The Pennsylvania State University

The Graduate School

Department of Education Policy Studies

GENDER EQUALITY AND THE GENDER ACHIEVEMENT GAP: A CROSSNATIONAL STUDY USING PISA 2015

A Dissertation in

Education Theory and Policy & Comparative and International Education

by

Hee Jin Chung

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

August 2018

The dissertation of Hee Jin Chung was reviewed and approved* by the following:

Soo-yong Byun
Associate Professor of Education Theory and Policy, and Demography
Dissertation Adviser
Chair of Committee

David P. Baker
Professor of Education Theory and Policy, Sociology, and Demography

Katerina Bodovski
Associate Professor of Education Theory and Policy

Nancy Luke
Associate Professor of Sociology and Demography

Kevin P. Kinser
Professor of Education
Head of Department of Education Policy Studies

* Signatures are on file in the Graduate School.

ABSTRACT

The topic of gender gap in academic achievement is a topic worth investigating because not only is the gender achievement gap closely associated with the gender differences in labor market outcomes but also the gap itself reflects a crucial facet of educational inequality. Despite a body of literature reports the progress of women in many realms of education, it does not imply that the issues related to gender inequality in education are completely resolved. Rather, gender equality should be approached from a new perspective, which investigates the developmental, cognitive, and environmental sources of both genders in relation to their educational outcome.

In order to understand how country-level environmental sources, especially the structural and ideological features regarding gender equality across countries, are associated with gender gap in achievement, the current study investigated three research questions: (1) How do the gender gaps in mathematics, science, and reading achievement vary in size and direction by country? (2) How many classes do the countries cluster into in terms of gender-related attitudes and how are the classes different from each other? (3) How do gender achievement gaps in mathematics, science and reading, respectively, associate with the structural and ideological aspect of gender inequality in each country?

Utilizing the 2015 Programme for International Student Assessment (PISA), the current study revealed substantial cross-national variations in the gender gap across all subject areas as expected. However, the patterns of the variations were different for each subject. For example, for mathematics achievement, a greater number of countries demonstrated a statistically significant female disadvantage. Meanwhile, there was a balance between the number of countries that demonstrated a female advantage and disadvantage in science achievement. Unlike the case of mathematics and science, female students scored ahead of their male counterparts on the reading examination across every single country included in the analyses.

In a subsequent latent class analysis four classes of gender ideology were identified: (1) liberal egalitarian, (2) partial traditionalist, (3) familial egalitarian, and (4) egalitarian essentialist. Most of the classes that surfaced from the analyses did not fall along the one-dimensional traditional-egalitarian line. Besides the class labeled as the liberal egalitarian which resembles the characteristics of a typical individual with egalitarian gender role attitudes, partial traditionalist, familial egalitarian, and egalitarian essentialist were demonstrating different combinations of the various dimensions of gender attitudes.

Finally, multilevel analyses were conducted in order to examine the relationships between gender achievement gap in mathematics, science and reading, and national gender equality, respectively. In terms of the gender gap, there was a significant female disadvantage in mathematics no gender difference in science, and a significant female advantage in reading after controlling for all other variables. The relationship between achievement and national gender equality was not robust. The statistically significant association appeared once interaction terms with gender were included in the models. The structural gender equality measured by Gender Inequality Index (GII) was the predictor with the greatest effect on students' achievement across all subjects however. On the other hand, the national gender ideology, the other important measure of national gender equality, generally did not have a statistically significant influence on students' achievement except for a few exceptions. Gross Domestic Product (GDP) per capita was also a significant predictor of students' academic achievement in mathematics, science, and reading.

The current study aimed to contribute to the existing body of literature in four ways. This study intended to incorporate national gender ideology in addition to structural traits regarding gender equality into the analysis. The second aim of this study was to reflect the multidimensionality of egalitarianism instead of using one-dimensional measurement of gender ideology. Third, this study aimed to investigate this topic utilizing a more recent and relevant

dataset. Finally, the current study sought to explore the achievement gap in a more comprehensive manner by incorporating students' performance on all subject areas available in PISA, namely math, science, and reading.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

My heart is racing with happiness and excitement as I complete my doctoral studies at Penn State and start the next chapter of my life back in Korea! Being a novice researcher and mother, this accomplishment may not have been possible without the support of many individuals. First of all, I would like to express my deepest gratitude to my advisor, Professor Soo-yong Byun, for being such a great mentor and friend from day one. Your guidance and encouragement not only inspired me to become a better researcher but also helped me get through difficult times during my Ph.D. study. It has been a great pleasure and honor to have been your student, and I hope our collaborations continue after my graduation. I would also like to thank my dissertation committee members: Dr. Katerina Bodovski, Dr. David Baker, and Dr. Nancy Luke for sharing their experiences and insights regarding my study. Katerina, I will truly miss your enthusiasm, energy and sense of humor! Dr. Baker, thanks for your valuable comments and suggestions which really helped improving my research. Finally, Nancy, I had really enjoyed taking your gender class, which inspired the development of my dissertation.

I would also like to thank all the staff members at Penn State, colleagues and friends who made my life as a graduate student much better. A special shout-out to Linda Grant-Oishi, who has always been a great supporter and administrator, I cannot thank you enough for taking care of all the administrative work when I had to, unexpectedly, postpone my proposal defense. Yifan, and Olga, my only cohort members, thanks for the wonderful memories at State College, and I will miss our late-night chats and drinks! Although life had led us to different cities and countries, I hope we stay in touch. Hyewon and Joeun, I am forever grateful for your moral support and advice throughout this process. Finally, thank you, Haram and Jiyun, for providing me with the resources that I needed for my doctoral thesis.

I am also greatly indebted to my family for their love and support. Jin, thanks for your constant encouragement and compassion. Your words of wisdom and encouragement kept me going. I am so blessed to have such a wonderful sister and will miss our days together in the United States. Furthermore, special thanks to my parents and parent-in-laws, Bo Young Chung, Miran Lim, Gyu Seon Ko, Jeong Hwan Lee, for taking care of my little man so that I can focus on writing my dissertation. This work would have been simply impossible without your support and devotion.

Last but not at all the least, I would like to say thank you to my husband, Chang Seok Ko, for his support and encouragement in pursuing my dreams, and our son, Liam Youngbin a.k.a. Beanie, for bringing so much happiness and joy in my life. Chang Seok, thanks for being my best friend, my strength, and my rock. I look forward to the many years we will have together!

Chapter 1

Introduction

The topic of gender gap in academic achievement has attracted many scholars from various disciplines, and has been studied for an extended period of time. It is a topic worth investigating because not only is the gender achievement gap closely associated with the gender differences in labor market outcomes but also the gap itself reflects a crucial facet of educational inequality (Lai, 2010). Prior studies have suggested a positive relationship between academic performance and long-term outcomes, such as educational attainment and future earnings (Jencks & Phillips, 1999; Rose, 2006). Some studies emphasize that this association has been consistently documented and is especially pronounced in relation to the academic performance of certain academic subjects such as mathematics although the reported magnitude of the relationship varies by study (Murnane, Willett, Duhaldeborde & Tyler, 2000; Murnane, Willet & Levy, 1995; Weinberger, 2001)

Historically, the body of literature on this area of study highlights the female disadvantages in education (Buchmann, DiPrete & McDaniel, 2008; Jacobs, 1996). However, as a matter of fact, recent evidence shows that females made substantial progress in many realms of education, and even outperform males in various dimensions (DiPrete & Buchmann, 2006). According to Freeman (2004), higher percentage of females dropped out of high school compared to males (15.1% vs. 14.1%) in the early 1970s. Nonetheless, by 2000, not only was this pattern reversed but also the gap became wider (female: 9.3% vs. male: 12.2%), with females attaining higher education levels. Higher academic achievement in early childhood also has been noted in females. Specifically, using the Early Childhood Longitudinal Study-Kindergarten (ECLS-K) class of 1998-99 data, Husain and Millime (2009) found that female students tend to outperform their male counterparts in reading in a sizeable manner over the first four-years of schooling. This gap was especially evident for Hispanics and it further widened for low income black and

Hispanic students who attended public schools. Moreover, females are also standing out in higher education enrollment. Compared to 1970 when males constituted more than half of the total undergraduate (57.7%) and graduate (61.2%) students, the enrollment of female students consisted of 56.1% of the undergraduates and 57.9% of the graduate student body in 2001 across the United States. All in all, "in school and in college, females are now doing as well as or better than males on many of the indicators of educational attainment, and the large gaps in educational attainment that once existed between men and women have in most cases been eliminated" (U.S. Department of Education, 2000, p.9).

Despite the progress noted in many realms of education, such gender gap reversal in education does not imply that the issues related to gender inequality in education are completely resolved. Rather, new questions are being raised amid the quickly changing discipline of gender inequality (Buchmann et al., 2008). For instance, researchers sought to understand why the gender gap in education is shifting, and what would be the consequences of the change across the life course. Furthermore, Buchmann and her colleagues (2008) accentuated that it is important to refrain from the old way of thinking – that gender differences in education are caused by impediments to female education – and started approaching gender inequality from a new perspective that it is important to learn the developmental, cognitive, and environmental sources of both genders in relation to their educational outcome. Following their suggestion, this study attempted to understand how country-level environmental sources, especially the structural and ideological features regarding gender across countries, are associated with the gender gap in achievement. The findings may yield important insights into how boys and girls are positioned by the environment in different ways.

Statement of the Problem

A line of cross-national studies examining the relationship between gender inequality and gender achievement gap has been carried out over a prolonged duration of time (Baker & Jones, 1993; Fryer & Levitt, 2010; Guiso, Monte, Sapienza & Zingales, 2008; Van Langen, Bosker & Dekkers, 2006). However, when considering country-level features that are associated with gender inequality, a number of studies focus largely on structural traits without consideration of ideological aspects. For example, Baker and Jones (1993) incorporated information of the percentage of females in higher education, the ratio of female university to nonuniversity, the percentage of females in the labor force, and the percentage of females in the industrial, service, and agricultural sector, as measures of structural traits of gender stratification. Similarly, in their comparative study of investigating the gender gaps in education, Van Langen and her colleagues (2006) included country-level indicators such as female economic activity rate, gender empowerment index and female enrolment in science, technology, engineering and mathematics (STEM) courses as a proxy for women status. Likewise, Fryer and Levitt (2010) replicated the study conducted by Guiso et al. (2008) by only incorporating the gender gap index (GII), a measure that reflects the structural aspect of gender inequality. As can be seen from the preceding examples, previous studies neglected to take culture, or any ideological components in relation to gender, into account.

Although the social-economic structure of a country is an important factor to consider when studying social phenomena not to mention gender equality, I insist that the ideological aspect is also a pivotal component that must be considered in order to get a full realization of gender equality. The ideological perspective is important because gender-related attitudes can influence the beliefs and in turn behaviors of individuals (McDaniel, 2008). For example, studies revealed that gender-egalitarian values tend to reduce gender disparities in time devoted to family care (Bianchi, Milkie, Sayer & Robinson, 2000), and that parents with more egalitarian beliefs seek to provide equal level of education for their children, regardless of their gender (Buchmann

& DiPrete, 2006). Furthermore, gender egalitarianism is also associated with women's representation in various fields such as in the labor force, education, and politics at the national level (McDaniel, 2008). Despite the significant consequences of gender ideology on human behavior, a number of studies failed to integrate such construct into their studies.

The few studies that have incorporated gender-related values in their research tend to root their studies on the assumption that such values configure a single dimensioned, bipolar spectrum, between traditional and egalitarian attitudes. Guiso and his colleagues (2008), for instance, have investigated the relationship between culture, gender, and mathematics achievement. In doing so, they have created an index of cultural attitudes toward women by averaging the level of disagreement in statements from the World Value Surveys (WVSs), where high values indicated better position of women within a society, or greater gender egalitarianism. As previously stated, such operationalization assumes that gender ideology falls along a line where traditional attitudes are placed on one end of the spectrum and egalitarian values are located at the other.

However, Knight and Brinton (2017) pointed out that such presumption might be too simplistic to capture the complexity of egalitarianism. Under such assumption where liberal egalitarianism is on the ultimate end of the one end of the spectrum, there is limited space for egalitarianism to grow (Knight & Brinton, 2017). Consequently, after a major shift to an egalitarian direction in the late 1960s and 1970s, scholars have detected profoundly slow changes in gender-related attitudes since the mid-1990s (Cotter, Hermsen, & Vanneman, 2011; England, 2010).

Not surprisingly, the constructs of gender egalitarianism have varied across studies. For instance, Gornick and Meyers (2008) defined gender egalitarianism as a gender symmetric society where both men and women equally engage in caregiving as well as in employment. On the other hand, Orloff (2009) envisioned gender egalitarianism as a society where impediments to freedom are removed, so that individuals, men or women, can make their own choices about what

role they would assume. Knight and Brinton (2017) also asserted the need for a different approach to the notion of egalitarianism by acknowledging the multidimensionality of this construct that may have been influenced by new cultural frameworks.

Finally, many comparative international studies on this topic utilized outdated data that were garnered more than a decade ago (Baker & Jones, 1993; Chiu & McBride-Chang, 2006; Guiso et al., 2008; Marks, 2008). For instance, aforementioned study conducted by Baker and Jones (1993) used the Second International Mathematics Study (SIMS) collected in 1981 and 1982, while Marks (2008), whose work investigated the gender gaps in mathematics and reading, used the Programme for International Student Assessment (PISA) administered in 2000. In addition, studies conducted by Chiu and McBride-Chang (2006), and Guiso et al. (2008) all utilized the 2003 PISA dataset. Thus, there is a need to examine this topic using more relevant, and up-to-date data. Further, studies that investigated the relationship between gender and academic achievement, have only examined this by taking into account a single academic content area, such as reading (Chui & McBride-Chang, 2006) and mathematics (Guiso et al., 2008). Needless to say, these studies only portray a fragment of the gender achievement gap, as is based on two subject areas that are different in nature.

Purpose of the Study

The current study aimed to address the limitations found in prior research as mentioned in the previous section. First of all, this study intended to incorporate national gender ideology in addition to structural traits regarding gender equality into the analysis. So far, very few studies that have contributed to the literature of gender equality and gender achievement gap considered the gender ideological aspect into their analyses. Utilizing information from the sixth wave of the WVS and fourth wave of the European Value Survey (EVS), this study included national gender-related attitudes into the analysis. By taking into account the role of culture (i.e., gender

egalitarianism), it is the hope that this study would provide a more accurate understanding of the condition of gender equality in a country. In order to maximize the number of countries included in the analysis, the questions that were present in both surveys were utilized.

The second aim of this study was to reflect the multidimensionality of egalitarianism. As suggested by Knight and Brinton (2017), there needs to be an alternative approach in defining egalitarianism. Instead of adopting the traditional way of operationalizing egalitarianism, which is creating an index that assumes gender-related attitudes fall along a single dimensional line, a latent class analysis was conducted to empirically characterize a set of latent types using gender-related information from the WVS and EVS. By including different dimensions of egalitarianism into the study, it is expected to not only help to understand the concept in a more sophisticated manner, but also provide a subtler nuance in its relationship to gender achievement gap.

The third aim of the study was to utilize a more recent and relevant dataset in examining this relationship. Thus, the current study utilized PISA 2015, the most recent international database that was released by the Organisation for Economic Co-operation and Development (OECD). Finally, the current study sought to explore the achievement gap in a more comprehensive manner by incorporating students' performance on all subject areas available in PISA, namely math, science, and reading.

Chapter 2

Review of the Literature

This chapter consists of three major parts. First, I reviewed works on gender egalitarianism, as it is an essential notion of this study. Specifically, I discussed differing views on the concept of gender egalitarianism, summarized how prior research documents the change of gender-related attitudes overtime, and reviewed measurement issues of the construct. Second, I examined theoretical and empirical studies on gender achievement gap. To be specific, I attempted to review theories that account for the gender gaps in educational outcomes, and summarized empirical findings on gender achievement gap in mathematics, science and reading. Finally, I reviewed past research that has studied the relationship between gender equality and gender achievement gap, and attempted to demonstrate the mechanism of how gender equality might influence gender achievement gap, which is the overarching framework of this study.

Gender Egalitarianism

The concept of gender egalitarianism

Gender egalitarianism is a popular topic that has been a subject of research in various scientific and scholarly fields. Although there is a general agreement on the underlying qualities of gender egalitarianism, however, there seems to be lack of consensus on how it should be externalized in reality. For instance, Gornick and Meyers (2008) envisaged a fundamentally gender egalitarian society where both men and women "symmetrically" engage in paid work and unpaid caregiving, a society where family caregiving as well as paid labor is equally respected. They based their argument on the belief that gender inequality is deeply intertwined with the gender disparities in the amount of time engaged in family life as opposed to paid work (Orloff, 2009). Along with their vision of a utopian society, they further emphasized the institutional

arrangements through which it could be realized such as granting equal shares of paid leaves between mothers and fathers in order to meet caregiving demands, offering publicly financed high-quality care options, and reducing work time. As seen, their policy suggestions aimed to achieve parity between men and women by unburdening working mothers' caregiving duty via drawing fathers into more family care work.

Orloff (2009), on the other hand, demonstrated a different version of gender egalitarianism. She maintained that it is critical to recognize the heterogeneity within the gender categories (men and women) not to mention across categories. Instead of trying to make men and women more similar (or "symmetric") to each other as do the approach made by Gornick and Meyers (2008), she argued that more attention should be paid to the "differences" and that freedom of choice should be encouraged for both men and women by removing obstacles to their autonomy and providing expanded options. Although she echoed Gornick and Meyers' (2008) need for policy solutions to achieve greater gender equality, she criticized that their gender symmetric approach maybe too simple to accommodate the diverse demands of individuals. For example, Gornick and Meyers' (2008) approach may promote the engagement of fathers in care work, however, it may imply limited time for mothers who may desire to spend more time with their child. In such conversation, Orloff (2009) emphasized the availability of practical options that could enable individual choice – men or women – about their at-home care alternative to their paid work.

The two disparate perspectives suggested by Gornick and Meyers (2008), and Oloff (2009) fundamentally differ in the way they conceptualize gender equality: the former focusing more on how women and men could become more comparable in terms of their employment and caregiving pattern, and the latter highlighting the differences not only between men and women but also within the same gender category. It is important to understand these different approaches because the answers, or policy solutions, to achieve gender parity may sharply differ based on the

underlying assumptions of gender egalitarianism. For instance, one who envisions gender equality as enabling women to support her own family and herself by her own earning, as do men, may advocate the "universal bread winner model" (Fraser, 1994). The fundamental goal of this approach is to promote employment of women by shifting the care work burden to the state via services such as day care and elder care. On the other hand, the "caregiver parity model" (Fraser, 1994) may receive support from those who believe that gender equality could be achieved by supporting unpaid domestic work. This model aims to boost the value of informal care work equal to that of formal paid work through caregiver allowances, which provide monetary compensation for domestic work equivalent to a breadwinner salary. Instead of making the lives of women similar to that of men, as does the universal breadwinner model, this approach recognizes the differences between men and women and attempts to "make such differences costless" (Littleton, 1987).

Change of gender attitudes overtime

How did gender-related attitudes change overtime? There has been a line of research documenting the transition of gender attitudes over the course of time. The majority of researchers have witnessed a striking rise in support for gender equality in the late 1960s and early 1970s in the United States (Brewster & Padavic, 2000; Cherlin & Walter, 1981; Cotter et al., 2011; Ferree, 1974; Mason & Lu, 1988; McBroom, 1986; Thornton & Freedman, 1979). However, there are contrasting findings on how the attitudes changed from the mid 1970s. On one hand, Mason and Lu (1988) reported a continued increase toward a more egalitarian gender ideology across all social and demographic subgroups throughout the early 1980s. They suggested that such change was largely due to the actual changes in individual opinions rather than cohort succession. Cherlin and Walter (1981), on the other hand, maintained that gender

attitudes have not changed between 1975 and 1978. They pointed out various conservative protests against the Equal Rights Amendment as plausible causes to the lack of change.

According to subsequent studies tracing changes in the beliefs about gendered behavior, however, it seems like the movement toward gender egalitarian attitudes has been slowing down compared to the past. Brewster and Padavic (2000), for example, reported attitude changes to a more egalitarian gender ideology until the mid 1990s, however, admitted that the change was noticeably less dramatic than the change documented by Mason and Lu (1988). They stated that cohort succession played a greater role in understanding the attitude trends than it did in the past however. In a more recent study by Cotter et al. (2011) investigating changes in gender-related attitudes from 1977 to 2008 identified minimum alteration in gender attitudes since the mid 1990s. They insisted that such stagnation could not be explained by forces that used to contribute to gender attitude changes such as cohort differences, social structural changes and the shift in the overall ideological climate. Rather, they claimed that the lack of change should be understood within a new cultural frame, referred to as "egalitarian essentialism," which is a combination of egalitarianism and traditionalism. Under this new frame, stay-at-home mothers are not conforming to traditional gender values, but rather making an autonomous and sensible choice to become a housewife for their children and themselves. Therefore, strictly speaking, the new gender frame is egalitarian in the sense that it acknowledges individual's autonomous decision, however, traditional at the same time since it approves gender-typical paths.

The diffusion of gender egalitarian attitudes is also observed outside of the United States (Seguino, 2007). In her analysis of the WVS of over seventy countries, Seguino (2007) found that gender norms are becoming less traditional worldwide among both men and women. Her findings suggest that women's economic participation was the main driving force of this shift, and she speculated that the expansion of the economic pie has reduced men's resistance and hostility against women's growing economic empowerment. Along with the spread of support for gender

equality, studies found increased presence of women in education, economic activities and politics (Charles, 2011; Dorius & Firebaugh, 2010; Schwab, Samans, Zahidi, Leopold & Ratcheva, 2017; Yu & Lee, 2013). According to the Global Gender Gap Report (Schwab et al., 2017), the gender gap in education has significantly narrowed over the past ten years, and the gap in the political sphere has also demonstrated much progress although it remains the most gender segregated dimension. Despite such improvement in women status across the globe, Dorius and Firebaugh (2010) reported that the growth of global gender inequality has been slowing down over the past several decades. They suggested uneven population growth as one of the main reasons for such decline. To elaborate, the authors maintained that the growth of gender equality is decelerating because the population growth is faster in regions where gender inequality is greater. In fact, according to their projections, global gender equality could have been sped up had their been uniform population growth across countries.

Persistence of gender inequality

Despite empirical evidence that demonstrated changes to less traditional gender attitudes, researchers still found great levels of gender inequality across the globe in various dimensions of the society. For example, Dorius and Firebaugh (2010) found that women were politically underrepresented even though the degree varied substantially across nations. Further, a number of studies revealed that women were also disadvantaged in relation to economic activities (Charles & Grusky, 2004; Dorius & Firebaugh, 2010; Schwab et al., 2017). According to the Global Gender Gap Report (Schwab et al., 2017), after reaching its highest in 2013, the gender economic gap in 2017 had regressed back to where it was back in 2008. They predicted that it would take roughly 217 more years to achieve economic gender equality with current rates of progress. In their book, in addition, Charles and Grusky (2004) also focused on sex segregation in the labor market. Among the many spheres that were being equalized by the proliferation of egalitarianism,

the authors argued that the workplace remains resistant to the equalizing forces. Specifically, they identified two types of segregation: horizontal and vertical segregation. Horizontal segregation refers to a phenomenon where women are matched into jobs in certain field, usually service oriented, while men tend to enter manual jobs. The authors argued that such segregation is maintained and reproduced through gender essentialism, stereotypes about characteristics of men and women. On the other hand, vertical segregation refers to the situation where men tend to occupy highest-paying and prestigious jobs. The authors argued that vertical segregation stems from male primacy, a cultural belief that men are better than women in multiple ways, therefore, men are entitled to enjoy authority and power over women (Charles & Grusky, 2004). They predicted that male primacy would eventually loose support due to the diffusion of gender egalitarianism, while gender essentialism will sustain since it is compatible with egalitarian values (i.e. egalitarian essentialism). Consequently, vertical segregation will slowly attenuate, while horizontal segregation continues.

Gender inequality can also be observed in private spheres. In her study of fifty couples in the United States, Hochschild and Machung (1989) found that women were more likely to be in charge of domestic work and child care duties, referred to as the "second shift," than men despite their participation in the labor market. In other words, the division of household labor was far from being equal between men and women. According to Hochschild and Machung (1989), women spend about an additional month each year compared to men by assuming the second shift. Moreover, her research revealed that there was a striking leisure gap between men and women. Unlike the male interviewees who had certain hobbies and had time to enjoy them, all of the female participants had no hobby at all. As illustrated, in spite of the growing gender egalitarian ideology across the world, gender inequality is still prevalent within the public and private spheres of lives.

Multidimensionality of gender egalitarianism

As demonstrated in previous sections, gender egalitarianism is a rather complex notion. As a result, it is an analytic challenge to operationalize the concept in empirical works. Nevertheless, gender egalitarianism remains a popularly studied topic across various fields such as but not limited to business (e.g. Adam & Funk, 2012), politics (e.g. Inglehart & Norris, 2003), education (e.g. Guiso et al., 2008), and sociology (e.g. Bianchi et al., 2000; Stickney & Konrad, 2007). However, how the construct was measured in each study varies. For example, in their study of investigating the division of housework labor between men and women, Bianchi and her colleagues (2000) created a gender ideology index by summing the responses of participants so that higher value represents a more egalitarian perspective. Stickney and Konrad (2007) assessed gender attitudes in relation to earnings by generating an index averaging responses from four gender-related items. Guiso et al. (2008), as pointed out previously, also created an index averaging the responses to gender-related questions. Similar to that or Bianchi et al. (2000), greater values of this index indicated higher levels of egalitarianism. As can be seen, the indices of gender ideology generated in studies are based on the assumption that the attitudes fall along a single line where traditional attitudes stand on one end of the continuum while egalitarian stands on the other.

A major criticism on such approach is that it does not fully capture the complexity of gender equality. This index assumes that the attitudes of individuals will fall somewhere between traditional and egalitarian. However, some researchers question this assumption. Fraser (1994), for instance, suggested that gender equality should not be conceptualized as something that can be adhered to a single value or norm. Similarly, Knight and Brinton (2017) pointed out that the unidimensional approach to gender egalitarianism may be a contributing factor to the stalled gender attitude change. That is, postulating liberal egalitarianism as the optimal destination for everyone may make the change seem as if it came to an end, especially when the national gender

attitude reached a high level of egalitarianism but not necessarily the endpoint. However, if we conceptualize gender egalitarianism as a multifaceted notion, what may be really happening is that the gender ideology is transitioning in relation to other dimensions of gender egalitarianism (Cotter et al., 2011; Knight & Brinton, 2017).

Grounded on such criticism, Knight and Brinton (2017) attempted to empirically identify a more nuanced classification of gender egalitarianism. To be specific, utilizing latent class analysis, they grouped countries into different classes based on the responses to questions that tap into different aspects of gender-related attitudes. For example, the indicators used for the analysis include questions measuring attitudes toward male primacy, gender essentialism, and normative imperative regarding gender roles. The authors emphasized the inclusion of the normative component because unlike gender essentialism, which is a general idea of the difference between genders, the normative component is more about the action, about what each gender should do (Knight & Brinton, 2017)

Their findings identified four distinct types of egalitarianism, which they respectively labeled as 1) traditionalism, 2) liberal egalitarianism, 3) egalitarian familism, and 4) flexible egalitarianism. The *traditionalism* group comprised of individuals who supported traditional values on gender roles. Those who belonged to this class generally approved of the hierarchy between men and women, and believed that woman's place is within the home, not at work. The largest group was the *liberal egalitarianism*, which consisted of individuals who strongly advocated for women's economic participation. Specifically, they denied gender essentialism, and believed that men and women are equal beings. The third group was *egalitarian familism*, which is characterized as supporting women labor force participation but also emphasizing the importance of home and family to women at the same time. They resembled the liberal egalitarianism in the sense that they upholded women employment, but different in that they held traditional values that family and home is essential to complete a woman's identity. Finally, the

last group was the *flexible egalitarianism*. Individuals who belonged to this group supported female participation in the workforce, but did not think that it was something necessary for women to pursue to fulfill their identity. They denied male primacy, but approved of gender-typical roles. The biggest difference between egalitarian familism was that they did not think familial roles were crucial in completing a woman's identity.

The study of Knight and Brinton (2017) provided a more sophisticated understanding of gender egalitarianism. Had there only been a scale from traditionalism to liberal egalitarianism to understand gender attitudes, it would have been extremely difficult to place the four classes of egalitarianism on a one-dimensional line, especially regarding egalitarian familism and flexible egalitarianism. Thus, in order to fathom the concept in a subtler manner, there is a need to reconceptualize gender egalitarianism by incorporating different aspects of the notion.

Gender Achievement Gap: Theoretical and Empirical Approaches

The other crucial concept of this study is the achievement gap between genders. There has been abundant scholarship on this topic for an extended period of time inside the United States (e.g. Benbow & Stanley, 1980, 1983; Freeman, 2004; Lee, Grigg, & Dion, 2007) and beyond (e.g. Chiu & McBride-Chang, 2006; Contini, Tommaso, & Mendolia, 2017; Shafiq, 2013). Further, many scholars tried to identify factors contributing to this disparities. In the following chapter, I introduced major theories that attempted to explain the gender achievement gap, and summarized empirical findings.

Theoretical approach to gender achievement gap

A large array of theories has been explored to understand why there are differences in achievement by gender. Although the theories bring in different perspectives of the gender gap, they are based on two broad ideological perspectives: 1) that human behavior is driven by

biological traits, and 2) that human behavior is socially constructed (Skelton, Francis & Valkanova, 2007).

The biological argument

The biological argument suggests that the gender difference in achievement in general, is due to the disparities in biological traits. The supporters of this perspective argue that girls and boys are born with certain inbuilt genetic inclinations that attract them to certain curriculum while make them refrain from others (Skelton et al., 2007). Scholars like Gurian (2002) and Head (1999) also argue that male and female has different brain structure, and that there are gender differences in how hormones influence the brain. According to Baron-Cohen (2003), such differences lead to differentiate academic skills and abilities at school. For example, boys are believed to have better spatial ability than girls, which yield higher achievement in mathematics (Wilder & Powell, 1989). From this perspective, the fact that boys are outperforming girls in mathematics and girls exceeding boys in literacy is viewed as a "natural" phenomenon, which stemmed from differences in biological characteristics.

However, there are criticisms that the biological argument is too deterministic, and it is problematic in the sense that it neglects the possibilities of change (Skelton et al., 2007). As stated by Halpern (2012), it may not only be the innate biological characteristics that are driving the gender gap, but the socialization practices may have to be considered simultaneously. Skelton and her colleagues (2007) emphasize that human brain is not unchangeable. It is an organ that develops by responding to external stimuli. Also, Rogers (2001) has insisted that hormones may be influenced by environmental elements. In short, the differences in brain structure and hormonal effects that is used to advocate the biological argument may be a result of social factors. Besides, if biological factors were the only ones that determined the learning process of boys and girls, then there will be a wider gap in mathematics achievement favoring male students.

The socialization argument

The socialization argument considers gender differences as a product of socialization, which starts at birth and continue throughout the life. Advocates of this theory emphasize that gender is socially constructed, thus the experience and structure an individual is exposed to is crucial in shaping one's gender identity. From this perspective, girls do not perform as well as boys because they are socialized to believe that mathematics is not important, useful, or part of the female identity (Wilder & Powell, 1989).

The construction of gender. Gender is a fundamental component of identity (Davies, 2003), and the behaviors of an individual are likely to be interpreted in gendered ways (Skelton et al., 2007). Boys and girls are often positioned in opposition to each other, and that the achievement gap is a produce of such "oppositional" ideas of gender. Supposing that boys are good at mathematics, for example, girls, the opposite identity of boys, are automatically assumed to be not good at mathematics.

Boys and girls develop their idea of gender, which guide their subsequent behaviors, in a larger societal context (Leaper & Friedman, 2007). To be specific, the gendered ideas and behaviors are likely to be influenced by their economic class, race, and ethnic group the individuals belong to. In this context, peer groups play a pivotal role. In fact, prior studies have reported that children tend to be more engaged in objects, activities, behaviors, and social roles that are believed to be related to their own gender, while refraining from what is associated with the other gender (Banerjee & Lintern, 2000; Guillet, Sarrazin, & Fontayne, 2000; Martin, Ruble & Szkrybalo, 2002).

Doing well at school is often associated with being feminine. In other words, it is not masculine to work hard for good grades at school. Therefore, some boys may seek to disengage

from academic work in order to maintain their idea of masculinity (Legewie & DiPrete, 2012). As a result, boys lag behind girls in terms of academic achievement in general. Such behavior was particularly apparent among working class boys, however, some researchers insist that boys from middle class are adopting these attitudes to a greater extent (Skelton et al., 2007).

Yet, this logic may work differently for the achievement of specific subjects. Traditionally, 'the sciences' have been perceived as masculine and 'the arts' as feminine (Cvencek, Meltzoff, & Greenwald, 2011; Lummis & Stevenson, 1990; Nosek et al., 2009). As mathematics is considered to be a boy- subject, girls may disassociate themselves from mathematics to boost their femininity, and focus on girl-subjects such as reading and writing. For similar reasons, boys will be encouraged to focus on mathematics instead of literacy. As a result, there will be a growing gender gap in achievement in both mathematics and literacy, where the former favors males and the latter does females.

The stereotype threat hypothesis. Stereotype threat is a concept that is widely used to account for gender and racial achievement gaps. It postulates that a conscious or unconscious belief that a group is stereotypically known for specific performance deficits will adversely affect the performance of the task specific to the stereotype among those who belong to the group (Stoet & Geary, 2012; Wei, 2009). That is to say, when an individual faces judgment based on societal stereotypes about his/her group regarding a particular skill or ability, in other words, experience ''stereotype threat," the individual may feel extra pressure, which in turn, hinders his/her performance (Spencer, Steel, & Quinn, 1999). In their study, Spencer, Steel, and Quinn (1999) conducted a series of experiments to test this hypothesis, to investigate whether they find evidence for stereotype threat in relation to women's mathematics achievement to be specific. They found that women greatly underperformed relative to men when the participants were explicitly told that the test demonstrated gender differences prior to the exam. On the other hand,

when the test was claimed not to yield gender differences, female participants accomplished at the same level as the male counterparts.

Despite all the attention it has received among scholars, the stereotype threat hypothesis also faces various criticisms. To start with, many researchers have questioned the external validity of the findings from the experiment conducted by Spencer, Steel, and Quinn (1999). The sample consists of only 67 highly selective participants, 31 men and 36 women. Thus, their findings may not be applicable to the larger population. In fact, Wei (2009) found counter evidence to the stereotype threat effect. Using the micro data from the 1978 to 1999 National Assessment of Educational Progress (NAEP), he revealed that girls actually performed significantly better in relation to boys when they were asked whether they feel mathematics is more for boys than girls, a question that is supposed to prompt stereotype threat. Wei (2009) points out that the question is asked in a stereotypical direction so that it may have increased the prominence of the negative stereotype while "empowering girls by allowing them to reject the stereotype" and encouraging them to make additional efforts to "prove everyone wrong".

Gender achievement gap: Empirical evidence

A large body of literature has been carried out to understand the gender achievement gap. To be specific, studies focused on not only the direction of the discrepancies but also the magnitude of the gap. This section displays empirical findings on gender achievement gap by subject matter, namely mathematics, science, and reading.

Gender gap in mathematics achievement

The evidence on whether or not females underperform in mathematics is not conclusive. On one hand, the majority of studies investigating this topic is consistently documenting that females are, in fact, not performing as well as their male counterparts in mathematics. Benbow

and Stanley (1980, 1983) found that there were significantly more boys than girls at the highest levels of mathematics performance. Although the sex ratio at the highest performing level had declined overtime, evidence show that boys are still outperforming girls (Wai, Cacchio, Putallaz, & Makel, 2010). Similarly, Lee and colleagues (2007) reported that male students in fourth and eighth grade outperformed their female counterparts on the 2007 NAEP, and such gap has been observed over the last two decades. Furthermore, Xie and Shauman (2003) found that twice as many males than females performed in the top five percent in high school mathematics over the past 20 years.

On the other hand, some researchers argue that the gender differences in mathematics achievement are negligible, and that the differences become noticeable only when the test material becomes advanced (Spencer, Steele & Quinn, 1999). In their meta-analysis of 242 studies published between 1990 and 2007, Lindberg et al. (2010) concluded that there is no difference in mathematics performance by gender, which is consistent with what Hyde, Lindberg, Linn, Ellis, and Williams (2008) found in their empirical study. As stated, by analyzing the state assessment results of 50 states, Hyde and her colleagues (2008) also concluded that girls were performing just as well as boys on standardized test in the United States. Despite evidence indicating minimal differences in mathematics achievement by gender in general, some studies suggest that male outperform female counterparts under certain conditions. The experiment conducted by Spencer, Steele and Quinn (1999) indicated that women, indeed, demonstrated lower performance on difficult tests but scored just as well on easier test compared to men.

Mixed findings in relation to gender mathematics achievement gap are observed outside of the United States, as well. In Italy, for instance, Contini and her colleagues (2017) found that girls consistently scored lower grades on mathematics than boys, even after holding individual and family characteristic variables constant. Utilizing national assessment data, they also found that the gap became more apparent as the students grew older. Further, their analyses

demonstrated that the disparities in achievement between genders were more prominent among high-achievers than the low performing children. This finding coincides with what Stoet and Greary (2013) found in their study. Specifically, they found that high-achieving girls consistently underperformed their high-achieving male counterparts. In his study of seven Muslim countries, Shafiq (2013) found that girls generally fell behind in mathematics compared to their male counterparts. There were some exceptions where girls achieved complete parity with boys, however. His results showed that girls in Azerbaijan and Jordan performed as well as boys. He insisted that the gender gap did not appear to be significantly related to key national-level characteristics such as economic development, fertility rates and labor market disadvantages as Azerbaijan and Jordan, where the achievement of girls are quite superb, are considered to be in a rather unfavorable condition in that regard. Baker and Jones (1993), too, concluded that it is not a universal phenomenon that boys do better in mathematics than girls.

Some studies argued that gender achievement gap in mathematics did not exist from the beginning. Indeed, substantial evidence that has been accumulated for the past 50 years demonstrated that the gender gap in mathematics was not observable before the students enter formal schooling (Fryer & Levitt, 2010). The gap, however, became apparent and significant when students proceeded to middle school and beyond (Fryer & Levitt, 2010; Lindberg et al., 2010). Utilizing ECLS-K, a nationally representative data of the United States, Fryer and Levitt (2010) found no discernible differences in mathematics abilities among kindergarten-aged boys and girls. Yet, by the end of fifth grade, girls were doing more than 0.2 standard deviations worse than their male counterparts in mathematics, which is equivalent to 22.5 months of schooling according to the authors. In an aforementioned study by Contini, Tommaso, and Mendolia (2017), the authors also found similar patterns in Italy. They demonstrated that the gender gap in mathematics appeared in as early as second grade, and the gap increased in older grades.

Gender gap in science achievement

Despite the growing emphasis on scientific literacy nowadays, science achievement gap has been receiving relatively less attention from researchers and policy makers compared to mathematics and literacy gaps (Beller & Gafni, 1996; Quinn & Cooc, 2015). However, as one of the disciplines of STEM, like mathematics, studies have been documenting male advantage in science for an extended period of time (Beller & Gafni, 1996; Jones, Mullis, Raizen, Weiss, & Weston, 1992; Keeves, 1973; Muller, Kinzie, Stage, 2001; National Center for Education Statistics [NCES], 2012; 2016). In some studies (Beller & Gafni, 1996), the gender gap in achievement was more pronounced in relation to science than mathematics. Empirical evidence supporting male advantage in science has been solid and consistent over the course of time. In their analysis of 1990 NAEP, Jones et al. (1992) reported that boys, on average, were more proficient in science than their female counterparts. To be specific, boys pointed approximately 10 scale points higher than girls in twelfth grade. However, they noted that such discrepancies between boys and girls did not exist at fourth grade. Rather, the gap emerged among eighth-grade students and became larger among twelfth graders.

Similarly, many countries beyond the United States also demonstrated similar patterns. Studies analyzing more recent data suggested similar results. In 2011, boys did better on the science assessment in both NAEP and ECLS-K, receiving approximately 0.15 standard deviations higher grades than girls (NCES, 2012; 2016). Analyzing 34 countries from the TIMSS, Nosek et al. (2009) found that boys scored significantly higher on the science exam than girls across 65% of the countries. In Muslim countries, the results were inconsistent. In Indonesia, the Kyrgyzstan Republic, and Tunisia, male advantage was observed, while girls performed as well as boys in Azerbaijan, Jordan, Qatar, and Turkey.

According to the analyses, which attempted to investigate whether this male advantage was observed across different subdivisions of the science subject, there were indeed some

discrepancies by the science content areas. Young and Fraser (1994) reported that boys tend to outperform their female counterparts in physics and chemistry among 12 year-old Australian students. Similarly, Beller and Gafni (1996) not only found significant gender differences in science achievement favoring boys on the 1991 International Assessment of Educational Progress but also identified that the largest gender gaps were observed in earth and space science, and physics. However, they found nearly no gender discrepancies in the topic regarding nature in science. Interestingly, comparable results were documented after more than two decades in a report using a more recent data, namely 2015 NAEP (NCES, 2016). Specifically, male students outperformed female students in physical science and Earth and space sciences, while there was no gender score gap in life science in twelfth grade.

There is also an ongoing debate on when the gender achievement gap in science emerges. Through their analysis of ECLS-K:99, Quinn and Cooc (2015) maintained that the science gender gap appeared from as early as third grade, which substantially diminished as students got older. A subsequent study conducted by Curran and Kellogg (2016) investigating the gender science gap in the early years of formal schooling insisted that the gap actually emerged much earlier than what previous research found. Specifically, they reported no gender science gap during kindergarten, however, the gap started to surface at the end of first grade. Their findings imply the possibility of female students going through various experiences during the early years of formal schooling that may lead to greater gender disparities in science later in life (Curran and Kellogg, 2016; Zakaib, 2011).

Although science and mathematics are interconnected as part of the STEM-field, somewhat less attention has been paid to science achievement gap compared to mathematics (Beller & Gafni, 1996; Quinn & Cooc, 2015). With one exception of Quinn and Cooc's (2015) study that found that gender science gap in eight grade no longer persist once the previous mathematics achievement was controlled for, documentation of gender achievement gap in

science seems to be persistently supporting male advantage unlike the case of mathematics, in which empirical evidence of gender difference is inconclusive.

Gender gap in reading achievement

There is a reasonably established consensus among the majority of studies that there is a female advantage in various types of reading assessments across the globe (Cloer & Dalton, 2001; Klecker, 2006; Lynch, 2002). In her study of investigating the gender gap in reading across years using multiple waves of NAEP data from 1992 to 2003, Klecker (2006) found that female advantage in reading was observed across all grades (fourth, eighth, and twelfth grade) and years. Furthermore, the discrepancies became much noticeable among older students. More recent data (i.e. 2015 NAEP) also demonstrated similar patterns. In 2015, girls from fourth and eight grade outperformed their male counterparts in reading, by seven scale points and nine scale points respectively (NCES, 2016). As seen, it has been well documented that girls have been exceeding boys in literacy attainment over a prolonged duration of time in the United States.

Female advantage in reading is also observed in countries outside of the United States. In their cross-national study of 43 countries, Chiu and McBride-Chang (2006) identified that girls outscored boys in the reading assessment in every participating country. As a matter of fact, Stoet and Greary (2013) reported that compared to girls, boys underperformed in reading in all nations in all four PISA assessments from 2000 to 2009. More importantly, the gap was greater among the low-performing students, which indicates that male students at-risk may be having greater problems in literacy attainment. Similarly, Mullis, Martin Kennedy, and Foy (2007) also reported that fourth-grade girls had significantly higher achievement than boys on the literacy reading exam administered by the International Association for the Evaluation of Educational Achievement (IEA), among all the participating countries but Luxemburg and Spain in which no gender gap in reading was observed. Shafiq (2013), who studied several Muslim countries,

reported that girls outperformed boys in reading assessment in all countries but one that were included in his analyses, namely Azerbaijan, Indonesia, the Kyrgyzstan Republic, Qatar, Tunisia, and Turkey. The only country that did not demonstrate female advantage on reading was Jordan, where there was no significant difference between genders.

As achievement gaps in other subject areas such as mathematics and science, the gender gap in reading performance did not appear from the beginning. Freeman (2004) suggested that the gap surfaces when students are in their early years of elementary school. Specifically, in her study utilizing the ECLS-K, Freeman (2004) found no gender differences in reading ability among students when they were in kindergarten, however, boys started to demonstrate lower reading scores compared to their female counterparts when they advanced to third grade. Such gap became wider as the students moved on to middle school.

Although it is a dominant finding that girls are better in literacy than boys in this field of study, Johnston and Watson (2005) found otherwise. The authors followed 300 students in the first year of primary school for seven years in order to evaluate the effectiveness of a synthetic phonics program in teaching reading and spelling. They not only found that the synthetic phonics program was effective in enhancing the literacy ability of children but also identified that boys exceeded girls in reading, spelling and reading comprehension after receiving instructions by the synthetic phonics program although the difference in reading comprehension was not statistically significant.

Gender Equality and the Gender Achievement Gap

Several studies have attempted to examine how gender equality – cultural, political and economic – associates with gender achievement gap across various countries. The following section summarized empirical findings of such studies, and strived to illustrate the mechanism of

how gender equality can influence the academic performances of boys and girls as suggested in prior literature.

Empirical findings

By and large, empirical analysis on this topic identified a negative relationship between national gender equality and gender achievement gap (Baker & Jones, 1993; Else-Quest, Hyde & Linn, 2010; Guiso et al., 2008; Hyde, Mertz & Schekman, 2009; Van Langen et al., 2006). The study by Baker and Jones (1993), one of the earliest to investigate this topic cross-nationally, for instance, explored how the variation in mathematics performance by gender associated with gender stratification of future opportunities, measured by various structural indicators of women educational access and labor force participation. They found smaller gender achievement gap in countries where equal opportunities were granted to males and females.

Similarly, in their cross-national study of 2003 PISA, Van Langen and her colleagues (2006) incorporated measures of gender equality at the country level such as female economic activity rate, gender empowerment index and female enrollment in STEM courses. There analysis revealed that the gender achievement gap had a negative association with economic activity of the females in each country. In other words, girls had the tendency to earn higher grades than their male counterparts in countries where greater proportion of women participated in the labor force.

Using different measures of gender equality, Guiso and his colleagues (2008) found comparable results. They examined whether indicators of gender equality such as the GGI, the index of cultural attitudes toward gender, which was constructed by the authors themselves, the rate of female economic activity, and the political empowerment index correlated with the gender gap in test performances across countries. There results displayed greater gender gap in low gender-equality countries such as Turkey, while the gap completely disappeared in more gender-

equal countries such as Norway and Sweden. Consistent results were demonstrated on all four indicators of gender equality.

However, the study of Fryer and Levitt (2010) contradicted findings from the previous body of literature. Utilizing the 2003 TIMSS, the authors find no relationship between the gender gap in mathematics achievement and gender equality. To be specific, they revealed that there was no gender gap in many Middle Eastern countries where there is high level of gender inequality. As a replicated study of Guiso et al. (2008), this result is quite striking. Although the authors failed to provide plausible explanations of why there is no association between gender equality and gender achievement gap in those countries, they did note that the discrepancies of findings between their study and that of Guiso et al. (2008) is primarily due to the composition of participating countries of the data they have each used. That is, 2003 TIMSS, which was used by Fryer and Levitt (2010), included a larger number of Middle Eastern countries than 2003 PISA, which was the primary data used in the analysis by Guiso et al. (2008). Once they restricted their analysis to countries that were included in 2003 PISA, they found similar results. Based on their analyses, the authors emphasized the sensitivity of previous cross-national research to the set of countries included in their study.

The mechanism: How gender egalitarianism affects gender achievement gap

As suggested in the previous section, various studies identified a negative relationship between gender equality and educational disparities by gender measured as student academic performances in various subject areas. That is, the gender achievement gap diminished or disappeared in countries where higher degree of gender equality was observed. Although there is limited empirical evidence toward the mechanism through which gender equality – structural and ideological – influences gender achievement gap, I attempted to suggest plausible explanations of the mechanism drawing on previous studies across disciplines.

At the individual-level, gender equality can influence academic performance through family. Family is the primary unit that makes the decision on how to invest in children's education, especially in developing countries where education is not affordable (Van der Vleuten, 2016). As a matter of fact, several studies have suggested that parents with different gender attitudes could demonstrate different behaviors in terms of their children's education. In India, for example, Gandhi Kingdon (2002) found substantial discrepancies in educational attainment between boys and girls. He suggested parents' values toward son preference, and beliefs about the gender division of labor as important determinants of such disparities. Such behavior could be applied to academic achievement as well. If the parents of a family believe that the education of boys is more important than that of girls, they might decide to make greater investments to their sons to improve their academic achievement. Parents with more egalitarian beliefs, on the other hand, seek to provide equal level of education for their children, regardless of their gender (Buchmann & DiPrete, 2006).

In developed countries, however, it is unlikely that families influence children's education in the same way as they do in developing countries. Especially since they generally have compulsory education that guarantees education for everyone up to a certain level. The family could impact the children in a subtler manner. Dryler (1988) revealed that children whose parents hold a more gender egalitarian attitude are likely to choose a gender-atypical program, a daughter choosing to study engineering for example, at upper secondary school. These children, in other words, are less likely to hold stereotypes about gender differences in certain fields of study. Considering that such stereotypes are important predictors of academic achievement (Steffens, Jelenec & Noack, 2010; Nosek, & Smyth, 2011), it is possible to suspect that the absent of such gender stereotypes could lead to better academic performance in that field.

Some researchers focus on the institutional-level factors such as future opportunities in the labor market in understanding the relationship between gender equality and gender

achievement gap (Baker & Jones, 1993; Chiu, 2007). Baker and Jones (1993) maintained that what individuals perceive as future opportunities could guide how well they perform. In other words, perceived future opportunities serve as kind of a motivator for individuals to make an effort to do better. Applying this to school settings, once students view academic achievement linked to their future success, it might lead to higher achievement.

Chiu (2007) made a similar argument. As suggested by Hofstede (2001), an egalitarian society has greater social mobility and is tend to be based on meritocracy. That is, one's skills and talents are important means to achieve mobility within the society. In such society, Chiu (2007) insisted, students are likely to perceive that their future success and their current academic achievement are connected, and in turn take their academic performance very seriously and work harder.

The important part is, however, gender could influence one's perception of this future success-academic achievement link. For example, similar to what has been suggested by Baker and Jones (1993) and Wiseman, Baker, Riegle-Crumb, and Ramirez (2009), male students could perceive mathematics and science more relevant to their future careers given current labor force conditions, which could shape their level of performances regarding those subjects. Furthermore, the social interactions they have with surrounding people, which take place within the existing institutional context, could also reinforce or discourage their perceptions. For instance, teachers and parents might guide female students to non-STEM fields as they see fewer opportunities in STEM fields for women. As a result, female students might view mathematics and science less essential for their future, which in turn could influence their academic performance in those subject areas (Baker & Jones, 1993). As demonstrated, the individual-level and institutional-level factors are closely knit together.

Guiding Research Questions and Hypotheses

As described in the previous chapter, prior research on gender equality and gender achievement gap has several limitations. These include: 1) the failure to include the ideological aspects when considering country-level features that associate with gender equality, 2) the failure to incorporate the multidimensionality of the national gender ideology among the studies that did include this construct, 3) the use of outdated data, and 4) the failure to investigate this topic regarding multiple subjects. This study attempts to contribute to the current literature by addressing these limitations. Specifically, the current study seeks to answer the following three questions. In addition, for each research question, I have formulated corresponding hypotheses of what I expect to observe based on prior empirical studies.

1. How do the gender gaps in mathematics, science, and reading achievement vary in size and direction by country?

HYPOTHESIS 1: Reflecting on a large body of literature that has utilized previous waves of PISA, there will be substantial cross-national variations in the gender achievement gap in all subject areas across participating countries. Particularly, a greater number of countries will demonstrate a male advantage in mathematics and science, and a female advantage in reading.

2. How many classes do the countries cluster into in terms of gender-related attitudes and how are the classes different from each other?

HYPOTHESIS 2: Although it is next to impossible to predict the number of classes and characteristics of each class, the classes of gender-related attitudes identified from the analysis will not fall along a single-dimensioned line between traditional and egalitarian values. Instead, they will demonstrate different combinations of attitudes toward male primacy, gender essentialism, and normative gender roles.

3. How do gender achievement gaps in mathematics, science and reading, respectively, associate with the structural and ideological aspect of gender inequality in each country?

HYPOTHESIS 3: There will be a negative relationship between gender equality and the gender gap in achievement for all subjects. That is, in countries where gender egalitarian values are more prevalent, the achievement gap by gender will be smaller. This association will hold for both structural and ideological aspects of gender inequality.

Chapter 3

Methodology

Data

The Programme for International Student Assessment (PISA)

The primary source of data of this study is PISA, a large-scale cross-national survey administered by OECD. Specifically, the current study utilized the latest survey, namely PISA 2015, in order to investigate the proposed research questions. PISA is an age-based survey that targets 15-year-old students of each country who are approaching the end of compulsory schooling in most of the participating countries. Students were assessed in mathematics, science and reading. Though, rather than assessing students' mastery of a specific curriculum, PISA aims to evaluate students' ability to apply their knowledge and skills to solve real-life challenges. Started in 2000, PISA has been carried out every three years, and the most recent survey took place in 2015. In addition to the academic assessments, PISA also provides background information of students, and schools. In the most recent survey of 2015, countries were given the option to distribute questionnaire to parents and teachers. 18 countries opted in to distribute the parent survey while 19 countries did so for the teacher survey.

Sample design

The intended target population of PISA in each country was 15-year-old students who attended educational institutions in grades 7 and higher (OECD, 2017). According to the technical report (OECD, 2017), the sampling of the survey participants was carefully selected through scientifically designed sampling strategies in order to assure that the full target population in each country has been represented.

The students were selected via a two-stage stratified sampling design in all participating countries besides the Russian Federation. The first-stage sampling units were schools that had 15-year-old students. These schools were sampled from a comprehensive national list of PISA eligible schools with Probability Proportional to Size (PPS) sampling. Students were sampled in the second-stage. Approximately 35 to 42 students were selected based on whether the countries administered paper- or computer-based surveys. All students who were selected to participate in the survey were done so with equal probability. However, all of the students who attend schools with fewer than 42 students were selected.

In case of Russian Federation, a three-stage design was employed. Instead of sampling schools at the first stage, geographic areas were sampled as first-stage units using probability proportional to size sampling. Then, as stated in the previous paragraph, the schools were sampled within the geographic areas with PPS sampling followed by the third-stage sampling where students were sampled from selected schools.

Participants

Despite efforts made to reduce exclusions during the data collection process, exclusions occurred at both school level and within-school level when necessary (OECD, 2017). The exclusions took place mainly due to special education needs or language barriers. At the school level, an entire school was ruled out if the school was exclusively attended by students who were not able to take the academic assessment for intellectual, functional, or linguistic reasons. In accessibility due to geographical isolation was also a valid reason to exclude a school. Within a school, students were excluded if they were intellectually and functionally unable to go through a valid assessment, or had language barriers. Yet, the OECD (2017) reported that the overall exclusion rate within a country was required to be managed below 5% of the PISA desired target population.

In the 2015 study, roughly 540,000 students completed the survey from 72 participating countries and economies. Any students who were 15-year-old and (1) enrolled full-time in educational institutions; (2) enrolled part-time in educational institutions; (3) participate in vocational training programs, or (4) attend foreign schools within the country were able to take part in the survey. However, it should be noted that any student who were 15 years old and was home schooled, schooled at workplace, or resided out of the country were not included in the target population.

Analytic Sample

The first analysis, which examined the gender achievement gap, was analyzed for 69 countries and economies. Cyprus, Kazakhstan, and Malaysia were excluded due to the unavailability of data in PISA. For the latent class analysis, ten countries that did not participate in the WVS or the EVS were further removed from the analysis. These countries included Canada, Costa Rica, Dominican Republic, Hong Kong, Indonesia, Israel, Macao, United Arab Emirates, and Vietnam. Argentina, which was a participating country of the sixth wave of WVS, had no responses to one of the four gender attitude survey question, which made it impossible to conduct a latent class analysis. Thus, Argentina was also dropped from the study. Finally, Chinese Taipei, Kosovo, Malta, and Trinidad and Tobago were further excluded due to insufficient country-level information for the multilevel analysis. This yielded a sample size of 384,947 students from 55 countries and societies. Table 3-1 demonstrates the analytic sample size for each country in addition to information of the country-level variables incorporated in this study.

Table 3-1. List of countries and country-level variables

Country	GII	GDP per capita	GINI	Sample (N)
Albania	0.267	3934.90	28.96	5,215
Algeria	0.429	4160.22	27.6	5,519
Australia	0.12	56408.34	34.7	14,530
Austria	0.078	44255.58	30.5	7,007
Belgium	0.073	40375.39	28.1	9,651
Brazil	0.414	8757.21	51.3	23,141
China	0.164	8069.21	42.2	9,841
Bulgaria	0.223	6993.48	37.4	5,928
Chile	0.322	13653.23	47.7	7,053
Colombia	0.393	6044.53	51.1	11,795
Croatia	0.141	11638.08	32.2	5,809
Czech Republic	0.129	17715.62	25.9	6,894
Denmark	0.041	53013.00	28.5	7,161
Estonia	0.131	17155.87	34.6	5,587
Finland	0.056	42419.57	26.8	5,882
France	0.102	36526.77	32.3	6,108
FYROM (Macedonia)	0.16	4834.14	35.6	5,324
Georgia	0.361	3764.64	38.5	5,316
Germany	0.066	41323.92	31.4	6,504
Greece	0.119	18070.78	35.8	5,532
Hungary	0.252	12483.87	30.9	5,658
Iceland	0.051	50734.44	25.6	3,371
Ireland	0.127	62544.63	31.9	5,741
Italy	0.085	30171.74	34.7	11,583
Japan	0.116	34474.14	32.1	6,647
Jordan	0.478	4096.10	33.7	7,267
Korea	0.067	27105.08	31.6	5,581
Latvia	0.191	13658.09	35.1	4,869
Lebanon	0.381	8452.44	31.8	4,546
Lithuania	0.121	14252.43	37.7	6,525
Luxembourg	0.075	101446.79	31.2	5,299
Mexico	0.345	9152.87	48.2	7,568
Moldova	0.232	1832.50	27	5,325
Montenegro	0.156	6514.27	31.9	5,665
Netherlands	0.044	44746.33	28.6	5,385
New Zealand	0.158	38201.57	34.9	4,520
Norway	0.053	74498.14	26.8	5,456
Peru	0.385	6030.34	44.3	6,971
Poland	0.137	12564.48	32.1	4,478
Portugal	0.091	19252.63	35.6	7,325
Qatar	0.542	66346.52	40.4	12,083
Romania	0.339	8978.39	27.5	4,876
Russian Federation	0.271	9329.29	37.7	6,036
Singapore	0.068	53629.74	40.9	6,115
Slovak Republic	0.18	16132.86	26.1	6,350

Slovenia	0.053	20873.16	25.7	6,406
Spain	0.081	25789.52	36	6,736
Sweden	0.048	50812.19	27.2	5,458
Switzerland	0.04	82016.02	32.5	5,860
Thailand	0.366	5814.86	37.8	8,249
Tunisia	0.289	3828.10	35.8	5,375
Turkey	0.328	10984.77	41.2	5,895
United Kingdom	0.131	44305.55	34.1	14,157
United States	0.203	56469.01	41	5,712
Uruguay	0.284	15524.84	41.7	6,062
Average	0.192	26948.40	34.4	-
Total	-	-	-	384,947

Measures

Dependent variables

Three dependent variables were considered in this study: (1) mathematics achievement, (2) science achievement, and (3) reading achievement. The information on students' mathematics, science, and reading achievement was obtained from the ten plausible values provided by PISA. PISA employs a combination of an IRT model and a latent regression model in order to produce plausible values (OECD, 2017). Instead of using raw scores from the assessment, PISA computes plausible values using responses from the test items as well as various background information of students collected from the questionnaire. It is maintained that the accuracy of the assessment of student academic performance could be increased through this process (OECD, 2017).

Independent variables

There were two independent variables considered in this study in order to answer the first and third research questions, respectively: (1) gender of the student, and (2) structural and ideological gender equality of each country.

Gender

This information was retrieved from the PISA student-level survey, and was used to estimate the gender achievement gap in each country. The gender of the student was coded 1 if female and otherwise 0.

Gender equality

The gender equality of each country has been measured by two key national-level variables: Gender Inequality Index (GII) and the national gender ideology.

Gender Inequality Index (GII). The GII is an inequality index that was constructed by the United Nations Development Programme (UNDP). Information on GII was retrieved from the Human Development Reports published by the UNDP. The Human Development Reports have been published every year since 1990, which aims to prompt discussions on topics related to human development at the national, regional, and global level.

In order to maintain high standards of data quality, the Human Development Report Office (HDRO) consults and communicates with the national and international statistical agencies on a regular basis, and frequently held various meeting, workshops, and conferences so that distinguished statisticians and economists across the world could gather and discuss potential problems and solutions for the collection, management, and use of data. It is also the HDRO's duty to mitigate deficiencies in the definitions, or classification of a concept across countries by bringing these issues to light and finding solutions by consulting relevant agencies.

According to the Human Development Report (2016), the calculation of GII was based on the disparities between male and female achievement in three social dimensions: health, empowerment and labor market participation as shown in Figure 3-1. As demonstrated, the construction of GII was based on female reproductive health, measured by maternal mortality ratio and adolescent birth rates; female empowerment, measured by the share of females in parliamentary seats and the proportion of females and males with at least some secondary

education who are 25 years of age or older; and economic status measured by female and male labor force participation rate who are 15 years or older. The GII ranges between 0 and 1, in which higher values indicate greater gender inequality within a society.

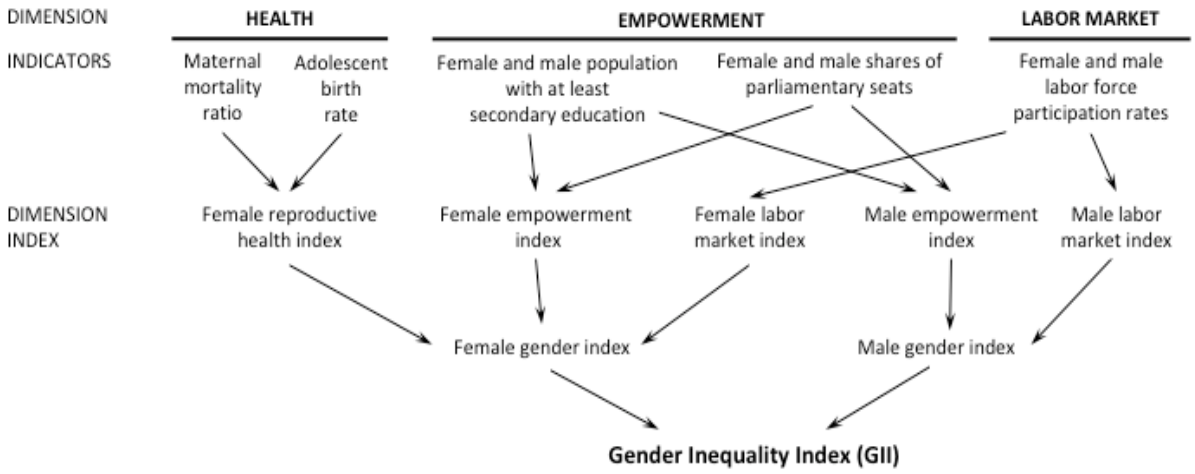

Figure 3-1. The calculation of the Gender Inequality Index (GII)
Note: The diagram was retrieved from the Technical notes of the Human Development Report 2016.

National gender ideology. Based on the results form the latent class analysis examining individuals' gender attitude, the countries were classified into different classes (further details of this analysis will be discussed in the following section). The information used in the analysis was mainly retrieved from the sixth wave of WVS, which was collected between 2010 and 2014. However, the corresponding information of the countries that were not included in this survey was garnered from the fourth wave of EVS (2008-09)[1]. Currently, the sixth and fourth wave is the most recent data released by WVS and EVS respectively. The surveys were selected to be utilized

[1] Information of the following countries were retrieved from the fourth wave of EVS: Albania, Austria, Belgium, Bulgaria, Croatia, Czech Republic, Denmark, Finland, France, FYROM(Macedonia), Greece, Hungary, Iceland, Ireland, Italy, Kosovo, Latvia, Lithuania, Luxembourg, Malta, Moldova, Montenegro, Norway, Portugal, Slovak Republic, Switzerland, and United Kingdom.

in this study due to their close proximity in time from which the main data, namely PISA 2015, has been collected.

However, some scholars have expressed their concerns with regard to the use of WVS and EVS survey questions to measure gender egalitarianism. To be specific, Braun (2008) pointed out that the use of gender-related attitude questions from various surveys might not be an adequate way to measure egalitarianism. He emphasized that the disagreement to a question that represents gender egalitarian view does not necessarily mean that the respondent holds a traditional view. Rather, it could be the case that the individual is indeed egalitarian but is merely disagreeing to a certain stance of egalitarianism, which is represented in the question. Similarly, traditional items do not have the ability to distinguish between non-traditional responders. This study adopted several strategies that had been suggested by Knight and Brinton (2017) in order to alleviate such concern. First of all, the gender ideology survey questions incorporated in the analysis are part of the list of items suggested by Davis and Greenstein (2009), which were considered valid and reliable measures of gender ideology. Furthermore, the current study included questions representing both traditional and egalitarian perspective. As stated by Braun (2008), the orientation of the gender ideology questions might play a role in the measurement of gender attitudes. Thus, this study attempted to incorporate traditional-slanted items (Q1, Q2, Q3) as well as egalitarian-slanted item (Q4). One may insist that it might not be enough to incorporate only one egalitarian-slanted item, however, given the items that were available for analysis in both WVS and EVS, this was the only option that could be considered. Finally, this study is interested in patterns of responses. As emphasized by Knight and Brinton (2017), a single agreement or disagreement to a certain question will not determine one's level of egalitarianism. Rather, the pattern of responses to the questions will serve as an indicator of the "type" of egalitarianism.

Following the precedents of Knight and Brinton (2017), this study attempted to select questions that were tapping different aspects of gender-related attitudes. However, due to the inconsistency of the set of questions measuring gender ideology between the 6th wave of WVS and 4th wave of EVS, the study could only incorporate part of the gender attitude questions presented in both surveys as follows:

Q1. When a mother works for pay, the children suffer.

Q2. Being a housewife is just as fulfilling as working for pay.

Q3. When jobs are scarce, men should have more right to a job than women.

Q4. Having a job is the best way for a woman to be an independent person.

These questions are believed to investigate one's opinion on maternal employment (Q1), women's role at home (Q2), gender equality in the work place (Q3), and women's employment. The responses to each question were dummy-coded where 0 representing disagreement and 1 representing agreement to the statement. This information was utilized to identify patterns of responses in relation to gender-related questions via latent class analysis. The country-level gender ideology was measured as follows: first, the types of gender attitudes was identified through latent class analysis of the pooled data. Then, the national gender ideology was measured as the percentage of individuals consisting each category of gender attitudes for each country. In other words, four variables (the number of gender ideology identified through latent class analysis) constituted the national gender ideology, which summed up to 100 percent. One variable was omitted in the multilevel analysis in order to avoid perfect collinearity among the gender ideology variables.

Control variables

Additional variables from the student-, and country-level were added as controls in the multilevel analysis.

Student level. The *socio-economic status (SES)* of each individual was controlled for in the multilevel analysis. Referred to as the PISA index of economic, social and cultural status (ESCS), PISA provided a construct of SES that can be used across countries. The index was inferred from the following three indicators: 1) the highest occupation status of parents, 2) the highest educational level of parents in years of education according to ISCED, and 3) home possessions. The index of home possessions incorporates family wealth, cultural possessions, and home educational resources as well as books in the home. According to the PISA Technical Report (OECD, 2016), the measure of family wealth was based on the students' responses on whether they had a room of their own, a link to the Internet, a dishwasher, a DVD player, and three other country-specific items at home. In addition, their responses on the number of cellular phones, televisions, computers, cars and the rooms with a bath or shower was also incorporated. Cultural possessions were measured by the students' responses to whether they had the following at home: classic literature, books of poetry and works of art. The existence of educational resources at home such as a desk and a quiet place to study, a computer that students can use for schoolwork, educational software, books to help with students' school work, technical reference books and a dictionary was an indicator of home educational resources. Finally, number of books was a four-level categorical variable demonstrated as follows: 0-10 books, 11-25 or 26-100 books, 101-200 or 201-500 books, more than 500 books. ESCS was standardized to a mean of zero and a standard deviation of one.

The *type of school* and the *location of school* attended by each individual were also controlled for. Obtained from the school questionnaire, school type was dummy coded as 1 if private, and 0 if public. In terms of school location, principals were asked whether the school was located in (1) a village, hamlet or rural area where there are fewer than 3,000 people, (2) a small town in which 3,000 to approximately 15,000 people reside, (3) a town of 15,000 to about 100,000 people), (4) a city where there are 100,000 to about 1,000,000 people, or (5) a large city

with over 1,000,000 people. Following the example of OECD (2013) in their analysis on the "urban advantage," school location was coded 1 if the responses indicated that the school was settled in a city of more than 100,000 people, while it was coded as 0 if the school was located in a less-populated area such as rural areas, or (small) towns with up to 100,000 inhabitants.

Country level. *GDP per capita* and *GINI coefficient* were added as country-level control variables. Obtained from the World Bank database (http://data.worldbank.org/), GDP per capita was included as a direct measure of the level of the national economic development in 2015. Natural logarithmic transformation was applied to this information, so that the distribution mirrors a normal distribution. This variable ranges from 1,833 US dollars for Moldova to 101,447 US dollars in Luxembourg (M = 26,948 US dollars). The GINI coefficient, a measure that demonstrates the deviation of the income distribution among individuals and households within a nation, was also collected from the World Bank database. Information of GINI coefficient was retrieved for the year 2015, however, countries that did not have information for this year was replaced with data from adjacent years. Data for New Zealand, Qatar, and Singapore were obtained from reports released by the respective national statistical institutions. Ranging between 0 and 100, greater GINI values indicate greater inequality in the distribution of income and wealth. The country with the greatest equality was Iceland (GINI = 25.6) while the country with the highest level of inequality was Brazil (GINI = 51.3).

Table 3-2 provides a summary of the variables used in this study with a brief description of how the variables were measured, and the sources of data.

Table 3-2. Summary of variables

Variable(s)	Source of data	Description
Student-level		
Academic achievement Mathematics Science Reading	PISA 2015	Computed plausible values of academic performance. 10 values for each subject area.

Gender	PISA 2015	Gender of individual (1 = female; 0 = male)
SES	PISA 2015	PISA index of economic, social and cultural status (ESCS). Standardized to a mean of zero and a standard deviation of one.
School location	PISA 2015	Location of school attended by the respondent (1 = private; 0 = public)
School type	PISA 2015	Type of school attended by the respondent (1 = urban; 0 = nonurban)
Country-level		
GII	Human Development Reports 2016	An inequality index constructed by UNDP. Ranges between 0 and 1 where higher values indicate greater gender inequality
National gender ideology	EVS & WVS	The percentage of individuals consisting each type of gender ideology in each country. The types were identified through latent class analysis of the following four gender attitude survey questions. Q1. When a mother works for pay, the children suffer. Q2. Being a housewife is just as fulfilling as working for pay. Q3. When jobs are scarce, men should have more right to a job than women. Q4. Having a job is the best way for a woman to be an independent person.
GDP per capita (logged)	Worldbank database	GDP per capita in current US dollars
GINI	Worldbank database; National statistical institutions	A measure demonstrating deviation of the income distribution among individuals and households within a nation. Range between 0 and 100.

Analysis Plan

First of all, in order to answer the first research question, the gender gap in mathematics, science, and reading achievement was examined by looking at the coefficient of *gender* estimated from the following ordinary least squares (OLS) regression for each country and each subject:

$$Y_i = a + \beta_1 gender_i + r_i,$$

where Y_i is the achievement score of each subject (i.e., mathematics, science, and reading) for student *i* in a country; a is the intercept; and r_i is the error term. The gender achievement gap – the coefficients of *gender* – in all three subjects was compared across countries in order to estimate the variations in achievement gap by country.

Then, a latent class analysis was conducted to identify distinct classes of gender attitudes using the aforementioned four gender-related survey questions to address the second research question. Latent class analysis is a technique that is used to group individuals into homogenous subgroups, often referred to as latent classes or latent types, based on their response patterns across a set of questionnaire items (Geiser, 2010; Vermunt & Magidson, 2002). In a latent class analysis, each latent class displays a class-specific response profile. The differences in such class membership explain individual differences in observed response patterns (Geiser, 2010; Knight & Brinton, 2017). Latent class analysis is similar with factor analysis in the sense that both can be used for data reduction (Geiser, 2010). Yet, the main difference between these two techniques is that the former uses discrete (categorical) data to explain the relationships between categorical indicators, whereas the latter analyzes sets of continuous indicators in order to characterize continuous latent variables (Geiser, 2010; Knight & Brinton, 2017; McCutcheon, 1987) Thus, for this study, which attempts to construct a typology of gender-related attitudes across countries, latent class analysis was implemented.

The latent class model with *r* observed dichotomous items, *u*, with a categorical latent variable *c* (*c = k; k = 1, 2, 3, … , K*), the marginal item probability for an individual *v* obtaining a score of 1 on item *i* ($u_{vi} = 1$) was specified as follows (Nylund, Asparouhov, & Muthén, 2007):

$$P(u_{vi} = 1) = \sum_{k=1}^{K} P(c = k)P(u_{vi} = 1|c = k).$$

The joint probability of all observed items *r* was,

$$P(u_{v1}, u_{v2,} \dots, u_{vr}) = \sum_{k=1}^{K} P(c = k)P(u_{v1} | c = k)P(u_{v2} | c = k) \dots P(u_{vr} | c = k).$$

Finally, hierarchical linear models were estimated to address the third research question. Such approach was selected instead of an ordinary least squares (OLS) regression because of the nested nature of the PISA data. Considering that PISA is a data with a hierarchical structure where students are nested within countries, the individual observations may not be independent. Using a typical OLS approach in examining such data might yield large biases in the standard errors and result in many spuriously "significant" results (Hox, 2010). In order to obtain precise estimation, Hox (2010) recommends using multilevel models to deal with hierarchical data. For mathematics, science and reading, the following two models were estimated respectively. At the student-level, the following model was incorporated for both models:

Level-1 (Student-level):

$$Y_{ij} = \pi_{0j} + \pi_{1j} SES_{ij} + \pi_{2j} gender_{ij} + \pi_{3j} school\ type_{ij} + \pi_{4j} school\ location_{ij} + e_{ij},$$

where Y_{ij} is the achievement of child *i* in country *j*; π_{0j} is the intercept for country *j*; π_{1j}, π_{2j}, π_{3j} and π_{4j} are the coefficient estimates for SES, gender, school type, and school location of child *i* in country *j* respectively; and e_{ij} is the level-1 random effect.

For the academic achievement of each subject area, country-level predictors were added to the model. The first model at the country-level could be specified as follows:

[Model 1] Level-2 (Country-level):

$$\pi_{0j} = \beta_{00} + \beta_{01} GII_{1j} + \beta_{02} Ideology\ Type_{\ 2j} + \beta_{03} GDP\ per\ capita_{3j} + \beta_{04} GINI_{4j} + r_{0j},$$

$\pi_{1j} = \beta_{10}$,

$\pi_{2j} = \beta_{20} + r_{2j}$,

$\pi_{3j} = \beta_{30}$,

$\pi_{4j} = \beta_{40}$,

where the outcomes π_{0j}, π_{1j}, π_{2j}, π_{3j}, and π_{4j} are the country effects which are the coefficient estimates from the level-1 model; β_{00}, β_{10}, β_{20}, β_{30}, and β_{40} are the intercepts for country *j*; and r_{0j}, and r_{2j} are the level-2 random effects for each corresponding outcomes of interest.

Then, as the second model, cross-level interaction terms between individual-level and country-level variables were added. Specifically, the model attempted to examine whether the relationship between academic achievement and national-level variables was moderated by gender. The model was specified as follows:

[Model 2] Level-2 (Country-level):

$\pi_{0j} = \beta_{00} + \beta_{01}GII_{1j} + \beta_{02}Ideology\ Type_{\ 2j} + \beta_{03}GDP\ per\ capita_{3j} + \beta_{04}GINI_{4j} + r_{0j}$,

$\pi_{1j} = \beta_{10}$,

$\pi_{2j} = \beta_{20} + \beta_{21}GII_{1j} + \beta_{22}Ideology\ Type_{\ 2j} + \beta_{23}GDP\ per\ capita_{3j} + \beta_{24}GINI_{4j} + r_{2j}$,

$\pi_{3j} = \beta_{30}$,

$\pi_{4j} = \beta_{40}$,

where the outcomes π_{0j}, π_{1j}, π_{2j}, π_{3j}, and π_{4j} are the country effects which are the coefficient estimates from the level-1 model; β_{00}, β_{10}, β_{20}, β_{30}, and β_{40} are the intercepts for country *j*; and r_{0j}, and r_{2j} are the level-2 random effects for each corresponding outcomes of interest.

For both models, individual-level variables were centered around the group mean and the country-level predictors were centered around the grand mean. Senate weights were applied to ensure that the countries included contributed to the analyses equally regardless of their population or sample size. The latent class analysis as well as the OLS regression was examined using STATA versions 15 and 14, respectively, and the multilevel analysis was carried out with the HLM software version 7.

Chapter 4

Results

In this chapter, the empirical findings of this study are presented in the following order. In the first section, a brief descriptive analysis of the individual-level variables is presented. In the following section, I attempt to illustrate what I have found in relation to the first research question: *How do the gender gaps in mathematics, science, and reading achievement vary in size and direction by country?* The next section describes the empirical results regarding the second research question: *How many classes do the countries cluster into in terms of gender-related attitudes and how are the classes different from each other?* In this section, I attempted to provide meaningful interpretations of the latent classes identified through the analysis. Finally, I present the findings for the final research questions: *How do gender achievement gaps in mathematics, science and reading, respectively, associate with the structural and ideological aspect of gender inequality in each country?*

Descriptive Statistics

Table 4-1 presents the weighted descriptive statistics of individual-level variables across the 55 countries included in this study. Not surprisingly, there were noticeable cross-national variations on all indicators across countries. First of all, the mean SES of each country ranged between -1.43 (Turkey) and 0.73 (Iceland) where the average was -0.24. There were also great disparities in the average academic performances across nations. Among the 55 countries, students from Algeria scored the lowest in mathematics (359.61) while their counterparts from Singapore scored the highest (564.19) on average. The average mathematics score was about 464, similar to what Croatian students roughly received. A similar pattern was observed with regard to science achievement. Algeria had the lowest average science achievement (375.75) whereas Singapore received the highest (555.57). However, the difference in the average achievement was

much smaller for science than for mathematics between these two countries. Finally, in relation to reading achievement, students from Lebanon had the lowest (346.55) while their counterparts from Singapore once again recorded the highest average score of 535.1.

Table 4-1. Weighted descriptive statistics of individual-level variables

Country	SES		Female (%)	City (%)	Private (%)	Math		Science		Reading	
	M	SE	M	M	M	M	SE	M	SE	M	SE
Albania	-	-	50.1	0.0	11.6	413.2	3.45	427.2	3.28	405.3	4.13
Algeria	-1.28	0.04	47.3	3.5	1.5	359.6	2.95	375.8	2.64	349.9	3.00
Australia	0.27	0.01	49.6	45.5	43.7	493.9	1.61	510.0	1.54	502.9	1.69
Austria	0.09	0.02	49.5	19.8	12.6	496.7	2.86	495.0	2.44	484.9	2.84
Belgium	0.16	0.02	49.2	11.1	54.5	507.0	2.35	502.0	2.29	498.5	2.42
Brazil	-0.96	0.03	51.5	18.9	14.5	377.1	2.86	400.7	2.30	407.4	2.75
Bulgaria	-0.08	0.03	47.2	16.4	1.2	441.2	3.95	445.8	4.35	431.7	5.00
Chile	-0.49	0.03	49.7	28.7	63.1	422.7	2.54	447.0	2.38	458.6	2.58
China	-1.07	0.04	46.7	18.6	10.6	531.3	4.89	517.8	4.64	493.9	5.13
Colombia	-0.99	0.04	52.7	27.5	24.1	389.6	2.29	415.7	2.36	424.9	2.94
Croatia	-0.24	0.02	51.9	4.1	2.3	464.0	2.77	475.4	2.45	486.9	2.68
Czech Republic	-0.21	0.01	48.7	11.5	8.2	492.3	2.40	492.8	2.27	487.3	2.60
Denmark	0.59	0.02	49.8	4.7	23.2	511.1	2.17	501.9	2.38	499.8	2.54
Estonia	0.05	0.01	49.0	0.0	4.2	519.5	2.04	534.2	2.09	519.1	2.22
Finland	0.25	0.02	48.2	0.0	4.5	511.1	2.31	530.7	2.39	526.4	2.55
France	-0.14	0.02	50.4	9.2	21.0	492.9	2.10	495.0	2.06	499.3	2.51
FYROM (Macedonia)	-0.23	0.01	47.3	0.3	1.9	371.3	1.28	383.7	1.25	351.7	1.41
Georgia	-0.33	0.02	47.4	23.9	7.4	403.8	2.78	411.1	2.42	401.3	2.96
Germany	0.12	0.02	49.1	5.3	7.3	506.0	2.89	509.1	2.70	509.1	3.02
Greece	-0.08	0.03	48.0	11.0	4.9	453.6	3.75	454.8	3.92	467.0	4.34
Hungary	-0.23	0.02	49.9	20.5	18.0	476.8	2.53	476.8	2.42	469.5	2.66
Iceland	0.73	0.01	51.5	0.0	0.6	488.0	1.99	473.2	1.68	481.5	1.98
Ireland	0.16	0.02	48.7	22.5	57.3	503.7	2.05	502.6	2.39	520.8	2.47
Italy	-0.07	0.02	50.3	7.3	4.1	489.7	2.85	480.6	2.52	484.8	2.68
Japan	-0.18	0.01	49.6	20.5	31.8	532.4	3.00	538.4	2.97	516.0	3.20
Jordan	-0.42	0.03	50.5	15.9	20.0	380.3	2.65	408.7	2.67	408.1	2.93
Latvia	-0.44	0.02	49.9	0.0	2.0	482.3	1.87	490.2	1.56	487.8	1.80
Lebanon	-0.60	0.04	53.5	3.2	50.3	396.3	3.69	386.5	3.40	346.6	4.41
Lithuania	-0.06	0.02	49.3	0.0	2.3	478.4	2.33	475.4	2.65	472.4	2.74
Luxembourg	0.07	0.01	50.3	0.0	15.6	485.8	1.27	482.8	1.12	481.4	1.44
Mexico	-1.22	0.04	49.3	22.4	12.5	408.0	2.24	415.7	2.13	423.3	2.58
Moldova	-0.69	0.02	49.8	0.0	1.5	419.7	2.47	428.0	1.97	416.2	2.52
Montenegro	-0.18	0.01	48.9	0.0	0.6	417.9	1.46	411.3	1.03	426.9	1.58

Netherlands	0.16	0.02	50.2	0.0	60.1	512.3	2.21	508.6	2.26	503.0	2.41
New Zealand	0.17	0.02	49.7	29.4	6.6	495.2	2.27	513.3	2.38	509.3	2.40
Norway	0.48	0.02	49.4	0.0	1.9	501.7	2.23	498.5	2.26	513.2	2.51
Peru	-1.08	0.04	49.8	1.7	31.4	386.6	2.71	396.7	2.36	397.5	2.89
Poland	-0.39	0.02	49.1	3.2	3.5	504.5	2.39	501.4	2.51	505.7	2.48
Portugal	-0.39	0.03	49.5	4.0	5.5	491.6	2.49	501.1	2.43	498.1	2.69
Qatar	0.58	0.01	48.8	20.2	41.8	402.4	1.27	417.6	1.00	401.9	1.02
Romania	-0.58	0.04	50.2	5.8	1.1	444.0	3.79	434.9	3.23	433.6	4.07
Russian Federation	0.05	0.02	50.8	21.3	1.0	494.1	3.11	486.6	2.91	494.6	3.08
Singapore	0.03	0.01	48.3	100	8.4	564.2	1.47	555.6	1.20	535.1	1.63
Slovak Republic	-0.11	0.02	48.5	0.0	11.6	475.2	2.66	460.8	2.59	452.5	2.83
Slovenia	0.03	0.01	48.3	0.0	2.6	509.9	1.26	512.9	1.32	505.2	1.47
South Korea	-0.20	0.02	47.8	47.6	34.7	524.1	3.71	515.8	3.13	517.4	3.50
Spain	-0.51	0.04	50.1	10.0	31.3	485.8	2.15	492.8	2.07	495.6	2.36
Sweden	0.33	0.02	49.5	-	-	493.9	3.17	493.4	3.60	500.2	3.48
Switzerland	0.14	0.02	47.9	0.0	6.1	521.3	2.92	505.5	2.90	492.2	3.03
Thailand	-1.23	0.04	56.7	6.6	14.8	415.5	3.03	421.3	2.83	409.1	3.35
Tunisia	-0.83	0.03	53.7	2.5	2.1	366.8	2.95	386.4	2.10	361.1	3.06
Turkey	-1.43	0.05	50.0	39.4	4.8	420.5	4.13	425.5	3.93	428.3	3.96
United Kingdom	0.21	0.02	49.2	8.0	56.0	492.5	2.50	509.2	2.56	498.0	2.77
United States	0.10	0.04	50.0	10.7	7.7	469.6	3.17	496.2	3.18	496.9	3.41
Uruguay	-0.78	0.02	52.4	28.8	15.4	418.0	2.50	435.4	2.20	436.6	2.55
Average	-0.24	0.02	50.0	13.2	16.4	463.8	2.59	468.5	2.47	464.1	2.79

In addition, to better understand the relationships among the national-level variables, the Pearson correlation was examined. As presented in Table 4-2, the correlations among the country-level variables were statistically significant. There was a small and negative correlation between logged GDP per capita and GINI, while stronger relationships were observed between GII and logged GDP per capita (r = -0.63), and GII and GINI (r = -0.52).

Table 4-2. Correlation of national-level variables

	GDP per capita (logged)	GINI	GII
GDP per capita (logged)	1		
GINI	-0.25*	1	
GII	-0.63***	0.52***	1

Research Question 1

Research question 1 inquired how the gender gaps in mathematics, science, and reading achievement varied in size and direction by country. In this section, the results for this question are displayed by subject.

Mathematics

As demonstrated in figure 4-1, a substantial cross-national variation in gender achievement gap in mathematics has been identified. As the bars represent the coefficient estimates for the female variable, positive values indicate that female students outperformed their male counterparts while negative values indicate the opposite. It is evident that a greater number of countries demonstrate a female disadvantage in mathematics achievement across countries. As a matter of fact, female students underperformed their male counterparts by approximately 4.5 points across countries on average. The country that showed the greatest gender gap in mathematics was Austria, in which female students scored approximately 27 points lower than their male counterparts, followed by Lebanon (male advantage, 22.07), and Argentina (male advantage, 21.28).

Interestingly, seven of the 71 countries, demonstrated a statistically significant female advantage in mathematics. For instance, female students from Trinidad and Tobago generally outperformed male students by 17.6 points. Similarly, 15 year-old girls from Jordan, Georgia, Qatar, Albania, Macao, Finland, Algeria, FYROM (Macedonia) surpassed their male counterparts by 14.38, 12.85, 11.57, 9.20, 8.20, 7.51, 6.58, and 6.5 points, respectively. However, about slightly over one-third of the countries did not demonstrate any gender gap in mathematics achievement.

Science

According to Figure 4-2, there were also great disparities in the gender science achievement gap. However, the noteworthy difference between the gender gap in mathematics is that there are greater number of countries where female students exceeded male students in science. At a glance of Figure 4-2, a roughly equal number of countries demonstrated female disadvantage and female advantage in science achievement. Furthermore, unlike the case of mathematics, the estimates of the majority of countries that demonstrated female advantage in science were statistically significant. In fact, on average, female students outperformed their male counterparts by 0.74 points across the countries.

Although results show that there is a balance between the number of countries that demonstrate female advantage and disadvantage in science achievement, the degree of the gender achievement gap is greater in countries with female advantage. For instance, female students from Jordan, the country that displayed the greatest female advantage in science achievement, scored approximately 40 points higher than their male counterparts on average. Meanwhile, female students from Austria, the country that showed the greatest female disadvantage in science achievement, generally underperformed their male counterparts by 18.84 points. There were no differences in science performance by gender in approximately one-third of the countries.

Reading

There was a substantial variation in gender reading achievement gap across the 71 countries. As demonstrated in Figure 4-3, the country with the smallest gap was Peru, where the gap was about 7.7 points, while the country with the greatest gap was Jordan, where the gap was about 72 points.

However, the gender gap in reading achievement illustrates a totally different pattern compared to that of mathematics and science. First of all, female students outperformed their

male counterparts in all countries included in the analysis. As can be observed from Figure 4-3, there is no negative value for the gender gap. What is even more astonishing is the size of the gap. Female students, on average, scored 30 points higher than their male counterparts. Female students in Jordan, scored as high as 72 points above male students. This is more than two times greater than the average gender gap in reading achievement. Second, the estimates were all statistically significant. Unlike the case of mathematics and science, in which one-third of the countries did not demonstrate any gender gap in achievement, there were no countries that did not demonstrate gender gap in favor of girls.

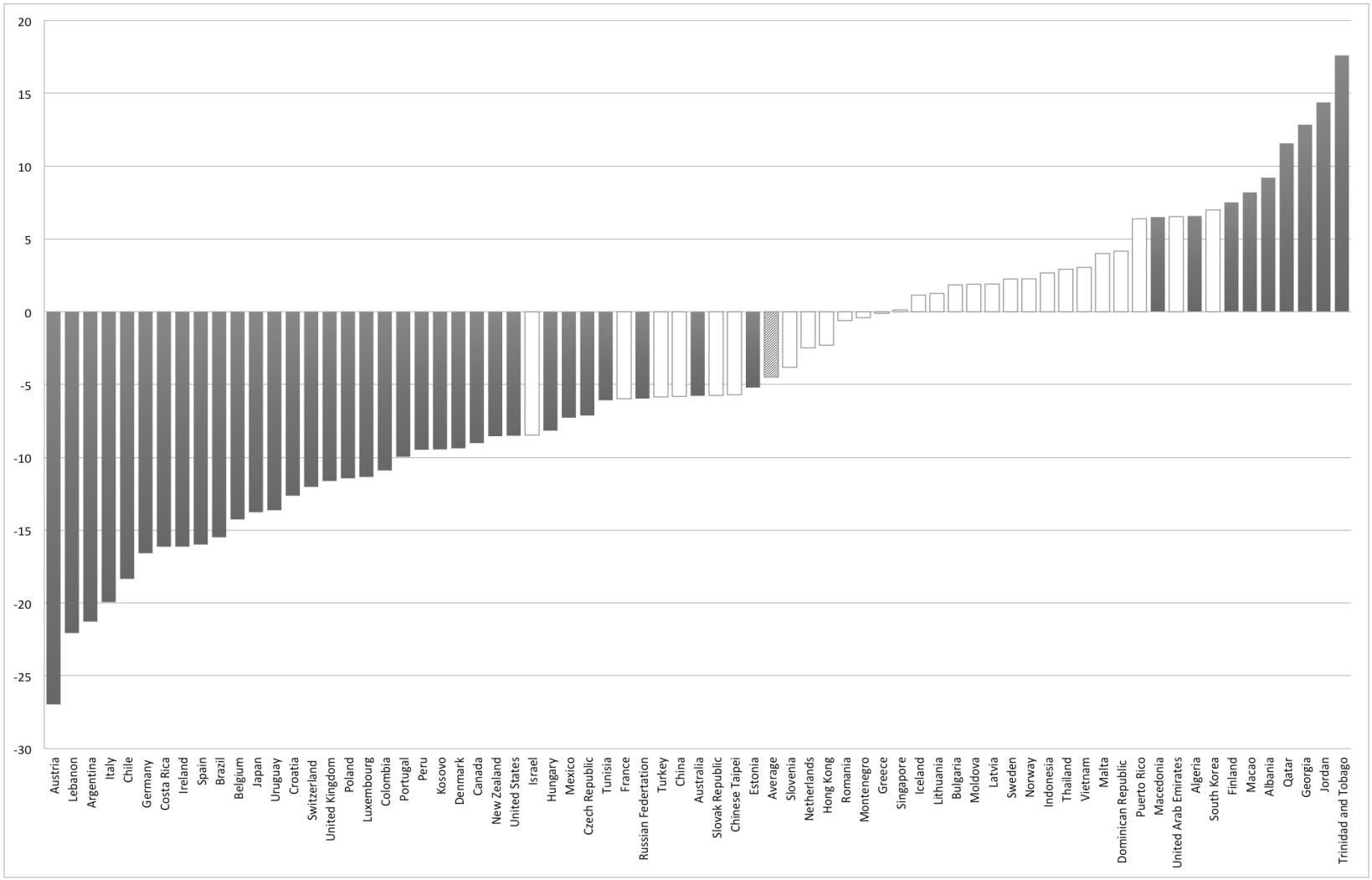

Figure 4-1. Cross-national variation of the gender gap in mathematics achievement

Note. The solid dark bars display that the coefficients are statistically significant at the .05 level. The white bars demonstrate that the coefficients are not statistically significant at the .05 level. Appropriate weights were used for the analysis.

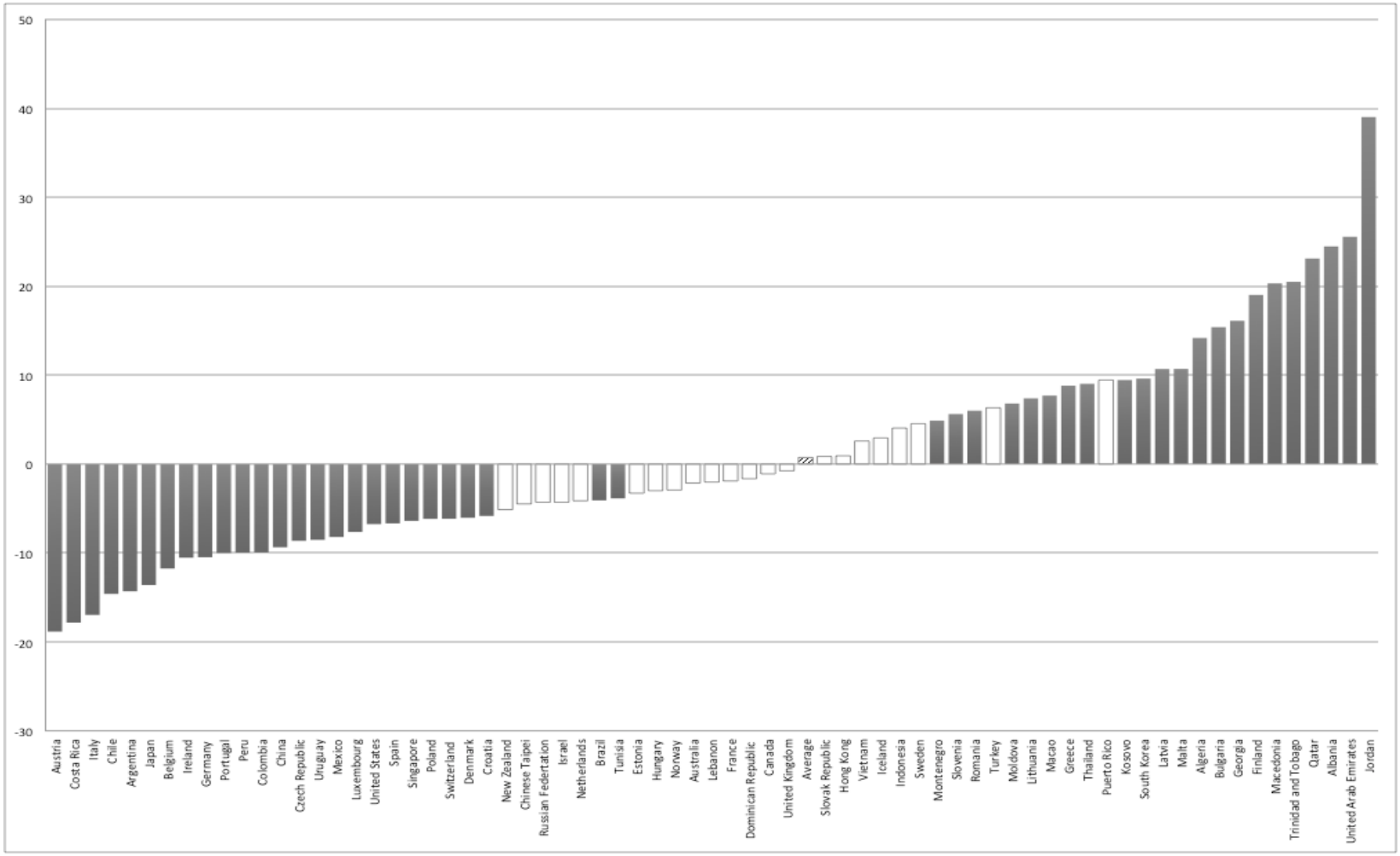

Figure 4-2. Cross-national variation of the gender gap in science achievement

Note. The solid dark bars display that the coefficients are statistically significant at the .05 level. The white bars demonstrate that the coefficients are not statistically significant at the .05 level. Appropriate weights were used for the analysis.

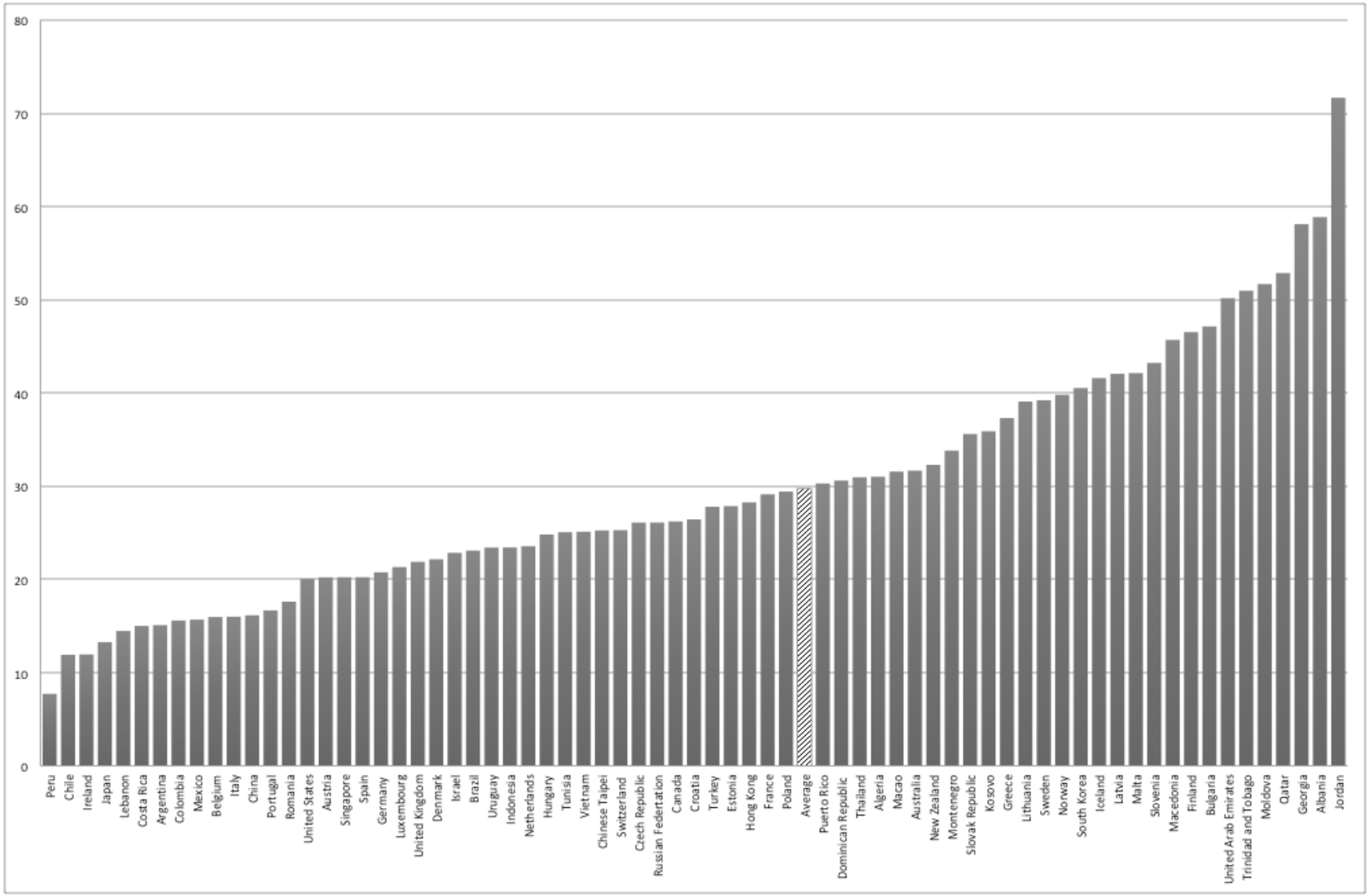

Figure 4-3. Cross-national variation of the gender gap in reading achievement

Note. The solid dark bars display that the coefficients are statistically significant at the .05 level. The white bars demonstrate that the coefficients are not statistically significant at the .05 level. Appropriate weights were used for the analysis.

Research Question 2

Research question 2 sought to answer how many classes the countries could cluster into in terms of gender-role attitudes and how the classes are different from each other. In this section, I illustrated the findings of the latent classes identified through the analysis followed by a description of each class.

Identification of latent classes

To address the second research question, a latent class model was estimated utilizing four indicators as described in the previous section. According to Geiser (2010), information criteria (IC) are used to compare models taking the number of estimated parameters as well as the goodness of fit into account. The best model is the one that not only fits well but also uses fewer parameters as possible (Geiser, 2010). In other words, the model with the smallest Akaike's information criterion (AIC), or Bayesian information criterion (BIC) should be considered. If there is an inconsistency between the AIC and BIC when determining the number of classes in latent class analysis, Nylund and her colleagues (2007) recommended using the BIC index over other ICs.

Table 4-3 demonstrates the values of ICs for the latent class analysis of this study. As can be seen, a four-class solution was the best fit for the current study. Although the value of AIC was the smallest for the five-class solution by a marginal difference, the best model was chosen based on the BIC index following the suggestion of Nylund and her colleagues (2007). The four classes were labeled as (1) liberal egalitarians, (2) partial traditionalists, (3) familial egalitarians, and (4) egalitarian essentialists. In the following section, I attempted to provide descriptions of the characteristics of each class based on the estimated latent class assignment probabilities.

Table 4-3. IC for different class solutions

Model	Number of class	Observations	AIC	BIC
One class	1	103,673	461827.5	461865.7
Two classes	2	103,673	454441.8	454527.8
Three classes	3	103,673	453818.7	453952.3
Four classes	4	103,673	453815.6	453947.5
Five classes	5	103,673	453813.5	454004.5

Description of the latent classes

Figure 4-4 presents the estimated latent class assignment probabilities conditional to class, which served as the basis for the interpretation of each class. The largest class was Class 3 (Familial egalitarian), consisted of 40 percent of the pooled sample, followed by Class 1 (Liberal egalitarian), Class 2 (Partial traditionalist), and Class 4 (Egalitarian essentialist), which consisted 26%, 21% and 12% of the pooled data, respectively.

Class 1. Liberal egalitarians

Class 1 resembles the characteristics of those who are generally known as liberal egalitarians. People who were classified into this class typically demonstrated high support for maternal as well as women employment, advocated equal treatment between men and women at the workplace, and disapproved gender roles at home. To be specific, respondents in Class 1 were the least likely among those in other classes to agree that children suffer when a mother works for pay (Q1, mean of 0.22), and about 87 percent of the members agreed that having a job is the best way for a woman to be an independent person (Q4). They also portrayed disapproval of differential treatment of men and women at the workplace. Only nine percent agreed that men should have more right to a job than women when jobs are scarce (Q3). Members of Class 1 were also the least likely to hold gender essentialist beliefs compared to those in other classes. They generally disagreed that being a housewife is just as fulfilling as working for pay (Q2, mean of 0.31). By and large, respondents were highly similar to what we refer to as egalitarians,

characterized as having strong support for women employment, and equal treatment of men and women at work, while rejecting gender-typical roles.

Class 2. Partial traditionalists

The characteristics of class 2 were similar to what we usually refer to as traditionalists. Respondents not only tended to hold traditional views on women's role at home and at work, but also disapprove economic participations of mothers. For example, respondents were highly supportive of gender-typical roles, agreeing that being a housewife is just as fulfilling as working for pay (Q2, mean of 0.84). Furthermore, they were also the most likely group to support male primacy at the work place. On average, over eighty percent of the respondents believed that men should have more right to a job than women when jobs are scarce (Q3). They were also concerned of the wellbeing of children whose mothers were working. Approximately 80 percent of respondents expressed that the children suffer when a mother works for pay (Q1).

However, the members of Class 2 were different from the typical traditionalists with regard to the support of women's employment. Typically, it would be assumed that an individual with traditional gender attitude would not support women's employment. Yet, respondents of Class 2 showed otherwise. Although it is the lowest among all classes, the respondents genderally agreed that having a job for women is the best way to become an independent person, displaying a relatively high support for women economic participation (Q4, mean of 0.72). Altogether, members of Class 2 tended to hold traditionalist gender attitudes, characterized as having strong gender essentialist beliefs, disapproving mother's employment and supporting male primacy at work. However, interestingly, they portrayed support toward women's engagement in economic activities.

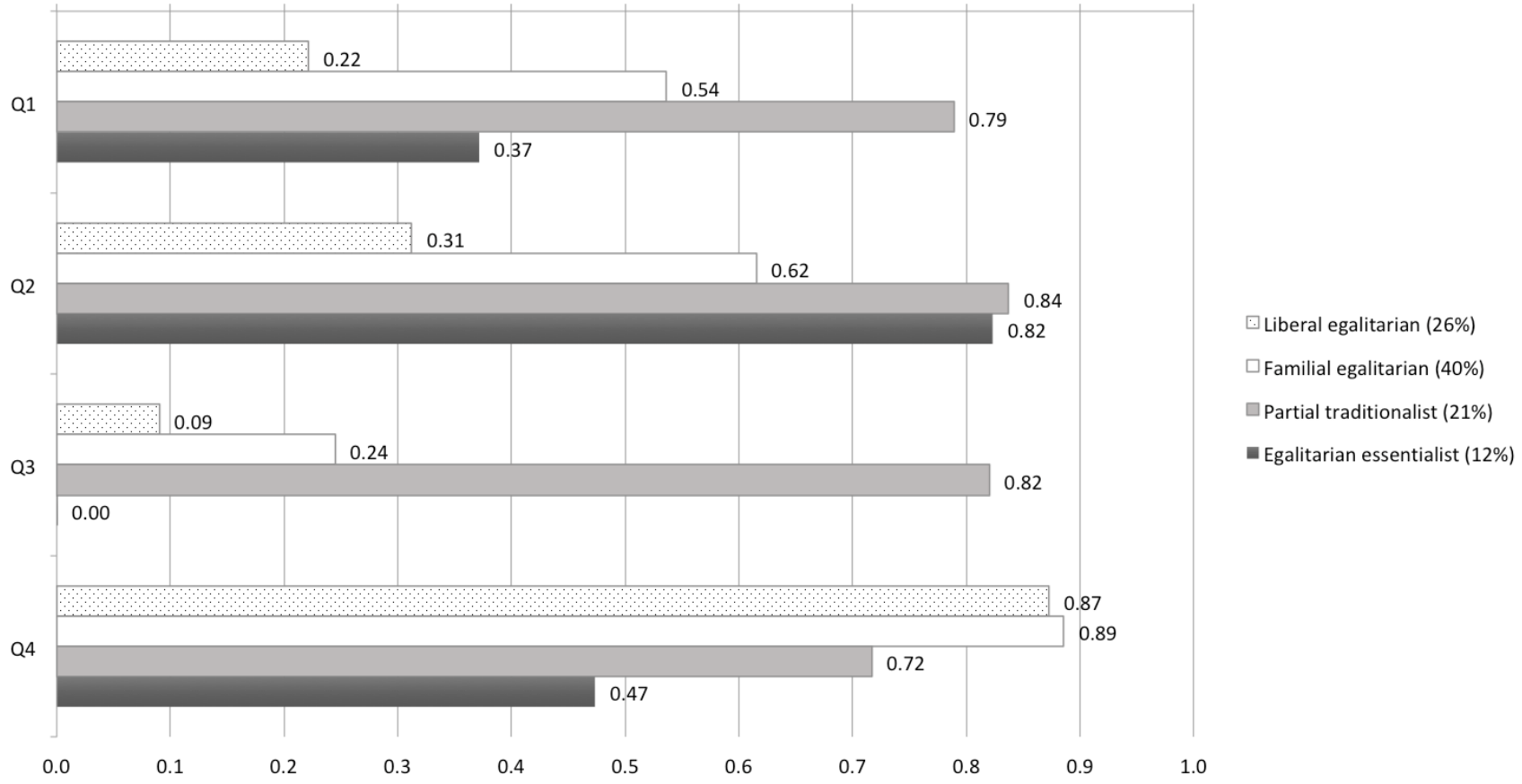

Figure 4-4. Conditional response probabilities of the pooled sample

Class 3. Familial egalitarians

The response pattern of members of Class 3 was similar to that of Class 1 regarding gender equality at work and women's employment. To be specific, respondents generally demonstrated disagreement to the statement, "when jobs are scarce, men should have more right to a job than women," (Q3, mean of 0.24) supporting equal treatment of men and women at the work place. They were likely to believe that having a job is the best way for a woman to be an independent person (Q4, mean of 0.89). As a matter of fact, the members of this group demonstrated the highest level of support for women's labor force participation.

Yet, Class 3 is different from Class 1 with respect to family-related items. The members of this class tended to be family-oriented, expressing concerns of the consequences of having a working mother to children, and supporting gender-specific roles at home. To be specific, respondents tended to agree that the children suffer when a mother works for pay (Q1, mean of 0.54). Similarly, they also responded that they believe that being a housewife is as fulfilling as working for pay (Q2, mean of 0.62), demonstrating support for gender essentialist beliefs. All things considered, members of Class 3 held gender egalitarian views in relation to equal treatment at work and women's engagement in economic activities, yet portrayed conservative attitudes on domestic issues related to children rearing and gender-specific roles at home.

Class 4. Egalitarian essentialists

Respondents of Class 4 could be characterized as supporting equal treatment of men and women at work, and simultaneously demonstrating strong support for gender-specific roles. For instance, this group was the least likely among all other groups to agree that men should have more right to a job than women when jobs are scarce (Q3, mean of 0.000003). Yet, the members tended to approve gender-specific roles at home by agreeing that being a housewife is just as fulfilling as working for pay (Q2, mean of 0.82).

However, members of Class 4 had a rather fluid view on women's employment. While it seemed like they were supportive of maternal employment by generally disagreeing that mother's employment can hurt the children (Q1, mean of 0.37), approximately less than half of the respondents agreed that having a job is the best way for a woman to be an independent person, showing the lowest support for women's economic activities among all other classes (Q4, mean of 0.47). In short, Class 4 could be characterized as rejecting male primacy at work, and showing high level of support for gender essentialist ideas. At the same time, they had a flexible opinion on women's employment.

Strictly speaking, the distinctions of gender role attitudes made in this study would not have been identified based on the typical unidemensional measurement of egalitarianism. Specifically, it would have been extremely difficult to pinpoint the position of Class 2, 3, and 4 along the line in which traditionalism stands on one end, and liberal egalitarianism stands on the other.

The national gender ideology

The national gender ideology was measured based on the proportion of people classified into each class in each country. Four figures were specified for each country, which represents the percentage of people constituting each of the four gender ideology classes. Due to the nature of these four figures, they sum up to 100 percent for each country. In the United States, for example, 20.74 percent of the country-specific sample was classified as Class 1, while 8.83 percent, 54.57 percent, and 15.86 percent were categorized into Class 2, Class 3 and Class 4, receptively. The four national gender ideology variables were coded accordingly. Table 4-4 shows the value of the national gender ideology by country. As demonstrated, the majority of the countries showed high representation of Class 3 in each country.

Table 4-4. National gender ideology by country

Country	Class 1	Class 2	Class 3	Class 4
Albania	24.05	17.34	50.39	8.21
Algeria	10.92	58.67	27.25	3.17
Australia	25.12	6.36	51.93	16.59
Austria	19.54	14.97	59.21	6.29
Belgium	24.06	9.15	56.06	10.74
Brazil	21.33	10.03	65.01	3.63
China	18.26	28.83	45.96	6.96
Bulgaria	32.87	10.40	52.80	3.93
Chile	39.70	12.30	44.10	3.90
Chinese Taipei	12.20	14.54	65.83	7.43
Colombia	24.27	10.85	58.07	6.81
Croatia	28.00	6.75	54.43	10.82
Czech Republic	25.86	12.52	56.12	5.49
Denmark	52.36	0.60	39.81	7.23
Estonia	13.47	11.64	67.98	6.92
Finland	26.90	3.44	39.68	29.98
France	31.45	7.26	54.36	6.93
FYROM	29.80	16.93	47.93	5.33
Georgia	10.62	29.02	57.29	3.07
Germany	38.87	10.00	46.86	4.27
Greece	9.27	19.53	65.80	5.40
Hungary	23.60	8.13	58.82	9.45
Iceland	30.45	1.73	27.72	40.10
Ireland	20.83	11.25	47.88	20.04
Italy	17.12	14.94	59.78	8.16
Japan	20.06	21.41	55.55	2.99
Jordan	2.75	74.17	21.08	2.00
Korea	18.83	23.42	55.92	1.83
Kosovo	14.68	19.80	62.34	3.19
Latvia	18.73	8.76	66.93	5.58
Lebanon	10.92	33.58	51.08	4.42
Lithuania	8.53	16.60	66.60	8.27
Luxembourg	22.61	7.89	63.23	6.27
Malta	5.00	26.60	56.73	11.67
Mexico	21.15	9.90	57.30	11.65
Moldova	17.73	17.34	59.51	5.42
Montenegro	26.52	13.13	52.70	7.65
Netherlands	38.86	7.18	40.48	13.48
New Zealand	24.02	6.78	48.87	20.33
Norway	44.04	1.47	45.50	8.99
Peru	26.61	11.90	53.72	7.77
Poland	16.44	13.85	61.67	8.04
Portugal	17.26	15.90	62.46	4.38
Qatar	2.83	58.21	33.77	5.19
Romania	25.67	12.77	57.89	3.68
Russian Federation	10.11	21.63	61.01	7.24

Singapore	17.75	24.14	51.42	6.69
Slovak Republic	27.63	13.59	53.88	4.90
Slovenia	33.68	8.83	42.67	8.83
Spain	35.33	10.15	48.64	5.88
Sweden	46.18	1.96	37.82	14.04
Switzerland	17.53	13.99	58.02	10.46
Thailand	15.58	24.33	57.58	2.50
Trinidad and Tobago	18.32	13.91	60.66	7.11
Tunisia	4.65	63.32	30.54	1.49
Turkey	8.00	48.08	39.31	4.61
United Kingdom	21.91	9.42	45.93	22.74
United States	20.74	8.83	54.57	15.86
Uruguay	23.70	14.30	54.60	7.40
Average	21.95	17.36	52.05	8.53

Research Question 3

Research question 3 inquired about the relationship between gender achievement gap and the structural and ideological aspect of gender inequality in each country regarding mathematics, science and reading, respectively. In this section, the empirical results of the multilevel analyses are illustrated by subject. Finally, the results of the supplementary analyses are also presented and discussed.

Gender achievement gap in mathematics and gender equality

Table 4-5 presents the result of the multilevel analysis for mathematics performance. As demonstrated by the coefficient estimates of the gender variable, there was a statistically significant gender gap in mathematics between boys and girls. In general, girls scored approximately five points behind their male counterparts controlling for SES and other variables.

The first model suggested that there is no statistically significant relationship between mathematics achievement and national gender equality. In other words, the average performance in mathematics did not systematically differ by the level of the nation's gender equality measured by GII and the percentage of people comprising each class of gender ideology. Yet, all

individual-level variables were statistically significant predictors of students' mathematics achievement. For instance, there was substantial SES influence on student's mathematics achievement as expected. To elaborate, a one-unit increase in the SES index was associated with roughly 31 points increase in achievement. Furthermore, students attending schools in the urban area and private schools received about 11, and 15 points higher than their counterparts attending non-urban schools and public schools, respectively.

The results of model 2 displayed that the association between national gender equality and academic achievement was not moderated by gender. However, GII, and GDP per capita were statistically significant predictors of the average national mathematics performance. Specifically, students from countries with lower levels of gender equality as measure by GII, and lower GDP per capita received lower scores on the mathematics assessment on average. Students from countries with higher proportion of familial egalitarians received higher mathematics scores controlling for other variables. Again, these relationships did not vary by gender.

Table 4-5. Multilevel regression estimates for mathematics achievement

	(1)		(2)	
Fixed effect	*B*	*SE*	*B*	*SE*
(Intercept)	465.26***	18.86	465.29***	3.52
Level 1: Student level (N =384,947)				
SES	30.94***	1.32	30.94***	1.32
Female	-4.84***	1.32	-5.04***	1.04
City	11.24***	2.56	11.23***	2.56
Private school	14.97***	4.30	14.98***	4.30
Level 2: Country level (N = 55)				
Gender equality				
GII	-205.00	262.36	-205.13***	63.63
Class of gender attitudes [a]				
Partial traditionalist	0.32	2.59	0.31	0.39
Familial egalitarian	1.12	2.63	1.12**	0.51
Egalitarian essentialist	0.24	3.19	0.21	0.42
GDP per capita (logged)	21.04	32.59	21.44***	7.16
GINI	-0.31	4.19	-0.28	0.84

Cross-level interactions				
Gender X				
Gender equality				
GII			1.76	16.81
Class of gender attitudes [b]				
Partial traditionalist			0.14	0.12
Familial egalitarian			-0.04	0.15
Egalitarian essentialist			0.29	0.19
GDP per capita (logged)			-3.56	1.74
GINI			-0.22	0.19
Random Effect	Variance Component	X^2	Variance Component	X^2
Country-level variance	635.1***	32092.2	634.68***	32151.7
Student-level variance	6941.2		6941.20	
Slope of female	67.40***	896.78	46.24***	619.54

Note: Robust standard errors reported.
The average of the ten analyses using plausible values is reported.
[a b] Reference group is Class 1, the liberal egalitarians.
* p < 0.10, ** p < 0.05, *** p < 0.01

Gender achievement gap in science and gender equality

Table 4-6 displays the results of the estimated effects of national gender equality and controls on science achievement. As shown, there were no statistically significant differences in science achievement between male and female students across countries.

According to the results of model 1, there were no statistically significant associations between science achievement and national gender equality. Nevertheless, at the individual-level, students with higher socioeconomic background, students attending urban schools, and those attending private schools received higher grades in science. family SES, school location and school type were successful predictors of students' science achievement. To elaborate, a unit increase in the SES index was associated with roughly 32 points increase in achievement. Moreover, students attending urban and private schools scored approximately 11, and 15 points higher than their counterparts attending non-urban schools and public schools, respectively.

Similar to the case of mathematics, GII and GDP per capita demonstrated a significant relationship to the average science achievement once interaction terms were introduced into the model. That is, students living in a structural arrangement that demonstrated greater level of gender inequality and lower GDP per capita generally scored behind their counterparts residing in a more gender-equal society and a more economically developed country by 189 points and 20 points, respectively. Results show that national gender ideology, GDP per capita and GINI were moderated by gender. To be specific, female students tended to perform slightly better than their male counterparts in countries with higher proportions of egalitarian essentialists (0.45 = 0.14 + 1*0.31). However, female students from a more economically developed country lagged behind male students by a marginal difference (-0.25 = 0.14 + 1*-0.39). Finally, female students residing in a more economically unequal society tended to underperform their male counterparts (-0.33 = 0.14 + 1*-0.47).

Table 4-6. Multilevel regression estimates for science achievement

	(1)		(2)	
Fixed effect	*B*	*SE*	*B*	*SE*
(Intercept)	469.48***	28.5	469.58***	3.57
Level 1: Student level (N =384,947)				
SES	31.65***	1.48	31.65***	1.48
Female	0.39	2.74	0.14	1.13
City	11.02***	2.8	11.01***	2.8
Private school	15.19***	4.21	15.2***	4.21
Level 2: Country level (N = 55)				
Gender equality				
GII	-181.95	425.9	-188.96***	55.69
Class of gender attitudes [a]				
Partial traditionalist	0.34	3.8	0.26	0.38
Familial egalitarian	0.86	3.81	0.87	0.53
Egalitarian essentialist	0.34	4	0.21	0.61
GDP per capita (logged)	18.84	47.99	20.36***	6.88
GINI	0.32	6.3	0.5	0.72
Cross-level interactions				
Gender X				

Gender equality				
GII			18.39	17.65
Class of gender attitudes [b]				
Partial traditionalist			0.22	0.15
Familial egalitarian			-0.02	0.16
Egalitarian essentialist			0.35*	0.19
GDP per capita (logged)			-3.99*	2.02
GINI			-0.47**	0.19
Random Effect	Variance Component	X^2	Variance Component	X^2
Country-level variance	650.35***	29538.6	643.77***	29553.2
Student-level variance	7222.7		72222.7	
Slope of female	109.13***	1418.3	59.13***	761.51

Note: Robust standard errors reported.

The average of the ten analyses using plausible value is reported.

[a b] Reference group is Class 1, the liberal egalitarians.

* p < 0.10, ** p < 0.05, *** p < 0.01

Gender achievement gap in reading and gender equality

Table 4-7 demonstrates the results of the multilevel regression for reading achievement. As shown in the table, there was a strong gender achievement gap in reading achievement. As a matter of fact, reading was the subject that demonstrated the greatest gender gap among the three subject areas. On average, female students scored approximately 30 points higher than their male counterparts holding all other variables constant.

As in the previous analyses, no statistically significant association between reading achievement and national gender equality was observed. However, family SES, school location and school type played a significant role in predicting reading achievement at the individual level. For example, student from higher SES family scored higher on reading tests (B = 32.67) and students attending urban and private schools, respectively received 14 and 17 points higher points on the reading assessment than their counterparts attending nonurban and public schools.

In relation to national gender equality, as shown in model 2, students exposed to social arrangements that demonstrated lower gender equality tended to receive lower reading score in general. In addition, students from economically more developed countries scored ahead of their

counterparts from economically underdeveloped countries (B = 21.87). Results demonstrated that the performance level of students was moderated by gender in relation to certain variables. For example, female students from countries with higher percentage of egalitarian essentialists scored ahead of their male counterpart by 30.26 points (= 29.76 + 0.5), which is higher than the average gender gap in reading achievement. On the other hand, female students from a more economically developed society tended to score approximately six points lower than the average gender gap. Still, female students outperformed male students by 23.79 points (=29.76 – 5.97) in these nations.

Table 4-7. Multilevel regression estimates for reading achievement

	(1)		(2)	
Fixed effect	*B*	*SE*	*B*	*SE*
(Intercept)	465.87***	27.5	465.82***	3.99
Level 1: Student level (N =384,947)				
SES	32.67***	1.45	32.67***	1.45
Female	30.28***	1.62	29.76***	1.52
City	13.98***	2.81	13.97***	2.81
Private school	17.1***	4.83	17.1***	4.83
Level 2: Country level (N = 55)				
Gender equality				
GII	-179.48	414.6	-180.66**	67.2
Class of gender attitudes [a]				
Partial traditionalist	-0.13	3.66	-0.11	0.42
Familial egalitarian	0.79	3.91	0.79	0.6
Egalitarian essentialist	0.33	5.3	0.38	0.6
GDP per capita (logged)	22.47	55.45	21.87***	8.05
GINI	0.94	6.82	0.89	0.82
Cross-level interactions				
Gender X				
Gender equality				
GII			-11.35	27.24
Class of gender attitudes [b]				
Partial traditionalist			0.25	0.19
Familial egalitarian			0.02	0.19
Egalitarian essentialist			0.5**	0.24
GDP per capita (logged)			-5.97**	2.7
GINI			-0.52	0.32
Random Effect	Variance Component	X^2	Variance Component	X^2

Country-level variance	768.22***	31352.5	767.16***	31191.8
Student-level variance	7806.4		7806.4	
Slope of female	153.45***	1769.3	107.73***	1284.2

Note: Robust standard errors reported.

The average of the ten analyses using plausible values is reported.

[a b] Reference group is Class 1, the liberal egalitarians.

* p < 0.10, ** p < 0.05, *** p < 0.01

Supplementary analyses

Due to the lack of robustness in the findings of the previous analyses, a series of supplementary analyses were conducted in order to further investigate the relationships between gender achievement gap, and national characteristics. Various models that included different combinations of the national gender ideology and other national-level variables were examined. Specifically, for each subject area, on top of the individual-level variables, national-level information was included as follows: (1) national gender ideology, (2) national gender ideology and GII, (3) national gender ideology and GDP per capita, and (4) national gender ideology and GINI. Subsequently, the same models were examined with interaction terms between gender and national-level variables. Table 4-8, Table 4-9, and Table 4-10 present the results of the analyses for mathematics, science, and reading, respectively.

As demonstrated in Table 4-8, national gender ideology had insignificant association with mathematics achievement. However, once interaction terms between national-level variables and gender were introduced to the model, significant relationship emerged for certain types of gender ideology. For instance, model 5 and model 8 indicated that students from countries with higher percentage of partial traditionalists tended to do worse on mathematics. Yet, the gender achievement gap was smaller in societies with higher percentage of partial traditionalists since female students scored slightly higher in these countries. For female students, the effect of a percentage point increase in partial traditionalists in a country was -0.94 points (= -1.18 + 1*0.24)

while the effect was -1.18 points for male students. Based on model 6, GII was the predictor that had the greatest effect on students' mathematics achievement. In general, students from countries with lower levels of structural gender equality demonstrated lower mathematics achievement. (*B* = -307.35). Generally, higher level of economic development of a country was associated with better mathematics performance. However, the economic development of a country was more strongly related to math achievement for male students than female students. Female students demonstrated a 30.77-point increase in mathematics achievement with each dollar increase in GDP per capita. Meanwhile, the score increased by 34.3 points for male students in such condition.

Similar to the case of mathematics, there was no statistically significant relationship between science achievement and national gender ideology (Table 4-9). Yet, the introduction of cross-level interaction terms revealed significant relationships between science achievement and a certain type of national gender ideology. To be specific, students from countries with higher percentage of partial traditionalists tended to do worse on science as displayed in model 5 and model 8. Yet, this relationship varied as a function of the students' gender. For instance, with each percentage point increase in the percent of partial traditionalists in a country, female students scored 0.68 points (= 0.31 + 1*0.37) higher, while it was 0.31 (= 0.31 + 0*0.37) points for male students. As displayed in model 6, the structural gender inequality was also the predictor with the greatest impact on students' science achievement (B = -307.35), however, this relationship was not moderated by gender. GDP per capita was also positively related to students' science achievement (model 7). Students from countries with higher levels of economic development scored higher than their counterparts from economically underdeveloped countries. Such relationship varied by the gender of students, where male students had science achievement scores that were 34.3 points higher for each dollar increase in national GDP per capita, whereas

female students had science scores only about 30.77 points higher for every dollar increase in GDP per capita.

The results for reading achievement demonstrated a similar pattern to those of mathematics and science. The national gender ideology did not have a statistically significant relationship with the average reading achievement. Once interaction terms were included in the analyses, significant associations emerged in relation to certain variables. As shown in model 5 and model 8, students from countries with higher percentage of partial traditionalists tended to do receive lower grades on reading. Yet, the magnitude of this relationship was moderated by gender. For example, female students tended to receive higher grades in a country with higher levels of partial traditionalists. Nevertheless, the overall effect of having higher levels of partial traditionalists in a country was negative for female students, where a one-percentage point increase in the percent of partial traditionalists in a country was associated with a 1.42 point-decrease (= -1.49 + 1 * 0.07) in reading. GII was also an important predictor of reading achievement (model 6). Higher levels of structural gender inequality were associated with lower reading achievement (B = -255.99). Finally, GDP per capita also had a positive relationship with students' reading achievement. On average, students from economically developed countries scored higher in reading. This association was moderated by gender as seen in model 7. Female students from countries that were more economically developed tended to receive lower grades on reading achievement. As a result, female students received 27.55 points higher for each dollar increase in GDP per capita while male students received 32.5 points higher for every dollar increase in GDP per capita.

Table 4-8. Multilevel supplementary analyses for mathematics achievement

	Without interactions							
	(1) Ideology		(2) Ideology + GII		(3) Ideology + GDP		(4) Ideology + GINI	
Fixed effect	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
(Intercept)	466.48***	45.17	464.63***	20.31	466.86***	28.87	465.14***	41.76
Level 1: Student level (N =384,947)								
SES	30.94***	1.32	30.94***	1.32	30.94***	1.32	30.94***	1.32
Female	-4.85	3.08	-4.84**	1.85	-4.84***	1.25	-4.84	3.18
City	11.24***	2.56	11.24***	2.56	11.24***	2.56	11.24***	2.56
Private school	14.97***	4.30	14.97***	4.30	14.97***	4.30	14.97***	4.30
Level 2: Country level (N = 55)								
Gender equality								
GII			-298.98	193.18				
Class of gender attitudes a								
Partial traditionalist	-1.16	6.53	0.33	2.75	-0.47	3.66	-0.84	5.82
Familial egalitarian	0.50	5.26	0.92	2.51	0.95	3.93	0.92	4.89
Egalitarian essentialist	1.18	5.35	0.85	2.66	0.01	4.41	1.19	5.40
GDP per capita (logged)					34.49	41.68		
GINI							-2.32	7.88
Cross-level interactions								
Gender X								
Gender equality								
GII								
Class of gender attitudes b								
Partial traditionalist								
Familial egalitarian								
Egalitarian essentialist								
GDP per capita (logged)								
GINI								
Random Effect	Variance Component	X^2	Variance Component	X^2	Variance Component	X^2	Variance Component	X^2
Country-level variance	1702.9***	88509.8	849.33***	40510.8	932.66***	54140.8	1580.4***	77663.3
Student-level variance	6941.2		6941.2		6941.1		6941.2	
Slope of female	67.48***	896.77	67.52***	896.79	67.39***	896.77	67.49***	896.78

Note: Robust standard errors reported.

The average of the ten analyses using plausible values is reported.

[a b] Reference group is Class 1, the liberal egalitarians.

* p < 0.10, ** p < 0.05, *** p < 0.01

Table 4-8. Multilevel supplementary analyses for mathematics achievement *(continued)*

	With interactions							
	(5) Ideology		(6) Ideology + GII		(7) Ideology + GDP		(8) Ideology +GINI	
Fixed effect	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
(Intercept)	466.56***	5.72	464.63***	4.09	466.85***	4.11	465.36***	5.73
Level 1: Student level (N =384,947)								
SES	30.94***	1.32	30.94***	1.32	30.94***	1.32	30.94***	1.32
Female	-4.91***	1.13	-4.84***	1.13	-4.94***	1.06	-4.99***	1.11
City	11.24***	2.56	11.24***	2.56	11.24***	2.56	11.24***	2.56
Private school	14.98***	4.30	14.98***	4.30	14.98***	4.30	14.98***	14.98
Level 2: Country level (N = 55)								
Gender equality								
GII			-307.35***	39.47				
Class of gender attitudes [a]								
Partial traditionalist	-1.43***	0.42	0.22	0.47	-0.47	0.41	-1.18***	0.42
Familial egalitarian	0.55	0.59	0.96*	0.50	0.95*	0.49	0.94	0.59
Egalitarian essentialist	0.94	0.76	0.71	0.62	0.02	0.49	0.93	0.71
GDP per capita (logged)					34.3***	6.48		
GINI							-2.10*	1.09
Cross-level interactions								
Gender X								
Gender equality								
GII			12.00	11.34				
Class of gender attitudes [b]								
Partial traditionalist	0.14	0.12	0.16	0.12	0.13	0.12	0.24**	0.12
Familial egalitarian	-0.04	0.15	-0.06	0.15	-0.08	0.15	-0.01	0.15
Egalitarian essentialist	0.29	0.19	0.20	0.21	0.29	0.20	0.19	0.20
GDP per capita (logged)					-3.53***	1.29		
GINI							-0.16	0.15
Random Effect	Variance Component	X^2	Variance Component	X^2	Variance Component	X^2	Variance Component	X^2
Country-level variance	1687.2***	89039.3	843.44***	41022.7	932.44***	54105.5	1557.65***	78270.9
Student-level variance	6941.2		6941.2		6941.2		6941.2	
Slope of female	55.64***	713.79	54.32***	696.97	47.59***	641.16	54.87***	704.35

Note: Robust standard errors reported.

The average of the ten analyses using plausible values is reported.

[a b] Reference group is Class 1, the liberal egalitarians.

* p < 0.10, ** p < 0.05, *** p < 0.01

Table 4-9. Multilevel supplementary analyses for science achievement

	Without interactions							
	(1) Ideology		(2) Ideology + GII		(3) Ideology + GDP		(4) Ideology + GINI	
Fixed effect	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
(Intercept)	470.18***	50.40	468.70***	33.74	470.52***	36.58	469.22***	65.50
Level 1: Student level (N =384,947)								
SES	31.65***	1.48	31.65***	1.48	31.65***	1.48	31.65***	1.48
Female	0.39	6.51	0.40	4.74	0.39	3.19	0.40	8.96
City	11.02***	2.80	11.02***	2.80	11.02***	2.80	11.02***	2.80
Private school	15.19***	4.21	15.19***	4.21	15.19***	4.21	15.19***	4.20
Level 2: Country level (N = 55)								
Gender equality								
GII			-243.76	306.63				
Class of gender attitudes [a]							0.72	6.4
Partial traditionalist	-0.68	6.49	0.39	4.10	-0.26	4.74	-0.42	7.78
Familial egalitarian	0.42	5.22	0.76	3.78	0.80	4.86	1.28	6.18
Egalitarian essentialist	1.25	4.76	0.93	3.57	0.21	4.79		
GDP per capita (logged)					29.39	46.53		
GINI							-1.66	10.56
Cross-level interactions								
Gender X								
Gender equality								
GII								
Class of gender attitudes [b]								
Partial traditionalist								
Familial egalitarian								
Egalitarian essentialist								
GDP per capita (logged)								
GINI								
Random Effect	Variance Component	X^2	Variance Component	X^2	Variance Component	X^2	Variance Component	X^2
Country-level variance	1560.8***	72897.2	884.17***	38558.4	855.11***	43602.0	1540.9***	69090.9
Student-level variance	7222.7		7222.7		7222.7		7222.7	
Slope of female	109.40***	1418.34	109.35***	1418.34	109.1***	1418.34	109.44***	1418.34

Note: Robust standard errors reported.

The average of the ten analyses using plausible values is reported.

[a b] Reference group is Class 1, the liberal egalitarians.

* p < 0.10, ** p < 0.05, *** p < 0.01

Table 4-9. Multilevel supplementary analyses for science achievement *(continued)*

	With interactions							
	(5) Ideology		(6) Ideology + GII		(7) Ideology + GDP		(8) Ideology +GINI	
Fixed effect	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
(Intercept)	470.31 ***	5.47	468.63 ***	4.16	470.58 ***	4.03	469.63 ***	5.55
Level 1: Student level (N =384,947)								
SES	31.65 ***	1.48	31.65 ***	1.48	31.65 ***	1.48	31.65 ***	1.48
Female	0.31	1.30	0.47	1.27	0.27	1.17	0.17	1.29
City	11.02 ***	2.80	11.02 ***	2.80	11.01 ***	2.8	11.02 ***	2.8
Private school	15.20 ***	4.21	15.20 ***	4.21	15.2 ***	4.21	15.19 ***	4.21
Level 2: Country level (N = 55)								
Gender equality								
GII			-267.04 ***	4.16				
Class of gender attitudes [a]								
Partial traditionalist	-1.27 ***	0.40	0.16	0.45	-0.38	0.4	-1.13 **	0.42
Familial egalitarian	0.47	0.56	0.83	0.51	0.84 *	0.49	0.69	0.58
Egalitarian essentialist	0.89	0.87	0.70	0.81	0.04	0.61	0.89	0.88
GDP per capita (logged)					31.74 ***	6.09		
GINI							-1.18	1.03
Cross-level interactions								
Gender X								
Gender equality								
GII			24.36 *	13.12				
Class of gender attitudes [b]								
Partial traditionalist	0.37 ***	0.13	0.24	0.16	0.24	0.15	0.41 ***	0.13
Familial egalitarian	-0.03	0.16	-0.07	0.17	-0.09	0.15	0.02	0.16
Egalitarian essentialist	0.23	0.23	0.25	0.24	0.36 *	0.2	0.23	0.22
GDP per capita (logged)					-4.83 ***	1.47		
GINI							-0.28 *	0.16
Random Effect	Variance Component	X^2	Variance Component	X^2	Variance Component	X^2	Variance Component	X^2
Country-level variance	1491.85 ***	71908.1	854.87 ***	38778.78	845.34 ***	43454.8	1451.08 ***	67821.89
Student-level variance	7222.7		7222.7		7222.7		7222.7	
Slope of female	79.43 ***	972.31	74.12 ***	918.52	64.51 ***	854.4	77.09 ***	945.8

Note: Robust standard errors reported.

The average of the ten analyses using plausible values is reported.

[a b] Reference group is Class 1, the liberal egalitarians.

* p < 0.10, ** p < 0.05, *** p < 0.01

Table 4-10. Multilevel supplementary analyses for reading achievement

	Without interactions							
	(1) Ideology		(2) Ideology + GII		(3) Ideology + GDP		(4) Ideology + GINI	
Fixed effect	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
(Intercept)	466.19***	23.75	464.59***	16.16	466.57***	14.44	465.55***	66.23
Level 1: Student level (N =384,947)								
SES	32.67***	1.45	32.67***	1.45	32.67***	1.45	32.67***	1.45
Female	30.28***	2.25	30.29***	2.37	30.28***	2.10	30.29***	4.91
City	13.98***	2.81	13.98***	2.81	13.98***	2.81	13.98***	2.81
Private school	17.10***	4.83	17.10***	4.83	17.10***	4.83	17.10***	4.82
Level 2: Country level (N = 55)								
Gender equality								
GII			-254.89	152.00				
Class of gender attitudes [a]								
Partial traditionalist	-1.43	3.50	-0.07	2.23	-0.71	1.77	-1.26	8.65
Familial egalitarian	0.45	3.00	0.79	2.00	0.86	1.93	0.65	8.72
Egalitarian essentialist	1.28	3.96	1.08	2.80	0.10	2.38	1.31	10.72
GDP per capita (logged)					33.70	22.35		
GINI							-1.10	17.81
Cross-level interactions								
Gender X								
Gender equality								
GII								
Class of gender attitudes [b]								
Partial traditionalist								
Familial egalitarian								
Egalitarian essentialist								
GDP per capita (logged)								
GINI								
Random Effect	Variance Component	X^2	Variance Component	X^2	Variance Component	X^2	Variance Component	X^2
Country-level variance	1622.0***	67644.6	1035.93***	41248.4	940.75***	42027	1610.14***	65464.8
Student-level variance	7806.4		7806.4		7806.41		7806.41	
Slope of female	153.55***	1769.32	153.57***	1769.35	153.46***	1769.31	153.56***	1769.33

Note: Robust standard errors reported.

The average of the ten analyses using plausible values is reported.

[a b] Reference group is Class 1, the liberal egalitarians.

* p < 0.10, ** p < 0.05, *** p < 0.01

Table 4-10. Multilevel supplementary analyses for reading achievement *(continued)*

	With interactions							
	(5) Ideology		(6) Ideology + GII		(7) Ideology + GDP		(8) Ideology +GINI	
Fixed effect	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
(Intercept)	466.26***	5.84	464.65***	4.73	466.54***	4.38	465.83***	5.89
Level 1: Student level (N =384,947)								
SES	32.67***	1.45	32.67***	1.45	32.67***	1.45	32.67***	1.45
Female	30.16***	1.70	30.18***	1.72	30.12***	1.58	29.88***	1.64
City	13.98***	2.81	13.98***	2.81	13.97***	2.81	13.98***	2.81
Private school	17.10***	4.83	17.10***	4.83	17.10***	4.83	17.10***	4.83
Level 2: Country level (N = 55)								
Gender equality								
GII			-255.99***	39.87				
Class of gender attitudes [a]								
Partial traditionalist	-1.58***	0.49	-0.21	0.51	-0.67	0.47	-1.49***	0.54
Familial egalitarian	0.46	0.62	0.80	0.58	0.84	0.52	0.60	0.67
Egalitarian essentialist	1.09	0.93	0.90	0.89	0.22	0.59	1.09	0.93
GDP per capita (logged)					32.50***	6.23		
GINI							-0.75	1.03
Cross-level interactions								
Gender X								
Gender equality								
GII			2.23	17.73				
Class of gender attitudes [b]								
Partial traditionalist	0.29	0.18	0.28	0.2	0.15	0.2	0.07*	0.21
Familial egalitarian	-0.03	0.22	-0.03	0.22	-0.08	0.21	0.35	0.18
Egalitarian essentialist	0.35	0.29	0.35	0.29	0.48*	0.27	0.35	0.27
GDP per capita (logged)					-4.95**	2.08		
GINI							-0.52**	0.25
Random Effect	Variance Component	X^2	Variance Component	X^2	Variance Component	X^2	Variance Component	X^2
Country-level variance	1616.9***	68192.0	1031.54***	41814.1	938.94***	41709.5	1600.3***	66328.8
Student-level variance	7806.41		7806.41		7806.41		7806.41	
Slope of female	134.63***	1492.16	134.48***	1491.03	119.06***	1409.81	126.3***	1382.92

Note: Robust standard errors reported.

The average of the ten analyses using plausible values is reported.

[a b] Reference group is Class 1, the liberal egalitarians.

* p < 0.10, ** p < 0.05, *** p < 0.01

Chapter 5

Discussion

The purpose of the final chapter is to provide an overview of the main findings of the study, and further discuss the meaning of these findings. In addition, I attempted to explain the limitations of the current study as well as suggest direction for future research that could be further developed in relation to this topic.

Summary of findings

The current study has investigated three research questions: (1) How do the gender gaps in mathematics, science, and reading achievement vary in size and direction by country? (2) How many classes do the countries cluster into in terms of gender-related attitudes and how are the classes different from each other? (3) How do gender achievement gaps in mathematics, science and reading, respectively, associate with the structural and ideological aspect of gender inequality in each country?

For the first research question, the analysis revealed substantial cross-national variations in the gender gap across all subject areas as expected. However, the patterns of the variations were different for each subject. For example, for mathematics achievement, a greater number of countries demonstrated statistically significant female disadvantage. Furthermore, the magnitude of the gap was greater for the countries demonstrating female disadvantage compared to those showing female advantage. Meanwhile, contrary to my hypothesis that expected to see a male advantage in science achievement, there was a balance between the number of countries that demonstrated a female advantage and disadvantage in science achievement. In terms of the magnitude, except for Jordan, which demonstrated the largest gender gap (female advantage), there also seemed to be symmetry between the gender gaps that favor girls and gaps that favor boys. Reading achievement also demonstrated a substantial gender gap across countries. Yet, the gender gap in reading achievement

was unique in the sense that all coefficients were positive. In other words, female students scored ahead of their male counterparts on the reading examination across every single country included in the analyses. In terms of the magnitude, the gender reading gap varied from less than ten points in Peru, to over seventy points in Jordan.

These findings aligned with my expectations to a certain degree. As expected, the empirical results demonstrated substantial cross-national variations across subject areas. Nevertheless, as subjects classified as STEM field, I expected to see similar patterns between mathematics and science. However, the results displayed higher levels of science performance by female students compared to male students on average, while females fell behind their male counterparts in mathematics, consistent with what have been reported in a variety of previous research. It is unclear, through this descriptive analysis, what is exactly happening however. As subjects of the STEM field, instead of treating mathematics and science in a similar manner, more attention should be paid to the differences of these two subjects.

The result of the latent class analysis, which aimed to address the second research question, identified four classes of gender ideology: (1) liberal egalitarian, (2) partial traditionalist, (3) familial egalitarian, and (4) egalitarian essentialist. Most of the classes that surfaced from the analyses did not fall along the traditional-egalitarian line. The liberal egalitarian were highly similar to what we typically refer to as those who have egalitarian gender attitudes. They demonstrated strong support for maternal as well as women employment, and low support for gender essentialist beliefs and male primacy at work. Partial traditionalist held strong gender essentialist beliefs, disapproved mother's employment and supported male primacy at work. Interestingly, however, this group displayed a relatively high support for women's employment, which is rather different from those who are typically considered as traditionalist. Familial egalitarian, on the other hand, is similar to liberal egalitarian in that they support women's labor force participation and deny male primacy at work. Yet, they held conservative positions in relation to household matters such as child

rearing, and gender roles at home. Finally, egalitarian essentialist had strong gender essentialist beliefs while rejecting unequal treatment of men and women at work. Fluid opinion on women's employment was also an important characteristic of this class. Based on this classification of gender ideology, the national gender ideology was established. Each country was assigned four figures, which represented the percentage of individuals that constituted each class that sums up to 100 percent. The results indicated that the majority of countries included in the analysis demonstrated high levels of Class 3 (familial egalitarian) representation.

As expected, the results of the latent class analysis yield gender attitude typologies that could not be explained by the typical traditional-egalitarian dichotomy. Besides the class labeled as the liberal egalitarian which resembles the characteristics of a typical individual with egalitarian gender role attitudes, partial traditionalist, familial egalitarian, and egalitarian essentialist were demonstrating different combinations of the various dimensions of gender attitudes.

Finally, multilevel analyses were conducted in order to examine the relationships between gender achievement gap in mathematics, science and reading, and national gender equality, respectively. In terms of the gender gap, there was a significant female disadvantage in mathematics; no gender difference in science; and a significant female advantage in reading after controlling for all other variables. The relationship between achievement and national gender equality was not robust. Statistically significant association appeared once interaction terms with gender were included in the models. The structural gender equality measure by GII was the predictor with the greatest effect on students' achievement across all subjects. On the other hand, the national gender ideology, the other important measure of national gender equality, generally did not have a statistically significant influence on students' achievement except for a few exceptions. GDP per capita was also a significant predictor of students' academic achievement in mathematics, science, and reading.

Limitations of the study

However, as any other studies, the present study is not without limitations. First, the two datasets utilized to measure national gender ideology, namely EVS and WVS, were not collected in the same time period. To be specific, the fourth wave of EVS was collected between 2008 and 2009 while WVS was done so between 2010 and 2014. There are two issues that could emerge from this time discrepancy in data collection. The first issue is that the measurement of national gender ideology of each country may not be an accurate one since the analysis was conducted using a pooled data of EVS and WVS, which, as previously mentioned, was collected in different time periods. Had this study used data that were collected in the same time period, different typologies of gender ideology might have emerged from the latent class analysis.

In a similar vein, the second issue surfaces in relation to the main analytic dataset of this study, PISA 2015. As demonstrated by a body of literature that investigated the change of gender-related attitudes (Brewster & Padavic, 2000; Cherlin & Walter, 1981; Cotter et al., 2011; Ferree, 1974; Mason & Lu, 1988; McBroom, 1986; Seguino, 2007; Thornton & Freedman, 1979), gender ideology is a fluid concept that could change overtime. In other words, between 2008, the earliest year when information on gender attitudes was collected, and 2015, when PISA was collected, there could be myriads of factors that could trigger change in gender attitudes in a country. Thus, analyses using the three datasets may not yield accurate estimates across countries in relation to gender-related attitudes.

Secondly, only four gender attitude questions were utilized to measure the classes of gender ideology in this study. However, in each survey of EVS and WVS, there are eight respective questions that inquired participants' gender attitudes. For instance, the fourth wave of EVS asked the following questions: (1) When jobs are scarce, men have more right to a job than women, (2) a working mother can establish just as warm and secure a relationship with her children as a mother

who does not work, (3) a pre-school child is likely to suffer if his or her mother works, (4) a job is alright but what most women really want is a home and children, (5) being a housewife is just as fulfilling as working for pay, (6) having a job is the best way for a woman to be an independent person, (6) both husband and wife should contribute to household income, (7) in general, fathers are as well suited to looking after their children as mothers, and (8) men should take as much responsibility as women for the home and children. In the sixth wave of WVS, the gender-related questions were as follows: (1) When jobs are scarce, men should have more right to a job than women, (2) if a woman earns more money than her husband, it's almost certain to cause problems, (3) having a job is the best way for a woman to be an independent person, (4) when a mother works for pay, the children suffer, (5) on the whole, men make better political leaders than women do, (6) a university education is more important for a boy than for a girl, (7) on the whole, men make better business executives than women do, and (8) being a housewife is just as fulfilling as working for pay.

As demonstrated, both surveys contain questions that are rich in quality as well as in quantity in revealing gender attitudes. Yet, as this study pooled the datasets to measure the same construct, the questions had to be limited to the four items that were present in both surveys. Had all of these questions were used in the analyses, the results of the analyses, especially the latent class analysis, might have yield different patterns of responses. Further, the use of limited number of gender attitude questions impacted the quality of the strategies used to mitigate concerns about using EVS and WVS survey questions. To be specific, following the footsteps of Knight and Brinton (2017), this study has incorporated questions that were asked from the traditional perspective as well as egalitarian perspective. Nevertheless, due to the limitation of the number of available gender attitude items in both surveys, only one egalitarian-slanted questions was included in the analysis.

Despite such limitations, the decision to use EVS and WVS to measure national gender ideology was made in order to secure the maximum number of countries in the analyses. As the magnitude of the sample size affects the power of the test, obtaining as many countries as possible was an important factor to consider in conducting a multilevel analysis. Moreover, as demonstrated in chapter 2, a body of empirical studies indicated that gender attitudes have not undergone significant change since the 1980s (Brewster & Padavic, 2000; Cotter et al., 2011). Thus, I assumed that there were minimal alteration in the nation's gender ideology between 2008 and 2015.

Directions for future research

Contrary to my expectations, the effect of national gender ideology appeared to be not significantly associated with students' academic performance across countries. One possible explanation of such result might be that the four classes of national gender ideology measured in this study may not be distinctive enough to be transferred into different levels of academic achievement. As previously mentioned, the aim of the use of latent class analysis was to investigate the subtle differences of gender attitudes across countries. The four classes were different in quality as demonstrated in chapter 4. However, the differences might be too subtle to be captured in the multilevel analysis, thus yield insignificant estimates. Yet, this does not necessarily mean that gender ideology is irrelevant to one's academic performance altogether. Perhaps it could be useful to investigate this topic at the family-level, and explore how parents' gender attitudes affect children's educational outcome. There is a large body of literature that studied the intergenerational transmission of gender attitudes between parents and their children (Booth, & Amato, 1994; Cunningham, 2001; Kulik, 2004; Moen, Erickson, & Dempster-McClain, 1997), which usually find a strong and positive relationship between parents' and children's gender attitude. A study that had looked into the relationship between the family's nontraditionalism and the offspring's academic

achievement conducted by Booth and Amato (1994) found no significant association. Still, this is a topic that could be further investigated in future studies.

Furthermore, this line of research should be expanded to longer-term outcomes. To be specific, studies should investigate how gender equality at the country-level as well as individual-level affect one's major choice in higher education, and in turn their occupation selection. A number of studies explored the gender gap in college major selection (Riegle-Crumb, King, Grodsky, & Muller, 2012; Zafar, 2013) and gender difference in the choice of occupation (Correll, 2001; Jacobs & Lim, 1992) and in earnings (Petersen & Morgan, 1995; Weichselbaumer & Winter-Ebmer, 2005). However, most of the studies focus on micro-level variables that affect one's choice. Thus, future research needs to examine how macro-level gender inequality shape the decisions of individuals in relation to their college major and occupation selection.

Finally, the concept of gender equality should go beyond the male-female, and heterosexual- non-heterosexual dichotomy. The majority of studies based their study on the assumption that gender is a binary concept, and people are in heterosexual relationships. Especially, the gender attitude survey items provide little information of what is actually going on among those who are outside these dichotomies. I admit that there is limited data to utilize, especially in terms of large-scale data, in order to expand the notions. Yet, efforts should be made to include this population that are systematically marginalized from surveys, as diversity in sexual orientation is being increasingly emphasized in relation to social inclusion across the globe (Reynolds, 2013).

Reference

Adams, R. B., & Funk, P. (2012). Beyond the glass ceiling: Does gender matter? *Management Science, 58*(2), 219-235.

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

Banerjee, R., & Lintern, V. (2000). Boys will be boys: The effect of social evaluation concerns on gender-typing. *Social Development, 9*(3), 397-408.

Baron-Cohen, S. (2003). The essential difference: Men, women and the extreme male brain. London: Allen Lane, Penguin Books.

Beller, M., & Gafni, N. (1996). The 1991 international assessment of educational progress in mathematics and sciences: The gender differences perspective. *Journal of Educational Psychology, 88*(2), 365-377.

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

Benbow, C. P., & Stanley, J. C. (1983). Sex differences in mathematical reasoning ability: More facts. *Science, 222*(4627), 1029-1031.

Bianchi, S. M., Milkie, M. A., Sayer, L. C., & Robinson, J. P. (2000). Is anyone doing the housework? trends in the gender division of household labor. *Social Forces, 79*(1), 191-228.

Booth, A., & Amato, P. R. (1994). Parental gender role nontraditionalism and offspring outcomes. *Journal of Marriage and the Family, 56*(4), 865.

Braun, M. (2008). Using egalitarian items to measure men's and women's family roles. *Sex Roles, 59*(9-10), 644-656.

Brewster, K., & Padavic, I. (2000). Change in gender-ideology, 1977-1996: The contributions of intracohort change and population turnover. *Journal of Marriage and Family, 62*(2), 477-487.

Buchmann, C., & DiPrete, T. A. (2006). The growing female advantage in college completion: The role of family background and academic achievement. *American Sociological Review, 71*(4), 515-541.

Buchmann, C., DiPrete, T. A., & McDaniel, A. (2008). Gender inequalities in education. *Annual Review of Sociology, 34*(1), 319-337.

Charles, M. (2011). A world of difference: International trends in women's economic status. *Annual Review of Sociology, 37*(1), 355-371.

Charles, M., & Grusky, D. B. (2004). *Occupational ghettos: The worldwide segregation of women and men*. Stanford, CA: Stanford University Press.

Cherlin, A., & Walters, P. B. (1981). Trends in united states men's and women's sex-role attitudes: 1972 to 1978. *American Sociological Review, 46*(4), 453-460.

Chiu, M. M. (2007). Families, economies, cultures, and science achievement in 41 countries: Country-, school-, and student-level analyses. *Journal of Family Psychology, 21*(3), 510-519.

Chiu, M. M., & McBride-Chang, C. (2006). Gender, context, and reading: A comparison of students in 43 countries. *Scientific Studies of Reading, 10*(4), 331-362.

Cloer, T., & Dalton, S. (2001). Gender and grade differences in reading achievement and self-concept as readers. *Journal of Reading Education, 26*(2), 31-36.

Contini, D., Tommaso, M. L. D., Mendolia, S. (2017). The gender gap in mathematics achievement: Evidence from Italian data. *Economics of Education Review, 58*, 32-42.

Correll, S. (2001). Gender and the career choice process: The role of biased self-assessments. *American Journal of Sociology, 106*(6), 1691-1730.

Cotter, D., Hermsen, J. M., & Vanneman, R. (2011). The end of the gender revolution? gender role attitudes from 1977 to 2008. *The American Journal of Sociology, 117*(1), 259-289.

Cunningham, M. (2001). The influence of parental attitudes and behaviors on children's attitudes toward gender and household labor in early adulthood. *Journal of Marriage and Family, 63*(1), 111-122.

Curran, F. C., & Kellogg, A. T. (2016). Understanding science achievement gaps by Race/Ethnicity and gender in kindergarten and first grade. *Educational Researcher, 45*(5), 273-282.

Cvencek, D., Meltzoff, A. N., & Greenwald, A. G. (2011). Math–gender stereotypes in elementary school children. *Child Development, 82*(3), 766-779.

Davies, B. (2003). *Frogs and snails and feminist tales: Preschool children and gender*. Cresskill, NJ: Hampton Press.

Diprete, T. A., & Buchmann, C. (2006). Gender-specific trends in the value of education and the emerging gender gap in college completion. *Demography, 43*(1), 1-24.

Dorius, S. F., & Firebaugh, G. (2010). Trends in global gender inequality. *Social Forces, 88*(5), 1941-1968.

Dryler, H. (1998). Parental role models, gender and educational choice. *The British Journal of Sociology, 49*(3), 375-398.

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

England, P. (2010). The Gender Revolution. *Gender & Society, 24*(2), 149-166.

Ferree, M. M. (1974). A woman for president? changing responses: 1958-1972. *The Public Opinion Quarterly, 38*(3), 390-399.

Fraser, N. (1994). After the family wage: Gender equity and the welfare state. *Political Theory, 22*(4), 591-618.

Freeman, C. E. (2004). *Trends in educational equity of girls & women, 2004*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

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

Gandhi Kingdon, G. (2002). The gender gap in educational attainment in India: How much can be explained? *The Journal of Development Studies, 39*(2), 25-53.

Geiser, C. (2010). *Data analysis with mplus*. New York, NY: The Guilford Press.

Gornick, J. C., & Meyers, M. K. (2008). Creating gender egalitarian societies: An agenda for reform. *Politics & Society, 36*(3), 313-349.

Guillet, E., Sarrazin, P., & Fontayne, P. (2000). "If it contradicts my gender role, I'll stop": Introducing survival analysis to study the effects of gender typing on the time of withdrawal from sport practice: A 3-year study. *European Review of Applied Psychology*, *50*(4), 417–421.

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

Gurian, M. (2002) *Boys and girls learn differently! A guide for teachers and parents.* San Francisco: Jossey-Bass.

Halpern, D. F. (2012). *Sex differences in cognitive abilities*. New York, NY: Psychology Press.

Head, J. (1999). *Understanding the boys: Issues of behaviour and achievement*. New York: Falmer Press.

Hochschild, A. R., & Machung, A. (1989). *The second shift*. New York, NY: Avon Books.

Hofstede, G. H. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations.* Thousand Oaks, CA: Sage Publications.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications*. New York, NY: Routledge.

Husain, M., & Millimet, D. L. (2009). The mythical "boy crisis"? *Economics of Education Review, 28*(1), 38-48.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science, 321*(5888), 494-495.

Hyde, J. S., Mertz, J. E., & Schekman, R. (2009). Gender, culture, and mathematics performance. *Proceedings of the National Academy of Sciences of the United States of America, 106*(22), 8801-8807.

Inglehart, R., & Norris, P. (2003). *Rising tide: Gender equality and cultural change around the world*. New York, NY: Cambridge University Press.

Jacobs, J. A., & Lim, S. T. (1992). Trends in occupational and industrial sex segregation in 56 countries, 1960-1980. *Work and Occupations, 19*(4), 450-486.

Jacobs, J. A. (1996). Gender inequality and higher education. *Annual Review of Sociology, 22*(1), 153-185.

Jencks, C., & Phillips, M. (1999). Aptitude or achievement: Why do test scores predict educational attainment and earnings? In S. E. Mayer, & P. E. Peterson (Eds.), *Earning and learning: How schools matter* (pp. 15–47). Washington, D.C: Russell Sage Foundation.

Jones, L. R., Mullis, I. V., Raizen, S. A., Weiss, I. R., & Weston, E. A. (1992). *The 1990 science report card: NAEP's assessment of fourth, eighth, and twelfth graders*. Washington, DC: Educational Testing Services.

Johnston, R., & Watson, J. (2005). *A seven year study of the effects of synthetic phonics teaching on reading and spelling attainment.* Edinburgh: Scottish Executive Education Department

Keeves, J. (1973). Differences between the sexes in mathematics and science courses. *International Review of Education / Internationale Zeitschrift für Erziehungswissenschaft / Revue Internationale de l'Education, 19*(1), 47-63.

Klecker, B. M. (2006). The gender gap in NAEP fourth-, eighth-, and twelfth-grade reading scores across years. *Reading Improvement, 43*(1), 50-55.

Knight, C. R., & Brinton, M. C. (2017). One egalitarianism or several? two decades of gender-role attitude change in europe. *The American Journal of Sociology, 122*(5), 1485-1532.

Kulik, L. (2004). Predicting gender-role attitudes among mothers and their adolescent daughters in israel. *Affilia, 19*(4), 437-449.

Lai, F. (2010). Are boys left behind? the evolution of the gender achievement gap in beijing's middle schools. *Economics of Education Review, 29*(3), 383-399.

Leaper, C., & Friedman, C. K. (2007). The socialization of gender. In J. E. Grusec & P. D. Hastings (Eds.), *Handbook of Socialization: Theory and Research* (pp. 561-587). New York, NY: Guilford Publications.

Lee, J., Grigg, W., and Dion, G. (2007). *The Nation's Report Card: Mathematics 2007* (NCES 2007–494). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Legewie, J., & DiPrete, T. A. (2012). School context and the gender gap in educational achievement. *American Sociological Review, 77*(3), 463-485.

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

Littleton, C. A. (1987). Reconstructing sexual equality. *California Law Review, 75*(4), 1279-1337.

Lummis, M., & Stevenson, H. W. (1990). Gender differences in beliefs and achievement: A cross-cultural study. *Developmental Psychology, 26*(2), 254-263.

Lynch, J. (2002). Parents' self-efficacy beliefs, parents' gender, children's reader self-perceptions, reading achievement and gender. *Journal of Research in Reading, 25*(1), 54-67.

Marks, G. N. (2008). Accounting for the gender gaps in student performance in reading and mathematics: Evidence from 31 countries. *Oxford Review of Education, 34*(1), 89-109.

Martin, C. L., Ruble, D. N., & Szkrybalo, J. (2002). Cognitive theories of early gender development. *Psychological Bulletin, 128*(6), 903-933.

Mason, K. O., & Lu, Y. (1988). Attitudes toward women's familial roles: Changes in the united states, 1977-1985. *Gender and Society, 2*(1), 39-57.

McBroom, W. H. (1986). Changes in role orientations of women: A study of sex role traditionalism over a five-year period. *Journal of Family Issues, 7*(2), 149-159.

McCutcheon, A. L. (1987).*Quantitative Applications in the Social Sciences: Latent class analysis* Thousand Oaks, CA: SAGE Publications Ltd

McDaniel, A. E. (2008). Measuring gender egalitarianism: The attitudinal difference between men and women. *International Journal of Sociology, 38*(1), 58-80.

Moen, P., Erickson, M. A., & Dempster-McClain, D. (1997). Their mother's daughters? the intergenerational transmission of gender attitudes in a world of changing roles. *Journal of Marriage and the Family, 59*(2), 281-293.

Muller, P. A., Stage, F. K., & Kinzie, J. (2001). Science achievement growth trajectories: Understanding factors related to gender and racial-ethnic differences in precollege science achievement. *American Educational Research Journal, 38*(4), 981-1012.

Mullis, I. V., Martin, M. O., Kennedy, A. M., & Foy, P. (2007). *PIRLS 2006 international report: IEA's progress in international reading literacy study in primary schools in 40 countries.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Murnane, R. J., Willett, J. B., & Levy, F. (1995). *The growing importance of cognitive skills in wage determination* (Working Paper No. 5076). Cambridge, MA: National Bureau of Economic Research.

Murnane, R. J., Willett, J. B., Duhaldeborde, Y., & Tyler, J. H. (2000). How important are the cognitive skills of teenagers in predicting subsequent earnings? *Journal of Policy Analysis and Management, 19*(4), 547-568.

National Center for Education Statistics [NCES]. (2012). *The nation's report card: Science 2011*. Washington, DC: Institute of Education Sciences, U.S. Department of Education.

National Center for Education Statistics [NCES]. (2016). *The nation's report card: 2015 science at grades 4, 8, and 12*. Retrieved from the Nation's Report Card website: https://www.nationsreportcard.gov/science_2015/#?grade=4

Nosek, B. A., & Smyth, F. L. (2011). Implicit social cognitions predict sex differences in math engagement and achievement. *American Educational Research Journal, 48*(5), 1125-1156.

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., … Samhällsvetenskapliga fakulteten. (2009). 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*(26), 10593-10597.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A monte carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(4), 535-569.

OECD (2013). *What Makes Urban Schools Different?* (PISA in Focus No. 28). Paris, France: OECD Publishing.

OECD (2016). *PISA 2015 Technical Report*. Paris, France: OECD Publishing.

Orloff, A. S. (2009). Should Feminists Aim for Gender Symmetry?: Why the Dual-Earner/Dual-Carer Model May Not Be Every Feminist's Utopia. In J. Gornick, & M. Meyers (Eds.), *Institutions for Gender Egalitarianism: Creating the Conditions for Egalitarian Dual Earner / Dual Caregiver Families* (pp. 129-160). Madison: University of Wisconsin Press.

Petersen, T., & Morgan, L. A. (1995). Separate and unequal: Occupation-establishment sex segregation and the gender wage gap. *American Journal of Sociology, 101*(2), 329-365.

Quinn, D. M., & Cooc, N. (2015). Science achievement gaps by gender and Race/Ethnicity in elementary and middle school: Trends and predictors. *Educational Researcher, 44*(6), 336-346.

Reynolds, A. (2013). Representation and rights: The impact of LGBT legislators in comparative perspective. *The American Political Science Review, 107*(2), 259-274.

Riegle-Crumb, C., King, B., Grodsky, E., & Muller, C. (2012). The more things change, the more they stay the same? prior achievement fails to explain gender inequality in entry into STEM college majors over time. *American Educational Research Journal, 49*(6), 1048-1073.

Rogers, L. (2001). *Sexing the brain*. New York, NY: Columbia University Press.

Rose, H. (2006). Do gains in test scores explain labor market outcomes? *Economics of Education Review, 25*(4), 430-446.

Schwab, K., Samans, R., Zahidi, S., Leopold, T. A., & Ratcheva, V. (2017). The global gender gap report 2017. Geneva, Switzerland: World Economic Forum.

Seguino, S. (2007). PlusÇa change? Evidence on global trends in gender norms and stereotypes. *Feminist Economics, 13*(2), 1-28.

Shafiq, M. N. (2013). Gender gaps in mathematics, science and reading achievements in muslim countries: A quantile regression approach. *Education Economics, 21*(4), 343-359.

Skelton, C., Francis, B., & Valkanova, Y. (2007). *Breaking down the stereotypes: Gender and achievement in schools* (Working Paper Series No. 59). Manchester, UK: Equal Opportunities Commission.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*(1), 4-28.

Steffens, M. C., Jelenec, P., & Noack, P. (2010). On the leaky math pipeline: Comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. *The Journal of Educational Psychology, 102*(4), 947-963.

Stickney, L. T., & Konrad, A. M. (2007). Gender-role attitudes and earnings: A multinational study of married women and men. *Sex Roles, 57*(11), 801-811.

Stoet, G., & Geary, D. C. (2012). Can stereotype threat explain the gender gap in mathematics performance and achievement? *Review of General Psychology, 16*(1), 93-102.

Stoet, G., & Geary, D. C. (2013). Sex differences in mathematics and reading achievement are inversely related: Within- and across-nation assessment of 10 years of PISA data. *Plos One, 8*(3), 1-10.

Thornton, A., & Freedman, D. (1979). Changes in the sex role attitudes of women, 1962-1977: Evidence from a panel study. *American Sociological Review, 44*(5), 831-842.

U.S. Department of Education. (2000). *Trends in educational equity of girls & women* (NCES 2000-030). Washington, DC: National Center for Education Statistics.

Van der Vleuten, L., (2016). Mind the gap! The influence of family systems on the gender education gap in developing countries 1950–2005. *Economic History of Developing Regions, 31*(1), 47-81.

Van Langen, A., Bosker, R., & Dekkers, H. (2006). Exploring cross-national differences in gender gaps in education. *Educational Research and Evaluation, 12*(2), 155-177.

Vermunt, J. K., & Magidson, J. (2002). Latent Class Cluster Analysis. In Hagenaars, J. A. & McCutcheon (Eds.), *Applied Latent Class Analysis* (pp. 89-106). New York, NY: Cambridge University Press.

Wai, J., Cacchio, M., Putallaz, M., & Makel, M. C. (2010). Sex differences in the right tail of cognitive abilities: A 30 year examination. *Intelligence, 38*(4), 412-423.

Wei, T. (2009). *Stereotype threat, gender, and math performance: Evidence from the National Assessment of Educational Progress*. Working Paper, Harvard University.

Weichselbaumer, D., & Winter-Ebmer, R. (2005). A Meta-Analysis of the international gender wage gap. *Journal of Economic Surveys, 19*(3), 479-511.

Weinberger, C. J. (2001). Is teaching more girls more math the key to higher wages? In King, M. C. (Ed.), *Squaring Up: Policy Strategies to Raise Women's Incomes in the U.S.* Ann Arbor: University of Michigan Press.

Wilder, G. Z., & Powell, K. (1989). *Sex differences in test performance: A survey of the literature* (ETS RR No. 89-4). New York, NY: Entrance Examination Board.

Wiseman, A. W., Baker, D. P., Riegle-Crumb, C., & Ramirez, F. O. (2009). Shifting gender effects: Opportunity structures, institutionalized mass schooling, and cross-national achievement in mathematics. In D. P. Baker, & A. W. Wiseman (Eds.), *Gender, equality and education from international and comparative perspectives* (pp.395-422). Bingley, UK: Emerald Group Publishing Limited.

Xie, Y., & Shauman, K. A. (2003). *Women in science: Career processes and outcomes*. Cambridge, MA: Harvard University Press.

Young, D. J., & Fraser, B. J. (1994). Gender differences in science achievement: Do school effects make a difference? *Journal of Research in Science Teaching, 31*(8), 857-871.

Yu, W., & Lee, P. (2013). Decomposing gender beliefs: Cross-national differences in attitudes toward maternal employment and gender equality at home. *Sociological Inquiry, 83*(4), 591-621.

Zafar, B. (2013). College major choice and the gender gap. *Journal of Human Resources, 48*(3), 545-595. doi:10.1353/jhr.2013.0022

Zakaib, G. D. (2011). Science gender gap probed: Overt sexism is no longer the norm, but societal barriers remain for women in science. *Nature, 470*(7333), 153.

VITA

Hee Jin Chung

EDUCATION

Ph.D. in Education Theory and Policy / Comparative and International Education
Minor in Demography
Pennsylvania State University August 2018

M.A. in Education and Social Policy
New York University May 2013

B.A. in English Language Education
Seo*ul National University* February 2010

FELLOWSHIPS AND AWARDS

2014 – 2016 UCEA Barbara L. Jackson Scholar
2013 – 2015 University Graduate Fellowship (UGF)

PUBLICATIONS

Byun, S., **Chung, H**., & Baker, D. (forthcoming). Global Patterns of the Use of Shadow Education: Student, Family, and National Influences. *Research in the Sociology of Education.*

Byun, S., **Chung, H.**, & Ahn, J. (forthcoming). The relationship between attendance in English kindergarten and parents' plans for their children's educational transition. The Korean Society for Sociology of Education. *Korean Journal of Sociology of Education.*

Bodovski, K., Byun, S., Chykina, V., & **Chung, H.** (2017). Searching for the golden model of education: Cross-national analysis of math achievement. *Compare: A Journal of Comparative and International Education*.

Byun, S., Bai, Y, & **Chung H.** (2017). Consequences of educational assortative marriage for children's academic achievement in South Korea.

Byun, S., Ahn, J., & **Chung, H.** (2017). Differences in academic achievement between private and public elementary school graduates during the middle school years. *Korean Journal of Educational Research, 55*(4), 167-187.

Chykina, V., **Chung, H.**, & Bodovski, K. (2016). Great expectations? Variation in educational plans of students in post-socialist Eastern Europe. *European Education, 48*(1), 43-62.

Byun, S., & **Chung, H.** (2016). Characteristics of parents who desire to send their children to specialized middle schools–Focusing on international middle schools. *Korean Journal of Educational Research, 54*(4), 195-232.