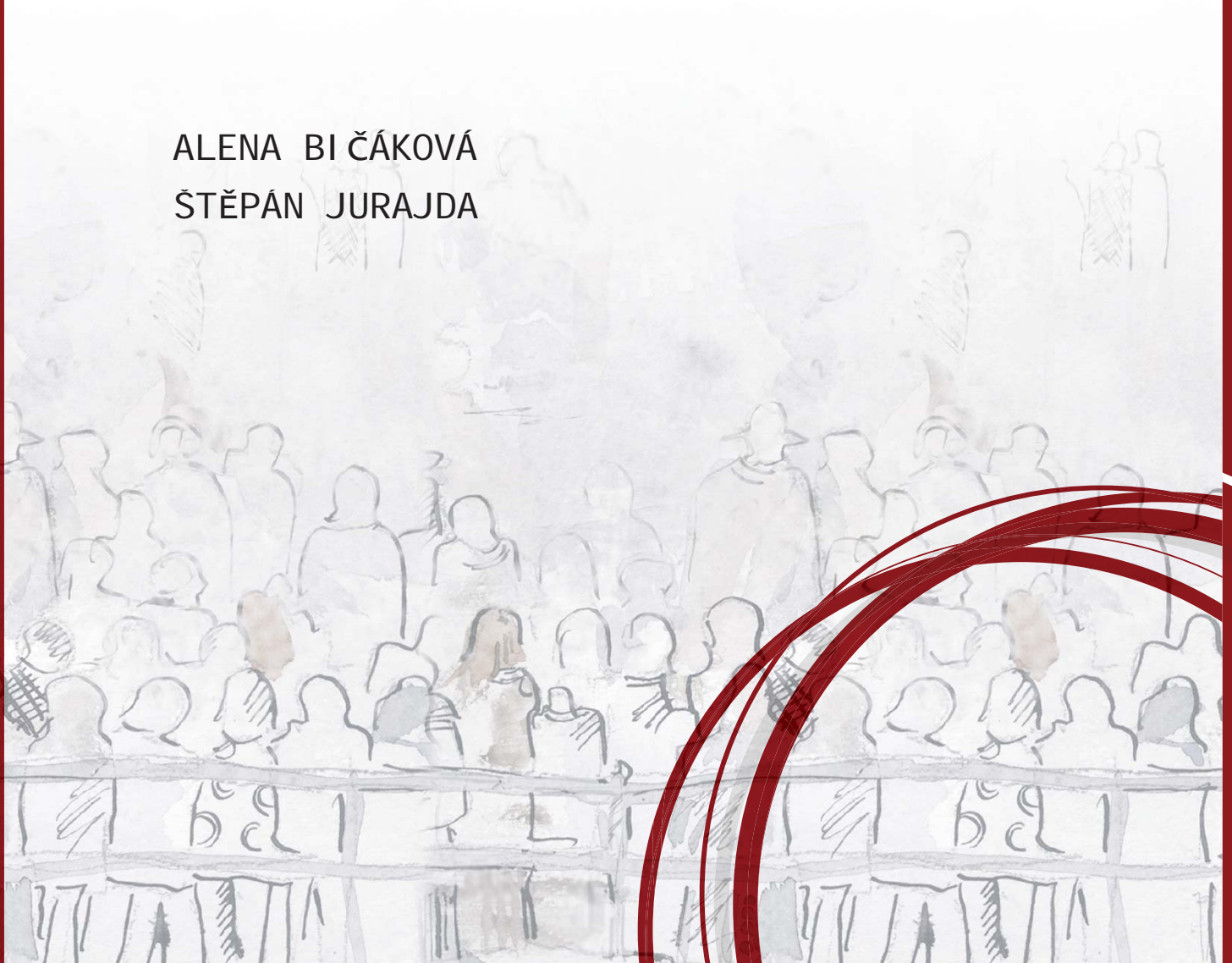# THE QUIET REVOLUTION AND THE FAMILY: GENDER COMPOSITION OF TERTIARY EDUCATION AND EARLY FERTILITY PATTERNS

May 2014

ALENA BIČÁKOVÁ

ŠTĚPÁN JURAJDA

The Quiet Revolution and the Family: Gender Composition of Tertiary Education and Early Fertility Patterns

CELSI Discussion Paper No. 22

May 2014

Alena Bičáková

CERGE-EI, Charles University and the Academy of Sciences

Štěpán Jurajda

CERGE-EI, Charles University and the Academy of Sciences

ABSTRACT

The Quiet Revolution and the Family:
Gender Composition of Tertiary Education and Early Fertility Patterns

It is well known that highly 'female' fields of study in tertiary education are characterized by higher fertility. However, existing work does not disentangle the selection-causality nexus. We use variation in gender composition of fields of study implied by the recent expansion of tertiary education in 19 European countries and a difference-in-differences research design, to show that the share of women on study peer groups affects early fertility levels only little. Early fertility by endogamous couples, i.e., by tertiary graduates from the same field of study, declines for women and increases for men with the share of women in the group, but non-endogamous fertility almost fully compensates for these effects, consistent with higher early fertility in highly 'female' fields of study being driven by selection of family-oriented students into these fields. We also show that the EU-wide level of gender segregation across fields of study has not changed since 2000, despite heterogeneous country-level evolution.

Keywords: Field-of-Study Gender Segregation, Tertiary Graduates, Fertility

JEL Classification: I23, J13, J16

Corresponding Author:

Štěpán Jurajda

CERGE-EI, Charles University Prague and Academy of Sciences of the Czech Republic
Politickych veznu 7, Prague, Czech Republic.
e-mail: stepan.jurajda@cerge-ei.cz

1 Introduction

The worldwide expansion of tertiary education in recent decades has coincided with a dramatic increase in the share of female students, reversing the historic male advantage in higher education. This 'quiet revolution' (Goldin, 2006) has received much attention, but little is known about the gender composition of education expansion across fields of study even though it is generally recognized that gender segregation by fields of study is the main remaining axis of gender differences in tertiary education across developed countries (Charles and Bradley, 2002).

A separate line of evidence suggests that the gender composition of peer groups affects marriage and fertility patterns (e.g., Angrist, 2002). Since assortative mating based on education is strong (Bruze, 2011; Bredemeier and Juessen, 2013; Greenwood, Guner, Kocharkov and Santos, 2014) and since a major part of the first births of the tertiary educated takes place within a few years of graduation (Herr, 2012), one may expect gender segregation at college to affect fertility outcomes. Yet, existing studies of this issue are based on cross-field comparisons such that they cannot disentangle the selection-causality nexus. It is typically found that highly 'female' fields of study are characterized by higher fertility (van Bavel, 2010), but whether this is due to selection of family-oriented students into these fields or whether it is the consequence of the gender mix in one's field of study remains unknown.

We begin to fill these two gaps in the literature by combining population statistics on the gender composition of eight fields of study in higher education with over a decade of harmonized labor force survey data on the early fertility of tertiary graduates from nineteen EU countries. First, we track the evolution of the representation of women, and we measure the implied changes in gender segregation by field of study. Second, we quantify the importance of the gender composition of field-of-study groups for fertility and its endogamy structure, corresponding to whether child births occur to couples graduating from the same field of study or not. We do so, unlike the existing work (in demography and sociology), by employing panel data, which allow us to invoke conservative identification assumptions.

1

2   Related Literature

Although historically men accounted for the vast majority of tertiary education degrees, things have changed thanks to the worldwide boom in higher education that started around 1970. Since 2005, females represent over 55% of enrollment in US degree granting programs.[1] According to Eurostat 2011 data, women make up the majority of university graduates in all EU countries with the exception of Liechtenstein, and similar increases in female representation in higher education have been achieved all over the world (Becker, Hubbard and Murphy, 2010).

There is a growing, mainly US literature asking about the driving forces behind this secular trend. While Jacob (2002) finds that the rise in the female share of college graduates can be largely explained by the lower costs of attending a college for women due to the gender difference in non-cognitive abilities, Hubbard (2011) suggests that there are no major differences across the two genders in their returns to education. Goldin, Katz, and Kuziemko (2006) highlight the importance of high-school study achievements and work expectations[2] while Ge and Yang (2013) point to rising divorce probabilities as an explanation for the reversal of the gender gap in college attainment.

As women now form the majority of higher-education graduates in developed countries, gender segregation across fields of study represents the main remaining axis of gender differences in tertiary education (Charles and Bradley, 2002). There is much discussion in sociology and economics about the sources of gender segregation by field of study (e.g., van Bavel, 2010) or, relatedly, by occupation (e.g., Dolado, Felgueroso and Jimeno, 2002). Mastekaasa and Smeby (2008) and Morgan, Gelbgiser and Weeden (2013) are examples of studies that attempt to quantify the importance of individual-level determinants of the gender-specific

---

[1]This share corresponds to all degree granting institutions and is similar for full-time and part-time students. Source: U.S. Department of Education, NCES, Digest of Education Statistics, 2012, Table 230.

[2]McDaniel (2010) asks about the role of education expectations of 15-year-olds across several countries.

choices of college major in the EU and the US, respectively.[3] Charles and Bradley (2009) use cross-country comparisons to ask about the importance of cultural gender beliefs, education systems, and labor markets for gender segregation by field of study.

However, there is surprisingly little evidence available to answer the basic question of whether the historical increase in the share of female tertiary graduates has been accompanied by a decline in segregation by field of study or not. The two major sociological studies of international differences in college field-of-study gender segregation, Charles and Bradley (2002, 2009), both use data from the 1990s and offer only cross-country comparisons. We know of no studies tracking the recent evolution of gender segregation in college across a wide set of countries.[4]

There is a separate, large literature, starting with Becker (1973), studying the implications of the gender composition of peer groups for marriage and fertility patterns. Angrist (2002) studies the natural experiment of the flow of US immigrants from different nationality groups at the beginning of the 20th century and finds that a higher ratio of men to women in a group increases the likelihood of female marriage, consistent with an increase in female bargaining power in the marriage market. Svarer (2007) and McKinnish (2007) focus on the gender composition of the workplace and present evidence suggesting that those who work with

---

[3]Beffy, Fougère and Maurel (2012), who do not focus on gender differences in detail, imply that in the French context, expected earnings are not a driving force of tertiary field-of-study choices. There is also a growing literature on gender differences in the college admission process (e.g., Jurajda and Münich, 2011).

[4]Charles and Bradley (2002) document gender composition by field of study for 12 industrialized countries while Charles and Bradley (2009) provide cross-country comparisons covering 44 countries. Similar to our study, Charles and Bradley (2009) use UNESCO population data on the gender composition of tertiary graduates, but due to their extensive country coverage, they are able to distinguish only four fields of study (we distinguish eight fields), and they do not track the evolution of segregation over time (while we do so). Barone (2011) is the only study we are aware of to follow the gender composition of higher education by field of study over time, but he does so only for four countries, and he uses measures based on the EU Labor Force Survey such that sampling error is a major concern in smaller fields. Barone also provides cross-country (but no time) comparisons for eight EU countries based on the small Reflex survey, where response rates varied across countries from 20 to 42 percent.

a larger fraction of workers of the opposite sex are more likely to get divorced. Lafortune (2013) highlights the reaction of pre-marital investments to varying gender shares.

Parenthood decisions by the tertiary educated are concentrated after graduation. For example, Herr (2012) uses the NLSY79 data to show that in the US about half of first births by tertiary educated women occur within 6 years of graduation. There is also a strong tendency towards matching of couples based on similar educational attainment (e.g., Bruze, 2011). These tendencies are clearly visible in our EU data as well (see Section 7). The field-of-study peer group is thus likely to represent an important pool of candidates for matching into parenthood, and the gender mix at college may have implications for marriage and fertility. In this paper, we study the timing of first childbirth (parenthood) separately for both genders.[5]

Demographers and sociologists have studied the issue: Van Bavel (2010) measures the relationship between field of study and the postponement of motherhood in 21 European countries using about three hundred graduates per country from the 2004 round of the European Social Survey. His results suggest that a higher share of women among graduates is related to less postponement, i.e., earlier female fertility. However, since his analysis is based on a cross-section of fields of study, it cannot disentangle the causality-selection nexus, which may explain the contrast between his and Angrist's findings.[6] Surprisingly, there is no work in economics examining the effects of gender segregation across college fields of study for marriage, fertility, or labor market outcomes.[7] Stevenson and Wolfers (2007) notice the gap and propose that future research is needed on these topics.

----

[5] Although we focus on parenthood, we build the interpretation of our fertility estimates in part based on exploring the corresponding marriage/cohabitation patterns.

[6] Van Bavel (2010) also discusses the few earlier published papers studying this question, which are all based on a one-country research design. Begall and Mills (in press) and Michelmore and Musick (in press) are recent additions to this one-country work based on cross-sectional field-of-study comparisons employing Dutch and US data, respectively.

[7] There is much research on the effect of schooling on fertility, e.g., Amin and Behrman (2014), but this work does not ask about the fertility effects of gender composition of study groups within education levels.

We follow their suggestion and begin to fill these gaps by providing evidence on the changing gender structure of fields of study and by asking about its effect on early fertility.[8] Unlike van Bavel (2010), we rely on panel data tracking fields of study in several countries over time, and we can thus employ a difference-in-differences identification strategy. Existing sociological work focuses on cross-sectional interpretations for the finding of a positive relationship between fertility and the share of women in a study group: highly 'female' fields of study have more supportive work-family culture in subsequent highly female occupations, often in the public sector, or these fields of study attract women and men with strong pro-family preferences, and there is a higher share of such women than men in the population leading to a higher share of women in these fields.

In contrast, our difference-in-differences research design allows us to generate causal evidence on the effect of the gender mix in one's field of study on one's early fertility decisions conditional on the stable part of family-preference differences across fields of study. Our identification strategy, explained in detail in Section 6, controls for the country-specific evolution of pro-family preferences (by gender) among college graduates by conditioning on country-year fixed effects, but we assume that changes in the gender composition of fields of study occurring during the recent education expansion and experienced by our sample countries since 1998 did not systematically affect the sorting of men and women into fields of study based on their pre-tertiary-education family formation preferences. This is a strong assumption and we return to discussing it in the concluding section. One could argue that women and men choose their fields of study based on the changing marriage prospects related to the changing gender composition of these fields. Future research is needed on this issue.[9]

---

[8] In a related line of research, Fletcher, Ross and Zhang (2013) and Arcidiacono, Aucejo, Hussey and Spenner (2013) study the implications of the demographic and academic background composition, respectively, of cohorts of US students on friendship formation.

[9] There is work studying the effect of marriage expectations on the decision to attend and graduate from college (e.g., Ge, 2011), but none of it focuses on the choice of the field of study.

3 Theoretical Considerations

That women marry faster and have children earlier when the share of men on the marriage market is high has been recognized at least since Becker (1973). Mortensen (1988) provided a search theoretical perspective for such patterns. Having a larger pool of potential partners increases one's chances of matching with a partner of desirable qualities, and the general prediction of matching theory is the short(er) side of the market is better off. Whether this implies higher fertility depends on gender-specific family-formation preferences. If women have stronger preferences for having children than men and if the female bargaining power is high when such is the share of men on the marriage market (as suggested by, e.g., Angrist, 2002), this will lead to higher fertility.

In contrast, increasing the share of women in a predominantly female marriage market (peer group) may lead to two opposing effects on male fertility choices: First, as the share of women increases, so may the heterogeneity of characteristics of the female pool of potential partners, which may allow men to find a highly desirable partner. On the other hand, men may realize that the costs of drawing another match is lower and may be less willing to commit to parenthood. Further, if the share of either gender on the peer group is close to one, the potential for endogamous (within-group) matches is small. This will likely imply a search strategy focused on out-of-group search. In our case, women in highly 'female' groups of graduates and men graduating from almost fully 'male' fields of study are likely to look outside of the group to compensate – to find non-endogamous partners – with bargaining power implied by cohort-wide gender shares.

The shape of the relationship between fertility and gender shares in tertiary education are thus likely to be different for men and for women and may also be non-linear, i.e., exhibiting a different slope on either side of the 50 percent share and also towards the extreme ends of the gender share range. Unlike existing analyses of the effect of field-of-study gender segregation on early fertility (discussed in Section 2), we therefore distinguish endogamous and non-endogamous fertility and use a non-linear empirical model (in Sections 6 and 7).

4 Data Sources

To measure the gender composition of tertiary-level graduates by field of study, we use country population statistics for ISCED education levels 5 and 6 compiled by the Data Centre of the UNESCO Institute for Statistics (UIS). The data cover eight fields of study (Education, Humanities, Social Sciences, Science, Engineering, Agriculture, Health, and Services) for 29 European countries from 1998 to 2010. After excluding countries with less than 6 years of information, the data allow us, in Section 5, to describe the recent evolution of gender segregation across fields of study in 23 EU countries using more than two thousand country-year-field of study group observations.[10]

Next, we merge the UIS data on the gender composition of fields of study groups with information from the EU Labour Force Survey (LFS) on tertiary graduates and their fertility in 19 EU countries. Specifically, we employ the 2012 release of the EU LFS covering reference years 2003 to 2011, when information on field of study is available in the data, and use information on individuals with ISCED education levels 5 and 6 who graduated between the ages of 20 and 44. The share of sampled individuals with missing data values generally does not exceed 5% in any of the country-year data cells. Fertility information for the sampled individuals is merged with UIS population statistics on the gender composition of tertiary education at the level of graduation year, country, and field of study, and the merged data are used in Section 7 to study the effect of the gender composition of one's group on fertility.

Several features of our data deserve to be mentioned. First, this appears to be the first data combining extensive panel information on the gender composition of fields of study with marriage and fertility measures. Second, we use data on graduates, which means that

---

[10]The Data Appendix provides details on all our data sources and procedures, in particular on how missing data cells were treated in our analysis. We also note that the same data submitted by the national authorities to UNESCO, which harmonizes these statistics over time, are also submitted to the OECD. However, as the data department of the OECD (OECD.Stat) confirmed in private conversations with us, the OECD data should not be used to track the time evolution of segregation by field of study since the OECD makes no attempt to harmonize statistics submitted by national authorities over time.

observed patterns of sex segregation reflect gender differences in both initial choices of field of study and in the completion rates (Alon and Gelbgiser, 2011). Third, unlike in, for example, Wei and Zhang (2011), our gender shares are not estimated off survey data and are therefore not affected by sampling error that could lead to attenuation biases in regression estimation (Aydemir and Borjas, 2011).

Our analysis could be affected by a different source of measurement error, however, namely by potential differences in the coding of education in the two data sources. To verify that both data collection efforts use the same coding of education fields, we have correlated the UIS population shares of women in each year-country-field cell with those measured with sampling error in the LFS. The correlation (measured at the year-country-field cell level) is 0.97. When we measured this correlation separately for our eight fields of education, the small field of Services (800) was a clear outlier with a correlation of only 0.33 (the median correlation across the other fields being 0.83). In a robustness check, we therefore drop this field from the regression analysis.

5   Segregation Evolution

During the first decade of the 21st century, the share of women in higher education graduates has increased in most, but not all, of the 23 EU countries covered by the UIS data, as Fig. 1 attests.[11] The share of female graduates varies from a low of 40 percent in Switzerland in the early 2000s to about 70 percent in the Baltic countries recently. While the share rose by about ten percentage points in seven countries including, e.g., Germany and Slovakia, it declined in Portugal and changed little in France, Spain, or the UK. The overall share of female graduates across all of these countries rose from 55% to 58% between 2000 and 2010.

As shown in Fig. 2, the increasing share of female graduates means that most fields of study with above-parity shares of women (at the level of all of our 23 countries) as of 2000

---

[11]The figure is based on 2,208 country-year-field observations, of which 220 missing values and outliers were imputed using neighboring years. See the Appendix for details.

Figure 1: Percentage of Women in Tertiary Graduates by Country

Note: Graduates with tertiary-level education (ISCED level 5 and 6) from eight fields of study (see Fig 2).
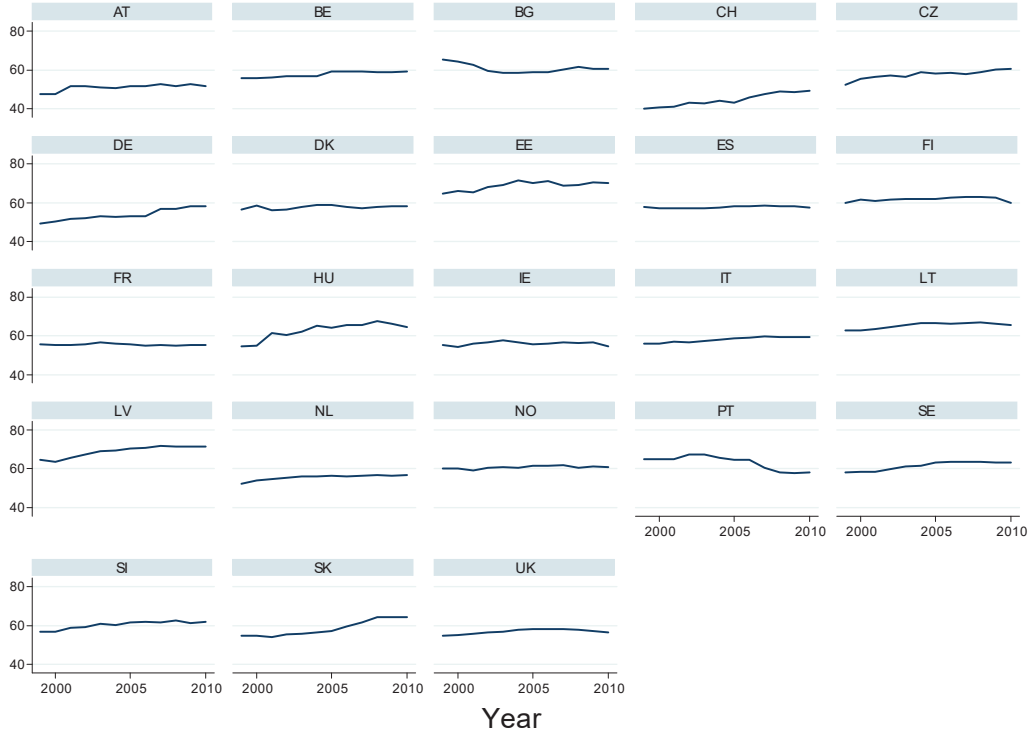
have actually witnessed further increases in their share of female students. Among fields with low initial shares of women, Engineering and Agriculture move toward parity while no such change occurs in Science. These simple statistics suggest segregation across fields of study is not generally declining.

To quantify the change in the extent of higher-education field-of-study gender segregation at the country level, we apply the widely used Duncan segregation index (Duncan and Duncan, 1955). For a given country c and year t, the index is defined as follows:

$$D_{ct} = \frac{100}{2} \sum_{f=1}^{8} \left| \frac{M_{fct}}{M_{ct}} - \frac{F_{fct}}{F_{ct}} \right|,$$

where M_{fct} denotes the number of males in a field of study f, F_{fct} is the corresponding

9

Figure 2: Percentage of Women in Tertiary Graduates by Field

Note: The share of women by field on the total sum of graduates from the 23 EU countries of Fig 1.
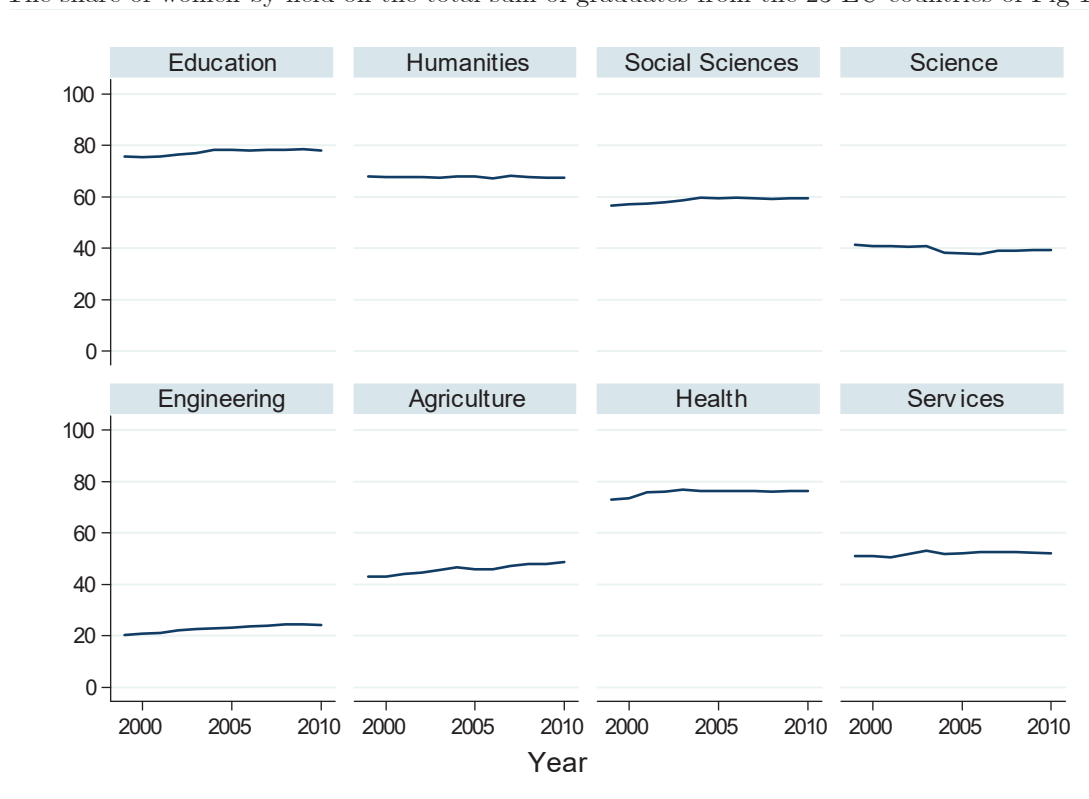
number of females in a group, and where M_{ct} and F_{ct} represent the total number of males and females in higher education, respectively. The index can be interpreted as reflecting the percentage share of the total body of graduates that would have to change the field of study in order to equalize the gender composition across fields; it ranges between 100 (complete segregation) and 0 (complete integration).

Figure 3 shows Duncan segregation index values calculated for each country in 2000 and in 2010 (or in the nearest available year) against the 45-degree line. It implies that field-of-study higher-education segregation changed by over 5 percentage points in seven EU countries. It declined by about 10 percentage points in Switzerland, where the share of women increased from a very low initial level, and in Denmark, where it stagnated at a high level of about 60 percent. The other three countries experiencing sizeable declines in segregation were similar

Figure 3: Field-of-Study Gender Segregation, 2000 and 2010

Note: The Duncan segregation index in 2010 (2009 in BE, FR, PT, SI and 2008 in IT) and 2000 (1999 in IE, IT, SI and 2001 in HU, UK).
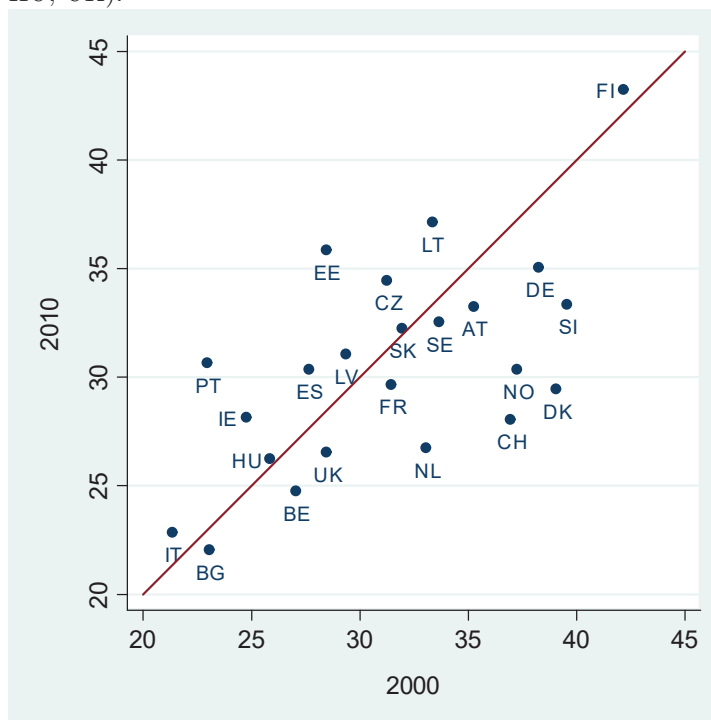
to Denmark in this respect. On the other hand, segregation increased rapidly in Portugal, where the share of women declined, and in Estonia, where it kept on increasing from a high initial level. The other two Baltic countries where the share of women also grew much above 60 percent also experienced increases in segregation. These statistics suggest that improvements in segregation are achieved by either increasing the share of women from low levels or by altering the allocation of students to fields of study once a high share of women in aggregate has been achieved, and that very high levels of the female share on tertiary education come at the cost of women increasingly enrolling in highly 'female' fields of study.

While there is clearly much heterogeneity in segregation evolution at the country level, Figure 3 also implies that there has been little change in the extent of segregation at the EU level over this decade. Summarizing the segregation index changes across the 23 countries

using an average of country-specific index values weighted by the country-specific number of tertiary-level graduates, the overall extent of field-of-study gender segregation changed from 30% in 2000 to 29% in 2010.

Our next goal is to exploit the rich variation in the country-specific changes in the gender composition of fields of study to ask whether it can explain early fertility patterns among recent tertiary graduates. In the next section, we introduce the econometric specifications used in the fertility analysis, which is presented in Section 7.

6   Econometric Methodology

We use individual-level data to measure the impact of the changing gender composition of fields of study on early fertility choices of tertiary graduates. Yet, the mechanism we test for operates at the group level defined by country, field of study, and graduation year. In order to reflect the group-level nature (degrees of freedom) of the estimation, we follow the simple two-stage procedure suggested by Donald and Lang (2007) separately for each gender.[12]

In the first step, we aggregate the individual-level EU LFS fertility data to the relevant country-field-graduation year group level whilst controlling for any age compositional and survey year effects, thus abstracting in our analysis from any EU-wide trends in fertility as well as potential survey-round-specific data issues. Specifically, we aggregate individual fertility measured as the presence of first childbirth at the most one year prior to and within five years after graduation[13] using the following least squares regression

$$y_{icfts} = \omega_{cft} + \beta age_{icfts} + \gamma_s + \varepsilon_{icfts}, \tag{1}$$

---

[12] All of our regression analysis is conducted separately for each gender. However, to simplify the exposition of the analysis, we drop the gender subscript from the regression specifications presented in this section.

[13] We do not observe the exact age of the child, only a three-year age range and use the mid-point of the age interval as our age proxy to compare with the year of graduation. This implies that a small random part of the children we classify as born within five years after and one year before graduation are, in fact, born in the two years adjacent to this interval.

where $y_{icfts} = 1$ if a first child was born to individual i from country c, who graduated from field of study f in year t, and who was interviewed in a survey conducted in year s and $y_{icfts} = 0$ otherwise; ω_{cft} are the country by field of education by year of graduation fixed effects, age_{icfts} is the age of the individual at graduation, and γ_s are the survey-year fixed effects.[14] As the outcome variable is censored for individuals who graduated less than five years ago, we focus only on individuals who are at least 5 years out of school such that the last graduation year in our estimation-ready sample is 2006.

In the second stage, we relate the fertility aggregates from the first stage, i.e., the estimated fixed effects $\widehat{\omega_{cft}}$, to the share of women and other control variables measured at country by field of education by graduation year level. The theory discussion in Section 3 implies that the share of women may have a non-linear effect on early fertility. We therefore estimate the second-stage relationship using a semi-parametric regression in which the key variable of interest – the share of women on a field-of-study group – enters non-parametrically. We also control (parametrically, using a linear specification) for the size of the group (the total number of graduates) and a set of fixed effects corresponding to our identification strategy.[15]

Our country-year-field of study panel data allow us to impose conservative identifying assumptions. We can ask about the effect of the gender composition within the field of study on early fertility whilst assuming that the sorting into fields of study based on pro-

---

[14]The regression is estimated without a constant (with a full set of fixed effects) and with age demeaned.

[15]We use the `semipar.ado` command in `Stata` developed by Verardi and Debarsy (2012), which implements the double-residual estimator proposed by Robinson (1988). Robinson's estimator starts by partialing out of both sides of the equation the effects of the variables entering linearly and, in the second step, uses kernel-weighted local regressions to estimate the non-parametric function using the residuals from the first stage. The preferred estimator for panel data applications, `xtsemipar.ado` in `Stata` developed by Libois and Verardi (2013), implements the Baltagi and Li (2002) extension of the Robinson estimator for panel data applications. We cannot use `xtsemipar.ado` as it does not allow for multiple sets of fixed effects. However, in the one specification we estimate below without field fixed effects (with only country-year fixed effects), we obtain results based on `xtsemipar.ado` that are almost identical to those based on `semipar.ado`.

family preferences is the same across our 19 EU countries but allow for country-specific time evolution of fertility, which could be potentially correlated with the country-specific time evolution of the share of women in tertiary education. Alternatively, we can assume that there are no country-specific time shocks to fertility (on top of the EU-wide common evolution, which we control for in equation (1)) and allow for country-specific sorting into fields of study based on pro-family preferences that could be potentially correlated with field-of-study differences in the representation of women across countries. The two alternative specifications corresponding to these two identification assumptions are presented in the following two equations: In equation (2), we condition on country by field of education (δ_{cf}) and graduation year (δ_t) fixed effects, while in equation (3), we use country by graduation year (δ_{ct}) and field of education (δ_f) fixed effects:

$$\widehat{\omega_{cft}} = \alpha_0 + \alpha_1 ln(graduates_{cft}) + \delta_{cf} + \delta_t + \Gamma(fshare_{cft}) + \nu_{cft}, \tag{2}$$

$$\widehat{\omega_{cft}} = \alpha_0 + \alpha_1 ln(graduates_{cft}) + \delta_{ct} + \delta_f + \Gamma(fshare_{cft}) + \nu_{cft}. \tag{3}$$

In both equations, $ln(graduates)$ is the logarithm of the total number of graduates in a given country, field of education, and graduation year; $fshare_{cft}$ is the share of women among graduates in a given country, field of education, and graduation year; and where $\Gamma(.)$ is an unknown function estimated non-parametrically.

Equations (2) and (3), together with equation (1), can be thought of as corresponding to a production function aggregating the number of men and women in a group into a single factor affecting fertility. We measure the gender composition of peer groups using the share of women as opposed to the sex ratio—the ratio of men to women. Angrist (2002) is a prominent example of a study of matching on marriage markets that uses the sex ratio. He analyzes situations where the share of men and women is not too far from balanced. In his specification of a matching function, the logarithm of the sex ratio can thus be approximated with a linear term. Given the wide variation in gender shares across the field-of-study groups presented in Section 5 and the theoretical arguments supporting non-linear effects, this strategy is not attractive in our case. Studies of workplace segregation effects, which

14

also work with variation in the share of women that ranges almost from 0 to 1, also typically condition on the share of women, not on the sex ratio.[16] Similar to Angrist (2002), we also condition on the logarithm of the group size to allow the probability of finding one's preferred match to depend on the pool of potential partners, as in any standard matching function.

7 Fertility Analysis

In Section 5, we illustrated that there is much variation in the gender composition of peer groups across the three principle dimensions of our data, namely year, country, and field of study.[17] In this section, we ask about the potential effect of this variation on early fertility patterns. We measure early fertility based on the presence of first childbirth at most one year prior to and within five years after graduation. In our EU LFS data covering nineteen countries, 60% of children born to a parent with tertiary education who graduated in 2000 were born within this interval.[18] Our sample countries differ dramatically in the evolution of early fertility of tertiary graduates. In three (five) out of the nineteen countries covered by the merged UIS-LFS data, female (male) early fertility rates have changed on average by more than 1.4 of a percentage point a year with several countries experiencing strong growth and several others sizeable declines.[19]

---

[16]See, e.g., Svarer (2007) and McKinnish (2007) for studies of divorce patterns and Macpherson and Hirsh (1995) or Baker and Fortin (2001) for work on the gender wage gap. There are also matching studies that use the shares of demographic groups as their main explanatory variable (e.g., Fletcher et al., 2013).

[17]Figures 6 and 7 in the Appendix document that much of this variation is independent of the increasing size of higher education. In other words, there is variation in gender shares conditional on the growing size of the field-of-study peer groups.

[18]This share is similar to that found by Herr (2012) for the U.S. Similar to, e.g., van Bavel (2010), we exclude from this calculation, and our subsequent analysis, those who already had children before our fertility window, in which one can expect fertility decisions to be affected by gender composition of peer groups.

[19]That fertility rate evolutions differ dramatically follows both from estimated time trends for each country and gender and from the corresponding 2010-2000 change in fertility.

Our goal in this section is to find out the extent to which the positive association between the 'femaleness' of a field-of-study group and fertility uncovered in earlier work is driven by an impact of the gender shares in tertiary education peer groups on fertility. Specifically, to provide insight into this nexus, we estimate specifications (2) and (3), which are non-nested with respect to each other such that we cannot formally test their validity. Given the first-order differences in country-specific fertility trends, our preferred specification corresponds to equation (3), which allows for country-specific evolution of fertility. We therefore ask how robust the positive association between fertility and gender shares is to conditioning on the stable part of the selection of students with strong pro-family preferences into highly 'female' fields of study. Within robustness analysis, we compare the key estimated parameters across the two available identification approaches to assess the sensitivity of our results to either the allowing for different pro-family preferences of students in a given field of study in different countries or for different evolutions of overall fertility across different countries.

The (2nd stage) outcome measures correspond to fertility group aggregates (at country-year-field of study level) estimated in equation (1) using data on 92,154 female graduates and 72,795 male graduates from 156 country-reference year LFS samples graduating between 1998 and 2006.[20] We have available 2,103 country-graduation year-field of education data cells: 1,058 for women and 1,045 for men.

In our first analysis, we confirm the earlier findings of a positive relationship between the share of women on a field of study in tertiary education and fertility. This is born out in the two left-side graphs of Figure 4, which rely, in large part, on cross-field of study comparisons as they are based on estimating equation (3) separately for each gender without the field-of-study fixed effects (δ_f). The plotted lines correspond to the estimated non-parametric effects of the share of women in a field-of-study peer group on fertility;[21] they suggest fertility is high in highly 'female' fields of study. In contrast, the right two graphs of the Figure, which are

---

[20] The coefficients for age at graduation from equation (1) are presented in the Appendix Table 2.

[21] These estimated effects would correspond to predicted fertility levels if one were to add the effect of mean groups size and the estimated fixed effect coefficients from equation (3).

Figure 4: Effects of Share of Women in Group on Fertility Conditional on Country-Year and Field-of-Study Fixed Effects

Note: Based on groups (field-of-study by country by year) with at least 10 individuals. The share of women corresponds to one's year of graduation. 95% confidence intervals are plotted together with the non-parametrically estimated effects.
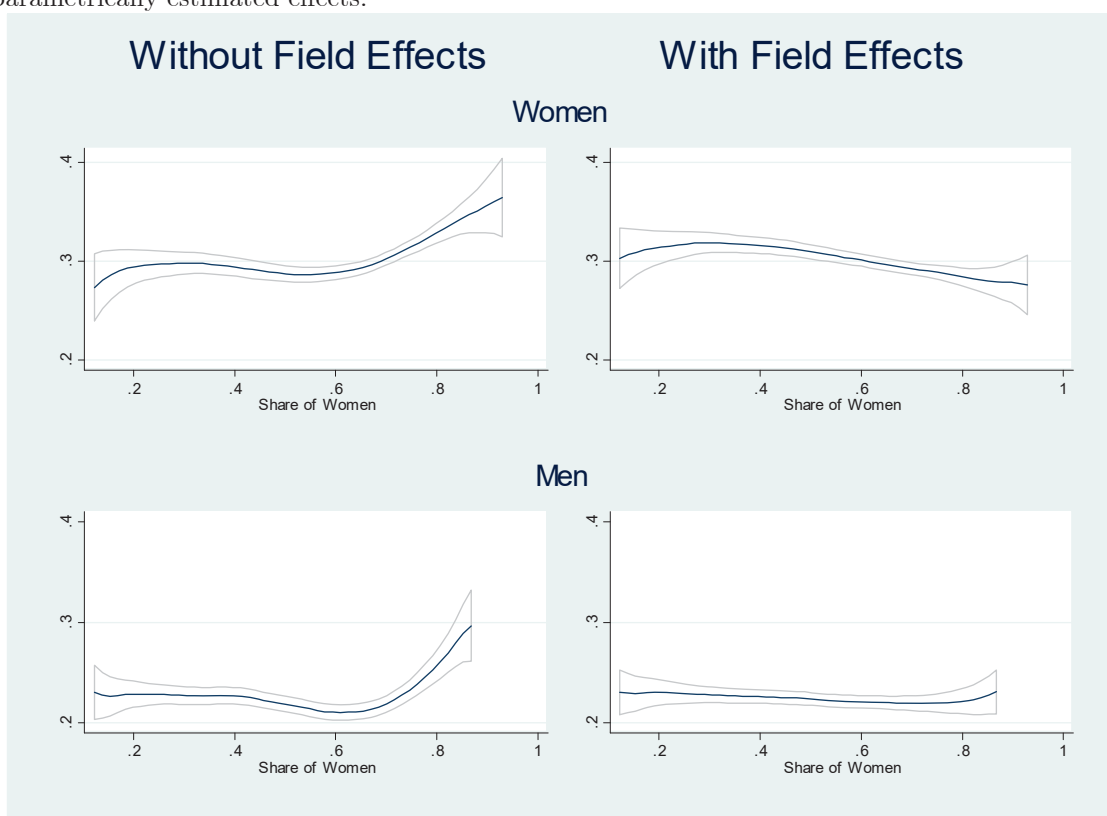
also based on estimating equation (3) but this time with field-of-study fixed effects included, i.e., based on difference-in-differences comparisons, suggest a starkly different picture—one of only a limited effect of gender composition of field-of-study groups on fertility.[22]

Next, we decompose the effects on overall fertility from the two right-side graphs of

---

[22]The top-right graph suggests early female fertility is decreasing as the share of women increases, but the magnitude of this decline is small, and it is not robust to some of the robustness checks we perform below.

Figure 4 into their endogamous and non-endogamous parts.[23] Figure 5, which is also based on estimating equation (3) separately for each gender, plots the estimated non-parametric effects of the share of women in a field-of-study peer group on two types of fertility outcomes: endogamous first childbirth and non-endogamous first childbirth. An endogamous child is a child born to a tertiary-educated couple[24] who graduated from the same field of study within at most five years from each other. A non-endogamous child is born either to a couple where one of the parents is not tertiary educated or to a tertiary-educated couple who graduated from different fields of study or from the same field of study more than five years apart. The summary fertility outcome used in Figure 4 includes not only childbirths identified as endogamous or non-endogamous, but also any children born to individuals without a partner and (the few cases of) individuals with a partner but with missing information about spouses' level of education, field of study, or year of graduation.[25]

The estimates shown in Figures 4 and 5 are based on defining peer groups (pools of potential partners for matching into parenthood) as corresponding to all those who graduated in the same year (in the same field of study and country). Finally, since some of our country-graduation year-education field data cells contain only few individuals such that the corresponding fertility rates are quite noisy, our preferred specifications shown in Figures 4 and 5 are based on a sub-sample of data cells with at least 10 graduates.[26] There are 773

---

[23]We again aggregate fertility of each type by re-estimating the corresponding version of equation (1).

[24]We define partners (couples) using the EU LFS data, which record the presence of "spouses or cohabiting partners in the same household."

[25]Table 1 in the Appendix shows average fertility rates in our data for the 2000 graduation cohort (based on the sample used in the estimation presented in Figures 4 and 5) for the three types of fertility and each gender. About one fourth (sixth) of early fertility is endogamous for men (women).

[26]The standard efficiency-improving procedure would be to weight the second-stage regression either with the inverse of the square-root of the variance of the fixed effects estimated in the first stage (Donald and Lang, 2007) or, preferably, by the exact variance of the outcome variable, which in this case is binomial (Aydemir and Borjas, 2011). However, the latter approach would result in dropping those data cells, small or large, that have fertility rates equal to zero or one. Not only would this lead to a substantial reduction

Figure 5: Effects of Share of Women in Group on Fertility by Type Conditional on Country-Year and Field-of-Study Fixed Effects

Note: Based on groups (field-of-study by country by year) with at least 10 individuals. The share of women corresponds to one's year of graduation. 95% confidence intervals are plotted together with the non-parametrically estimated effects.
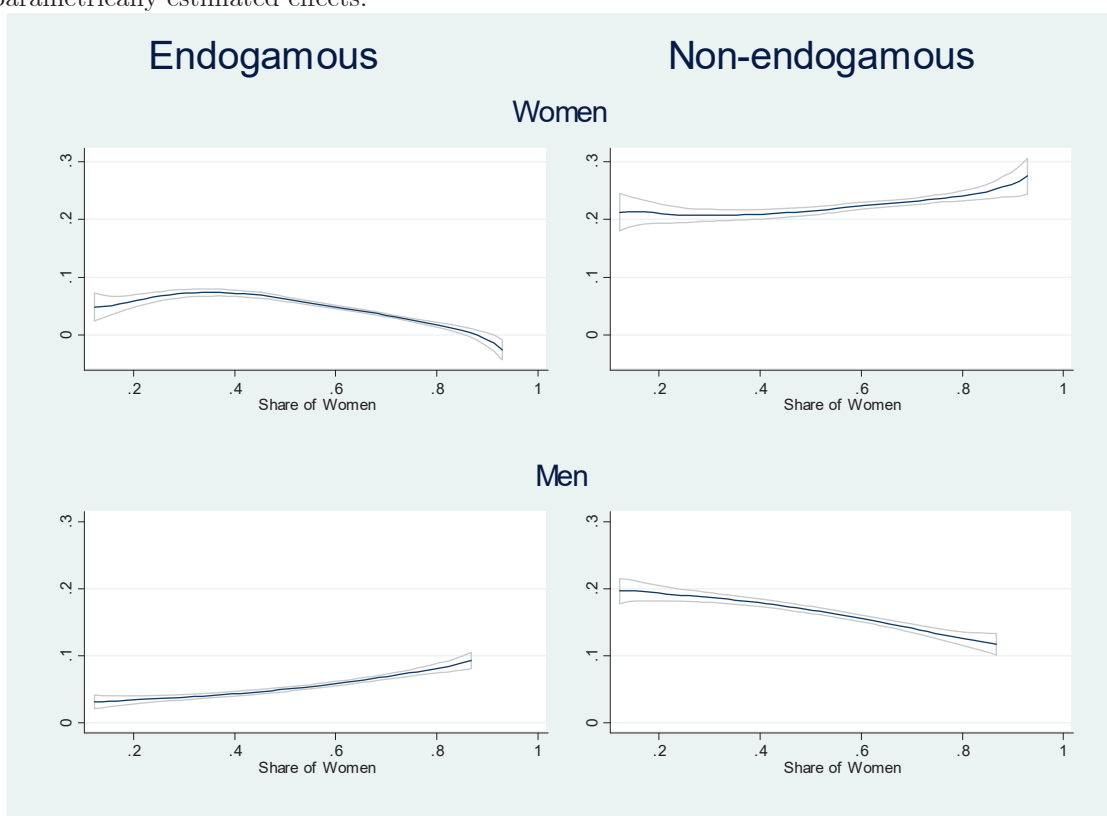
such data cells for women and 717 for men.[27]

The two graphs in the left column of Figure 5 paint a consistent story: For men, endogamous fertility is higher when the share of men is lower, and endogamous fertility of women declines, as expected, when the share of women increases, at least when the share of women

in the sample size, this would induce sample selection based on the outcome variable. Hence, we choose to reflect variance in the outcome measure using cell size only.

[27]Within this sub-sample, the mean (median) size of a data cell is 89 (46) individuals for women and 75 (37) for men.

is above one-third. These 'mirror' effects can be simply explained as corresponding to the availability of potential partners in the group. Interestingly, when the share of women is below one-third, endogamous fertility of women actually increases with the share of women although the estimated fertility slope is noisier here.

If women generally had higher preference for (endogamous) fertility than men, one may expect endogamous fertility of women to increase faster as their share declines to low levels, i.e., as they presumably gain bargaining power, compared to the change in male endogamous fertility corresponding to the shares of men declining to similarly low levels. However, our results are not supportive of this hypothesis. First, when the share of each gender is above one-third, the slopes of the endogamous fertility functions are fully similar across the two genders. Second, endogamous fertility of women is actually not higher when their share is particularly low. It may be that their bargaining power does not increase when their shares are low because most men in the group have focused their search for partners outside of the group. However, if male fertility preferences were lower than female, one would expect the male endogamous fertility effect to be flat when the share of men declines to very low levels, and this is not the case either. On the other hand, the increasing endogamous fertility of men when the share of women approaches 100 percent is consistent with women continuing their search inside group even in almost fully 'female' groups. In short, our evidence is consistent with similar pro-family preferences of men and women and with partner availability being the key driver of endogamous fertility for both genders.

Given the effects of the availability of partners on endogamous fertility and the low impact of gender composition of peer groups on overall fertility shown in Figure 4, it is natural to expect that those faced with a highly gender unbalanced peer group manage to compensate for the lack of suitable potential partners in the group by successfully forming non-endogamous parenthood matches. The right column graphs of Figure 5 bear this out for both genders. Our fertility decomposition thus provides an underlying mechanism for the low impact of gender composition of peer groups on overall fertility shown in Figure 4.

In sum, we find that the high fertility rates in highly 'female' fields of study observed

20

in earlier work are not robust to difference-in-differences comparisons. We uncover only a limited effect of changing gender composition on fertility. For women, overall early fertility is highest when women represent about a third of the group, and fertility may be particularly low in almost fully 'female' groups although the estimates are noisier at both extremes of the 'female' share where there are fewer observations in the data. These effects are small in any case: for women, the maximum/minimum fertility difference (in Figure 4) of about five percentage points corresponds to about one-sixth of the mean overall fertility rate for women in our data.[28]

Next, we perform several types of robustness checks. First, there is little sensitivity to using data cells with more than five or more than fifteen (as oppose to more than ten) individuals.[29] Second, as highlighted in Section 4, there may be measurement error in the assignment of students to field of study in the small field of Services. Hence, we ask whether dropping this field affects the results and find that it has only a small effect. A related issue is whether we may be defining peer groups (pools of potential partners) too broadly in some large fields of study, which may in fact correspond to several effectively distinct sub-fields. We therefore alternatively drop the largest fields of study. Specifically, we order country-field-of-study groups (summed up across all years and normalized by the total number of graduates in a given country) by size and exclude from the analysis the largest decile. Again, we find only limited sensitivity.[30]

---

[28]Other than the share of women in the group, equation (3) contains one other slope parameter: that of the group size. Appendix Table 3 shows the estimated coefficients for the logarithm of the size of each group from specifications corresponding to all three types of fertility. As expected, the group size coefficients are positive for endogamous fertility specifications, where the larger the pools of individuals to be matched, ceteris paribus, the higher the probability that the match occurs.

[29]These results are available in the Appendix Figure 8. All qualitative features of the main results are preserved. Sample size increases to 922 (861) data cells for women (men) when only the cells with fewer than five individuals are dropped, and it declines to 668 (606) data cells for women (men) when we alternatively drop cells with fewer than 15 individuals.

[30]These results are available in the Appendix Figure 9. The number of data cells used in the estimation is

Third, we consider alternative definitions of peer groups in terms of the year of graduation. Our preferred and simplest choice (in Figures 4 and 5) was to pool into a peer group all those who graduated in the same year (and field of study and country). To assess sensitivity to this definition of peer groups, we also define the share of women using a three-year moving window centered around the same graduation year for both genders and, alternatively, using a two-year window shifted forward for men by one year to allow for the fact that within couples, men may be older.[31] None of these alternatives resulted in quantitatively large changes in the estimates for women, but there is now a discernibly higher overall fertility predicted for men in fields where the share of women is high.[32]

Fourth, we apply the alternative identification strategy based on equation (2). Given the dramatically different evolution of fertility across our sample countries (discussed at the start of this section), we attempt to homogenize fertility trends by dropping countries with the highest change (in absolute value) of overall fertility during the sample years: We drop the three countries for women and the five countries for men where the average annual fertility change exceeds 1.4 of a percentage point. The alternative identification approach leads to qualitatively similar results with the exception of a higher overall fertility of women and a lower overall fertility of men in almost fully 'female' groups, where the estimates are noisier.[33] Given the heterogeneity in fertility evolution across sample countries and the general robustness of our preferred specification, we take these results as a confirmation of

---

similar for both of these robustness checks at about 700 for women and 650 for men. Dropping large fields of study works further against the traditional finding of higher fertility in higly 'female' fields of study.

[31] The median age gap of partners with endogamous children in our sample (i.e., in partnerships formed within peer groups) is zero years. Further, about 50 % of both women and men with an endogamous partner have a partner whose age is within one year of their own.

[32] These results, which are presented in the Appendix Figure 10, are, however, based on a substantially smaller number of data cells: 554 for women and 512 for men, i.e., about 30 percent less data compared to our favoured specification. The data loss is the result of the fact that we now need population statistics on the gender composition of graduates for each country and field of study for three consecutive years.

[33] These results are presented in the Appendix Figure 11. The sample size is 661 (624) for women (men).

22

the two main findings, namely of a small effect of the gender composition of field-of-study groups on overall early fertility and of clear 'mechanical' effects on the endogamy composition of fertility.

To assess the magnitudes of the fertility structure implications of our estimates, we use the estimated relationship presented in Figure 5 together with the (UIS) observed change in the gender composition of fields of study between 2000 and 2010 to predict fertility change by type separately for each country. These simulations confirm that predicted increases or declines in endogamous fertility are largely compensated by balancing changes in non-endogamous fertility. The largest country-specific implied change of endogamous fertility (in absolute value) is 0.01 for both women and men, i.e., about one-fourth and one-fifth of the sample-average endogamous fertility rate. Similarly, none of the implied country-specific changes of non-endogamous fertility were higher than 0.01 (in absolute value) for both genders, which means they were only one-twentieth and one-fifteenth of the sample-average non-endogamous fertility rates.[34] In short, the implied changes in fertility are small in magnitude.

8   Discussion

The fertility analysis provided in the previous section (in Figure 5) is consistent with little gender difference in pro-family preferences and with partner availability being the key driver of endogamous fertility for both genders. To provide a check on this interpretation, we supplement the fertility evidence with a brief analysis of couple formation.[35] Estimating equation (3) with marriage/cohabitation replacing fertility as the outcome variable implies that endogamous marriage/cohabitation depends on the availability of endogamous partners similarly as fertility does, only more strongly and with similar slopes for both genders. The estimated effects for non-endogamous marriage/cohabitation and for any form of mar-

---

[34]The country-specific calculations are available upon request.

[35]I.e., we study the incidence and endogamy composition of couples as defined in the EU LFS by the presence of "spouses or cohabiting partners in the same household."

23

riage/cohabitation are again fully consistent with the pattern uncovered for fertility.[36] It is therefore likely that the availability of endogamous partners drives the structure of match formation (couples) and that within couples bargaining power implied by group-specific gender shares plays little additional role for fertility.

We also found that non-endogamous fertility compensates for endogamous fertility when gender shares are highly unbalanced. It is natural to ask, however, whether the structure of non-endogamous fertility depends on the availability of endogamous partners.[37] Parents from the same field of study but more than five years apart in terms of graduation year are responsible for only about 5% of non-endogamous children in our data. We therefore ignore them and ask whether the education composition of non-endogamous parental couples – tertiary with non-tertiary vs. tertiary with tertiary from different field – depends on the share of women in one's field of study. We find that among tertiary educated men who are observed in a non-endogamous parental couple, the share of those who have children with less-than-tertiary educated women is lowest when the gender composition of field of study is close to balanced and it grows at both extremes of gender composition. Similarly, among tertiary-educated men observed in a non-endogamous marriage/cohabitation, the share married to less educated women is lowest in gender balanced fields of study. When men can easily match with endogamous women in gender-balanced groups, those who do not do so are likely to marry and to parent with other tertiary educated women and vice versa. In contrast, among women in non-endogamous parental couples, the share of those who have children with non-tertiary educated men is highest when the share of women in the field of study is high and it is low otherwise. Furthermore, the share of non-endogamous marriage/cohabitation couples of tertiary educated women with less educated men is lower in fields of study where the share of women is high. When there are relatively few men in

---

[36]These estimates are presented in Appendix Figure 12. They are based on the same sample and econometric specifications that was used to generate Figure 5 and the right two graphs in Figure 4.

[37]Our evidence here is complementary to that offered by Mansour and McKinnish (in press), who argue that highly educated individuals are especially likely to marry similarly-aged peers.

a field of study, such that it is difficult to form endogamous matches, women are likely to form couples with educated men from other fields, but they are likely to have children with less educated men, signalling a tradeoff between matching with similarly educated men and having children (with less educated men).[38] In sum, even though the gender composition of tertiary education does not affect overall fertility, it drives the education structure of (parental) couple formation and it does so differently for men and women.

9 Conclusions

Gender segregation by field of study in higher education has not changed much in majority of EU countries since the start of the 21st century despite the continuing expansion of university education driven by an influx of women into universities. In other words, the 'additional' women typically make field-of-study choices that are similar to those made by the earlier smaller cohorts of women in higher education. However, segregation declined dramatically in five countries countries, including Denmark and Norway, and grew strongly in two. It may be that increasing the share of women on tertiary education much above 60 percent comes at the cost of increasing segregation.

We use variation in the changes in gender composition of country-field-of-study groups to ask about the impact of field-of-study gender mix on early fertility of the tertiary educated. Our results based on a decade of data covering 19 EU countries suggest that the effects on fertility levels are generally small with some evidence of somewhat higher fertility for women when their share on a group is close to one-third. However, we uncover significant shifts in the *structure* of fertility driven by one's university peer group composition. Endogamy, i.e., parenthood by couples graduating from the same field of study, is clearly driven by gender shares for both genders. Our findings are not consistent with theories where female fertility preferences are higher than those of men and where women gain bargaining power in the marriage when their share declines to very low levels. Important for the overall

---

[38]The estimates underlying this summary are available upon request.

stability of fertility levels is the fact that non-endogamous fertility almost fully compensates for the varying ability to form endogamous parenting couples except, perhaps, for women in extremely gender unbalanced groups. In sum, while previous cross-sectional comparisons highlighted the higher fertility in highly 'female' fields of study, our evidence, based on a difference-in-differences research design, implies that the effects of gender shares in tertiary field-of-study peer groups on early fertility levels are minor and that most of the cross-field differences in fertility are driven by selection of students with different pro-family preferences and/or by the work-family culture of occupations linked to these fields of study.

Our difference-in-differences approach is based on the assumption that changes in the gender composition of fields of study did not systematically affect the sorting of men and women into fields of study based on their pre-tertiary-education family formation preferences. If some fields, in fact, became more family-friendly over time relative to others and this attracted both more women and more family-oriented types across both genders into such fields, one would expect the differences-in-differences approach to overestimate the effect of the female shares on fertility. It is therefore unlikely that violations of our identification assumption would mask an underlying positive effect of the share of women on field of study on overall early fertility.

The importance of gender composition of peer groups in higher education for endogamy fertility as well as for the education structure of marriage/cohabitation suggests that future research focus on marriage 'quality' and durability across the education endogamy dimension. Assortative mating on educational attainment level is clearly strong (Bruze, 2011), but less is known about the implications of assortative mating on one's field of study (and, presumably, occupation). There are other important avenues of future research. While we extend the study of fertility effects of gender segregation across college field of study, much work remains to be done in this area as noted by Stevenson and Wolfers (2007). In the study of fertility and marriage markets, we focus on parenthood effects within the same education level and abstract from the larger question of how the fact that more women than men obtain a tertiary degree affects the balance on the whole marriage market including the less-than-

tertiary educated. Changing gender shares of tertiary graduates are also likely to affect mating patterns for those with less than tertiary education. Further, there is little work thus far examining the importance of field-of-study gender segregation for labor market outcomes. One recent exception is Lindley (2012) who studies the implications of the gender differences in the labor supply structure due to gender segregation in degree subjects relative to the labor demand evolution driven by technical change, but future work could also ask, for example, about the importance of gender field-of-study segregation for gender differences in youth unemployment rates, thus extending the literature on overall unemployment rates (e.g., Núñez and Livanos, 2010).

Bibliography

Alon, Sigal, and Dafna Gelbgiser (2011) "The Female Advantage in College Academic Achievements and Horizontal Sex Segregation," *Social Science Research*, 40: 107–119.

Amin, Vikesh, and Jere R. Behrman (2014) "Do More-schooled Women Have Fewer Children and Delay Childbearing? Evidence from a Sample of US Twins," *Journal of Population Economics*, 27: 1–31.

Angrist, Josh (2002) "How Do Sex Ratios Affect Marriage and Labor Markets? Evidence from America's Second Generation," *Quarterly Journal of Economics,* 117(3): 997–1038.

Arcidiacono, Peter, Aucejo, Esteban, Hussey, Andrew, and Kenneth Spenner (2013) "Racial Segregation Patterns in Selective Universities," CEP Discussion Paper No. 1219.

Aydemir, Abdurrahman, and George J. Borjas (2011) "Attenuation Bias in Measuring the Wage Impact of Immigration," *Journal of Labor Economics*, 29(1): 69–113.

Baltagi, Badi H., and Dong Li (2002) "Series Estimation of Partially Linear Panel Data Models with Fixed Effects," *Annals of Economics and Finance* 3, 103–116.

Baker, Michael, and Nicole M. Fortin (2001) "Occupational Gender Composition and Wages in Canada, 1987-1988," *Canadian Journal of Economics*, 34(2): 345–376.

Barone, Carlo (2011) "Some Things Never Change: Gender Segregation in Higher Education across Eight Nations and Three Decades," *Sociology of Education,* 84(2): 157–176.

Becker, Gary S. (1973) "A Theory of Marriage: Part I." *Journal of Political Economy,* 82(4): 813–46.

Becker, Gary S., Hubbard, William H. J., and Kevin M. Murphy (2010) "The Market for College Graduates and the Worldwide Boom in Higher Education of Women," *American Economic Review: Papers & Proceedings,* 100: 229–233.

Beffy, Magali, Fougère, Denis and Arnaud Maurel (2012) "Choosing the Field of Study in Postsecondary Education: Do expected Earnings Matter?" *Review of Economics and Statistics*, 94(1): 334–347.

Begall, Katia, and Melinda C. Mills (in press) "The Influence of Educational Field, Occupation, and Occupational Sex Segregation on Fertility in the Netherlands," *European Sociological Review.*

Bredemeier, Christian, and Falko Juessen (2013) "Assortative Mating and Female Labor Supply," *Journal of Labor Economics*, 32(3): 603–631.

Bruze, Gustaf (2011) "Marriage Choices of Movie Stars: Does Spouse's Education Matter?" *Journal of Human Capital*, 5(1): 1–28.

Charles, Maria, and Karen Bradley (2002) "Equal but Separate? A Cross-national Study of Sex Segregation in Higher Education," *American Sociological Review*, 67(4): 573–599.

Charles, Maria, and Karen Bradley (2009) "Indulging Our Gendered Selves? Sex Segregation by Field of Study in 44 Countries," *American Journal of Sociology*, 114(4): 924–76.

Dolado, Juan J., Felgueroso, Florentino, and Juan F. Jimeno (2002) "Recent Trends in Occupational Segregation by Gender: A Look Across the Atlantic," chapter 4 in Argandoña, A. and J. Gual (eds.) *The Social Dimensions of Employment: Institutions, Reforms and Labour Markets,* Edward Elgar: 81–106.

Donald, Stephen G., and Kevin Lang (2007) "Inference with Difference-in-Differences and Other Panel Data," *Review of Economics and Statistics*, 89(2): 221–233.

Duncan, O.T. and B. Duncan (1955) "A Methodological Analysis of Segregation Indexes," *American Sociological Review*, 20: 210–217.

Fletcher, Jason M., Ross, Stephen L., and Yuxiu Zhang (2013) "The Determinants and Consequences of Friendship Composition," NBER Working Paper No. 19215.

Ge, Suqin (2011) "Women's College Decisions: How Much Does Marriage Matter?" *Journal of Labor Economics*, 29 (4): 773–818.

Ge, Suqin, and Fang Yang (2013) "Accounting for the Gender Gap in College Attainment," *Economic Inquiry*, 51 (1): 478–499.

Goldin, Claudia (2006) "The Quiet Revolution That Transformed Women's Employment, Education, and Family," *American Economic Review: Papers & Proceedings*, 96(2): 1–21.

Greenwood, Jeremy, Guner, Nezih, Kocharkov, Georgi, and Cezar Santos (2014) "Marry Your Like: Assortative Mating and Income Inequality," IZA Discussion Paper No. 7895.

Herr, Jane Leber (2012) "Measuring the Effect of the Timing of First Birth," mimeo, NBER.

Hubbard, William H.J. (2011) "The Phantom Gender Difference in the College Wage Premium," *Journal of Human Resources,* 46(3): 568–586.

Jacob, Brian A. (2002) "Where the Boys Aren't: Non-cognitive Skills, Returns to School and the Gender Gap in Higher Education," *Economics of Education Review,* 21: 589–598.

Jurajda, Štěpán, and Daniel Münich (2011) "Gender Gap in Performance under Competitive Pressure: Admissions to Czech Universities," *American Economic Review: Papers & Proceedings*, 101(3): 514–518.

Lafortune, Jeanne (2013) "Making Yourself Attractive: Pre-Marital Investments and the Returns to Education in the Marriage Market," *American Economic Journal: Applied Economics,* 5(2): 151–178.

Libois, Francois, and Vincenzo Verardi (2013) "Semiparametric Fixed-effects Estimator," *Stata Journal,* 13(2): 329–336.

Lindley, Joanne (2012) "The Gender Dimension of Technical Change and the Role of Task Inputs," *Labour Economics*, 19: 516–526.

Macpherson, David, and Barry T. Hirsh (1995) "Wages and Gender Composition: Why Do Women's Jobs Pay Less." *Journal of Labor Economics*, 13: 426–471.

Mastekaasa, Arne, and Jens-Christian Smeby (2008) "Educational Choice and Persistence in Male- and Female-Dominated Fields," *Higher Education,* 55(2): 189–202.

McDaniel, Anne (2010) "Cross-National Gender Gaps in Educational Expectations: The Influence of National-Level Gender Ideology and Educational Systems," *Comparative Education Review*, 54(1): 27–50.

McKinnish, Terra G. (2007) "Sexually Integrated Workplaces and Divorce: Another Form of On-the-job Search," *Journal of Human Resources*, 42(2): 331–352.

Michelmore, Katherine, and Kelly Musick (in press) "Fertility Delay and Childlessness among College Graduates: What Can We Learn from Variation by Field of Study?" *Population Studies.*

Mansourm Hani, and Terra McKinnish (in press) "Who Marries Differently-Aged Spouses? Ability, Education, Occupation, Earnings and Appearance," *Review of Economics and Statistics.*

Morgan, Stephen L., Gelbgiser, Dafna, and Kim A. Weeden (2013) "Feeding the Pipeline: Gender, Occupational Plans, and College Major Selection," *Social Science Research*, 42: 989–1005.

Mortensen, Dale T. (1988) "Matching: Finding a Partner for Life or Otherwise," *American Journal of Sociology*, 94: s215–s240.

Núñez, Imanol, and Ilias Livanos (2010) "Higher Education and Unemployment in Europe: an Analysis of the Academic Subject and National Effects," *Higher Education*, 54: 475–487.

Robinson, P.M. (1988) "Root-N-consistent Semiparametric Regression," *Econometrica*, 56(4): 931–954.

Stevenson, Betsey, and Justin Wolfers (2007) "Marriage and Divorce: Changes and their Driving Forces," *Journal of Economic Perspectives*, 21(2): 27–52.

Svarer, Michael (2007) "Working Late: Do Workplace Sex Ratios Affect Partnership Formation and Dissolution?" *Journal of Human Resources*, 42(3): 583–595.

Van Bavel, Jan (2012) "Choice of Study Discipline and the Postponement Of Motherhood in Europe: the Impact of Expected Earnings, Gender Composition, and Family Attitudes," *Demography*, 47(2): 439–458.

Verardi, Vincenzo, and Nicolas Debarsy (2012) "Robinson's Square Root of N Consistent Semiparametric Regression Estimator in Stata," *Stata Journal*, 12(4), 726–735.

Wei, Shang-Jin, and Xiaobo Zhang (2011) "The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China," *Journal of Political Economy*, 119(3): 511–564.

10 Appendix

10.1 Data Appendix

10.1.1 UNESCO Data Description, Missing Data and Outliers

To measure the gender composition of tertiary-level graduates by field of study, we use data from public reports provided by the Data Centre of the UNESCO Institute for Statistics (UIS, www.uis.unesco.org). Specifically, we have downloaded data corresponding to Table 16 ("Graduates by broad field of education in tertiary education"), which covers graduates with ISCED education levels 5 and 6, from the October 2012 release through the following link: http://stats.uis.unesco.org/unesco/ReportFolders/ReportFolders.aspx. The following eight fields of study are recorded in the data (with their ISCED codes and descriptions):

Education	100	Teacher training and education science
Humanities	200	Humanities, languages and arts
Social sciences	300	Social sciences, business and law
Science	400	Science, mathematics and computing
Engineering	500	Engineering, manufacturing and construction
Agriculture	600	Agriculture and veterinary
Health	700	Health and welfare
Services	800	Services

The UIS data cover a total of 29 countries. We have excluded Luxembourg, Iceland, Malta, Greece, Poland, and Romania, for which less than 6 years of information on all eight fields is available. We also do not use information on the number of graduates from unknown fields of study. With two exceptions (CZ in 2003 and the UK in 2000, where it reaches 10%), the share of graduates with missing fields of study never exceeds 6%.

Out of the maximum possible total of 2,392 country-year-field observations (for 23 countries, 8 fields, and 13 years from 1998 to 2010), there were 382 (15%) missing values, and we have further dropped 38 outlier data cells (1.5%). For this purpose, we defined outlier values as those corresponding to hard-to-verify large changes in the total number of graduates (i.e., not in terms of changes in the share of women). Specifically, we dropped a data cell whenever the number of all graduates in a given country-year-field differed from one of the two neighboring years by more than 50%. Excluding 1998 and focusing on the 2,208 maximum possible number of data cells from 1999-2010, the number of missing and dropped data cells decreases to 220 (10%).

Missing data cells were treated in our analysis as follows: For the purpose of Figures 1 and 2, we have excluded the year 1998 and imputed the values for all 220 missing and

dropped data cells from neighboring year values, starting with the previous year, continuing with the following year when the previous year was not available, and using information from two years ago (ahead) when no neighboring year was available. For the purpose of Figures 3, 6 and 7, we have replaced missing country-year information with neighboring years as indicated in the graph legend. In addition, there are country-year cells where we have valid information for some but not all fields; these are used in the regression analysis presented in Figures 4 and 5, but are excluded from the descriptive graphs (Figures 1 and 2) or the segregation index calculations (Figure 3).

10.1.2   EU LFS Data Description, Missing Data and Outliers

We use the 2012 release of the anonymised EU Labour Force Survey (LFS) for the reference years 2003-2011. More specifically, we use the annual samples ("yearly files") except for Finland, where the annual sample does not contain information about children, so we use the specific household data file where this information is available. We do not use data from before 2003 since no information about the field of education was asked until then.

The EU LFS is a collection of national labor force surveys from EU countries. While most of the underlying surveys are collected as short rotating panels, the publicly available version of the data does not allow linking of individuals within surveys. In order to ensure that we do not use repeated observations for the same individuals, we use data from a single annual interview wave (wave 1 in all cases when multiple waves are available in the data).

We exclude country-year samples with missing information on graduation year, graduation field, or the presence of children. We also exclude Denmark where the annual sample does not contain information about children, and the information on educational attainment in the specific household data file is limited only to the reference person in the household. The share of missing values in educational attainment does not exceed 5% of prime-aged individuals in any of the country-year data cells. There are only few exceptions where the share of missing graduation year or field exceeds 5%.

10.2   Appendix: Figures and Tables

Table 1: Average Fertility Rates

Endogamous	Non-endogamous	Any child
	Women	
0.041	0.193	0.261
	Men	
0.045	0.145	0.199

Notes: Fertility rates of the 2000 graduation cohort based on the sample used in Figure 4 and 5 corresponding to the presence of first childbirth at most one year prior to and within five years after tertiary graduation. Endogamous couples graduated from the same field of tertiary education within at most five years from each other. 'Any child' fertility covers endogamous (same education) and non-endogamous (different education) couples and also children born to individuals without a partner.

Table 2: Equation (1) Estimation: Fertility Aggregation

	Age (demeaned)		R^2	N
Fertility Outcome	coeff.	s.e.	adjusted	
Women				92,154
Endogamous first childbirth	0.0008	0.0002	0.06	
Non-endogamous first childbirth	0.0049	0.0003	0.24	
Any first childbirth	0.0071	0.0003	0.31	
Men				72,795
Endogamous first childbirth	0.0023	0.0002	0.08	
Non-endogamous first childbirth	0.0135	0.0003	0.20	
Any first childbirth	0.0166	0.0004	0.26	

Notes: OLS estimates with survey year and country by field-of-study by graduation year fixed effects included.

Table 3: Equation (3) Estimation Corresponding to the Right Column of Figure 4 and to Figure 5

Fertility Outcome	ln(# of graduates) coeff.	s.e.	R^2 adjusted	N
Women				773
Endogamous first childbirth	0.0091	0.0039	0.32	
Non-endogamous first childbirth	-0.0246	0.0076	0.47	
Any first childbirth	-0.0196	0.0079	0.62	
Men				717
Endogamous first childbirth	0.0021	0.0043	0.41	
Non-endogamous first childbirth	0.0040	0.0070	0.46	
Any first childbirth	0.0076	0.0080	0.55	

Notes: OLS estimates with country by graduation year and field-of-study fixed effects included.

Figure 6: 2010-2000 Change in Share of Women and Total Number of Graduates by Country

The percentage-point change in the share of women among graduates against the change in the logarithm of total graduates between 2010 (2009 in BE, FR, PT, SI and 2008 in IT) and 2000 (1999 in IE, IT, SI and 2001 in HU, UK).
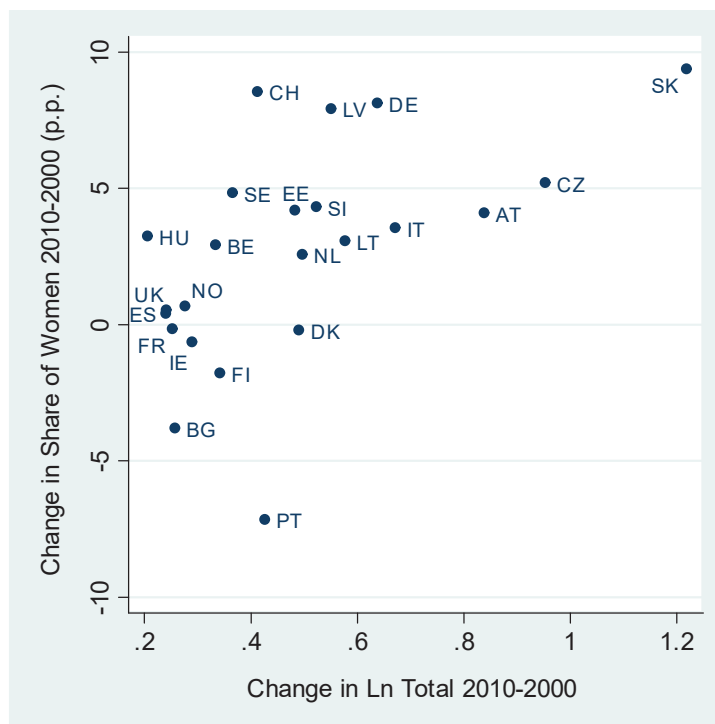
Figure 7: 2010-2000 Change in Share of Women and Total Number of Graduates by Field

Note: The percentage-point change in the share of women among graduates against the change in the logarithm of total graduates between 2010 and 2000 for the 23 countries (and year exceptions) of Fig 6. The relative size of the circles corresponds to the field-specific sum of graduates across these countries in 2000.
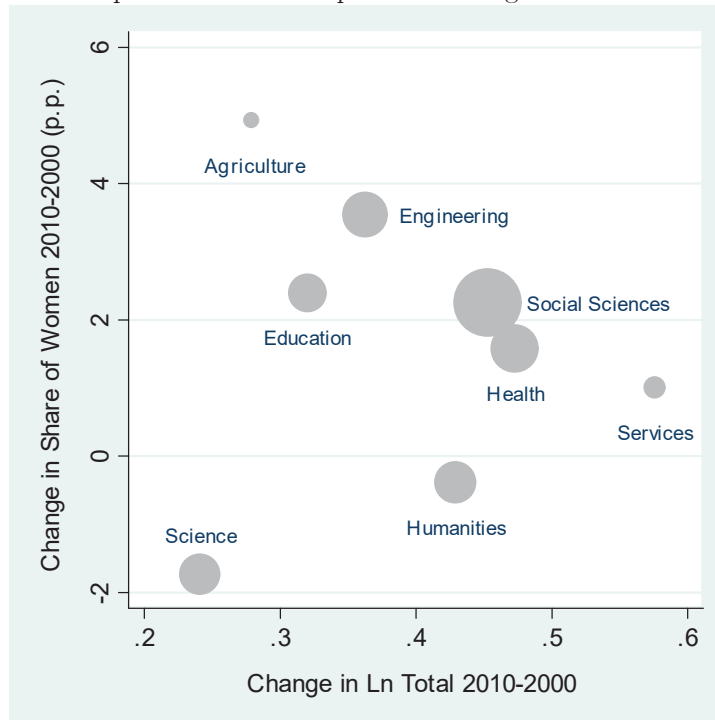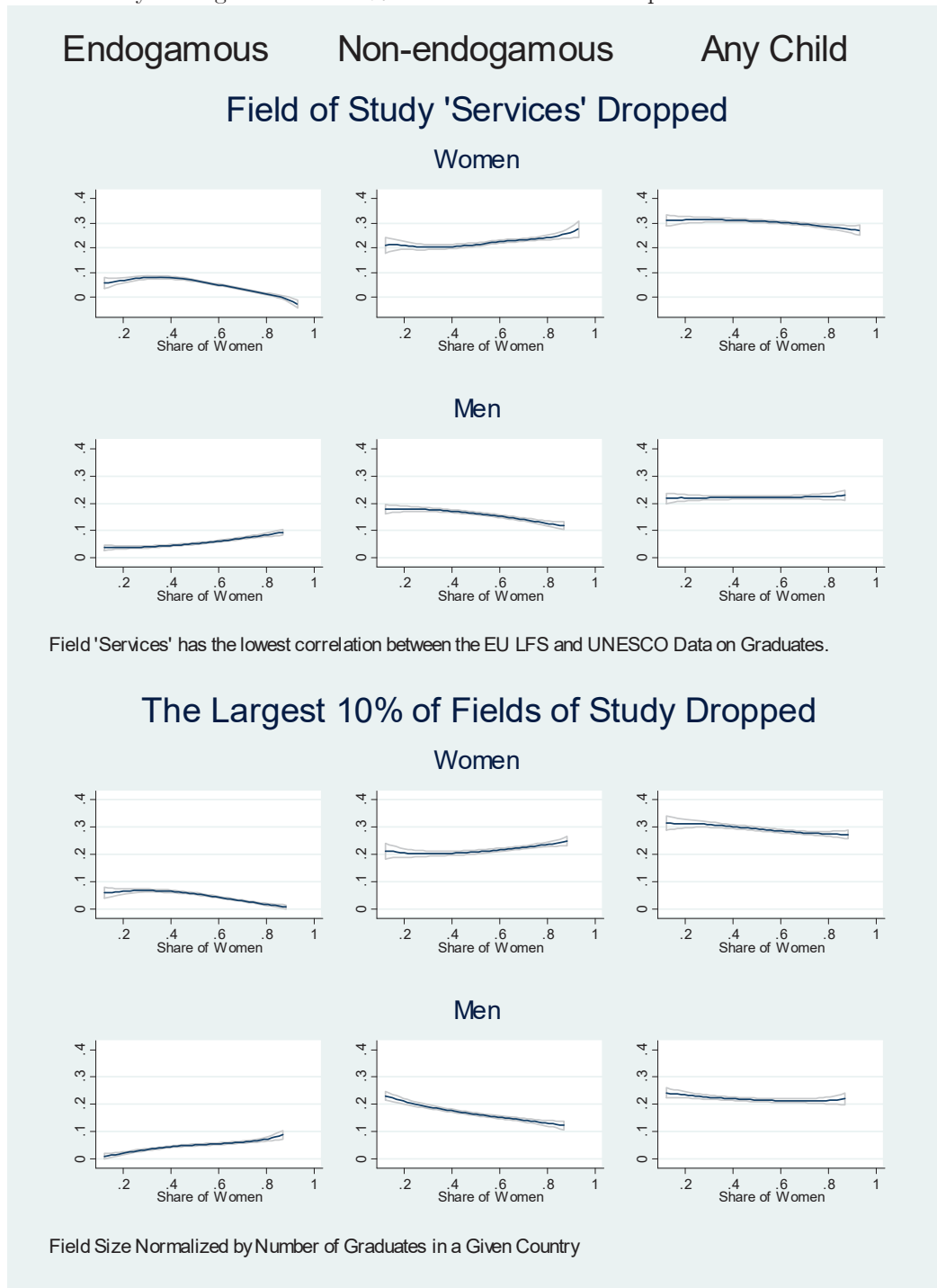
Figure 8: Effects of Share of Women in Group on Fertility by Type Conditional on Country-Year and Field-of-Study Fixed Effects

Note: Based on groups (field-of-study by country by year) with at least 5 or 15 individuals. The share of women corresponds to one's year of graduation. 95% confidence intervals are plotted.
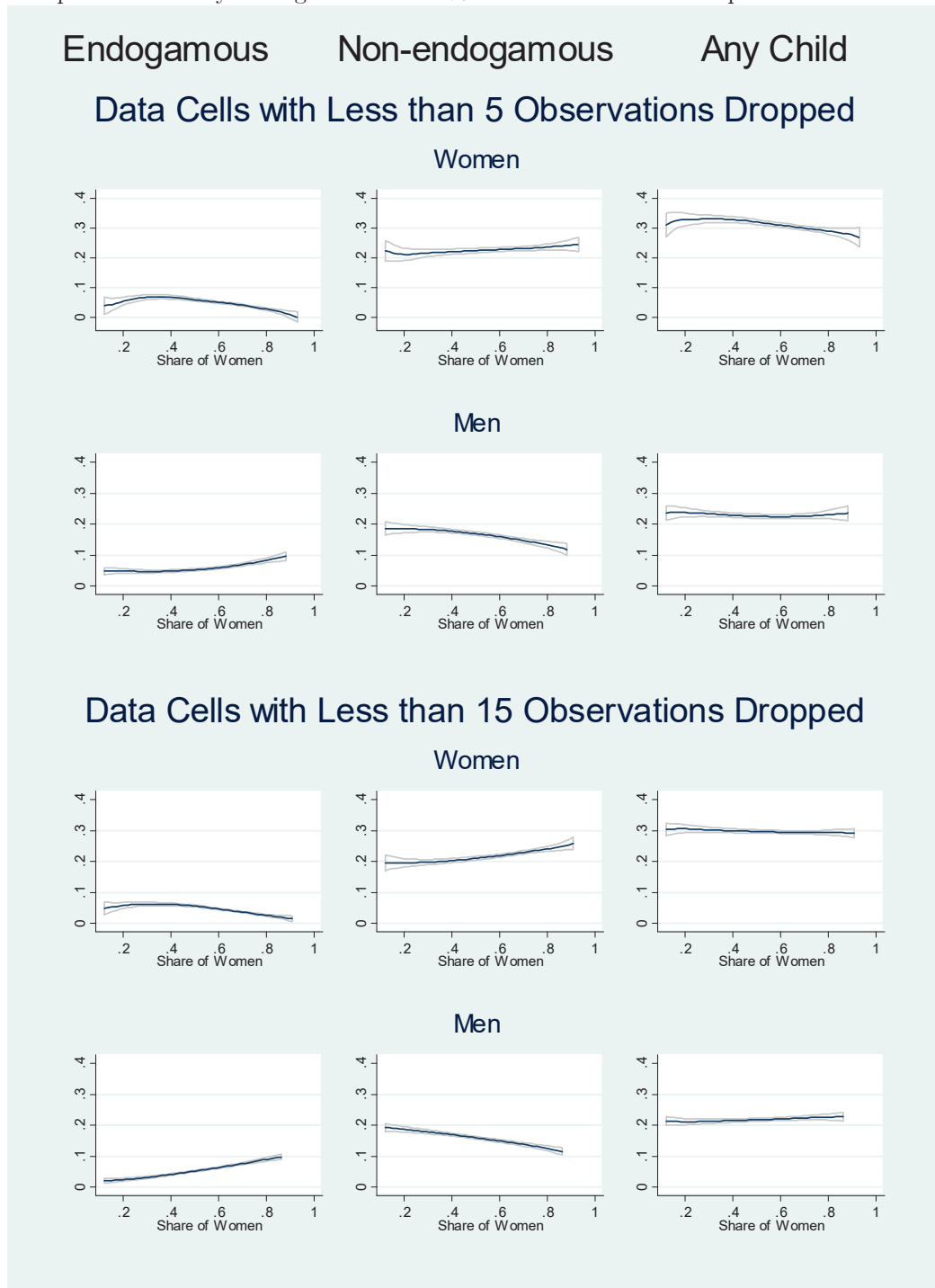
Figure 9: Effects of Share of Women in Group on Fertility by Type Conditional on Country-Year and Field-of-Study Fixed Effects

Note: Based on groups (field-of-study by country by year) with at least 10 individuals. The share of women corresponds to one's year of graduation. 95% confidence intervals are plotted.

Figure 10: Effects of Share of Women in Group on Fertility by Type Conditional on Country-Year and Field-of-Study Fixed Effects

Note: Based on groups (field-of-study by country by year) with at least 10 individuals. The share of women corresponds to peer groups as defined below. 95% confidence intervals are plotted.
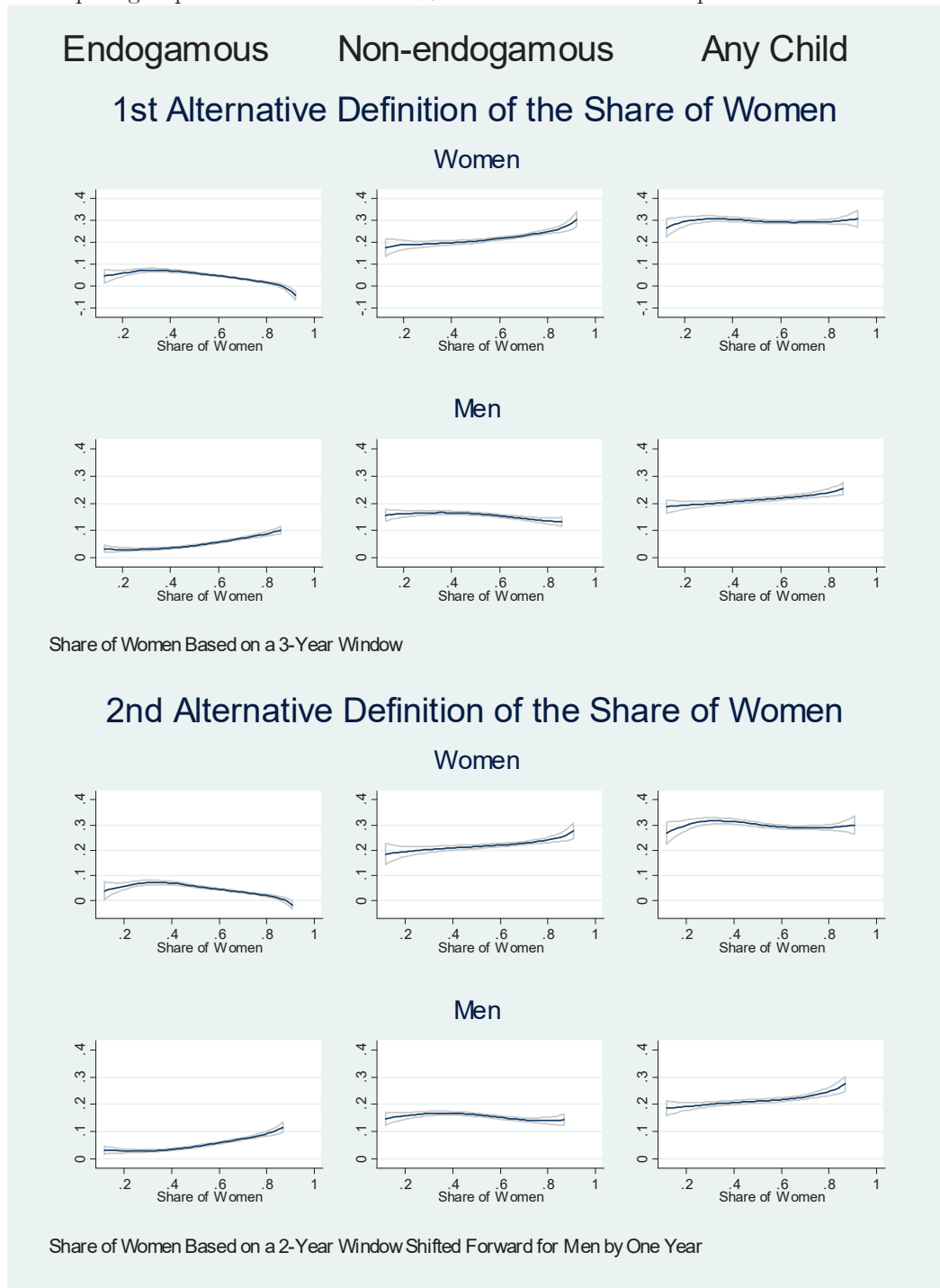
Figure 11: Effects of Share of Women in Group on Fertility by Type Conditional on Country-Field-of-Study and Year Fixed Effects

Note: Based on groups (field-of-study by country by year) with at least 10 individuals. The share of women corresponds to one's year of graduation. Countries (three for women and five for men) with average annual fertility change exceeding 1.4 of a percentage point are dropped. 95% confidence intervals are plotted.
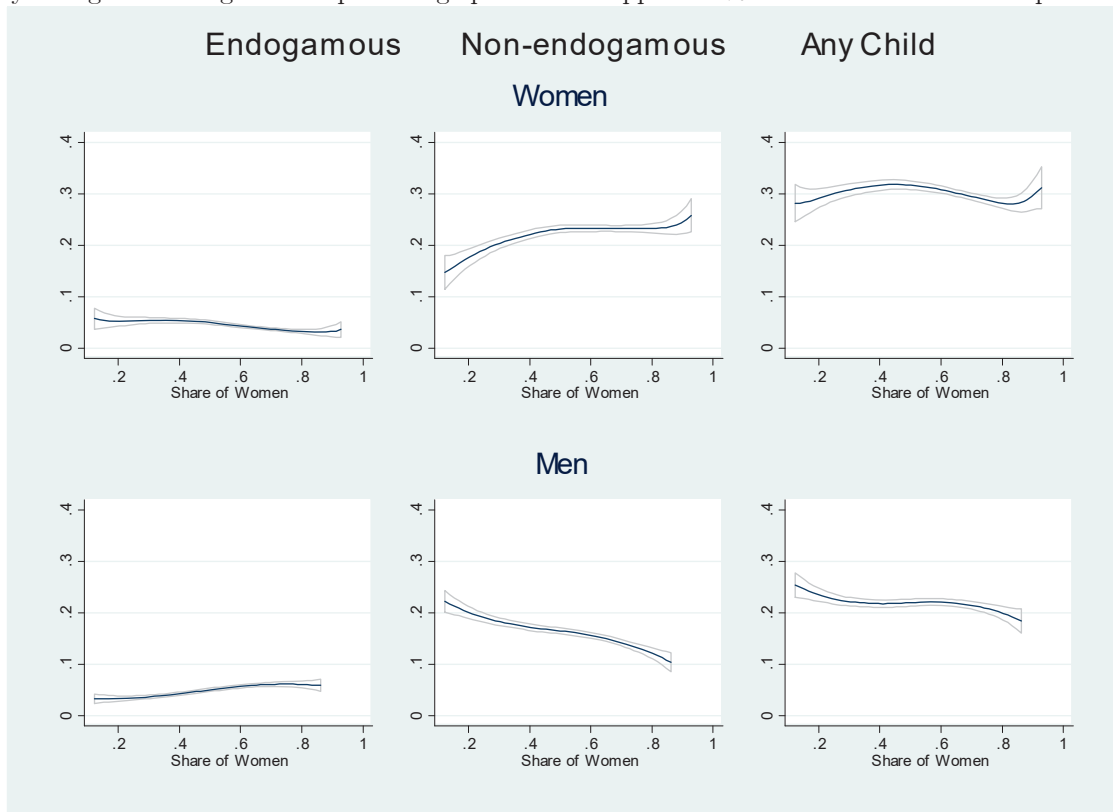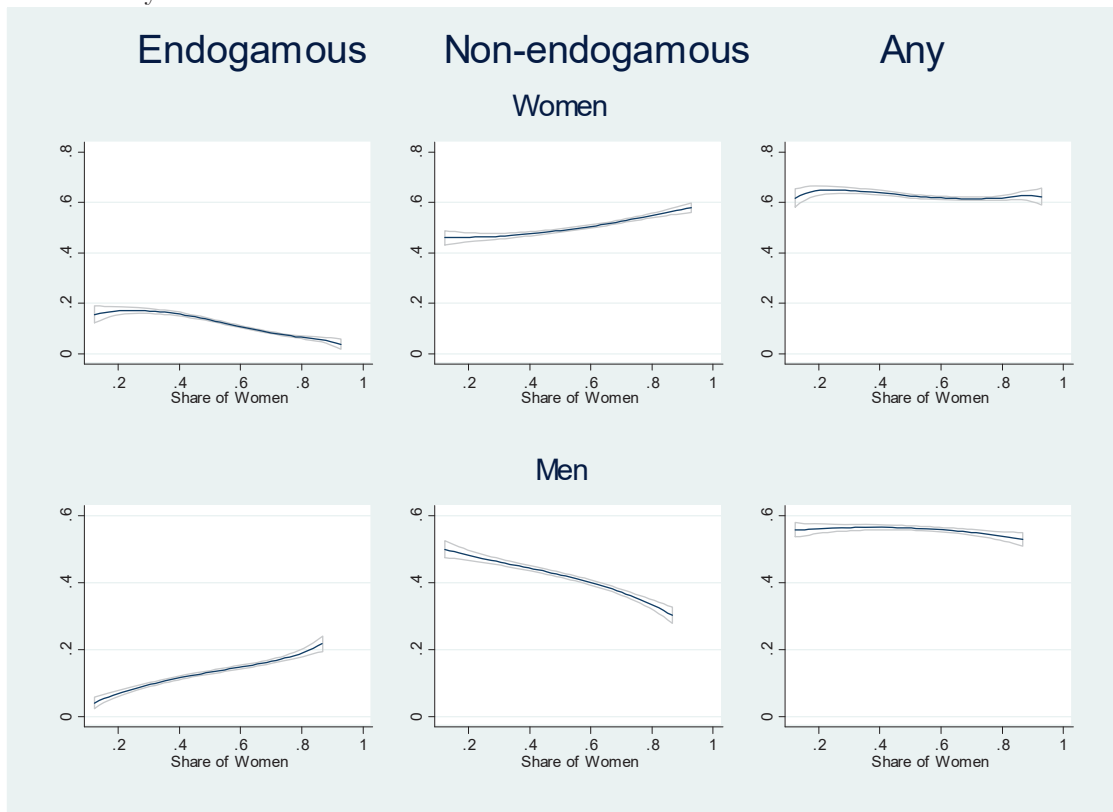
Figure 12: Effects of Share of Women in Group on Marriage/Cohabitation by Type Conditional on Country-Year and Field-of-Study Fixed Effects

Note: Based on groups (field-of-study by country by year) with at least 10 individuals. The share of women corresponds to one's year of graduation. 95% confidence intervals are plotted together with the non-parametrically estimated effects.

CELSI