Mixing or Separating Genders in Higher Education?
Evidence from a Natural Experiment in Iran

Working Paper for Attendance in
American Economic Association Annual Meeting 2019

Zahra Kamal

Friedrich-Bauer Strasse 27,
91058 Erlangen
Tel.: +49 (0) 157 74 13 35 47
Email: zahra.kamal@uni-bamberg.de

Supervisor:
Prof. Dr. Guido Heineck

Chair of Empirical Microeconomics
Feldkirchenstrasse 21,
96052 Bamberg
Tel.: +49 (0)951 863 2600
Email: guido.heineck@uni-bamberg.de
Abstract
Numerous research on single-sex education has produced inconsistent results, mainly due to methodological issues and selection biases. The current research attempts to contribute to the vast literature on single-sex education by addressing the dearth of research on the effect of gender separation policy in the context of higher education, as well as in the context of Muslim-majority countries where, ironically, single-sex education is more prevalent.

Although higher education is mainly not separated by gender in Iran, following a resurgence of the debate over single-sex education in recent years, some Iranian public universities implemented the policy of gender separation at classroom level. In 2011, one of the largest public universities in Tehran started to exclusively offer single-sex classrooms without a pre-announcement to the public. As a result, students who selected to study in this university in 2011 were not aware that they would participate in classrooms merely with those of their own sexes. Exploiting such a natural experiment, this paper investigates the causal impact of single-sex versus mixed classrooms in higher education on students’ achievements.

Initial naïve estimates show an ostensibly large effect of mixing genders in classrooms (+0.30 standard deviation for males and -0.23 standard deviation for females). However, using appropriate controls for socioeconomic status and incoming test score as a measure of ability reveals that participation in mixed classrooms improves males’ first-year GPA by 0.14 standard deviation, while it decreases females’ performances by 0.12 standard deviation. Using a difference-in-difference approach increases the effect in size and significance for males (0.17 standard deviation), while it decreases the impact on females’ attainments to non-significant. Findings also show that the positive effect of mixed classrooms is larger for males from lower socioeconomic status and with higher levels of ability. In contrast, females from higher social class and with lower ability levels are more negatively affected by attending mixed classrooms. Interestingly, coming from a more religious family seems not to affect the impact of the policy on educational outcomes.

Although the current paper found that the policy of separating classrooms by gender at universities is beneficial for females and harmful for males in terms of their achievements, the question whether the Iranian government should widen the scope of the policy in terms of the number of public single-sex universities or mixed universities with separated classrooms is still open. Future studies with additional data on various aspects need to evaluate the impact of the policy in the long term, and on key social outcomes beyond educational performances.

Keywords
single-sex education, gender separation, coeducation, mixed education, gender equality, education policy
Introduction
Gender equality has always been a focal point of conversation among scholars, political figures and policymakers. Among different aspects of equality, access to education constitutes a prominent indicator of equality in any society. In an educational landscape marked by problems of low achievement, violence, drugs, poverty, sexism, racial and ethnic tension, the emergence of single-sex education has been regarded by some as a promise of a way out (Morse, 1998), and a mechanism to address existing gender gaps in test scores, in choices to study technical or so-called male-dominated courses and degrees, or in wages and occupations. (see for example Booth et al, 2018; Billger, 2009 )

Definition
The U.S. Department of Education defines “single-sex education” as “education at the elementary, secondary, or postsecondary level in which males and females attend school exclusively with members of their own sex”. (U.S. Department of Education, 2005) In contrast, “coeducation” or “mixed” setting is provided when students attend in a mixed environment alongside their counterparts of the opposite sex. A related though different phenomenon is “single-sex classroom”, whereby schools that are otherwise coeducational, offer separate classes for males and females. (Mael et al., 2005)

History
By 19th century, the most common way for girls to access education was at home, through private tutoring. Higher educational degrees were uncommon among females in most countries. During the 19th century, ideas about education started to change: modern ideas defined education as a right -rather than as a privilege available only to a small elite (Peitzman, 2000). Together with mass education, the practice of coeducation was universalized in many places as a solution to promote educational equality. In the 20th century, with the rise of secularism in the western world, single-sex education fell into deep disfavor, except in private or parochial schools. (Riordan, 2011)

The recent resurgence in the number of single-sex education environment in the western world is mainly associated with educators’ and policymakers’ rising concerns about gender (Sax, 2009) and their belief that single-sex environments are educationally beneficial, at least for some students (Salomone, 2006) As a result, single-sex education was not anymore viewed as an archaic system based mostly on a strict religious culture, rather as a mechanism to improve academic performance, especially among females, the poor, and disadvantaged students.

Today, in most western countries where single-sex education is still common, public schools are mostly coeducational, and single-sex education is primarily offered by private (usually religious) schools. In contrast, in Muslim countries, public schools are mainly separated by gender, and coeducation is usually associated with private sector. Both in Muslim and non-Muslim world, higher education is mostly not separated by gender, although there exist few single-sex universities as well.

Background and Context of Research
Education System in Iran
Iran’s education system was modeled on French Education structure in the 19th century. Formal education is highly centralized and divided into K-12 education plus higher (tertiary) education. K-12 education is supervised by the Ministry of Education, while higher education is under the supervision of the Ministry of Science, Research, and Technology and the Ministry of Health and Medical Education. Individual schools
have the authority to take their exams according to the regulations provided by the Ministry of Education. However, at the end of both elementary and secondary education (years 6 and 12 of the study), all students participate in the same final exams held at the national level. There are both public and private institutions in Iran at all levels from elementary school to university (UNESCO, 2002). In order to pursue higher education at a university in Iran, male and female students graduated from high schools have to participate in a National Examination for University Entrance –Konkur- administered each year by Sanjesh Organization under the supervision of the Ministry of Science, Technology, and Research. Having his/her Konkur test score announced by the Ministry, each student submits a list of up to one hundred university-majors to Sanjesh Organization. The choice lists and students’ test scores are then used by Sanjesh Organization to assign the students in successive rounds to universities until all university slots are filled.

**Single-sex Education in Iran**

K-12 education has always been separated by gender in Iran. Even before the Revolution, schools were basically either for male or female students, reflecting religious norms and culture of the society. Coeducation at this level of study merely existed in few international schools aimed at the children of foreigners residing in Tehran or other big cities. In contrast, higher education were mainly not separated by gender. There were few all-female universities at the time. After the Revolution, single-sex schooling was regulated and maintained, and higher education remained primarily as coeducation. Among public universities, very few have their enrollment limited to one gender. Nevertheless, the issue of gender separation in educational environments and single-sex universities have always been a controversial debate in post-revolution Iran.

Discussion over single-sex education in Iran has often focused more on ideological and political concerns rather than research results on the policy evaluation. The debate stems from the 1987 ratification of the Supreme Council of Cultural Revolution (SCCR), called “Ratification of Retaining Islamic Values in Universities and Higher Education Centers”; As an attempt to make universities’ environment more in line with religious values, the ratification required that universities with enough facilities and resources offer separate classrooms for male and female students (Supreme Council of Cultural Revolution, 2011). The ratification has not been enacted until a recent resurgence of the issue among authorities and political figures between 2009 and 2011. In 2011, the Minister of Science, Technology, and Research at the time–Kamran Daneshjou- supported the idea of gender separation in higher education in an official lecture for the Ministry’s managers and employees and Shahid Beheshti University’s students. He claimed that the SCCR’s ratification was binding and needed to be taken as a law. Thereafter, several associations and communities of university students supported the idea by signing a petition to ask the decision makers at the Ministry and authorities of universities to offer single-sex classrooms for male and female students at universities. They asked for the Theological Seminaries’ comment on the implementation of the policy, and most renowned grand religious figures such as Ayatollah Sobhani and Ayatollah Nouri Hamedani appreciated the idea. In practice, a few universities started to implement the policy of gender separation at classroom level. Despite the objections from the students and other social communities, the presidents of such universities supported the policy mostly on the ground that it resulted in relatively higher

---

1 - I would use the term “gender separation” as in the literature “gender segregation” is mainly used to address the policy of imposing restrictions on some fields of study to limit the access of female students, leading to “occupational segregation” in working environment [e.g. Rezaei-Rashti, 2012; Shirazi, 2014; Esfahani and Shajari, 2010].
performance of the students, particularly females, basically due to a decrease in distractions from studying for both genders in the new environment.

The implementation of gender separation policy by some large public universities, once again ignited the debate on gender separation; Proponents of the policy argue that by providing gender-separated environment in universities females’ participation in higher education increases as many conservative families would no longer be reluctant that their daughters acquire higher educational degrees. They believe that the policy helps to promote educational equality as it gives more options to those students who prefer to study in single-sex environment. Supporters of single-sex education also asserts that non-separation causes distractions from studying and reinforces gender biases against women, leading to a decline in educational outcomes, particularly for female students. They claim that in single-sex environment, educational outcomes for female students improve as they have more opportunities to express themselves and cooperate in same-sex study groups in the absence of male students.

Others are at odds with the positive effect of separated education as it has not been investigated by research so far. Opponents also weaken the argument for separation by pointing to potentially less-developed social skills of separated students who have also studied in single-sex environment in all school years beforehand and had no experience to communicate with the opposite sex in school. They believe that such lower social skills could lead to adverse socio-economic outcomes for separated students. Furthermore, they say that there could be other social costs of separation such as an increase in the age of marriage, less success in finding a job, or lower wages of the less-socially-developed and -linked graduates from such separated universities.

Although debates continue on this issue, the policy is still in practice at those separated universities. Since there has been no empirical research done to evaluate the causal impact of the policy on the potential outcomes in this context, the question whether the government should widen the scope of the policy in terms of the number of single-sex universities or mixed universities with separated classrooms is still open. In a conference held by the Iranian Association for Scientific Development in 2011, the participant authorities and experts emphasized on the crucial need for a rigorous research on the consequences of the policy (Iranian Association for Scientific Development, 2011). In this paper, I examine the effect of mixing versus separating genders at classroom level on achievements in higher education in Iran.

Theoretical Framework
The merits of single-sex education have been fiercely debated by scholars, policymakers, and authorities in education. In this section, I present the line of reasoning for the arguments of supporters and the counterarguments of critics of single-sex education, and review the most prevailing rationales and theories for and against the policy that each side mostly rely on.

Arguments for separated education
A variety of rationales for educational benefits of gender separation have usually been emphasized by the supporters; single-sex education is mainly implemented to address male-female biological differences in development and learning-related traits, sexism and stereotype threat, or sexual attraction among male and female students.

Biological Differences
Many supporters of single-sex education hold the essentialist view as they believe that separating boys and girls increases their academic performances because of the supposedly substantial biological
differences between the two sexes. (Pahlke & Hyde, 2016) A deluge of research work has been done in recent years on how boys and girls have different brains, and different developmental trajectories. Leonard Sax represents the essential-difference view arguing that some sex differences have an innate basis. Referring to studies on hearing and visual system, brain development, and nervous system of males and females, Sax asserts that the differences among genders are hard-wired, and the various regions of the brain develop in a different sequence in boys compared with girls. He also claims that many of the differences among genders are educationally relevant. For example, while boys covet risk, girls shy away. Therefore, Boys perform better under moderate stress and girls perform worse. Sax’s main conclusion is that male and female students learn differently for reasons of biology, and perform optimally when instruction targets such learning-related differences. For example, according to Sax, instructors should use louder voice, cool colors and constant movements in all-boy classes, while in all-girl classes they are recommended to lower their voice, and use warm colors and little movements (Pahlke & Hyde, 2016).

Nevertheless, the essential-difference theory is mostly addressed as a ground for separating children studying at elementary levels by gender. As children enter adolescence, they develop stereotype consciousness and awareness of others’ stereotypes (Pahlke & Hyde, 2016). Therefore, for adolescents some other theories and rationales are more emphasized;

**Sexism and Stereotype Threat**

Supporters of single-sex education at secondary and post-secondary educational environment focus on sexism in coeducational classrooms and biases particularly aimed at female students in mixed settings. Some researchers have indicated that in mixed classrooms, boys tend to call out answers or dominate in “hands-on” activities, and thereby to seek and receive most of the teacher’s attention (Askew & Ross, 1988; Francis, 2004; Howe, 1997; Pahlke & Hyde, 2016). These biases are believed to decrease girls’ performance and their interest in traditionally masculine fields such as STEM (Sadker et al., 2009). A number of theories illustrate the mechanisms of the negative influence of sexist approach on students’ performances:

**Identity Theory**

According to “Identity Theory”, perceived group status differences, perceived legitimacy and stability of those status differences, and perceived ability to move from one group to another affect one’s behavior and performance (Tajfel & Turner, 1979; Turner et al., 1999). Supporters of single-sex education believe that in coeducational contexts, status differences are probably endorsed by males making negative comments on females’ abilities and competencies in specific subjects. Therefore, removing the negative influence of males on females’ status perception increases females’ performance and confidence, particularly in STEM subjects. (Pahlke & Hyde, 2016).

**Expectancy-value Theory**

Expectancy-value theory posits that a student’s perception of others’ endorsement of traditional gender stereotypes may result in less interest for pursuing gender-atypical fields (Pahlke & Hyde, 2016). Thus,

---

3 - Sax, Boys Adrift, pp. 17 – 22.
4 - STEM (Science, Technology, Engineering, and Mathematics) subjects are recognized in the literature as male-dominated subjects, in which females are underrepresented or typically underperform (See for example, Park et al., 2018).
supports of single-sex education argue that in coeducational context, females’ academic goals and interests could be negatively affected by their perception of others’ beliefs.

Stereotype Threat
Stereotype threat is defined as the risk of negative evaluation and rejection. Such threat is heightened when one feels at risk of confirming negative stereotypes about one’s group or being negatively stereotyped (Cook et al., 2012). Research posits that stereotype threat contributes to underperformance in difficult tests (Steele et al., 2002). In the presence of stereotype threat, one feels anxious about being judged in terms of negative stereotypes about one’s group rather than personal merit.

Advocates of single-sex education believe that coeducational settings reinforce gender-stereotypes. They argue that in the absence of male counterparts, women feel no pressure to conform to negative stereotypes about women leading to a better performance on their tests and higher scores. (Jackson, 2012) They also believe that boys, particularly those from ethnic minorities are more likely to succeed in all-boy classrooms as stronger mentorship, better social and emotional support, and higher academic focus in such an environment is available (Fergus & Noguera, 2010). Therefore, while coeducational settings activate commonly held stereotypes, the elimination of stereotype threat in single-sex classrooms results in higher performance of females and ethnic minorities (Steele & Aronson, 1995).

Sexual Attraction
Some supporters of single-sex education point to sexual attraction and the distraction inherent in coeducational settings for adolescents to justify separating genders at classroom or institution level. In an early study, Coleman (1961) pointed to the strong emphasis on “rating and dating” in American high school culture, and attributed low achievement of girls in coeducational high schools to peer pressure for prioritizing relations with the opposite sex over schoolwork. According to Epstein (1997), administrators of VMI and The Citadel, two venerated all-male institutions for i.a. military-style training in the U.S., believed that the presence of women would undermine the learning and bonding of men who attended these institutions. By separating students in learning environments, supporters of single-sex education helps them to better focus on their academic activities.

Arguments against separated education
Lack of rigorous evidence
At the first place, critics of single-sex education point to the lack of rigorous research work to support some arguments for separation; For example, critics believe that there is little evidence of substantial differences in brain development between boys and girls, and that it is not clear whether there are implications for single-sex education. Epstein (1997) criticizes the essentialist view for “masquerading fictions as facts” and “irresponsibly and simplistically” claiming that men and women have distinct traits “to persuade women to think less highly of themselves”. Mark Liberman, a linguistics professor at the University of Pennsylvania, believes that the original research behind Sax’s claims about essential gender differences are exaggerated, misunderstood, or misrepresented. In particular, regarding Sax’s assertions about sex differences in hearing based on Corso’s 1959 study, Liberman says that the study found only between one-quarter and one-half of a standard deviation in male and female hearing thresholds.
Liberman also criticizes Sax’s views on gender differences on the ground that the sample sizes in the reference studies are far too small to make meaningful conclusions about gender differences in learning. (Liberman, 2008) Furthermore, opponents point to the lack of sufficient evidence for the relevance of any differences among girls and boys to classroom learning. (Halpern et al., 2011).

In contrast to essential-difference view, opponents emphasize on a bunch of similarities among males and females. Hyde (2005) emphasizes on gender similarities on most, but not all, psychological variables. She believes that “Gender differences can vary substantially in magnitude at different ages and depend on the context in which measurement occurs”. Reviewing 46 meta-analyses, Hyde concludes that her hypothesis that “men and women, as well as boys and girls, are more alike than they are different” is supported as “studies showing increased academic performance often involve other factors that cannot be disentangled from the effects of the single-gender component”. Epstein (1997) believes that far from relying on natural and obvious differences between sexes, society employs laws, rules, and social codes to create sexually divided educational, political, and social spheres.

Opponents also claim worries that gender-targeted classrooms do not address the needs of many students who are not similar to members of their own sex. (Epstein, 1997)

Regarding proponents’ assertion about educational benefits of separation for minorities, some empirical research results contradict the views on better performance of ethnic-minority boys in single-sex schools. (Else-Quest, 2015)

Development Intergroup Theory
Secondly, in opposition to those supporters reasoning on more sexist attitudes in coeducational settings, opponents of gender separation in education concern that dividing students by gender can reinforce gender biases and entrenched stereotypes. Epstein (1997) expresses worries that by denying the extraordinary diversity within each category of men and women, stereotypes are perpetuated. For such as argument, opponents refer to “Development Intergroup Theory” which assumes that increased psychological salience of gender leads to higher levels of essentialist thought, in-group favoritism, and out-group bias. The extent to which single-sex education causes greater gender salience is controversial though. (Pahlke & Hyde, 2016)

Beyond Educational Outcomes
Last but not least, opponents of single-sex education argue that separating educational environment by gender is harmful regardless of the underlying rationales. They believe that gender separation is problematic for the same reason that segregation by race, and social class is, that the diverse environment in education promote tolerance, and cooperation. (Ref. 22 from #4) They worry that by reducing cross-group communication in single-sex classrooms, students could hardly learn from and cooperate with one another. (Ref. 23 from #4) Hyde (2005) also concerns about overinflated claims of gender differences as they cause harm in numerous realms, including women’s opportunities in the workplace, couple conflict and communication, and self-esteem problems among adolescents.
Literature Review and Contribution

This section outlines some of the main findings of and trends in previous research on single-sex education in order to illustrate the knowledge gap in this field as well as the contribution of the current research.

Single-sex education and students’ outcomes

Numerous studies have been conducted to evaluate the effect of single-sex versus mixed education, predominantly at primary and secondary educational level (eg. Riordan, 1994; Campbell & Evans, 1997; Hoffman et al. 2008). Scholars and policymakers have examined the impacts of single-sex education on various outcomes of students. In their systematic review of single-sex education research, Mael et al. (2005) introduced 32 students’ outcomes as the main focus of empirical research in this field, categorized as concurrent/long-term academic accomplishments, concurrent/long-term adaptation and socioemotional development, perceived school culture, and subjective satisfaction.

Most research on the evaluation of single-sex education focused on the impact on students’ academic performance, either using a summary measure of overall achievement (eg. Carpenter & Hayden, 1987; Caspi, 1995; Malacova, 2007; Riordan, 1994; Riordan, et al. 2008; Robinson & Smithers, 1997; Sax et al., 2009; Sullivan et al., 2010; Woodward et al., 1999) or examining achievements in certain subjects such as mathematics (eg. Baker et al., 1995; Bradley, 2009; Daly, 1996; Daly & Defty, 2004; Gordon et al., 2009; Jackson, 2012; Laster, 2004; Riordan, et al. 2008; Stephens, 2009; Stotsky et al., 2010; Sutton, 2011; Wong et al., 2002), science (eg. LePore & Warren, 1997; Marsh, 1989; Robinson & Gillibrand, 2004; Roth, 2009; Shapka & Keating, 2003; Spielhofer et al., 2004), and verbal/English (eg. Daly, 1996; Laster, 2004; Harker, 2000; Riordan, 1985; Riordan, et al. 2008; Russotto, 2009; Woodward et al. 1999).

Other outcomes most frequently addressed by the research on single-sex education are students’ tracking and subject preferences (Billger, 2009; Jackson, 2012; Schneeweis & Zweimüller, 2012).

Main findings of the effect on students’ outcomes

In general, most studies found a positive effect of single-sex education on female students’ achievements and negative or no-effect on males’ (eg. Adkinson, 2008; Bradley, 2009; Bryk et al., 1993(from #11); Lee & Bryk, 1986), Laster, 2004; Sentos et al., 2013; Sax et al., 2009; Stotsky et al., 2010). In an early study at high school level education, Lee and Bryk (1986) found that single-sex education positively affected females’ grades in math and science, but had no significant effect on males’ scores in the same subjects. In another study, Bryk, Lee, and Holland (1993) found a clear advantage for girls’ achievements in single-sex schools, controlling for important contributing factors such as social background and prior achievement. Stotsky, Denny, and Tschepikow (2010) also controlled for potential confounders and found a positive impact of single-sex education on females’ grades in mathematics, and a negative effect on males’ scores in math.

Contradicting the general trend stated above, some research work found a positive effect of single-sex education on males’ outcomes or negative effect on females’ (eg. Brathwaite, 2010; Park et al. (2018); Riordan, 1994; Roth, 2009; ; Spielhofer et al., 2004; Stephens, 2009; Stotsky et al., 2010; Subotnik & Strauss, 1995; Sullivan et al., 2010); Subotnik and Strauss’s (1995) work suggested that single-sex schooling improved male students’ math scores, while it negatively affected females’ scores in math. In addition, some research found positive effect of all-male classrooms on achievements. Comparing the two types of schools without controlling for contributing factors, Brathwaite (2010) found a positive impact of single-sex education on both genders’ overall performance and math’s grades. Roth (2009) also found a
clear benefit in single-sex education, particularly for males, using no control for important confounders such as family backgrounds and educational competencies. Similarly, with no control for pre-differences among the students attending each type of high schools, Speielhofer Benton, and Schagen (2004) found a sizable positive impact of gender separation in education on males’ achievement in math and science. However, some scholars strictly favored single-sex education both for male and female students although they used a rigorous design with sufficient controls or random assignment (Bradley, 2009; Caspi, 1995; Dhindsa & Chung, 2003; Diehm, 2009; Doris et al., 2013; Riordan, 1985; Shapka & Keating, 2003; Stephens, 2009; Park et al., 2018). A recent study by Park, Behrman, and Choi (2018) found a significant positive effect of all-boy schools on students’ achievements in all STEM subjects utilizing the unique setting in Seoul where assignment to single-sex or coeducational high schools was random. Their results were not similar for girls though.

There also exist studies reporting null effect on either group (Baker et al., 1995; Edwards, 2002) or mixed evidence both in support of and against single-sex schooling in terms of the effect on students’ achievements (Stotsky et al., 2010; Vrooman, 2010). Although their sample size was small, Stotsky, Denny, and Tschepikow (2010) considered intake differences among the schools and found that single-sex setting benefitted students in terms of their math scores, while it negatively affected both male and female students’ overall performances.

Research findings were also inconsistent regarding other outcomes of students; for example, Schneider et al., 1988 and Crump (2004) found coeducation strictly beneficial for male and female students in terms of self-concept. However, regarding the same outcome, Cherney and Campbell (2011) and Conway (1997) found it beneficial for both male and female students to study in single-sex schools. Billger (2009) examined the effect of attending a single-sex high school on students' choice of college major. Her findings suggested that girls who studied in single-sex environment are more likely to receive scholarships, but not more likely to pursue college degrees. Schneeweis and Zweimüller (2012) looked the effect of single-sex versus mixed high schools on the students’ choices in higher education. Using the share of female students in adjacent cohorts, they found that females among a higher proportion of females were more likely to choose a technical course in their second year.

**Moderator Investigations**

Many studies on the evaluation of single-sex education recognized different sources of variation in the effect, leading to report heterogeneous effects of single-sex education on different subgroups of male and female students;

First, the mechanisms of the effect for single-sex education are primarily different on subgroups of male and female students;

First, the mechanisms of the effect for single-sex education are primarily different on students from different ages or levels of education. For example, regarding the effect on verbal test scores of students, advantage of single-sex education was medium in elementary school, small in middle school, and close to zero in high school. (Pahlke et al., 2014) Therefore, one of the main factors to be reported in research work on single-sex education is the age or level of education for the agents under study (Pahlke & Hyde, 2016).

The second factor is the dosage or level of exposure to single-sex environment. In their meta-analysis of previous research in the field, Pahlke, Hyde, and Allison (2014) suggested that larger effects were found among girls when single-sex versus coeducation occurred in classes rather than in schools.
Regarding the third source of heterogeneous effect, socioeconomic status, advantages have been claimed for students of lower social class. However too few studies exist to conclude on the moderating effect of socioeconomic factors (Pahlke et al., 2014).

As another moderator, race/ethnicity has been investigated, particularly in the U.S., to affect the size and direction of the effect (eg. Gordon et al., 2009; Riordan, 1994). Some studies reported particularly positive impact on male students from minorities and disadvantaged families (Gordon et al., 2009; Olson, 2010; Riordan, 1994; Russotto, 2009). Most of these studies did not take into account important confounders of family background and other relevant variables. For example, Russotto (2009) relied on small samples of around 20 students in each type of educational settings to conclude that in single-sex setting, Black male students achieved strictly higher grades in math and verbal courses. Few rigorous studies reported a positive effect on minority male students; in an analysis of a subsample of racial/ethnic minority students, Riordan (1994) found that the overall achievements of boys from Latino and Black ethnicity origins improved in single-sex settings. In contrast, some studies suggested somewhat different conclusions that the effects for racial minority male students were also negative (see for example Sanders, 1992; Singh et al., 1998; Sudler, 2009). According to Sudler (2009), math and science grades of male students from the minority Black families in the U.S. were lower in case of studying in single-sex context. Overall, according to Pahlke, Hyde, and Allison (2014), too few number of high-quality research exist to conclude on the impact of single-sex schooling on minorities.

Knowledge gap and contribution
The current study contributes to the research on single-sex education by addressing three main gaps in the literature; 1) inconclusive results of previous studies, 2) the dearth of studies at post-secondary and higher educational levels, and 3) the scarcity of research in the context of Muslim-majority countries.

Methodological issues and inconclusive results
Despite the theories suggesting benefits for single-sex education, little conclusive empirical evidence on the effect exists (Jackson, 2012) as many studies in this field suffer from methodological weaknesses; A primary criticism of previous work on evaluating single-sex education has been the confounding of single-sex effects with the effects of religious values, financial privilege, selective admissions or other advantages associated with the single-sex schools being studied. According to Jackson (2012), two major limitations caused most studies to suffer from considerable bias in their measurement of the effect; First, because students who decide to participate in single-sex education are potentially and probably different from those who decide to attend coeducation in important unobserved characteristics, comparisons between the two groups’ outcomes is likely subject to severe self-selection bias. Second, since single-sex schools often differ in relevant unobserved ways from mixed schools (eg. curriculum, selectivity, teachers’ motivation, extracurricular activities), the comparison confounds single-sex education effect with other institutional differences. A profound research in this field should include statistical controls to account for individual differences (eg. socioeconomic status, individual ability, and age) as well as school and class differences that might account for the differences between single-sex and coeducational schools. (Mael et al., 2005). A number of studies attempted to overcome such restrictions by randomly assigning students to either type of schools or using appropriate controls. However, their findings, as stated above, were not in the same direction.

As the research to evaluate single-sex education has produced inconsistent and sometimes contradictory results, several scholars attempted to integrate previous research in the field and conclude on the
ultimate impact of single-sex education on each group of students (Mael et al., 2005; Morse, 1998; Pahlke et al., 2014). However, the dominant approach in such reviews was not to quantify the ultimate impact, but to merely report the overall direction of the effects (Mael, et al., 2005). In order to synthesize research on the effect of single-sex education, Pahlke et al. (2014) used the method of meta-analysis by assigning a weight to each study's measured effect according to their sample size and quality of research design (controlled vs. uncontrolled). According to Pahlke et al., although uncontrolled studies showed modest advantage for single-sex education, the results from high-quality research did not provide support for single-sex schooling; controlled studies which used random assignments or controls for selection, showed only trivial differences between students in single-sex versus mixed schools in the domains such as mathematics performance or and science performance, and in some cases showed small differences favoring mixed settings.

The current research addresses both methodological limitations stated above by exploiting an abrupt change of policy in one university. Therefore, no selection bias based on students’ choice or institutions’ characteristics is expected to influence the results; The natural experiment setting in this case ensures that students did not select into the treatment group. Moreover, in my research, I do not compare educational outcomes of students at different universities, with different curriculum, different professors, different regulations and institutions. Both treatment and control group studied at the same university, by the same professors, with the same methods and contents. Therefore, in this context, systematic differences between the two groups and their educational experience - except for the gender composition of classes - are unlikely. Nevertheless, to account for coincidental differences or changes that come with the policy implementation, I use several controls regarding students’ inputs and the university/program’s related characteristics.

Higher levels of education

Although a bunch of rigorous research examined the impact of single-sex schooling, there is a dearth of controlled studies on evaluating the implementation of single-sex education policy at higher educational levels (Pahlke et. al. 2014). Pahlke and Hyde (2016) pointed to the crucial need for an investigation of the consequences and impacts of single-sex education at all levels of development. Due to more freedom of choice for adults and their higher tendency to participate in mixed education, conducting a field experiment to study single-sex higher education is usually prohibitively expensive. Few such studies tried to lower the costs by limiting their sample size to students in one major, or confined the exposure to treatment (single-sex classrooms) to merely a small proportion of instruction hours; Booth et al. (2018) conducted a field experiment at a high-ranked university in the UK to examine the effect of participating in single-sex versus mixed classrooms on students’ first-year grades and their course choices in the second year. However, their research design was at first restricted to students of economics. Secondly, the treated students were participating in single-sex classrooms for only one hour of their economics course per week. Booth et al. (2018) did not claim to generalize the positive effect that they found to more hours of single-sex instruction or for other subject areas.

The natural experiment approach in this research enables the investigation of the policy impact in higher education, without incurring the costs and issues of controlled trials. Thus, the sample in this study includes all students majoring in 18 fields of study, and the treatment group participated in single-sex classrooms for all their lectures and classes.
The different context of Muslim-majority countries

Single-sex education is more prevalent and in some cases mandatory in Muslim-majority countries due to religious values and cultural backgrounds. However, studies evaluating the impact of the policy in these contexts are rare.

Considerable research on single-sex schooling has been done in the context of countries such as Britain (eg. Daly, 1996; Robinson & Gillibrand, 2004; Spielhofer et al., 2004; Sullivan, 2009; Sullivan et al., 2010), Canada (eg. Crombie et al., 2002; Shapka, 2009; Shapka & Keating, 2003), and the United States (eg. Basilo, 2008; Calder, 2006; Campbell & Evans, 1997; Edwards, 2002; Phillipps, 2008; Roth, 2009; Sax et al., 2009; Spikes, 2009; Stotsky et al., 2010, Thompson, 2003; Vrooman, 2010), in which single-sex schools make up a small and selective group, or countries such as New Zealand (eg. Baker et al., 1995; Caspi, 1995; Harker, 2000), Australia (eg. Ainley & Daly, 2002; Carpenter & Hayden, 1987; Davey et al., 2011; Dyer & Tiggemann, 1996; Norton & Rennie, 1998), and Ireland (Doris et al., 2013; Hannan et al., 1996), which have a sizable number of single-sex schools (Smyth, 2010). Their results reveal that the mechanisms and thereby the size and direction of the effect heavily depend on the context. In a comparative cross-country study on the effect of single-sex education, Baker, Riordan, and Schaub (1995) found a small effect for the policy in Belgium, approximately no effect in New Zealand, and large positive effects on the achievements for the students from Thailand and Japan. They pointed to the role of national context to explain the differences among the effects. Moreover, the results from other countries are hardly applicable for Muslim-majority countries with extremely different culture, and on individuals who, in some cases/countries spent all their school years in single-sex environments.

Only a few studies have been conducted on the effect of single-sex education in developing countries, particularly Muslim-majority contexts such as the countries in MENA regions. In their meta-analysis, Pahlke, Hyde, and Allison (2014) referred to only one study in the context of Iran as a Muslim-majority country; Esfandiari & Jahromi (1989) compared the achievements and aspirations of students from a single-sex monolingual high school and a bilingual mixed high school in Tehran. However, as the schools in the comparison differed in various ways, the measured effect was not attributable to the gender composition of educational environment as the authors concluded on the effect of bilingualism versus monolingualism rather than single-sex versus coeducation.

In my research, I examine the effect of the policy in the context of Iran, a country which shares a lot of cultural aspects with other countries in the region. Thus, the results of the research could have many implications for the policy making of other countries in the Middle East.

Overall, to the best of my knowledge, this study is the first to uncover the causal effect of single-sex versus mixed education at higher educational level and in the context of a Muslim-majority country with a natural experiment approach.

---

5 - Pahlke et al. (2014) also included in their meta-analysis some studies from Nigeria, a country which is sometimes counted as part of MENA (eg. Banu, 1986; Egbochuku & Aihie, 2009; Lee & Lockheed, 1990; Mallam, 1993).
Mixing or Separating Genders in Higher Education?  
Evidence from a Natural Experiment in Iran  

Research Questions and Hypotheses
To uncover the effect of mixing versus separating classrooms by gender at universities, I particularly seek an answer to the following questions:

- How does participation in single-sex classrooms versus mixed classrooms affect educational outcomes of male and female students in the first year of university?
- How does participation in single-sex classrooms versus mixed classrooms affect males’ and females’ GPA improvement from high school to the first year of university?
- Does single-sex learning environment differently affect the academic success of students based on their socioeconomic background, field of study, or educational competency?

Based on the previous literature related to the research questions above, the following conjectures are tested:

**Hypothesis 1.** Participation in single-sex classrooms positively affects first-year GPA of female students, but has zero or negative impact on males’ first-year GPAs.

In their line of reasoning for single-sex education, supporters mainly rely on *biases against women* in mixed classrooms. When students are exposed to the stereotype that females perform worse, females might shy away from expressing their ideas and speaking up in classroom in order to avoid embarrassment and confirming the negative stereotype about their ability. As a result, male students catch most of instructors’ attention and better involve in class discussions, leading to their better learning experiences and higher performance in exams.

**Hypothesis 2.** Females and males from disadvantaged families (Quota 3) and more religious families (Quota 4) benefit more from participation in single-sex classrooms.

People residing in the countryside and rural areas of Iran have completely different cultural norms from those residing in big cities. Girls and women in small districts and rural areas wear more modest and plain clothing, use less makeup, and try to behave in a way not to attract males’ attention. However, more secularized culture in big cities, especially in the capital, leads to more sexual attraction among the youth. Students from rural areas are expected to more easily be distracted in a mixed classroom than those who have already been exposed to the culture of cities. In addition, interactions among men and women are more limited in families with more religious background. Thus, it is expected that students form such families experience more distractions by higher interaction levels with the opposite sex in mixed classrooms. Therefore, students from more conservative cultures are expected to benefit more from decreasing sexual attractions and the resulted distractions in single-sex classrooms.

**Hypothesis 3.** Students with lower level of ability benefit more from single-sex classrooms than students of higher ability.

When tasks are more difficult, people tend to feel higher anxiety, and stereotypes are easily activated in their mind. (Steele et. al., 2002) In mixed classrooms, where stereotype threat is present against women, females with lower ability feel less confident to express themselves, and avoid more from being involved in class activities. When the biases against women are removed in single-sex classrooms, those females of lower ability change their behavior more than do females of higher ability. Therefore, single-sex classrooms are expected to benefit low-performers of females more than high-performers.
Research Design

In 2011, one of the most venerated and large public universities in Humanities and Foreign Languages in Tehran - The University of Allameh Tabatabaei (UAT) - implemented the policy of gender separation at classroom level, *unexpectedly and without a pre-announcement to the public*. Therefore, the students admitted at the university right after the policy change simply did not know that they would attend classrooms merely with those of their own gender. In order to evaluate the impact of participation in single-sex classrooms, I compare the first-year GPA of these students, who participated in but *had not actually selected* single-sex education, with those of the previous cohort who attended mixed classrooms in their first year of study.

The educational experience of the two cohorts in the first year were otherwise the same; First-year courses were offered by the faculties as compulsory credits, and did not change between the two years. Both cohorts had almost all of their lectures and instructions with the same professors for each course, and same professors instructed all-male and all-female classrooms for the second cohort. Other characteristics of the program such as the assignments, tutorials, exams, and extracurricular activities were also fairly comparable for the two academic years. Therefore, assuming that no major change in the realm of higher education affected the students’ selection of the university, the variation in first-year GPA of the two cohorts is plausibly attributable to different gender composition of their classrooms. For the following years of study, both cohorts participated in single-sex classrooms.

The natural experiment setting in this context and the data at hand enable me to find the effect of mixing versus separating students in higher education with two methodological approaches. In this section, I explain my research design to test the hypotheses listed above.

Data and summary statistics

For my analysis, I combined two datasets taken from the UAT and Sanjesh Organization; The UAT data contained information on some of the basic demographic and socioeconomic characteristics of students as well as their educational program and achievements at university. To take into account potential confounders, I collected a complementary dataset including students’ pre-university test scores as a measure for their ability. Roughly 85 percent of UAT dataset were linked to Sanjesh data with no contradictory information on similar fields. For the problematic cases where the information provided by the two organizations differed, I contacted the authorities at both organizations to decide on the correct value for the variables. The merged and cleaned dataset contained individual-level information for 2672 students, the two cohorts admitted at the UAT in the last year before the policy change (2010), and the first year afterwards (2011). In particular, the data used in this research covers students’ year of birth (transformed to Age), gender (Female), Cohort, university field of study (FS), faculty (Fac), high school GPA.
Mixing or Separating Genders in Higher Education? Evidence from a Natural Experiment in Iran

(HSGPA), as well as their Quota6, exam group7 (ExGr), and test score (TSK) in Konkur, and first-year, second-year and total GPA (GPA1, GPA2, and GPAT) at university.

In addition, to account for potential difference in the level of difficulty in Konkur exam and have a more precise sorting of abilities between cohorts, I used the mean and standard deviation of Konkur test scores for each exam group in each year in order to standardize individual Konkur test scores. I also standardized the university GPAs within each field of study for each cohort, so that the achievements could better be compared among students taking the same exams. I also collected information on the number of professors at UAT in each faculty for each year in order to control for basic institutional factors.

6 - To remove the educational discrepancies among regions of Iran and provide support for students of lower social classes and disadvantaged families in the competition of university entrance, a policy of “quota system” has been implemented since 1983; In the procedure of ranking and admission to universities according to students’ test scores in Konkur, a preferential treatment is considered by Sanjesh Organization, that is, a larger multiplier is considered for the students from undeveloped and disadvantaged areas or those belonging to martyrs’ and veterans’ families in order to at least partially compensate for the disadvantages (Mehr, 2014; Malekshahi, 2015). Therefore, quota is regarded as a measure for socioeconomic status of individuals.

7 - Exam Group (ExGr) of the students in Konkur is mainly related to their field of study in high school, except for the exam group 4. “Foreign Languages” is not regarded as a major in high school, and all students from all fields could participate in Konkur as the fourth exam group. This variable is also controlled as it distinguishes the students taking the same exam in each year. However, there is no overlap between the students in our data as the test scores and exam group of students are recognized according to their admission at UAT.
The following table defines each value for the categorical variables.

**Table 1: Definition of each category for categorical variables.**

<table>
<thead>
<tr>
<th>Variable/Value</th>
<th>Category Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Quota</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Highly-developed areas</td>
</tr>
<tr>
<td>2</td>
<td>Medium-developed areas</td>
</tr>
<tr>
<td>3</td>
<td>Least-developed areas</td>
</tr>
<tr>
<td>4</td>
<td>Families of Martyrs and Veterans</td>
</tr>
<tr>
<td><strong>ExGr (in Konkur)</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Mathematics and Physics</td>
</tr>
<tr>
<td>2</td>
<td>Natural Sciences</td>
</tr>
<tr>
<td>3</td>
<td>Humanities</td>
</tr>
<tr>
<td>4</td>
<td>Foreign Languages</td>
</tr>
<tr>
<td><strong>FS</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Theology and Islamic Knowledge</td>
</tr>
<tr>
<td>2</td>
<td>Statistics and Mathematics</td>
</tr>
<tr>
<td>3</td>
<td>Accounting</td>
</tr>
<tr>
<td>4</td>
<td>Laws</td>
</tr>
<tr>
<td>5</td>
<td>Guidance and Counseling</td>
</tr>
<tr>
<td>6</td>
<td>Public Relations</td>
</tr>
<tr>
<td>7</td>
<td>Psychology</td>
</tr>
<tr>
<td>8</td>
<td>Journalism</td>
</tr>
<tr>
<td>9</td>
<td>Languages and Literature</td>
</tr>
<tr>
<td>10</td>
<td>Social Sciences</td>
</tr>
<tr>
<td>11</td>
<td>Economics</td>
</tr>
<tr>
<td>12</td>
<td>Educational Sciences</td>
</tr>
<tr>
<td>13</td>
<td>Political Sciences</td>
</tr>
<tr>
<td>14</td>
<td>Philosophy</td>
</tr>
<tr>
<td>15</td>
<td>Library and Information Science</td>
</tr>
<tr>
<td>16</td>
<td>Social Work</td>
</tr>
<tr>
<td>17</td>
<td>Management</td>
</tr>
<tr>
<td>18</td>
<td>ECO College of Insurance</td>
</tr>
</tbody>
</table>
Figure 1 presents descriptive statistics for some of the key characteristics of the students by cohort. As shown in the figure, the two cohorts were comparable in terms of socio-economic status and exam groups of the entrants. However, the share of female to male students changed between the cohorts. Therefore, controlling for gender would be of paramount importance to evaluate the impact of the policy. In addition, the university did not admit any student in certain fields of study in the second cohort (FS = 2, 13, 14, 15, 16). I will discuss about how this exclusion could influence the results in the discussion section.

Table 2 describes the summary statistics for the students’ educational competencies broken up by cohort and gender. Females in the second cohort, i.e. who attended single-sex classrooms, on average performed 0.263 standard deviation higher in first-year exams than females in the previous cohort participated in mixed classrooms. However, males who attended all-male classrooms underperformed those who participated in mixed classes by 0.310 standard deviation in the first-year exams. As the table shows, mere comparison of achievements could lead one to conclude that single-sex classrooms had a large positive effect on females’ outcomes, while it negatively affected males’. Comparing high school GPAs also strengthens the support for single-sex education as females and males of second cohort had lower high...
Mixing or Separating Genders in Higher Education?
Evidence from a Natural Experiment in Iran

school GPA on average. However, comparing the means for Konkur test scores (TSK) tells a different story; Females of cohort 2 had significantly higher initial test scores, and males of cohort 2 had significantly lower performance in Konkur. Thus, higher ability of females and lower ability of males in the second cohort, could at least be partially responsible for the ostensibly large effect of separation in the first year. To produce more reliable results for the impact of single-sex classrooms, a controlled analysis of the data follows in the next section.

Table 2: Summary statistics and mean differences for continuous variables for cohort 1 and cohort 2 by gender (Numbers in parentheses are standard errors.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cohort 1 (mixed classrooms)</th>
<th>Cohort 2 (single-sex classrooms)</th>
<th>Mean Difference between the two cohorts</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPA1 Total</td>
<td>15.88 (0.05)</td>
<td>015.97 (0.06)</td>
<td>0.10*</td>
</tr>
<tr>
<td>GPA1 Females</td>
<td>15.92 (0.05)</td>
<td>16.43 (0.07)</td>
<td>0.51***</td>
</tr>
<tr>
<td>GPA1 Males</td>
<td>15.67 (0.11)</td>
<td>15.24 (0.09)</td>
<td>-0.43***</td>
</tr>
<tr>
<td>GPAT Total</td>
<td>16.11 (0.04)</td>
<td>15.98 (0.05)</td>
<td>-0.13**</td>
</tr>
<tr>
<td>GPAT Females</td>
<td>16.22 (0.05)</td>
<td>16.46 (0.06)</td>
<td>0.24***</td>
</tr>
<tr>
<td>GPAT Males</td>
<td>15.63 (0.10)</td>
<td>15.21 (0.08)</td>
<td>-0.42***</td>
</tr>
<tr>
<td>HS GPA Total</td>
<td>16.51 (0.05)</td>
<td>15.65 (0.08)</td>
<td>-0.863***</td>
</tr>
<tr>
<td>HS GPA Females</td>
<td>16.79 (0.06)</td>
<td>16.62 (0.07)</td>
<td>-0.17**</td>
</tr>
<tr>
<td>HS GPA Males</td>
<td>15.28 (0.14)</td>
<td>14.10 (0.13)</td>
<td>-1.18***</td>
</tr>
<tr>
<td>TSK (standardized) Total</td>
<td>1.940 (0.026)</td>
<td>1.969 (0.030)</td>
<td>0.028</td>
</tr>
<tr>
<td>TSK (standardized) Females</td>
<td>1.943 (0.030)</td>
<td>2.206 (0.035)</td>
<td>0.263***</td>
</tr>
<tr>
<td>TSK (standardized) Males</td>
<td>1.930 (0.058)</td>
<td>1.591 (0.052)</td>
<td>-0.339***</td>
</tr>
</tbody>
</table>
Identification Strategy
With the data at hand, I used two identification strategies to find the effect of mixing versus separating classrooms by gender at the university. In this section, I describe the two approaches for identification.

Model 1: Simple Cross-sectional Analysis
Given that the second cohort did not select the university based on the policy implementation, I assume that the two cohorts had no systematic difference in unobservable factors. Therefore, in the first approach I simply compare first-year GPA of the two cohorts whose learning experiences were the same except for gender composition of their classes. In this approach, I use the baseline specification as the following equation:

\[ GPA_{1ijk} = \beta_0 + \sum_n \beta_{in} X_{in} + \sigma_{m1} Female_i \times Mix_j + \delta SP_k + \epsilon_{ijk} \]

In equation (1), \( GPA_{1ijk} \) denotes the first-year GPA of student \( i \) in cohort \( j (j = 1, 2) \) and faculty \( k (k = 1, 2, 3, ..., 11) \). \( X_i \) is a matrix of student \( i \)'s basic characteristics (age, gender), socioeconomic status (quota), education (field of study, exam group), and test score in Konkur. In addition, as a main confounding institutional factor, student-to-professor ratio for each faculty \( (SP_k) \) is controlled. The policy variable \( Mix_j \) takes the value of 1 for the cohort with mixed classrooms \( (j = 1) \) and equals 0 for the cohort with separated classrooms \( (j = 2) \). The coefficient \( \sigma_{m1} \) estimates the effect of participation in mixed classrooms on first-year GPA of students.

Model 2: Difference-in-Difference Analysis
As stated earlier, all primary and secondary schools in Iran are separated by gender. Therefore, both cohorts were studying in single-sex classrooms before entering the university. Then in the first year of higher education, the first cohort participated in mixed classrooms, while the second remained separated. In such a classical setting for using a difference-in-difference approach, I compare the first-year GPA of the students controlling for their high school GPA, their achievements at the initial same status of single-sex education. In equation (2), the added term includes a control for high school GPA of student \( i \) \( (HSGPA_i) \) and its interaction with the policy variable \( Mix_j \).

\[ GPA_{1ijk} = \beta_0 + \sum_n \beta_{in} X_{in} + \sigma_{m2} Female_i \times Mix_j + f(HSGPA_i, Mix_j) + \delta SP_k + \epsilon_{ijk} \]

The OLS estimate of parameter \( \sigma_{m2} \) shows the effect of mixing rather than separating students in the first year of university according to model 2 using a difference-in-difference approach.

Figure 1 presents a schematic diagram to illustrate model 1 and model 2 approaches.

![Figure 2: Schematic representation of Model 1 and Model 2 approaches. The dashed and dotted lines represent Model 1 and 2 approaches respectively.](image-url)
Heterogeneity and Moderator Analysis

Among several sources of variation in the effect of mixed versus single-sex classrooms on students’ outcomes, I examine the moderating effect of socioeconomic status (SES) and the level of ability as the most relevant in the context. ⁸

To test the second hypothesis, I used an interaction term between the policy variable (Mix) and the students’ SES defined by the variable “Quota”. To test the last hypothesis about the moderating effect of ability, I classified students according to their ability level using percentiles of their performance in Konkur. All students whose test scores in Konkur (TSK) were above the 75 percentile, were classified as of high ability. Those performing between 25 and 50 percentile, and those scoring below 25 percentile were classified as of average- and low-ability respectively. Then I used an interaction term between the new categorical variable “Abil” and the policy variable “Mix” to see whether single-sex classrooms have different effects on the outcome of students with different levels of ability.

The results are presented in the following section.

Results

To illustrate the importance of addressing main confounding factors, I firstly present the naïve estimates of the effect of gender separation without controlling for the students’ age, quota, education type, and test score in Konkur. Simple regression of first-year GPA on the policy variable produces an ostensibly large effect for the policy both for male and female students. In particular, without accounting for pre-differences among the two cohorts, the authorities would conclude that the separation policy improved females’ outcomes by 0.23 standard deviation and decreased males’ by 0.29 standard deviation. However, once the differences in demographic and socioeconomic status, and educational competency of students are controlled the policy effect remarkably reduces. More reliable estimations for the effect of mixing versus separating genders in university classrooms, are provided by using model 1 and 2 described above.

Table 2 shows the estimated coefficients for each variable and the marginal effect of the policy on male and female students according to all three estimation strategies.

In both approaches, besides controlling for the field of study, I standardized GPA scores within the fields so that the relative difficulty of exams is taken into account and the students who took the same exams are better compared.

According to model 1 results, while female students benefitted from participating in single-sex classrooms, the outcome of male students decreased by attending separated classrooms. Female students who attended single-sex classrooms (cohort 2) performed 0.12 standard deviation better on average than comparable females in the previous cohort who participated in mixed classrooms. However, in single-sex classrooms males’ performances decreased by 0.14 standard deviation on average. While the

---

⁸ - The University of Allameh is predominantly specialized on humanities, language studies, and social sciences, the fields mostly recognized as female-dominated majors. The share of female students was quite high (between 60 to 70 percent) for nearly all fields. Regarding how different fields are perceived in the culture of society as a male- or female-dominated major and the classic STEM categorization, only the second field “Statistics and Mathematics” in the Faculty of “Mathematics” could be considered as male-dominated. Unfortunately the number of students in this field was not enough (only 27 students in the first cohort) to perform the analysis on the moderating effect of the type of subject.
result for females is highly significant, the negative effect of single-sex classrooms on males’ outcomes is hardly significantly different from zero.

Interestingly, by using a difference-in-difference approach in model 2 the positive effect of single-sex classrooms on females’ outcomes vanishes, while the negative effect on males’ outcomes increases both in size and significance level; Model 2 results posit that the effect of mixed classrooms on females’ first-year GPA is small (-0.07 standard deviation) and insignificant. However, male students who attended mixed classrooms in the first year of university performed 0.17 standard deviation better than their counterparts who participated in single-sex classrooms.

In line with the literature, stereotype threat and biases against females students in mixed classrooms constitute the predominant justification for the results. Male students probably feel more self-confident in the presence of women in classrooms while female students feel higher levels of anxiety and stress about confirming negative stereotypes against women in mixed setting.

The coefficients for control variables are also presented in the table.
Mixing or Separating Genders in Higher Education?  
Evidence from a Natural Experiment in Iran  

Table 3: Regression results and marginal effect of the policy “Mix” on male and female students, according to model 1 and 2.

<table>
<thead>
<tr>
<th></th>
<th>Naïve Model</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>0.69***</td>
<td>0.47***</td>
<td>0.23***</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td>0.04***</td>
<td>0.04***</td>
</tr>
<tr>
<td>FS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Theology and Islamic Knowledge</td>
<td>-</td>
<td>0.71***</td>
<td>0.70***</td>
</tr>
<tr>
<td>2. Statistics and Mathematics</td>
<td>-</td>
<td>-0.69***</td>
<td>-0.72***</td>
</tr>
<tr>
<td>3. Accounting</td>
<td>-</td>
<td>-0.26***</td>
<td>-0.30***</td>
</tr>
<tr>
<td>4. Laws</td>
<td>-</td>
<td>-0.01</td>
<td>-0.13</td>
</tr>
<tr>
<td>5. Guidance and Counseling</td>
<td>-</td>
<td>0.57***</td>
<td>0.61***</td>
</tr>
<tr>
<td>6. Public Relations</td>
<td>-</td>
<td>0.31**</td>
<td>0.42**</td>
</tr>
<tr>
<td>7. Psychology</td>
<td>-</td>
<td>-0.13</td>
<td>-0.12</td>
</tr>
<tr>
<td>8. Journalism</td>
<td>-</td>
<td>0.33**</td>
<td>0.46***</td>
</tr>
<tr>
<td>9. Languages and Literature</td>
<td>-</td>
<td>-0.22**</td>
<td>-0.20**</td>
</tr>
<tr>
<td>10. Social Sciences</td>
<td>-</td>
<td>0.44***</td>
<td>0.46***</td>
</tr>
<tr>
<td>11. Economics</td>
<td>-</td>
<td>-0.25***</td>
<td>-0.24***</td>
</tr>
<tr>
<td>12. Educational Sciences</td>
<td>-</td>
<td>0.55***</td>
<td>0.64***</td>
</tr>
<tr>
<td>13. Political Sciences</td>
<td>-</td>
<td>0.42**</td>
<td>0.36**</td>
</tr>
<tr>
<td>14. Philosophy</td>
<td>-</td>
<td>0.44**</td>
<td>0.45**</td>
</tr>
<tr>
<td>15. Library and Information Science</td>
<td>-</td>
<td>0.04</td>
<td>0.15</td>
</tr>
<tr>
<td>16. Social Work</td>
<td>-</td>
<td>0.25</td>
<td>0.19</td>
</tr>
<tr>
<td>17. Management (Base category&lt;sup&gt;9&lt;/sup&gt;)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>18. ECO College of Insurance</td>
<td>-</td>
<td>-0.22**</td>
<td>-0.21**</td>
</tr>
</tbody>
</table>

**Quota**

1. Highly-developed areas (Base category) | - | - |
2. Medium-developed areas                   | - | 0.11***  | 0.14***  |
3. Least-developed areas                    | - | 0.18***  | 0.27***  |
4. Families of Martyrs and Veterans         | - | -0.34*** | -0.18*   |

**ExGr**

1. Mathematics and Physics                  | - | 1.63***  | 1.52***  |
2. Natural Sciences                          | - | 1.02***  | 0.89***  |
3. Humanities (Base category)                | - | -        | -        |
4. Foreign Languages                         | - | 1.40***  | 1.27 *** |

**TSK**

<table>
<thead>
<tr>
<th></th>
<th>Naïve Model</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.46***</td>
<td></td>
<td>0.33***</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>0.31**</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-0.03</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>0.01</td>
<td>-0.00</td>
</tr>
</tbody>
</table>

**Marginal Effect on**

Males             | 0.30***     | 0.14*    | 0.17**   |
Females           | -0.23***    | -0.12*** | -0.07    |

<sup>9</sup> Base categories are selected according to the largest-group criterion.
Table 4 presents the results for moderator analysis. 

Interestingly, the results show different trends for male and female students from areas with different levels of development: while males from disadvantaged and rural areas (Quota = 3) are more positively affected by attending mixed classrooms, females from such regions performed nearly indifferent under the policy. However, females from highly-developed regions significantly performed worse in mixed setting, while their male counterparts from similar districts did not perform differently under the policy.

According to the literature we do not expect that the policy affect males’ outcomes significantly, except for those from minorities and disadvantaged families (see for example Billger, 2009 & Gordon et al., 2009). The current results are in line with the literature in this respect. However, the ironic part is the direction of the effect; I found a large and significant positive effect of mixed classrooms on the outcome of students from least-developed areas while according to the literature, male students from disadvantaged families should have benefited from participation in single-sex classrooms. A possible explanation could be that in small towns and rural areas, males who are exceptionally successful in school stand out and get a large support and encouragement from their small society to pursue higher education at good universities in Tehran such as the UAT. The attention and admiration they get from the family and society increase their self-confidence, which is better preserved in the presence of female classmates rather than in all-male setting. Also, female students from rural areas who could enter a venerated university such as the UAT, probably outperformed their classmates in school years and received immense attention from their society. Their performance is not significantly affected in mixed classrooms as they rely more on their abilities and believe less in stereotypes against women. However, female students from developed areas are more aware of negative biases against their abilities and feel more stressed about confirming such stereotypes in mixed setting. As a result, they feel higher level of stereotype threat in mixed context and are more likely to shy away from expressing themselves in mixed classrooms. Moreover, females from highly-developed regions tend to dress more attractively and use more makeup and thereby attract more attention from male students. Therefore, they could better focus on their class activities in single-sex environment, and their performance is more affected by the gender composition of classroom.

Ironically, coming from a more religious family (Quota=4) seems not to be a source of variation in the effect of the policy. First-year GPAs of male and female students from relatively more religious families were not affected significantly by the gender composition of classrooms. As these students receive a large privilege from the Ministry for the assignment process in Konkur, they generally have lower test scores in Konkur. Therefore, the no-effect result for this subgroup could be partly explained by the small impact of the policy on low-performers.

As predicted in the last hypothesis, single-sex classrooms are more beneficial for female students with lower levels of ability. High-ability women with higher level of self-confidence are almost indifferent to the gender composition of their classrooms. According to the table, participation in mixed classrooms hardly affect the GPA of high-ability female students (-0.13 standard deviation only at the 10% level of significance by model 1, and not at all significant by model 2). However, low-performers among females would have performed 0.28 standard deviation better if they had studied in all-female classrooms.

---

10 - To explain the results in moderator analysis, I would rather rely mostly on cultural differences among subgroups as the first-year students are typically more under the influence of their family and their society’s culture.
Regarding the results for males, high-ability males benefitted more by attending mixed classrooms. Males of Lower ability were almost indifferent to gender composition of the classrooms.

This result is also expected from the literature (eg. ... ) and supported by the cultural background that low-ability males perform better in the absence of female classmates as they feel less anxious and express themselves better. Therefore, the group of low-ability males benefit less from attending mixed classrooms.

Table 4: Marginal effect of “Mix” for each subgroup by gender.

<table>
<thead>
<tr>
<th>Variable/Value</th>
<th>Subgroups</th>
<th>Gender Group</th>
<th>Policy Effect by Model 1</th>
<th>Policy Effect by Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quota</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Highly-developed areas</td>
<td>Males</td>
<td>0.08</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Females</td>
<td>-0.16***</td>
<td>-0.11**</td>
</tr>
<tr>
<td>2</td>
<td>Medium-developed areas</td>
<td>Males</td>
<td>0.18*</td>
<td>0.21**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Females</td>
<td>-0.06</td>
<td>-0.01</td>
</tr>
<tr>
<td>3</td>
<td>Least-developed areas</td>
<td>Males</td>
<td>0.24**</td>
<td>0.23**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Females</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>4</td>
<td>Families of Martyrs and Veterans</td>
<td>Males</td>
<td>0.07</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Females</td>
<td>-0.17</td>
<td>0.03</td>
</tr>
<tr>
<td>Abil</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Low-ability</td>
<td>Males</td>
<td>0.05</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Females</td>
<td>-0.28***</td>
<td>-0.12</td>
</tr>
<tr>
<td>2</td>
<td>Average-ability</td>
<td>Males</td>
<td>0.16**</td>
<td>0.17**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Females</td>
<td>-0.16***</td>
<td>-0.11**</td>
</tr>
<tr>
<td>3</td>
<td>High-ability</td>
<td>Males</td>
<td>0.20**</td>
<td>0.19**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Females</td>
<td>-0.13*</td>
<td>-0.09</td>
</tr>
</tbody>
</table>

Robustness Check

In this section, I test the robustness of the estimated effects to ensure that the variation in the outcome variable is indeed attributable to the gender composition of classrooms. To do so, the policy variable (Mix) in the previous analysis is substituted with a relevant, though different variable, female ratio, calculated within each field of study.

The results are consistent with the estimated effect for participation in mixed versus separated classrooms: An increase in female ratio in classrooms gives rise to both male and female students’ performances. This translates into the benefit of single-sex classrooms for females and mixed classrooms for males, which is in line with our results in the previous section.

The consistency of results are also maintained regarding the significance level and the discrepancies of estimations by the two models; As shown in table 3, the marginal effect of “Mix” on the first-year GPA of male students increases both in size and significance when model 2 is used rather than model 1. In addition, while by model 1 the estimated marginal effect on females’ outcomes is highly significant, using model 2 decreases the estimated effect for females to non-significantly different from zero. Table 5 below shows that both trends are preserved when female ratio is employed rather than the policy variable (Mix) in the regressions; the effect of female ratio on males’ outcomes is higher and more significant in model
2 estimation than model 1. Also, while the effect estimated by model 1 for females is significant, model 2 posits no-effect on females’ outcomes.

Table 5: Marginal effect of female ratio in classroom on students’ first-year GPA by gender

<table>
<thead>
<tr>
<th>Fe (Female Ratio)</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>0.18*</td>
<td>0.28***</td>
</tr>
<tr>
<td>Females</td>
<td>0.51**</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Discussion

In this section, I discuss about the threats to causality in the results and the direction for potential bias in the estimated effect, illustrating some of the limitations for this study.

The identification strategies in this research rely on the assumption that all of the difference in the first-year GPA of the students of the cohorts immediately before and after the policy change is attributable to the differences in gender composition of their classrooms. Besides no-selection by students, comparability of the cohorts requires that in the period of first-year study of the cohorts, no major coincidental transformation occurred 1) in the institutions and programs at the UAT, and 2) in the general realm of higher education in the country. This is more plausible when the time separating the two cohorts is less.

Regarding the program and institutions at the UAT, first-year mandatory courses did not change between 2010 and 2011. Moreover, according to the lists of human resources at the employment office of the UAT, the number of professors hired or retired between the two years was small (lower than 12%) and in most faculties the cases were not related to the first-year courses. Furthermore, interviews with professors revealed that in order to design assignments and exams, they usually rely on a pool of sample questions which is hardly influenced by non-scientific factors such as the university’s policy of gender separation in classrooms. Other potential contributing factors such as the number of tutorial hours for each course, extracurricular activities, etc. were also fairly comparable.

In the realm of higher education in Iran, no major change occurred in the period 2010 to 2011. For example, the ranking of the UAT in official ratings and in public imagination did not change, nor did the process of assigning students to universities by the Ministry of Science. The latter issue points to the reason that I did not go further into the past to use the data from the previous cohorts (2009, 2008,...) for the comparison; In 2009, the admission process of universities coincided with the re-election of the President Mahmoud Ahmadinejad and the time of the demonstrations against the government in which many university students involved. At the time, rumor had it that the assignment of students to universities by the Ministry of Science was biased towards students from small cities to prevent the universities in the Capital, Tehran, from increasing social unrest. Although it is highly unlikely that there is truth in the rumor, it could still affect students’ selection for universities in Tehran such as the UAT. In addition, I could not use the data for subsequent cohorts after 2011 because prospective students were probably aware of the university’s new situation and could have selected the university due to their preferences for separated classrooms.

Consistency of the estimated effect in this paper would still be threatened if there exist relevant factors that have coincided with the implementation of gender separation policy at the UAT. For example,
Mixing or Separating Genders in Higher Education? Evidence from a Natural Experiment in Iran

although I controlled student-to-professor ratio as a measure for relevant institutional shift between the two years, data on the average class size for each cohort did not exist. Given that the share of male students in all fields at the UAT was considerably lower than the share of females, we expect that the size of all-male classrooms for the second cohort remarkably decreased. According to the positive impact of smaller class size on students’ performance (Angrist & Lavy, 1999), we expect that the negative effect of separating classrooms by gender on males’ outcome had actually been higher than the estimations in this paper. However, as the number of female students in most fields well exceeded the university’s class size threshold (25 individuals), the shift in the class size of female students between the two years was negligible. Therefore, I expect the size of estimations for male students to be smaller than the true impact.

According to descriptive statistics on the university fields of study (Figure 1), we see that the UAT did not admit any student in some fields for the second cohort. The OLS estimates of parameters $\sigma_{m1}$ and $\sigma_{m2}$ may suffer from bias if the students’ GPA in those no-entrant-majors were either higher or lower from the rest fields. To check for potential bias caused by this factor, separate regressions were done by both models excluding the students in those fields of study from the first cohort.

Table 6: Marginal effect of the policy “Mix” on male and female students, with reduced sample (excluding fields of study with no entrant in the second cohort).

<table>
<thead>
<tr>
<th>Marginal Effect on</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>0.13*</td>
<td>0.15*</td>
</tr>
<tr>
<td>Females</td>
<td>-0.12***</td>
<td>-0.06</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>2490</td>
<td>2490</td>
</tr>
</tbody>
</table>

As shown in the table 6 above, the marginal effect of mixing policy for females does not change and for males only slightly decreases. As the number of students excluded from the regression are only 147 individuals (sample size only 147 units decreased), even the standard errors did not increase from the first regressions of model 1 and 2. Thus, one could hardly attribute the estimated effect in the previous section to the fact that the UAT did not accept any student in some fields in 2011.

Another causal threat for the results stems from attritions and dropouts. A number of 78 individuals from the second cohort (2011) in the merged dataset, did not register or relinquished their study at the UAT in the first year. Accordingly, one could argue that the estimated effect is only valid for those who selected to continue studying in single-sex classrooms, and thereby suffers from selection bias. I rule out such a likelihood based on two grounds: 1) for Iranian high school graduates, the opportunity cost of taking Konkur exam and scoring high to be admitted at such a venerated university is huge. They would hardly abandon the fruit of all their endeavor based on the university’s policy regarding the gender composition of classrooms. 2) The number of students who did not register or gave up studying in the first year from cohort 2010 was nearly the same; given that also in 2010, 88 students whose requests to register at the UAT were accepted refrained from studying at the UAT while the university did not offer single-sex classrooms, it is reasonable to claim that those 78 cases of attritions and dropouts in 2011 are not due to selection against single-sex classrooms.

Last but not least, the university’s atmosphere was nevertheless under the influence of the abrupt implementation of the policy. There were objections against and supports for separating classrooms by
gender among the students and professors. Some students made strenuous effort to change the new policy of the university and gathered signatures for a petition opposing the scheme. All these objections and supports could influence professors’ motivation, and distract students’ focus on study in the new setting. Unfortunately, disentangling the policy effect from the distraction effect was not feasible due to unobservable factors. The positive effect of “Mix” on males’ outcome could be lower as the measured impact partly reflects the negative effect of distractions from study for the second cohort. Also, the negative impact of “Mix” on females’ outcomes could be even higher had the policy change did not distract students of second cohort from studying.

**Conclusion**

In 2011, the authorities at the UAT introduced the policy of gender separation at classroom level without a pre-announcement to the public. Exploiting such a natural experiment setting, I compared the first-year GPA of the 2011 cohort who participated in single-sex first-year classrooms, with those of students in the previous cohort (2010) who attended mixed classrooms.

The current research found a positive effect of mixed classrooms on males’ outcomes and a negative or no-effect on females’ outcomes. According to the simple cross-sectional analysis, although male students strongly benefit from participating in mixed classrooms, such an experience negatively affects the outcome of female students. By using a difference-in-difference approach, the positive effect of mixed classrooms on males’ outcomes increases, while the negative effect on females’ first-year grades becomes insignificant. In fact, the results show that in transition from high school to university, male students benefit from attending mixed classrooms and females do not significantly perform worse. Therefore, according to this research, separating classrooms by gender does more harm than good in terms of the impact on educational outcomes of the whole student body.

However, academic performance is not the only important outcome that could be affected by gender composition of a classroom. Whether this effect is positive or negative, for a thorough evaluation of gender separation policy, policymakers should also take into account developmental and social issues and investigate the specific consequences for each aspect. If single-sex education ends up having positive impact on academic performance but negatively affecting social and emotional development of students, decision makers who opt to implement gender separation policy, should seek for additional policies and plans to develop students’ social skills.

Application of the findings should be made bearing in mind that they are generated from a specific context;

First of all, K-12 education is completely separated by gender in Iran. Although the results are in line with previous findings in western countries, the current paper does not intend to extrapolate the estimated effects to such different contexts. After all, the mechanisms could vary. For instance, university students who were participating in mixed classrooms in all school years might not feel uncomfortable to express their ideas and involve in class discussions in the presence of the opposite sex. Therefore, the results of this research are mainly applicable in the context of the countries and societies with single-sex schools, particularly in Muslim-majority countries. In some western countries such as the UK, the U.S., Australia, Ireland, and Canada, that single-sex schooling is rather prevalent, the results could also be applied, with extra concerns about cultural norms and values.
Secondly, The University of Allameh is mainly specialized in humanities and language studies, fields that are mostly approached by women in Iran. The share of male students in none of the fields exceeded 40%.\(^{11}\) However, the effect of the policy could vary among these categories. The negative impact of mixed classrooms on females’ outcomes could disappear when they study male-dominated or so-called STEM subjects besides male counterparts as the cooperation among genders in assignments and educational tasks could help women improve their GPA in mixed classrooms. One could contradict the previous argument by claiming that mixed classrooms in STEM fields are more harmful to female students as they are more likely to shy away and refrain from being closely involved in class discussions due to a higher sense of stereotype threat. The current research does not refute or confirm either of the two statements based on the findings generated from a different domain.

To sum up, although in this paper the policy of gender separation at universities turned out to positively affect educational outcomes of female students, the question whether the Iranian government should widen the scope of the policy in terms of the number of public single-sex universities or mixed universities with separated classrooms, is still open. Future studies with additional data on various aspects need to evaluate the impact of the policy in the long term, and also on key social outcomes such as the average age of marriage, rate of divorce, time to find a job, wage, and life satisfaction of separated versus mixed university students.

\(^{11}\) - This is the reason that I did not conduct moderator analysis based on the type of the fields (male-dominated versus female-dominated). Based on the public perception of the majors and also the mostly used threshold of two-third (see for example Jackson, 2012) to define male-dominated fields, I only found one field, “Statistics and Mathematics”, to be regarded as male-dominated. However, the number of students in this fields were not sufficient (Only 27 individuals) to derive meaningful conclusions.
Bibliography


Mixing or Separating Genders in Higher Education?
Evidence from a Natural Experiment in Iran