10.73)	(9.67)	(9.55)	(9.34)	(11.24)
Observations	15886	15886	15886	15886	15886	15886

* p < 0.1, ** p < 0.05, *** p < 0.01. The data comes from the German National Evaluation of Pisa. The sample is restricted to individuals who were born in Germany. Individual controls: a quadratic in age, class, a quadratic in wealth, education and employment status of both parents. School controls: Size of location area and share of girls in the school and Länder fixed-effects. The rate of unemployment is computed from the answers given by the students on their parent's job status and aggregated at the Länder level. The variable wealth is developed by the OECD based on the answers given by the students on the size of their house, the equipment of their house, the number of books that their parents have. We use this variable and aggregate the answers given by the students at the Länder level. This variable is scaled from -6.17 to 3.15, a higher score meaning a higher plausible family wealth.

C.3 Are the Results Driven by Different Religion?

Figure A1: The Geographic Distribution of Religious Affiliation in Germany, in the 17th and 20th Century



Source: Spenkuch (2017).

Table A7: Religion and the Gender Gap in Mathematics - GSOEP Self-Declared Grade. Within West Germany

Dependent Variable: High Self-Declared Math Grade (0-1 scale)			
	(1)	(2)	(3)
Female	-0.06**	-0.07***	-0.07***
	(0.03)	(0.02)	(0.01)
Share of Protestants	-0.29		
	(0.23)		
Female * Share of Protestants	-0.04		
	(0.16)		
Share of Catholics		0.24	
		(0.19)	
Female * Share of Catholics		0.02	
		(0.12)	
Protestant/Catholic Ratio			-0.01
			(0.00)
Female * Protestant/Catholic Ratio			0.00
			(0.00)
Observations	11020	11020	11020

* p < 0.1, ** p < 0.05, *** p < 0.01. Linear probability model. The data comes from the German Socio-Economic Panel. Standard errors clustered at the household level are given in parentheses. Sample restricted to individuals born before 1971. East=1 if the household head lived in the GDR before 1990. Controls: age, household Income and Länder fixed-effects. Question: Can you remember your last report card ? What grade did you have in mathematics (1-6 scale with 1 being the highest grade)? Answer originally on a 1-6 scale recoded on a 0-1 scale, 1 corresponding to a grade of 1 or 2. The share of religious individuals is computed from the answers given by individuals and aggregated at the Länder level.

D PISA Europe

D.1 Country List

We examine the gender gap in mathematics in Western and Eastern Europe. We selected all the countries who participated at least once in the PISA international assessments from 2000 until 2012. Countries included in the Western group are: Austria, Belgium, Denmark, Finland, France, Iceland, Ireland, Italy, Luxembourg, Netherlands,

Norway, Portugal, Spain, Switzerland, Sweden and United-Kingdom. Countries included in the Eastern group are: Albania, Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Macedonia, Moldova, Montenegro, Poland, Romania, Serbia, Slovakia, Slovenia, Russia and Yugoslavia.

D.2 Reading Results

Table A8: Reading Scores - PISA Europe 2000 to 2012

Dependent Variable: Global Reading Score	
	(1)
Female	33.06***
	(0.39)
Former Socialist	-36.90***
	(3.65)
Female*Former Socialist	10.41***
	(0.67)
Observations	724784

* p < 0.1, ** p < 0.05, *** p < 0.01. The data comes from the 2000, 2003, 2006, 2009 and 2012 waves of PISA. Controls: quadratic in age, class, quadratic in household wealth, education of both parents as well as country and year fixed effects. Sample restricted to national born individuals.

D.3 Quantile Regression

We ran quantile regressions and computed the p-value from a Wald Test between the 10th and 90th percentile. As compared to the 10th percentile, the gender gap in math score is significantly greater at the 90th percentile (p-value = 0.000 with a t-stat of 19.55). The attenuation of this gap, related to the coefficient $Female * East$ in Eastern European countries is significant at every percentile but slightly smaller at the 90th percentile as compared to the 10th one (p-value=0.093).

Table A9: Maths Scores by Quantiles - PISA Europe 2000-2012

Dependent Variable: Global Mathematics Score					
	(1)	(2)	(3)	(4)	(5)
Quantile	10	30	50	70	90
Female	-9.81***	-14.34***	-17.58***	-20.31***	-22.39***
	(0.48)	(0.38)	(0.32)	(0.36)	(0.51)
East	-1.71	-5.31**	-8.96***	-9.59***	-7.47***
	(3.39)	(2.58)	(2.37)	(2.4)	(2.62)
Female*East	8.68***	10.9***	11.01***	10.75***	7.19***
	(0.71)	(0.59)	(0.53)	(0.62)	(0.75)
Observations	726142	726142	726142	726142	726142

* p < 0.1, ** p < 0.05, *** p < 0.01. The data comes from the 2000, 2003, 2006, 2009 and 2012 waves of PISA. Standard errors are calculated with a bootstrap method. Controls: quadratic in age, class, quadratic in household wealth, education of both parents as well as country and year fixed effects. Sample restricted to national born individuals.