The Link between Gender Gaps in School Enrollment and School Achievement

ISA STEINMANN & LESLIE RUTKOWSKI

This is the postprint version of the manuscript published in *Comparative Education Review* in August 2023: https://doi.org/10.1086/725395

Author Notes

We have no conflicts of interest to disclose. Correspondence concerning this article should be addressed to Isa Steinmann, Department of Primary and Secondary Teacher Education, Oslo Metropolitan University, Pilestredet 42, 0167 Oslo, Norway, isa.steinmann@oslomet.no, ORCID: https://orcid.org/0000-0002-9940-4413. We thank Roisin Cronin for copyediting the manuscript.

Abstract

Usually, gender gaps in school enrollment and achievement are investigated as separate facets of gender inequality in education. This study suggests that they might be linked by design. If boys and girls differ in their school enrollment rates, the student population will be gender biased. Since out-of-school adolescents tend to be less advantaged than in-school ones, school-based large-scale assessments effectively compare a more fully represented group of one gender with a less represented and more advantaged group of the other one. This should shift student gender achievement gaps so that they favor the latter group. In country-level regression models using data from all PISA cycles, we indeed found evidence of a small, negative association between gender gaps in secondary school enrollment and gender gaps in student achievement. This finding is robust across different achievement domains and specifications. We discuss the study's limitations and implications for applied research with PISA data.

The Link between Gender Gaps in School Enrollment and School Achievement

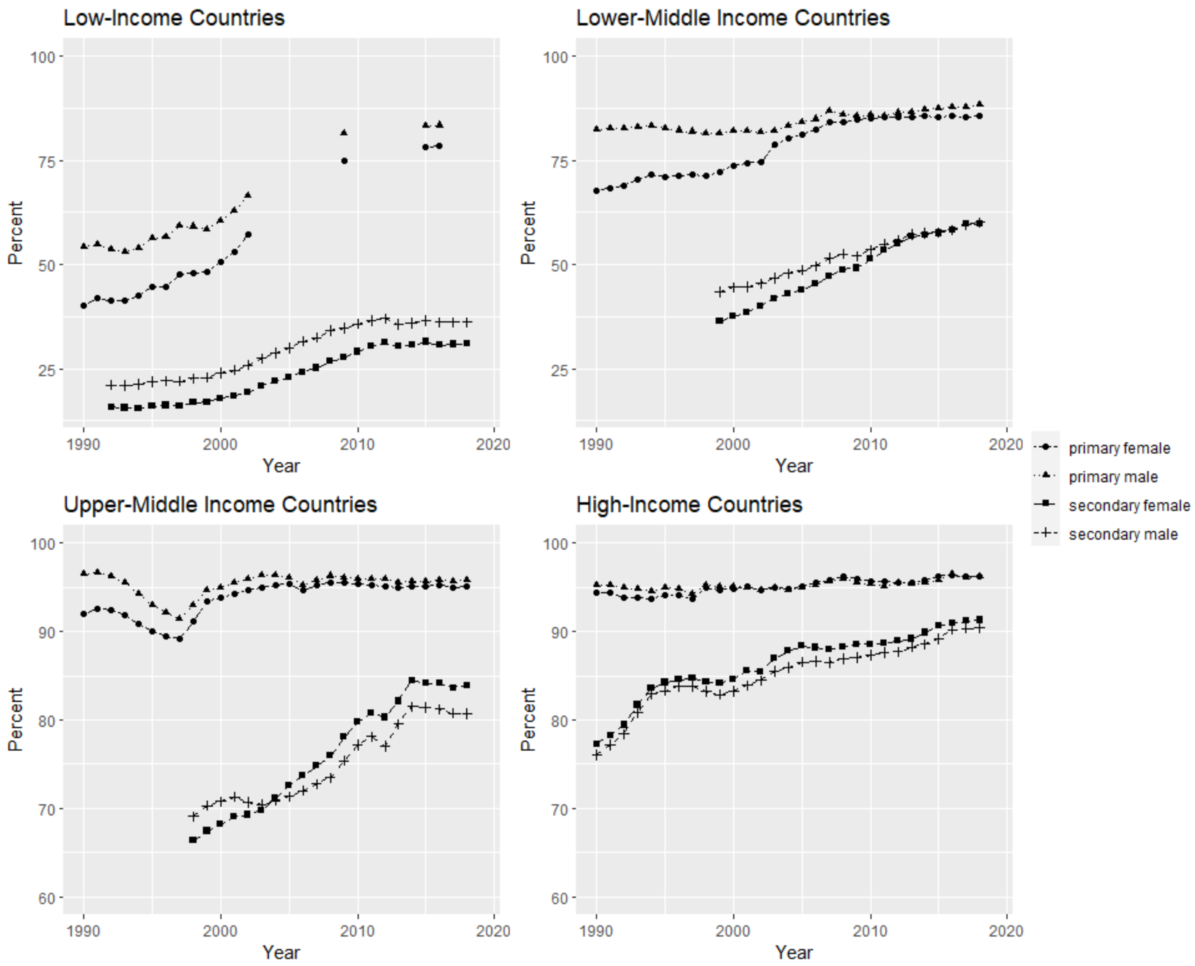
In the field of education, gender differences in school enrollment and in school achievement represent two key facets of gender inequality. Both are crucial, since school enrollment is a prerequisite for participating in formal education and school achievement is intrinsically linked to students' actual learning success. These two gender gaps are also addressed in the United Nations' Sustainable Development Goal of quality education, which promises to "[b]y 2030, ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes" (United Nations 2015, 19). In previous research, gender gaps in enrollment and achievement were usually investigated as two separate dimensions of gender-related educational inequalities. As outlined below, this study focuses on the link between these gender gaps.

Global School Enrollment Gender Gaps at the Primary and Secondary School Level

Since being enrolled in school is crucial for learning, school attainment, societal participation, and access to the labor market, policy makers across the world have long sought to increase enrollment rates and reduce gender differences in enrollment (e.g., OECD 2015; Quintini and Martin 2014; UNESCO 2019; World Economic Forum 2019). As illustrated in Figure 1, these political efforts appear to have been successful: Primary and secondary school enrollment rates increased over the last 30 years in lower and higher income countries. For the most part, gender gaps in school enrollment have also decreased. In low- and lower-middle-income countries, however, boys have retained average enrollment advantages at primary school level; in low-income countries this also applies at the secondary school level. Across upper-middle and high-income countries, boys' and girls' primary school enrollment rates are currently very high and almost equal. At the secondary school level, girls have higher average enrollment rates than boys, although the difference is small across high-income countries.

Figure 1

International Primary and Secondary School Enrollment Trends for Boys and Girls



Note. Illustration of net enrollment ratio data as reported by the UNESCO Institute for Statistics (2021) for all countries in the low-, lower-middle, upper-middle, and high-income groups, as classified by the World Bank.

Although enrollment rates have increased overall, not being enrolled in school is still more common in lower income countries (see also UNDP 2019; World Economic Forum 2019). Non-enrollment or early dropout are much more prevalent among socioeconomically disadvantaged children and adolescents in low-income (Inoue et al. 2015; Kazeem, Jensen and Stokes 2010; UNDP 2019), middle-income (Cárdenas, Hoyos and Székely 2015; Cardoso and Verner 2006; Hannum 2003; OECD 2020), and high-income countries

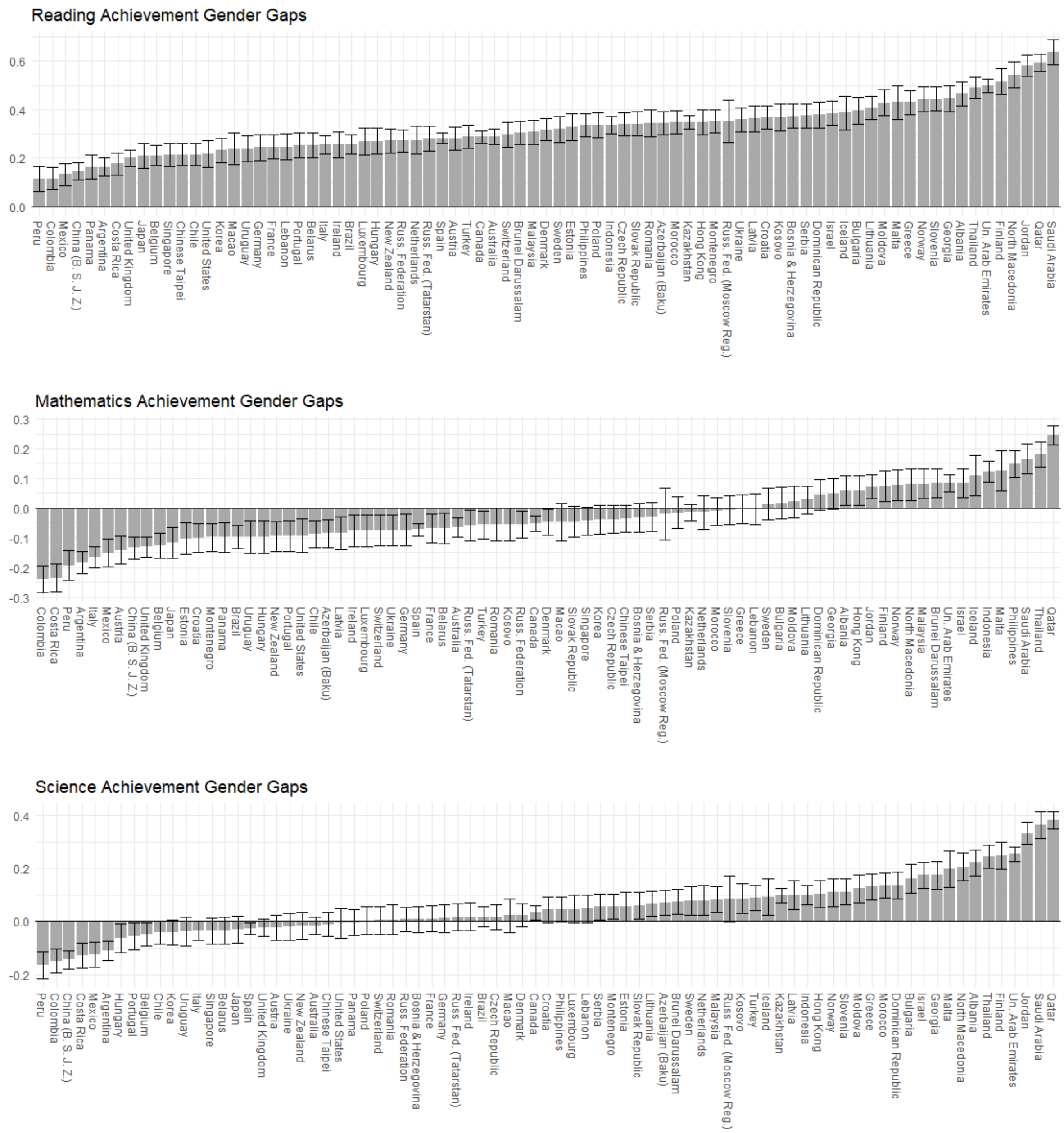
(Archambault et al. 2017; Hippe and Jakubowski 2018; Lavrijsen and Nicaise 2015; Lutz 2007). The reasons for dropping out of secondary school are manifold. For instance, adolescents might have to work to support the family instead of going to school, especially in low-income countries and families (Inoue et al. 2015; Kazeem, Jensen and Stokes 2010). Furthermore, low perceived returns of education appear to increase the risk of early dropout (Buchmann and Brakewood 2000; Cárdenas, Hoyos and Székely 2015; Cardoso and Verner 2006; Hippe and Jakubowski 2018).

Gender Gaps in School Achievement in International Large-Scale Assessment Studies

Comparative studies find large between-country differences in gender gaps in school achievement worldwide, both at primary and secondary school level. In reading, girls usually achieve higher scores than boys in almost all participating countries; this finding is especially evident in the Programme for International Student Assessment (PISA) study (OECD 2019b), which was conducted at secondary school level, but also emerges in studies conducted at the primary school level (Mullis et al. 2017). In mathematics and science, findings are mixed between boys and girls, with inconsistent gender gaps at both primary and secondary school level (Mullis et al. 2020; OECD 2019b). Figure 2 illustrates the standardized mean differences between girls' and boys' achievement scores in the three domains in PISA 2018. Internationally, these gender gaps are rather stable over time (Mullis, Martin and Loveless 2016; OECD 2009; OECD 2015; Steinmann, Strietholt and Rosén forthcoming) and correlate with each other at the country level (Guiso et al. 2008; Marks 2008; Stoet and Geary 2013). Countries with pronounced reading advantages of girls tend to show mathematics and science gender gaps to the girls' advantage, while countries with small reading gaps tend to show boys' advantages in mathematics and science (ibid.). Likewise, gender gaps at the primary and secondary school level correlate at the country level (Mullis, Martin and Loveless 2016; Steinmann, Strello and Strietholt under review). The gender gaps in achievement do not vary systematically between regions of the world. However, more pronounced male advantages are often observed in Middle and South American countries (see countries on the left-hand side in Figure 2) and more pronounced female advantages are often observed in Middle Eastern and Northern European countries (see countries on the right-hand side in Figure 2).

Figure 2

Gender Gaps in Student Achievement in PISA 2018



Note. Gender gaps are depicted as the standardized mean differences of girls' and boys' achievement scores and their confidence intervals in all countries or regions that participated in PISA 2018 (cf. OECD 2019b). Positive values imply that girls have higher mean scores and negative values indicate that boys have higher mean scores.

These results suggest that the international variation in gender achievement gaps is due to country characteristics rather than the properties of the large-scale assessment tests used or the time at which they are conducted. Previous research investigated associations between achievement gender-gap variation and numerous country-level factors (see literature review by Rosén, Steinmann and Wernersson 2022). For the most part, previous studies have not found consistent associations with the countries' economic situations (Guiso et al. 2008; Hamamura 2012; Hermann and Kopasz 2019; Tao and Michalopoulos 2018). As is also evident in Figure 2, pronounced gender gaps occur in countries with both relatively high income and low income levels. Other studies investigated the role of policy differences in explaining gender gaps, like early and late between-school tracking (Hermann and Kopasz 2019; Steinmann, Strello and Strietholt under review) or the existence of central exams and other sorts of standardization measures (Ayalon and Livneh 2013). Furthermore, some previous research has focused on gender-related cultural differences between countries (Else-Quest, Hyde and Linn 2010; Guiso et al. 2008; Hamamura 2012; Nosek et al. 2009; Penner 2008; Reilly 2012; Stoet and Geary 2013; Tao and Michalopoulos 2018). In the present study, we suggest that another type of country characteristic might explain between-country variation in gender achievement gaps.

Linked by Design: Gender Gaps in School Enrollment and School Achievement

The reasons for and correlates of both gender differences in school enrollment and school achievement can be assumed to be multifaceted and complex (e.g., historical, political, cultural, and economic factors; see e.g., Connell 2002; Maccoby 1998; Maccoby and Jacklin 1974). In this study, we suggest that they might also be linked by design. If boys and girls differ in their school enrollment rates in a given country, the country in question will necessarily have a gender-biased student population. Since disadvantaged children and adolescents will likely be overrepresented in out-of-school groups, as mentioned above (e.g., Lavrijsen and Nicaise 2015; OECD 2020; UNDP 2019), we can expect this gender bias in the student population to translate into gender achievement gaps in school-based assessments. In countries where more girls are enrolled than boys, the male student population should be represented by fewer and on average more advantaged boys. In other words, the male student population should be more positively selected than the female one. Since school-based assessments are conducted on representative samples of *students* and not on out-of-school children and adolescents, this implies that such assessments compare the achievement scores of almost all girls to those of a somewhat more privileged subset of boys. This should shift gender achievement gaps to the advantage of male students. By contrast, in countries with more enrolled boys than girls, we would expect to find an opposite shift in favor of female students.

To the best of our knowledge, no previous study has empirically investigated this hypothetical relationship between school enrollment and gender achievement gaps. Salchegger and Suchań (2018) did, however, follow the same argument in a simulation analysis. They proceeded from the observation that boys had a large mean advantage over girls in the PISA 2015 mathematics test in Austria and suggested that this might relate to the higher proportion of male out-of-school teenagers compared to female ones. Ultimately, they concluded that about one third of the mathematics gender gap could be explained by this gender difference in secondary school enrollment in Austria. Hermann and Kopasz (2019) also suspected that enrollment gender gaps might be associated with gender achievement gaps; they hence excluded countries where the total enrollment ratio was below 90%. Other studies also investigated associations between school enrollment and gender achievement gaps but expected to find positive associations. These studies interpreted the enrollment gender gap as one of many indicators of societal gender inequalities, similar to the gender wage gap indicator (e.g., Else-Quest, Hyde and Linn 2010; Penner 2008).

Apart from the gender gap issue, many studies have discussed the potential role of out-of-school populations for educational outcomes in international large-scale assessments (e.g., Hanushek and Woessmann 2008; OECD 2019a; Spaull and Taylor 2015; Taylor and Spaull 2015). For instance, the PISA reports contain league tables that contrast the countries' achievement distributions with the percentage of 15-year-olds who are not covered in the sampling frame (see e.g., Figure I.6.1 in OECD 2019a). These tables also contain adjusted estimates of achievement, indicating what would have occurred if the out-of-school adolescents had taken the tests. Assuming that out-of-school groups would have performed less well than average students (see, e.g., Figure I.9.5 in OECD 2019a), overall achievement levels would, by design, have been lower in countries with higher enrollment rates. The underlying argument, namely that the exclusion of on average disadvantaged out-of-school adolescents would impact on PISA estimates, is therefore similar to the one in our study. However, PISA does not report or address the coverage indicators separately for boys and girls or other social groups.

The Present Study

This study empirically investigated whether there is an association between gender gaps in secondary school enrollment rates and gender gaps in the mean reading, mathematics, and science achievement scores of 15-year-old students, as measured in PISA. We expected to find the following relationship: The higher the proportion of girls enrolled in secondary school, the more the gender achievement gaps shift to advantage male students, and vice versa.

We focused on secondary school enrollment and gender achievement gaps, since international comparative assessments mostly include upper-middle and high-income countries. In these countries, almost all primary-school-age children attend school, regardless of their gender (see Figure 1). We used PISA data, since it assesses 15-year-old students, who are usually legally permitted to leave school. Therefore, the PISA samples may exclude considerable proportions of out-of-school adolescents. Another international large-scale assessment study (Trends in International Mathematics and Science Study (TIMSS)), in contrast, samples grade 8 students. Attendance at school is usually still compulsory in grade 8 in upper-middle and high-income countries. We supplemented the PISA data with data from the United Nations Educational, Scientific and Cultural Organization (UNESCO) on boys' and girls' secondary school enrollment ratios.

Method

Sample

PISA has assessed representative samples of adolescent students in three-year cycles since 2000. In a two-stage stratified design, PISA first sampled educational institutions (e.g., schools, vocational training programs) in country-specific strata from all educational institutions that might be attended by 15-year-olds (either part or full time). In a second step, the study sampled 15-year-old students who attend grade 7 or higher in the selected schools

(OECD 2019c).[1] In order to maximize the number of observations, we included both countries and regions (e.g., United States and United States (Massachusetts) in 2015); we also used all available PISA cycles and incorporated the school-based part of the *PISA for Development* assessment.

Across all cycles, data for 439 country-by-year observations were available. In the first step, we excluded 12 observations for which serious data quality concerns were raised in the technical documentation, and where the OECD did not fully recommend using the data. In the second step, we excluded five observations from Liechtenstein, a European microstate with very small sample sizes.[2] Third, we excluded 86 observations for which no secondary school enrollment data were available (see measures section below). All included and excluded countries are listed in Appendix A. We therefore included 336 observations for countries or regions that participated in one or more of the PISA cycles between 2000 and 2018. For all but one country, data for all three achievement domains—reading, mathematics, and science—were available.[3] Data pertained to 83 unique countries or regions.

Measures

Gender Gaps in Student Achievement

Per country-by-year observation, we estimated the mean gender gaps in student achievement based on the five to ten plausible values for all students with gender information. The rate of missing gender values was negligible (0–2% in the country-by-year observations). The gender variable was assessed in the student questionnaires, where the adolescents were asked whether they were female or male. Consistent with the OECD

[1] Specifically, PISA samples students who are between 15 years and 3 months up to 16 years and 2 months old at the time of the international testing period. Countries can make exclusions from the target samples at the regional level (e.g., very remote areas), at the school level (e.g., special needs schools), and at the student level (e.g., insufficient mastery of test language).

[2] In Liechtenstein, the sample sizes ranged between 175 and 339 students in the PISA cycles. Therefore, even a few dozen individual boys or girls who attended school across the border in Austria or Switzerland could have distorted school enrollment statistics. In the other countries, the PISA sample sizes are four-digit.

[3] Due to data quality issues, PISA did not publish the reading data for the United States in 2006.

(2019c), we refer to this variable as *gender* instead of *sex*. In the first step, we estimated the girls' and boys' mean achievement scores and their standard deviations in reading, mathematics, or science per country-by-year observation. We applied Rubin's (1987) rules to combine the plausible values and used student sampling weights to account for the complex stratified sampling design. The means and standard deviations conform to those in the PISA international reports. The datasets are freely available online, along with extensive documentation (OECD 2021).

In the second step, we subtracted the mean achievement scores for boys from those of girls and divided the resulting figure by the pooled standard deviation for boys and girls. The resulting achievement gap values can be interpreted as Cohen's *d* effect size measures (Cohen 1988). A value of zero implies that boys' and girls' mean achievement scores are the same in a given country, a value of +1 implies that girls score 1 pooled standard deviation higher than boys, and a value of -1 implies that boys score 1 pooled standard deviation higher than girls. Note that we used within-country standard deviations and not the international standard deviation.

Gender Gaps in School Enrollment

We determined the gender gaps in school enrollment using data from the UNESCO Institute for Statistics. These annual statistics define boys' and girls' net enrollment as the "total number of students of the official age group for a given level of education who are enrolled in any level of education, expressed as a percentage of the corresponding population" (UNESCO Institute for Statistics 2021). The net enrollment rate may be lower than 100% if students never enroll in school or if they drop out early. This data is freely available online, alongside further documentation (UNESCO Institute for Statistics 2021).

We computed the gender gap in school enrollment at secondary school level as the percentage of enrolled girls minus the percentage of enrolled boys. A value of zero implies that boys and girls have the same school enrollment ratio. A (theoretical) value of +1 implies that 100% of girls but 0% of boys are enrolled; a (theoretical) value of -1 implies the opposite. A value of 0.01 hence reflects a 1% difference in enrollment rates in favor of girls.

As mentioned above, we only included a country-by-year observation when we could also find the associated enrollment gender gap information. In some instances, there was a substantial amount of missing data, at least for some years. Since the enrollment ratios could be expected to change, albeit slowly, over time (cf. Figure 1), we decided to pool the enrollment information in five-year intervals (2000–2004, 2005–2009, etc.) before merging them with the achievement datasets. This allowed us to strike a balance between including as many country observations as possible and using adequate school enrollment information. It meant that the enrollment gender gap value was identical for country observations within these five-year intervals, for instance, in PISA 2000 and 2003. Note that the enrollment information was not reported for separate regions—these were assigned the same enrollment gender gap value as the countries they belonged to.

Main Analyses

In order to estimate the association between gender gaps in achievement and enrollment, we ran two kinds of country-level regression analyses. In a first step, we simply regressed the gender achievement gaps on the enrollment gender gaps in separate models for reading, mathematics, and science achievement. By implication, every country-by-year observation had the same weight. In the second step, we added three control variables to the regression models: the countries' gross domestic product (GDP) per capita, as reported by the World Bank (2022a)[4], the assessment cycle in question (2000 to 2018)[5], and the mean

[4] As for the UNESCO enrollment data, we pooled the GDP data in five-year intervals (2000–2004, 2005–2009, etc.) per country to minimize the number of missing values. This way, the GDP data was complete for the 335 respectively 336 observations in PISA. In TIMSS, the GDP variable was missing for Syria in 2011.

[5] The school-based *PISA for Development* assessment was conducted in 2017. For the sake of simplicity, we added the observations to the main PISA assessment in 2018.

achievement levels (*z*-standardized). We included these control variables because gender gaps in school enrollment cannot be assumed to vary randomly between countries but instead differ by their economic situation and over time (see Figure 1). Since these factors can be expected to relate to gender gaps in student achievement (see above), this control strategy aims to address confounders of the association between enrollment and achievement gaps.[6] We also ran two types of robustness analyses to make sure that the results were not driven by violations of linear regression assumptions.

Further Analyses

In the main analyses, we regressed student achievement on school enrollment gender gaps; these were purely correlational in nature. By extension, this meant that other unanticipated mechanisms might be in place. Our argument stated that if one group (boys or girls) was more likely to leave school early in a given country, the members of this group who stayed in school would be numerically fewer and more advantaged than those of the other group. This would lead to gender achievement gap shifts in favor of the smaller, selected group. However, another explanation for the association could be that general gender-related cultural, political, or societal factors might affect both gender gaps in enrollment and achievement (i.e., third variable effect).

Although our design did not allow us to draw causal conclusions about these mechanisms, we ran three further analyses to provide additional evidence. First, we tested to determine whether gender gaps in school enrollment were positively associated with gender gaps in the sample sizes. We assumed that if boys or girls were underrepresented in the secondary school student population because they dropped out of school earlier on average,

[6] Another approach to address country-level confounders would have been to estimate country-fixed effects, which would base the estimation of the association between enrollment and achievement gaps on within-country comparisons. However, our limited sample size of 336 observations of 83 unique countries or regions (i.e., approximately four observations per country or region) prohibited the estimation of country-fixed effects in the present study.

they would have also been underrepresented in the student samples in PISA. In the two-stage sampling design in PISA, boys and girls should have had the same sampling probability if they were equally represented in the student population. We therefore expected to find a positive correlation between gender gaps in school enrollment and gender gaps in PISA sample sizes. We calculated the sample-size gender gap as the difference between the weighted number of girls and boys in the samples.

Second, we contrasted the main PISA findings with findings based on the TIMSS grade 8 assessment. While the 15-year-old students in PISA were typically no longer legally required to be enrolled in school, the opposite largely applied to the grade 8 students examined in TIMSS. We thus expected the association between secondary school enrollment and gender achievement gaps be less pronounced in TIMSS than in PISA. As a result, we replicated the main analyses for 161 country-by-year observations in the TIMSS grade 8 assessments between 2003 and 2019.[7] We regressed gender achievement gaps in mathematics and science on enrollment gender gaps with and without additional control variables.

Third, we extended the main PISA analyses in a model that included indicators for societal gender inequalities at the country level. We used the four sub-indices of the Gender Gap Index (GGI)[8] as reported by the World Bank (2022b) on political empowerment, economic participation and opportunity, educational attainment, and health and survival. These indices are assessed annually since 2006 and range between 0 (gender inequality) and 1 (gender equality). The aim of these analyses was to check whether the association between achievement and enrollment gaps was confounded with gender-related societal factors (i.e.,

[7] Just as for PISA, we excluded countries with flagged data quality issues (i.e., sample participation rates lower than 77% and/or national target populations not reflecting full international target populations (e.g., private schools or schools with certain instruction languages excluded)) and observations for which no secondary school enrollment information was available.

[8] As for the UNESCO enrollment data, we pooled the GGI subindex information in five-year intervals (2005–2009, 2010–2014, etc.) per country to minimize the number of missing values. In the PISA cycles since 2006, the GGI subindex data is missing for 12 observations.

third variable effect).

Results

Descriptive Findings: Gender Gaps in Student Achievement and School Enrollment

As depicted in Table 1, the gender gaps in reading, mathematics, and science varied considerably across the 335 or 336 country-by-year observations used. In reading, the gender gaps ranged between almost non-existent to an advantage of almost one standard deviation in favor of girls. In mathematics, the gender gaps ranged between almost half a standard deviation in favor of boys to more than a quarter of a standard deviation in favor of girls. The science gender gap ranged between about one quarter of a standard deviation in favor of boys to more than half a standard deviation in favor for girls. There were also variations in the gender gaps in secondary school enrollment rates between the country-by-year observations (see Table 1): These ranged from a 14-percentage-point advantage of boys over girls to an 11-percentage-point lead of girls over boys.

Table 1

Distributions of Achievement and Enrollment Gender Gaps

		Distribution of gender achievement gaps				Distribution of enrollment gender gaps			
Domain	*N* country-by-year obs.	Min	*M*	*SD*	Max	Min	*M*	*SD*	Max
Reading	335	0.02	0.37	0.14	0.90	-0.14	0.02	0.03	0.11
Mathematics	336	-0.44	-0.08	0.11	0.27	-0.14	0.02	0.03	0.11
Science	336	-0.26	0.01	0.12	0.53	-0.14	0.02	0.03	0.11

Note. Achievement gaps are standardized mean differences, with positive values indicating higher mean scores of girls and negative values pointing to higher mean scores of boys. Enrollment gaps are differences between the percentages of enrolled girls and boys; positive values imply that girls have higher enrollment rates than boys and negative values mean that boys have higher enrollment rates than girls.

Main Findings: Regression of Gender Gaps in Student Achievement on Gender Gaps in School Enrollment

In a first model, we regressed the student gender achievement gaps on the school enrollment gender gaps separately for reading, mathematics, and science achievement. Consistent with our expectations, we found a small but significant negative relationship between these variables in all three achievement domains (see Table 2). According to these models, increasing the enrollment gender gap by 10 percentage points in favor of female students (i.e., girls' enrollment rate 10 percentage points higher than the enrollment rate of boys) would shift the gender gap in student achievement by 5.2–6.4% of a standard deviation in the direction of male advantage. The regression coefficients were remarkably similar in the three achievement domains. Yet, these models explained less than 2% of the variation in the gender achievement gaps across the country-by-year observations, which implies that a lot of variance remained unexplained when regressing achievement on enrollment gender gaps.

In a second set of regression models, we controlled for the countries' GDP, for the assessment cycles to which the observations pertained, and for mean achievement levels in the respective achievement domains. As Table 2 shows, the enrollment gender gap regression coefficients were even more pronounced in these second models than in the first simple models. This finding emphasized that the analyses remained robust even after considering further characteristics of the country-by-year observations. In all three domains, the gender achievement gaps were not significantly associated with the countries' GDP. In reading, the gender gaps were more shifted to the advantage of girls in PISA 2006, 2009, and 2012 as compared to PISA 2000. In mathematics, the gender gaps were slightly more shifted to the advantage of girls in PISA 2018 as compared to PISA 2000. In science, the association between the assessment cycles and the gender achievement gap was not significantly different from zero. In reading, the mean achievement level was not significantly associated with the gender achievement gaps. In mathematics and science, higher mean achievement levels were associated with a slight shift in the gender achievement gaps to the advantage of male students.

Table 2

Results of Regressing Gender Achievement Gaps on Enrollment Gender Gaps

	Reading gender gap		Mathematics gender gap		Science gender gap	
Regression parameter	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)
Intercept	0.38***	0.36***	-0.07***	-0.09***	0.02***	0.02
	(0.01)	(0.03)	(0.01)	(0.02)	(0.01)	(0.02)
Enrollment gender gap	-0.64*	-0.86***	-0.52*	-0.61**	-0.60**	-0.86***
	(0.28)	(0.25)	(0.21)	(0.21)	(0.23)	(0.23)
GDP per capita		0.00		0.00		0.00
		(0.00)		(0.00)		(0.00)
Assessment cycle 2003		0.04		-0.01		-0.03
		(0.03)		(0.03)		(0.03)
Assessment cycle 2006		0.10**		0.00		0.02
		(0.03)		(0.03)		(0.03)
Assessment cycle 2009		0.13***		0.00		0.05
		(0.03)		(0.02)		(0.03)
Assessment cycle 2012		0.11***		0.01		0.04
		(0.03)		(0.02)		(0.03)
Assessment cycle 2015		-0.03		0.04		0.00
		(0.03)		(0.02)		(0.03)
Assessment cycle 2018		-0.02		0.06*		0.04
		(0.03)		(0.02)		(0.03)
Mean achievement		0.00		-0.01*		-0.02**
		(0.01)		(0.01)		(0.01)
N country-by-year obs.	335	335	336	336	336	336
R^2	.016	.239	.019	.093	.020	.127

Note. Dependent variables are gender achievement gaps (positive values imply higher scores for girls and negative values imply higher scores for boys). Main independent variable is enrollment gender gap (positive values imply that girls have higher enrollment rates; negative values imply higher enrollment rates among boys). Reference category to assessment cycle is 2000. Mean achievement is *z*-standardized.

* $p < .050$, ** $p < .010$, *** $p < .001$

Findings of Robustness Checks

We evaluated whether the data fulfilled the prerequisites for linear regression models. The findings are summarized in Appendix B and support running simple linear regression models overall. In a first set of robustness analyses, however, we reran the simple regressions of achievement on enrollment gender gaps after removing country-by-year observations that were either extreme outliers in the scatterplots or flagged in the homoscedasticity or quantile-quantile plots (see Appendix B). Furthermore, we observed some autocorrelation of residuals (see Appendix B) and ran a second type of robustness check that additionally added autoregressive errors. As depicted in Appendix C, the results of these two robustness checks qualitatively replicated the main findings. In all three domains, the coefficients for regressing achievement on enrollment gender gaps were negative and significantly different from zero. The coefficients were at least as large or even more pronounced than in the main analyses, which suggests that the main findings are conservative estimates.

Findings of Further Analyses

As a first further analysis, we calculated the correlation between gender gaps in school enrollment and in the PISA sample sizes. Consistent with expectations, we found a moderate to large positive correlation of $r = .497$. Second, we replicated the main regression models for 161 country-by-year observations in TIMSS grade 8. We expected to find a less pronounced association between mathematics and science gender gaps and secondary school enrollment gaps in TIMSS than in PISA. Indeed, the coefficients did not differ significantly from zero in the TIMSS-based models with and without further controls (see Appendix D). However, the absence of a significant parameter cannot be interpreted as evidence of the *absence* of an association. Third, we extended the main PISA regression models with additional control variables for societal gender inequality, namely the GGI sub-indices on political empowerment, economic participation and opportunity, educational attainment, and health and survival. As depicted in Appendix E, these analyses replicated the main findings. We observed negative associations between achievement and enrollment gender gaps that were even more pronounced as compared to the main findings. Note that the number of observations was reduced because these indicators were not available before 2006 or for all observations. All three further analyses strengthen the conclusions from the main results.

Discussion and Conclusion

This study assumed that gender gaps in school enrollment and student achievement are not just two important facets of gender inequalities in education (see e.g., OECD 2015; Quintini and Martin 2014; UNESCO 2019; World Economic Forum 2019) but that they are also linked by design. We argued that if one group (boys or girls) is underrepresented in the secondary school student population because they leave school earlier on average, the country's student population will consist of fewer and more advantaged individuals from this group. Therefore, this group should score higher in achievement tests on average than the other group, which is more completely represented in the student population. Similar arguments have been made in previous studies (e.g., Salchegger and Suchań 2018), but the hypothesis has, to our knowledge, not yet been empirically investigated.

Using data from all PISA cycles and as many countries as possible, we investigated the relationship between enrollment and gender achievement gaps in country-level analyses. Consistent with our expectations, we found small negative associations between the gender gaps in secondary school enrollment and the gender gaps in reading, mathematics, and science achievement. In other words, in countries with more enrolled girls than boys, gender achievement gaps slightly shifted to the advantage of male students, and vice versa. These estimates were remarkably robust across the three achievement domains, across models that included further control variables, and when subjected to robustness analyses that accounted for extreme outliers and slight violations of linear regression assumptions.

However, cross-sectional regression approaches of the kind used in this study do not allow us to draw causal conclusions. Another explanation for the association could be that countries with high enrollment gender gaps have characteristics that also bear upon gender gaps in student achievement (i.e., third variable effects). Our finding that the enrollment gender gap variable explained rather small proportions of the gender achievement gap variance emphasized the importance of additional country characteristics. However, if country characteristics (e.g., culture, society, education system) explained gender gaps in both secondary school enrollment and achievement, we would have expected enrollment advantages to be linked to achievement advantages (see arguments in Else-Quest, Hyde and Linn 2010; Penner 2008) and not to disadvantages, as found in this study. Furthermore, we found that gender gaps in school enrollment and in the PISA sample sizes correlated positively. In countries where girls were overrepresented in the school population because boys dropped out earlier, the PISA samples tended to contain more girls than boys, and vice versa. This supports our argument of exclusion mechanisms. Furthermore, we found no significant association between enrollment and gender achievement gaps in TIMSS, which assessed students during compulsory schooling in grade 8. Although the absence of a significant coefficient cannot be interpreted as the absence of an association, we believe this provides some evidence in favor of our argument. In the middle and high-income countries we examined, the secondary school enrollment gender gap should matter more for the 15-year-olds in PISA, who can decide to leave school, than for grade 8 students in TIMSS, who are still legally required to enroll. Additionally, we added indicators for societal gender inequalities to the main regression models to check whether this would reduce the observed association between enrollment and gender achievement gaps. The results supported the main argument of the present study, however. We observed even more pronounced negative associations between enrollment and gender achievement gaps after controlling for societal gender inequalities in terms of political empowerment, economic participation and opportunity, educational attainment, and health and survival. However, we would like to acknowledge that if more observations were available per country, we could have run country-fixed effects models, which would estimate the association between enrollment and achievement gaps within countries. This would be an even more effective control strategy for this study since it would control for all economic or societal between-country differences. Since every country or region participated in only approximately four PISA cycles on average this strategy could not be applied.

In summary, the findings of our main, robustness, and further analyses seem to support the argument that gender gaps in enrollment and achievement are linked negatively by design through selection mechanisms. However, it cannot be ruled out that other mechanisms are at play, potentially at the same time. If it was indeed true that both enrollment and achievement gaps related to underlying societal gender inequalities, as discussed above, the negative selection mechanisms might be subdued by such positive associations through third variables. In this case, our main findings would *underestimate* the negative association between gender gaps in enrollment and achievement. This assumption is supported by the finding that the negative associations were more pronounced in further analyses that included controls for societal gender inequalities. This could be an explanation for the robustly negative but small effect sizes that we observed. Another explanation for the small effect sizes is that the range of the enrollment gender gap variable was limited in the observed countries. In cases with larger gender gaps in enrollment, larger differences in gender gaps in achievement might be observed. This could be the subject of follow-up research that utilizes future cycles of PISA and PISA for Development, which will include a more heterogeneous set of countries.

Apart from the cross-sectional design, we would like to acknowledge a number of other central limitations of this study. First and most importantly, the enrollment and gender achievement gap variables do not align perfectly. To address research questions like ours, it would be ideal if PISA directly provided group-specific population coverage information for 15-year-old students, not just the coverage information across all students. In the present case, this would imply reporting coverage rates of 15-year-olds overall and separately for boys and girls. Instead, both the PISA reports and this study rely on the UNESCO indicator of net enrollment for secondary school as a whole. Likewise, the UNESCO enrollment indicator is based on administrative data, which might overstate actual school enrollment (see argument in Hanushek and Woessmann 2008). Second, since our study focused on PISA and TIMSS grade 8 data, the findings are not generalizable to primary or upper secondary school or to nonparticipating countries. We would expect countries to only differ slightly in school enrollment gender gaps in large-scale primary-school assessments (e.g., TIMSS grade 4), since the enrollment rates for both primary school boys and girls are nowadays very high in upper-middle and high-income countries (see Figure 1). However, future research could utilize regional large-scale assessments from lower income areas of the world (cf. UNESCO 2018) to explore associations between enrollment and gender achievement gaps. Third, we only investigated gender gaps in mean achievement levels; future research might also investigate associations with gender gaps in top-performing students or in the variability of test scores, for instance (cf. Meinck and Brese 2019; Rutkowski, Rutkowski and Plucker 2012). Another area of interest would be to explore associations between enrollment differences between socioeconomically advantaged and disadvantaged groups and related achievement inequalities between these social groups.

In summary, this study found the expected link between higher enrollment rates among girls and a small shift in the achievement gap to the advantage of boys, and vice versa, in PISA. The extent to which this is rooted in the assumed selection mechanisms or in others cannot be determined with the present cross-sectional study and remains a question for future research. Nevertheless, our study suggests that it might be important to consider the implications of out-of-school adolescents for between-group comparisons with PISA data (cf. OECD 2019a). This might be especially relevant in light of pandemic-related increases in (temporary) out-of-school children and adolescents (UN 2020) in future assessments. In this vein, our study does not only contribute to understanding between-country differences in student gender achievement gaps but also to the discourse on the scope of school-based large-scale assessments (see e.g., OECD 2020; Rutkowski and Rutkowski 2021). A general conclusion that can be drawn from this study is that “it is important to see test scores in the context of participation rates” (OECD 2020, 22).

References

Archambault, Isabelle, Michel Janosz, Véronique Dupéré, Marie-Christine Brault, and Marie Mc Andrew. “Individual, Social, and Family Factors Associated with High School Dropout Among Low-SES Youth: Differential Effects as a Function of Immigrant Status.” The British Journal of Educational Psychology 87, no. 3 (2017): 456–477.

Ayalon, Hanna, and Idit Livneh. “Educational Standardization and Gender Differences in Mathematics Achievement: A Comparative Study.” Social Science Research 42, no. 2 (2013): 432–445.

Buchmann, Claudia, and Dan Brakewood. “Labor Structures and School Enrollments in Developing Societies: Thailand and Kenya Compared.” Comparative Education Review 44, no. 2 (2000): 175–204.

Cárdenas, Mauricio, Rafael de Hoyos, and Miguel Székely. “Out-of-School and Out-of-Work Youth in Latin America: A Persistent Problem in a Decade of Prosperity.” Economía 16, no. 1 (2015): 1–40.

Cardoso, Ana R., and Dorte Verner. “School Drop-Out and Push-Out Factors in Brazil: The Role of Early Parenthood, Child Labor, and Poverty.”. IZA Discussion Papers 2515. https://www.iza.org/publications/dp/2515/school-drop-out-and-push-out-factors-in-brazil-the-role-of-early-parenthood-child-labor-and-poverty.

Cohen, Jacob. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Lawrence Erlbaum Associates Publishers, 1988.

Connell, Raewyn W. Gender. Polity Press, 2002.

Else-Quest, Nicole M., Janet S. Hyde, and Marcia C. Linn. “Cross-National Patterns of Gender Differences in Mathematics: A Meta-Analysis.” Psychological Bulletin 136, no. 1 (2010): 103–127.

Guiso, Luigi, Ferdinando Monte, Paola Sapienza, and Luigi Zingales. “Diversity. Culture, Gender, and Math.” Science 320, no. 5880 (2008): 1164–1165.

Hamamura, Takeshi. “Power Distance Predicts Gender Differences in Math Performance Across Societies.” Social Psychological and Personality Science 3, no. 5 (2012): 545–548.

Hannum, Emily. “Poverty and Basic Education in Rural China: Villages, Households, and Girls’ and Boys’ Enrollment.” Comparative Education Review 47, no. 2 (2003): 141–159.

Hanushek, Eric A., and Ludger Woessmann. “The Role of Cognitive Skills in Economic Development.” Journal of Economic Literature 46, no. 3 (2008): 607–668.

Hermann, Zoltán, and Marianna Kopasz. “Educational Policies and the Gender Gap in Test Scores: A Cross-Country Analysis.” Research Papers in Education (2019): 1–22.

Hippe, Ralph, and Maciej Jakubowski. “Immigrant Background and Expected Early School Leaving in Europe: Evidence from PISA.” JRC Technical Reports 109065.

Inoue, Keiko, Emanuela Di Gropello, Yesim S. Taylor, and James Gresham. Out-of-School Youth in Sub-Saharan Africa: A Policy Perspective. The World Bank, 2015.

Kazeem, Aramide, Leif Jensen, and C. S. Stokes. “School Attendance in Nigeria: Understanding the Impact and Intersection of Gender, Urban-Rural Residence and Socioeconomic Status.” Comparative Education Review 54, no. 2 (2010): 295–319.

Lavrijsen, Jeroen, and Ides Nicaise. “Social Inequalities in Early School Leaving: The Role of Educational Institutions and the Socioeconomic Context.” European Education 47, no. 4 (2015): 295–310.

Lutz, Amy. “Barriers to High-School Completion Among Immigrant and Later-Generation Latinos in the USA.” Ethnicities 7, no. 3 (2007): 323–342.

Maccoby, Eleanor E. The Two Sexes: Growing up Apart, Coming Together. Belknap Press, 1998.

Maccoby, Eleanor E., and Carol N. Jacklin. The Psychology of Sex Differences. Stanford, CA: Stanford University Press, 1974.

Marks, Gary N. “Accounting for the Gender Gaps in Student Performance in Reading and Mathematics: Evidence from 31 Countries.” Oxford Review of Education 34, no. 1 (2008): 89–109.

Meinck, Sabine, and Falk Brese. “Trends in Gender Gaps: Using 20 Years of Evidence from TIMSS.” Large-scale Assessments in Education 7 (2019).

Mullis, Ina V. S., Michael O. Martin, Pierre Foy, and Martin Hooper. PIRLS 2016 International Results in Reading. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and the International Association for the Evaluation of Educational Achievement (IEA), 2017.

Mullis, Ina V. S., Michael O. Martin, Pierre Foy, Dana L. Kelly, and Bethany Fishbein. “TIMSS 2019 International Results in Mathematics and Science.”. https://timssandpirls.bc.edu/timss2019/international-results/.

Mullis, Ina V. S., Michael O. Martin, and Tom Loveless. 20 Years of TIMSS: International Trends in Mathematics and Science Achievement, Curriculum, and Instruction. TIMSS & PIRLS International Study Center, Boston College, 2016.

Nosek, Brian A., Frederick L. Smyth, N. Sriram, Nicole M. Lindner, Thierry Devos, Alfonso Ayala, and Yoav Bar-Anan et al. “National Differences in Gender-Science Stereotypes Predict National Sex Differences in Science and Math Achievement.” Proceedings of the National Academy of Sciences of the United States of America 106, no. 26 (2009): 10593–10597.

OECD. PISA 2018 Results (Volume I): What Students Know and Can Do. Paris: OECD Publishing, 2019a.

———. PISA 2018 Results (Volume II): Where All Students Can Succeed. Paris: OECD Publishing, 2019b.

———. PISA 2018 Technical Report. 2019c. https://www.oecd.org/pisa/data/pisa2018technicalreport/.

———. Equally Prepared for Life? How 15-Year-Old Boys and Girls Perform in School. Paris: OECD Publishing, 2009.

———. The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence. OECD Publishing, 2015.

———. PISA for Development: Out-of-School-Assessment Results in Focus. PISA in Focus. 2020. https://www.oecd-ilibrary.org/education/pisa-for-development_c094b186-en.

———. “PISA: Programme for International Student Assessment.”. https://www.oecd.org/pisa/.

Penner, Andrew M. “Gender Differences in Extreme Mathematical Achievement: An International Perspective on Biological and Social Factors.” American Journal of Sociology 114 Suppl (2008): S138-70.

Quintini, Glenda, and Sébastien Martin. “Same Same but Different: School-to-Work Transitions in Emerging and Advanced Economies.”. OECD Social, Employment and Migration Working Papers 154.

Reilly, David. “Gender, Culture, and Sex-Typed Cognitive Abilities.” PLoS ONE 7, no. 7 (2012).

Rosén, Monica, Isa Steinmann, and Inga Wernersson. 2022. “Gender differences in achievement.” In International Handbook of Comparative Large-Scale Studies in Education: Perspectives and Findings, ed. T. Nilsen, A. Stancel-Piątak, and J.-E. Gustafsson. Springer.

Rubin, Donald B. Multiple Imputation for Nonresponse in Surveys. Hoboken, NJ, USA: John Wiley & Sons, Inc, 1987.

Rutkowski, David, and Leslie Rutkowski. “Running the Wrong Race? The Case of PISA for Development.” Comparative Education Review 65, no. 1 (2021): 147–165.

Rutkowski, D., Leslie Rutkowski, and Jonathan A. Plucker. “Trends in education excellence gaps: a 12-year international perspective via the multilevel model for change.” High Ability Studies 23, no. 2 (2012): 143–166.

Salchegger, Silvia, and Birgit Suchań. “Was Bedeutet Es Für Den Geschlechterunterschied in Der Mathematikkompetenz Bei PISA, Wenn Dem Schulsystem Leistungsschwache Jungen Verloren Gehen?” Zeitschrift für Bildungsforschung 8, no. 1 (2018): 81–99.

Spaull, Nicholas, and Stephen Taylor. “Access to What? Creating a Composite Measure of Educational Quantity and Educational Quality for 11 African Countries.” Comparative Education Review 59, no. 1 (2015): 133–165.

Steinmann, Isa, Andrés Strello, and Rolf Strietholt. “The Effects of Early Between-School Tracking on Gender Segregation and Gender Gaps in Achievement: A Differences-in-Differences Study.” (under review)

Steinmann, Isa, Rolf Strietholt, and Monica Rosén. “International reading gaps between boys and girls, 1970–2016.” Comparative Education Review (forthcoming).

Stoet, Gijsbert, and David C. Geary. “Sex Differences in Mathematics and Reading Achievement Are Inversely Related: Within- and Across-Nation Assessment of 10 Years of PISA Data.” PLoS ONE 8, no. 3 (2013).

Tao, Hung-Lin, and Christos Michalopoulos. “Gender Equality and the Gender Gap in Mathematics.” Journal of Biosocial Science 50, no. 2 (2018): 227–243.

Taylor, Stephen, and Nicholas Spaull. “Measuring Access to Learning over a Period of Increased Access to Schooling: The Case of Southern and Eastern Africa Since 2000.” International Journal of Educational Development 41 (2015): 47–59.

UN. The Sustainable Development Goals Report 2020. United Nations, 2020. https://unstats.un.org/sdgs/report/2020/The-Sustainable-Development-Goals-Report-2020.pdf.

UNDP. Human Development Report 2019: Beyond Income, Beyond Averages, Beyond Today: Inequalities in Human Development in the 21st Century. New York: Bernan, 2019.

UNESCO. SDG 4 Data Digest: Data to Nurture Learning. Quebec Canada: UNESCO Institute for Statistics, 2018.

———. Building Bridges for Gender Equality. Paris: UNESCO, 2019. https://en.unesco.org/gem-report/2019genderreport.

UNESCO Institute for Statistics. "Glossary: Total Net Enrolment Rate.". http://uis.unesco.org/en/glossary-term/total-net-enrolment-rate.

———. "UIS.Stat.". http://data.uis.unesco.org/.

United Nations. Transforming Our World: The 2030 Agenda for Sustainable Development: A/RES/70/1. 2015. https://sdgs.un.org/sites/default/files/publications/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf.

World Bank. "GDP Per Capita (Current US$)." 2022a. https://data.worldbank.org/indicator/NY.GDP.PCAP.CD

World Bank. "Overall Global Gender Gap Index." 2022b. https://tcdata360.worldbank.org/indicators/af52ebe9?country=BRA&indicator=27959&countries=ASM&viz=line_chart&years=2006,2021

World Economic Forum. The Global Gender Gap Report 2020. Geneva, Switzerland: World Economic Forum, 2019.

Appendix A: Overview of Included and Excluded Countries in Main Analyses

Table Appendix A

Overview of Included and Excluded Countries in Main Analyses

	Inclusion in PISA analyses						
Country or region	2000	2003	2006	2009	2012	2015	2018
Albania	Incl.			Excl.[c]	Incl.	Excl.[a]	Incl.
Algeria						Excl.[c]	
Argentina	Incl.		Incl.	Incl.	Incl.		Incl.
Argentina (Ciudad Aut. de Buenos)						Incl.	
Australia	Excl.[c]	Excl.[c]	Excl.[c]	Excl.[c]	Excl.[c]	Incl.	Incl.
Austria	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Azerbaijan			Excl.[c]	Excl.[c]			
Azerbaijan (Baku)							Incl.
Belarus							Incl.
Belgium	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Bosnia & Herzegovina							Excl.[c]
Brazil	Excl.[a]	Excl.[a]	Incl.	Incl.	Incl.	Incl.	Incl.
Brunei Darussalam							Incl.
Bulgaria	Incl.		Incl.	Incl.	Incl.	Incl.	Incl.
Cambodia							Excl.[c,e]
Canada	Excl.[c]	Excl.[c]	Excl.[c]	Excl.[c]	Incl.	Incl.	Incl.
Chile	Excl.[c]		Incl.	Incl.	Incl.	Incl.	Incl.
China (B. S. J. G.)						Excl.[c]	
China (B. S. J. Z.)							Excl.[c]
China (Shanghai)				Excl.[c]	Excl.[c]		
Chinese Taipei			Excl.[c]	Excl.[c]	Excl.[c]	Excl.[c]	Excl.[c]
Colombia			Incl.	Incl.	Incl.	Incl.	Incl.
Costa Rica				Excl.[c]	Incl.	Incl.	Incl.
Croatia			Incl.	Incl.	Incl.	Incl.	Incl.
Czech Republic	Incl.	Incl.	Excl.[c]	Excl.[c]	Incl.	Incl.	Incl.
Denmark	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Dominican Republic						Incl.	Incl.
Ecuador							Incl.[e]
Estonia			Incl.	Incl.	Incl.	Incl.	Incl.
Finland	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
France	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Georgia				Incl.		Incl.	Incl.
Germany	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Greece	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Guatemala							Incl.[e]
Honduras							Incl.[e]
Hong Kong	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Hungary	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Iceland	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
India (Himachal Pradesh)				Excl.[c]			

Country or region	Inclusion in PISA analyses 2000	2003	2006	2009	2012	2015	2018
India (Tamil Nadu)				Excl.[c]			
Indonesia	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Ireland	Incl.	Incl.	Incl.	Incl.	Incl.	Excl.[c]	Excl.[c]
Israel	Incl.		Incl.	Incl.	Incl.	Incl.	Excl.[a]
Italy	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Japan	Excl.[a]	Excl.[c]	Excl.[a]	Excl.[c]	Excl.[c]	Excl.[c]	Excl.[c]
Jordan			Incl.	Incl.	Incl.	Incl.	Incl.
Kazakhstan				Incl.	Incl.		Excl.[c]
Korea	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Kosovo						Excl.[c]	Excl.[c]
Kyrgyzstan			Incl.	Incl.			
Latvia	Excl.[c]	Excl.[c]	Incl.	Incl.	Incl.	Incl.	Incl.
Lebanon						Excl.[c]	Excl.[c]
Liechtenstein	Excl.[b]	Excl.[b]	Excl.[b]	Excl.[b]	Excl.[b]		
Lithuania			Incl.	Incl.	Incl.	Incl.	Incl.
Luxembourg	Excl.[a]	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Macao		Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Malaysia				Incl.	Incl.		Incl.
Malta				Incl.		Incl.	Incl.
Mauritius				Excl.[c]			
Mexico	Excl.[a]	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Moldova				Incl.		Incl.	Incl.
Montenegro			Excl.[c]	Excl.[c]	Excl.[c]	Incl.	Incl.
Morocco							Incl.
Netherlands	Excl.[a]	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
New Zealand	Excl.[c]	Excl.[c]	Excl.[c]	Excl.[c]	Incl.	Incl.	Incl.
North Macedonia	Excl.[c]					Excl.[c]	Excl.[c]
Norway	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Panama				Incl.			Incl.
Paraguay							Excl.[c,e]
Peru	Incl.			Incl.	Incl.	Incl.	Incl.
Philippines							Incl.
Poland	Excl.[a]	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Portugal	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Excl.[a]
Qatar			Incl.	Incl.	Incl.	Excl.[c]	Excl.[c]
Romania	Excl.[c]		Excl.[c]	Excl.[c]	Incl.	Incl.	Incl.
Russ. Fed. (Moscow Reg.)							Incl.
Russ. Fed. (Perm)					Incl.		
Russ. Fed. (Tatarstan)							Incl.
Russ. Federation	Excl.[c]	Excl.[c]	Excl.[c]	Excl.[c]	Incl.	Incl.	Incl.
Saudi Arabia							Incl.
Senegal							Incl.[e]
Serbia			Incl.	Incl.	Incl.		Incl.
Singapore				Excl.[c]	Excl.[c]	Incl.	Incl.
Slovak Republic		Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Slovenia			Incl.	Incl.	Incl.	Incl.	Incl.

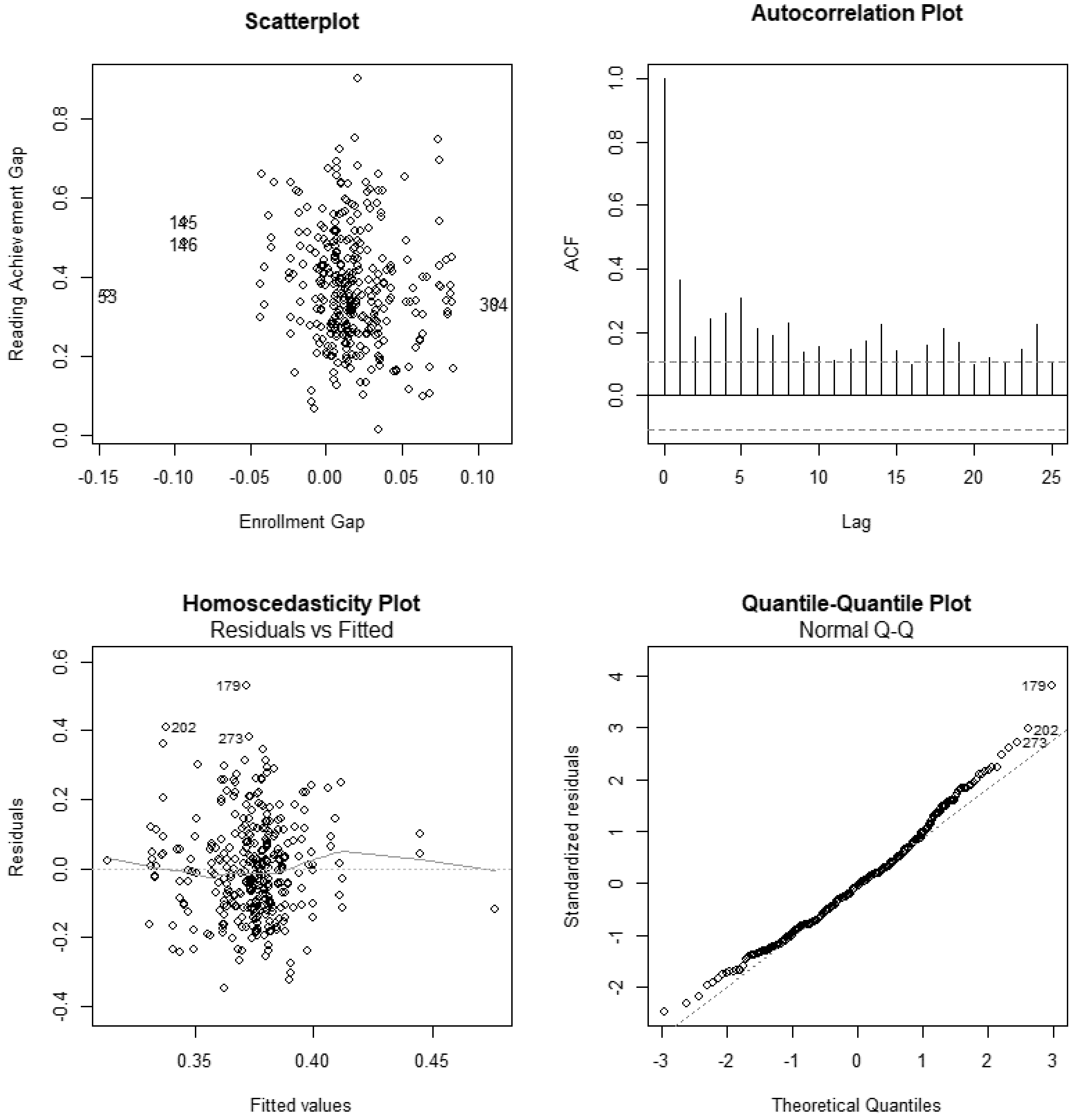
Country or region	Inclusion in PISA analyses 2000	2003	2006	2009	2012	2015	2018
Spain	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Spain (Regions)						Incl.	
Sweden	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Switzerland	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Thailand	Excl.[c]	Excl.[c]	Incl.	Incl.	Incl.	Incl.	Incl.
Trinidad & Tobago				Excl.[c]		Excl.[c]	
Tunisia		Excl.[c]	Excl.[c]	Excl.[c]	Excl.[c]	Excl.[c]	
Turkey		Incl.	Incl.	Incl.	Incl.	Incl.	Incl.
Ukraine							Excl.[c]
Un. Arab Emirates				Excl.[c]	Excl.[c]	Incl.	Incl.
United Kingdom	Incl.	Excl.[a]	Incl.	Incl.	Incl.	Incl.	Incl.
United States	Excl.[c]	Excl.[c]	Incl./Excl.[d]	Incl.	Incl.	Incl.	Incl.
United States (Massachusetts)						Incl.	
United States (North Carol.)						Incl.	
United States (Puerto Rico)						Incl.	
Uruguay		Excl.[c]	Incl.	Incl.	Incl.	Incl.	Incl.
Venezuela (Miranda)				Incl.			
Vietnam					Excl.[c]	Excl.[c]	
Yugoslavia		Excl.[c]					
Zambia							Excl.[c,e]

Note. Empty cells imply that the country or region did not participate in the PISA cycle or that the data was not published due to severe data quality issues. [a] Excluded in first step due to data quality concerns (i.e., usage of data is not fully recommended in the technical PISA reports). [b] Excluded in second step due to small sample size. [c] Excluded in third step due to missing secondary school enrollment data. [d] Included for mathematics and science, excluded for reading domain. [e] Participation in PISA for Development in 2017 not PISA 2018.

Appendix B: Regression Prerequisite Analyses

Appendix B.1

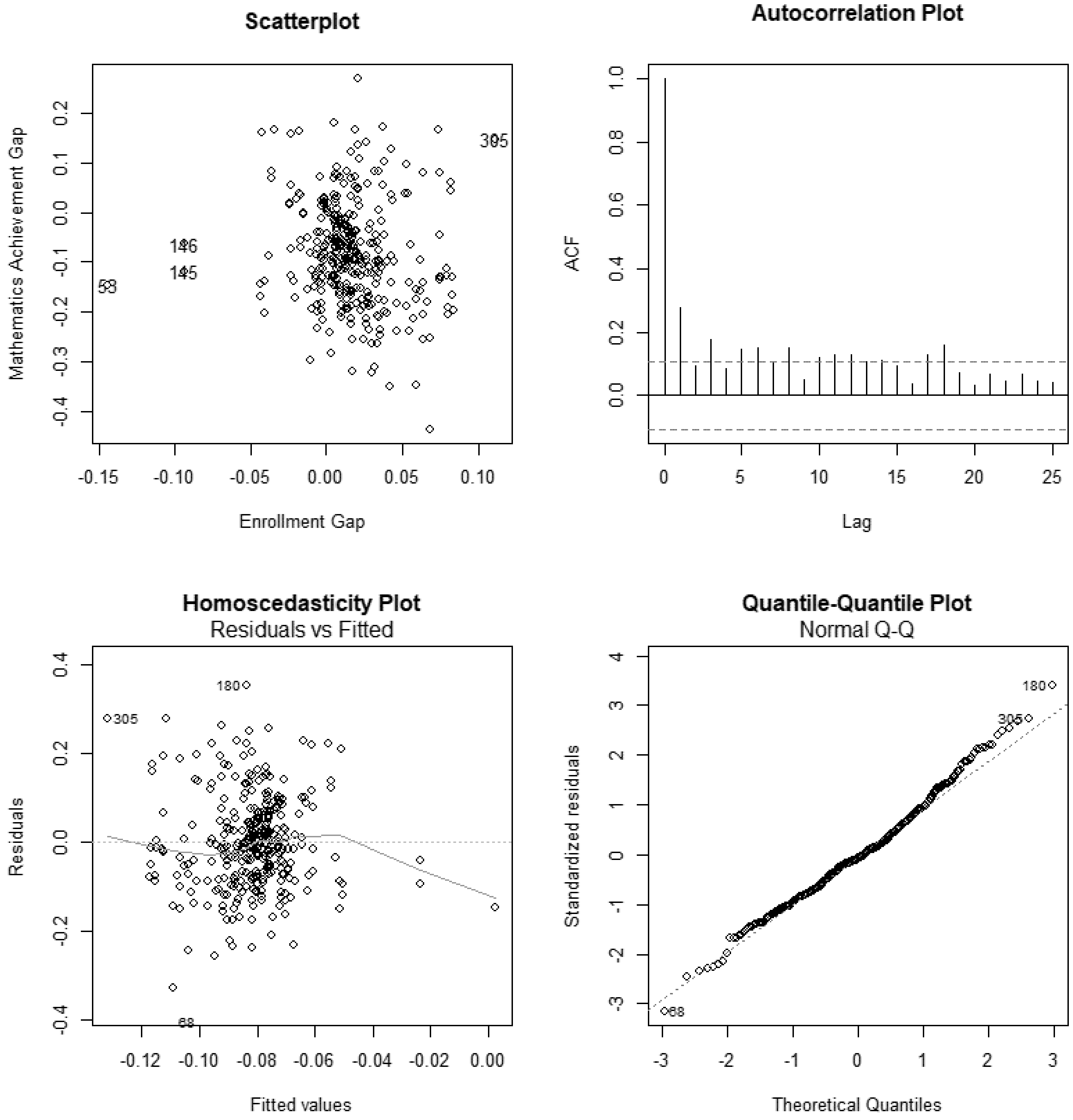
Results of Prerequisite Checks for Regressing Reading Gender Gaps on School Enrollment Gender Gaps



Note. Scatterplot (top left) displays gender achievement gaps (y-axis) by enrollment gender gaps. Extreme outliers (three times inter-quartile range) highlighted with observation numbers. Autocorrelation plot (top right) displays autocorrelation (ACF) of residuals (y-axis) by lags (x-axis). Homoscedasticity plot (lower left) displays residual (y-axis) by fitted values (x-axis). Flagged observations highlighted with observation numbers. Quantile-quantile plot (lower right) displays observed quantiles (y-axis) by theoretical quantiles for normal distributions (x-axis). Flagged observations highlighted with observation numbers.

Appendix B.2

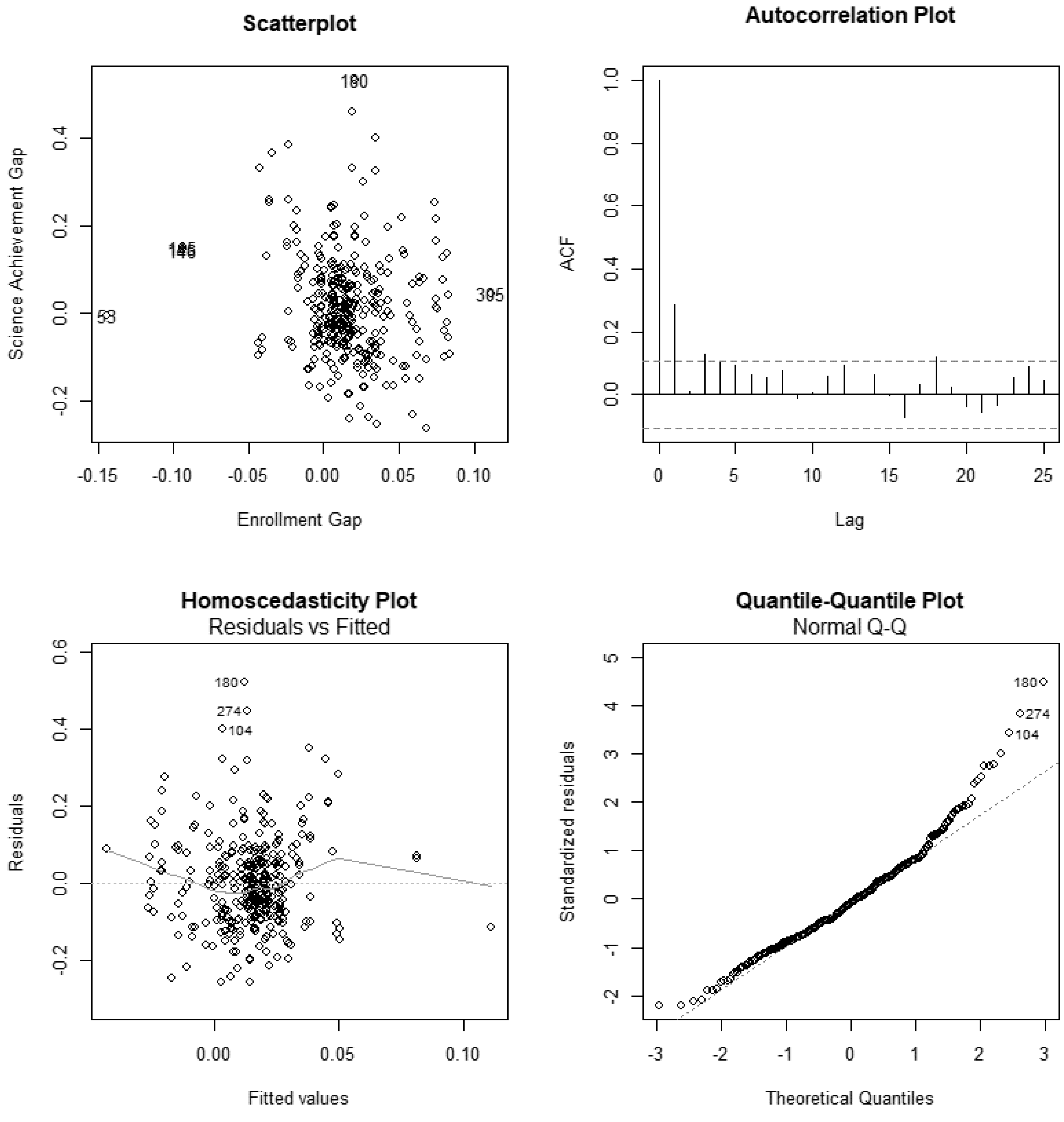
Results of Prerequisite Checks for Regressing Mathematics Gender Gaps on School Enrollment Gender Gaps



Note. Scatterplot (top left) displays gender achievement gaps (y-axis) by enrollment gender gaps. Extreme outliers (three times inter-quartile range) highlighted with observation numbers. Autocorrelation plot (top right) displays autocorrelation (ACF) of residuals (y-axis) by lags (x-axis). Homoscedasticity plot (lower left) displays residual (y-axis) by fitted values (x-axis). Flagged observations highlighted with observation numbers. Quantile-quantile plot (lower right) displays observed quantiles (y-axis) by theoretical quantiles for normal distributions (x-axis). Flagged observations highlighted with observation numbers.

Appendix B.3

Results of Prerequisite Checks for Regressing Science Gender Gaps on School Enrollment Gender Gaps



Note. Scatterplot (top left) displays gender achievement gaps (y-axis) by enrollment gender gaps. Extreme outliers (three times inter-quartile range) highlighted with observation numbers. Autocorrelation plot (top right) displays autocorrelation (ACF) of residuals (y-axis) by lags (x-axis). Homoscedasticity plot (lower left) displays residual (y-axis) by fitted values (x-axis). Flagged observations highlighted with observation numbers. Quantile-quantile plot (lower right) displays observed quantiles (y-axis) by theoretical quantiles for normal distributions (x-axis). Flagged observations highlighted with observation numbers.

Appendix C: Robustness Check Results

Table Appendix C

Results of Robustness Checks for Regressing Gender Achievement Gaps on Enrollment Gender Gaps

	Reading gender gap		Mathematics gender gap		Science gender gap	
Regression parameter	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)
Intercept	0.38***	0.38***	-0.07***	-0.07***	0.02**	0.02*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Enrollment gender gap	-0.80**	-0.80**	-0.78**	-0.67**	-0.71**	-0.57*
	(0.30)	(0.30)	(0.23)	(0.23)	(0.25)	(0.25)
N country-by-year obs.	328	328	330	330	329	329
Removed observations	yes	yes	yes	yes	yes	yes
Auto-correlated residuals	no	yes	no	yes	no	yes

Note. Dependent variables are gender achievement gaps (positive values imply higher scores for girls; negative values imply higher scores for boys). Independent variable is enrollment gender gap (positive values imply higher rates for girls; negative values imply higher rates for boys). In first models, observations flagged in regression prerequisite analyses (see Appendix B) were removed. In second models, auto-correlated residuals were considered.

* $p < .050$, ** $p < .010$, *** $p < .001$

Appendix D: Results of Further Analyses with TIMSS Data

Table Appendix D

Results of Regressing Gender Achievement Gaps on Enrollment Gender Gaps in TIMSS

	Mathematics gender gap		Science gender gap	
Regression parameter	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)
Intercept	0.03*	0.01	0.06**	-0.11*
	(0.01)	(0.03)	(0.02)	(0.05)
Enrollment gender gap	0.42	0.39	-0.01	0.28
	(0.41)	(0.41)	(0.59)	(0.57)
GDP per capita		0.00		0.00
		(0.00)		(0.00)
Assessment cycle 2007		0.01		0.15*
		(0.04)		(0.06)
Assessment cycle 2011		0.00		0.14*
		(0.04)		(0.06)
Assessment cycle 2015		0.01		0.21***
		(0.04)		(0.06)
Assessment cycle 2019		0.00		0.19**
		(0.04)		(0.06)
Mean achievement		-0.05***		-0.08***
		(0.01)		(0.02)
N country-by-year obs.	161	160	161	160
R^2	.007	.094	.000	.165

Note. Dependent variables are gender achievement gaps (positive values imply higher scores for girls; negative values imply higher scores for boys). Main independent variable is enrollment gender gap (positive values imply higher rates for girls; negative values imply higher rates for boys). Reference category to assessment cycle is 2003. Mean achievement is *z*-standardized.

* $p < .050$, ** $p < .010$, *** $p < .001$

Appendix E: Results of Further Analyses with Gender Gap Index Data

Table Appendix E

Results of Regressing Gender Achievement Gaps on Enrollment Gender Gaps and Gender Gap Sub-indices in PISA

	Reading gender gap		Mathematics gender gap		Science gender gap	
Regression parameter	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)
Intercept	0.27	1.29	4.19***	4.31***	4.2***	4.12***
	(1.47)	(1.37)	(1.02)	(1.06)	(1.16)	(1.22)
Enrollment gender gap	-0.97**	-1.55***	-0.62*	-0.88***	-0.73**	-1.19***
	(0.35)	(0.33)	(0.24)	(0.26)	(0.28)	(0.30)
GGI political empowerment subindex	-0.12	0.00	-0.12*	-0.06	-0.16**	-0.08
	(0.07)	(0.07)	(0.05)	(0.05)	(0.06)	(0.06)
GGI economic participation and opportunity subindex	-0.10	0.07	0.20*	0.24**	-0.08	0.00
	(0.11)	(0.10)	(0.08)	(0.08)	(0.09)	(0.09)
GGI educational attainment subindex	0.97	1.42**	0.11	0.35	0.25	0.68
	(0.59)	(0.52)	(0.41)	(0.41)	(0.47)	(0.47)
GGI health and survival subindex	-0.77	-2.29	-4.59***	-4.98***	-4.43***	-4.84***
	(1.42)	(1.34)	(0.99)	(1.04)	(1.12)	(1.19)
GDP per capita		0.00**		0.00*		0.00*
		(0.00)		(0.00)		(0.00)
Assessment cycle 2003		0.02		0.00		0.03
		(0.03)		(0.02)		(0.02)
Assessment cycle 2006		0.00		-0.01		0.01
		(0.03)		(0.02)		(0.02)
Assessment cycle 2009		-0.15***		0.02		-0.03
		(0.03)		(0.02)		(0.02)
Assessment cycle 2012		-0.14***		0.03		0.01
		(0.03)		(0.02)		(0.02)
Assessment cycle 2015		0.00		-0.01		-0.01
		(0.01)		(0.01)		(0.01)
Assessment cycle 2018	0.27	1.29	4.19***	4.31***	4.02***	4.12***
	(1.47)	(1.37)	(1.02)	(1.06)	(1.16)	(1.22)
Mean achievement	-0.97**	-1.55***	-0.62*	-0.88***	-0.73**	-1.19***
	(0.35)	(0.33)	(0.24)	(0.26)	(0.28)	(0.30)
N country-by-year obs.	270	270	271	271	271	271
R^2	.048	.305	.122	.187	.143	.203

Note. Dependent variables are gender achievement gaps (positive values imply higher scores for girls and negative values imply higher scores for boys). Main independent variable is enrollment gender gap (positive values imply that girls have higher enrollment rates; negative values imply higher enrollment rates among boys). Higher values on GGI sub-indices reflect greater gender equality. Reference category to assessment cycle is 2000. Mean achievement is *z*-standardized.

* $p < .050$, ** $p < .010$, *** $p < .001$