Social Science Research 97 (2021) 102558

Contents lists available at ScienceDirect

Social Science Research

journal homepage: http://www.elsevier.com/locate/ssresearch

Under the influence of our older brother and sister: The association between sibling gender configuration and STEM degrees

Roujman Shahbazian [a,b]

[a] Swedish Institute for Social Research (SOFI), Stockholm University, Sweden
[b] Department of Sociology, University of Munich (LMU), Germany

ARTICLE INFO

Keywords:
Administrative registers
Educational choice
Birth order
Sex composition
STEM

ABSTRACT

This study examines the association between sibling gender configuration and second-born siblings' choice of so-called STEM educational fields (i.e., science, technology, engineering, and mathematics) in Swedish two-child families. Using population data from administrative registers in Sweden, the findings show that younger siblings, net of parental characteristics, are more likely to choose a STEM field if their older sibling has attended or is already attending a STEM program. Moreover, a gender difference is clear concerning the choice of a STEM field among younger siblings: Girls are more likely to choose a STEM field if they have an older sister who has attended a STEM program, than if they have an older brother in a similar program. However, the corresponding results are not found for boys. Given that STEM fields are markedly male-dominated at tertiary level, this indicates an importance of a same-sex role model for young girls contemplating gender-atypical educational choices.

1. Introduction

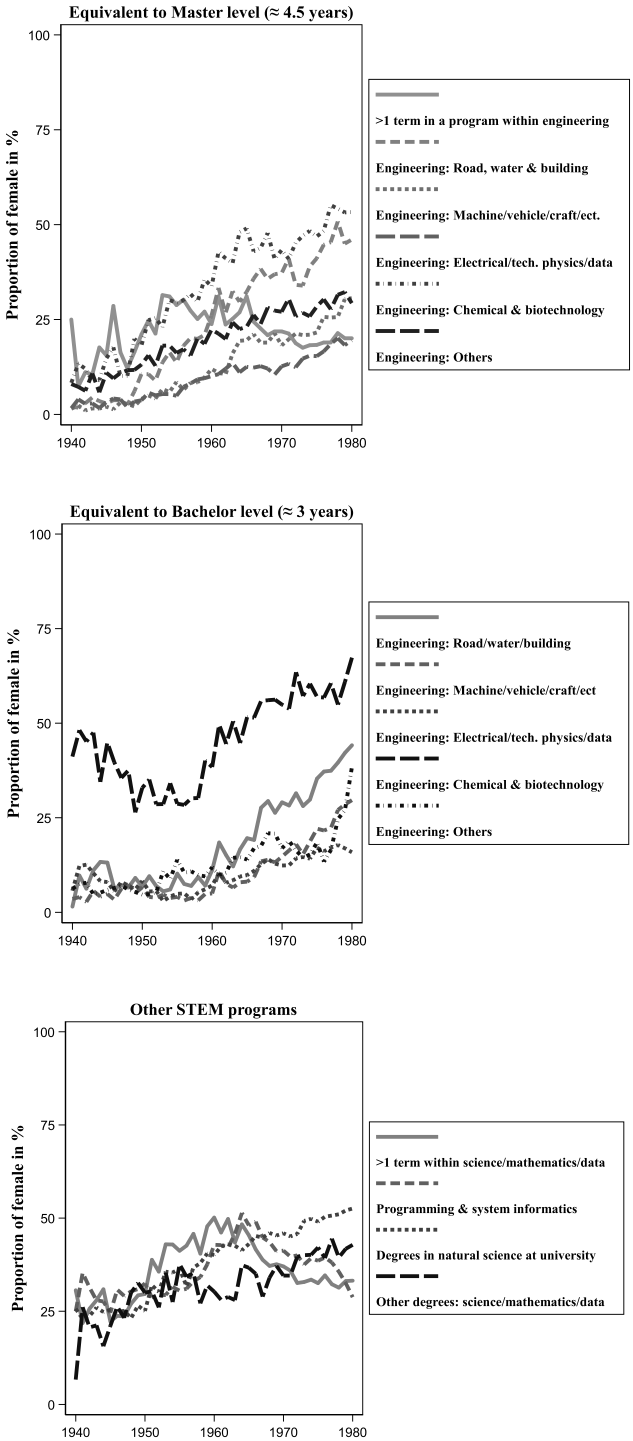
Choices of higher educational fields have historically been, and still are, notably horizontally segregated by gender (e.g., Charles and Bradley 2002; Alon and Gelbgiser 2011), even in countries subscribing to pronounced egalitarian policies such as Sweden (see Barone 2011). Women are especially underrepresented in the so-called STEM fields, i.e., science, technology, engineering, and mathematics, at the tertiary level (Charles and Bradley 2009; Mann and DiPrete 2013). These fields are argued to constitute the engine for a knowledge-based economy, and many countries and supranational organizations, such as the European Union, have boosted the emphasis on and resources allocated to these fields in particular to increase the number of STEM students over the past decades (Caprile et al., 2015). Combined with the fact that lifetime earnings are higher for STEM graduates than for most other graduates (Webber 2014), STEM programs have become increasingly interesting for students.

Some argue that gender differences in the number of STEM graduates can be attributed to disparate mathematic performance between girls and boys (Boli et al., 1985; Halpern et al., 2007); however, recent research indicated that mathematic performance only explains the observed gender segregation to a very small extent (Mann and DiPrete 2013). Gender differences in educational choices should instead be viewed from a social and cognitive approach, with an emphasis on institutions that contribute to generating these differences (Fouad et al., 2010). One important institution that generates these differences is the family of origin.

The family provides an excellent avenue for studying the socializing role of small-group interactions as well as differential

E-mail address: roujman.shahbazian@sofi.su.se.

https://doi.org/10.1016/j.ssresearch.2021.102558
Received 15 December 2018; Received in revised form 18 February 2021; Accepted 9 March 2021
Available online 8 April 2021



Note: The figure was compiled by the author from Swedish registers (total population). Birth cohorts between years 1940-80 are on the x-axis, while proportion of female is on the y-axis (from 0 to 100%). The figure shows the trend for 15 educational categories. Individual's education is measured as their highest level of education between ages 32-42. See Table A1 in the Appendix for the detailed description of these educational fields according to the Swedish national educational classification (SUN).

(caption on next page)

Fig. 1. Gender proportion of 15 different educational categories within STEM, according to individuals birth cohorts (from 1940 to 1980). Note: The figure was compiled by the author from Swedish registers (total population). Birth cohorts between years 1940- 80 are on the x-axis, while proportion of female is on the y-axis (from 0 to 100%). The figure shows the trend for 15 educational categories. Individual's education is measured as their highest level of education between ages 32–42. See Table A1 in the Appendix for the detailed description of these educational fields according to the Swedish national educational classification (SUN).

opportunities for individual group members (e.g., Brim 1958). Although parents are the primary agents in the socialization of their children, few studies have focused on the siblings' role, in particular on the combined effect of birth order and sibling sex composition, hereafter referred to as *gender configuration*, on their (younger) siblings' educational outcomes. Two exceptions are Joensen and Nielsen (2018), whose research focused on math and science courses in upper secondary school, and Gabay-Egozi et al. (2016), who focused on the selection of STEM programs and who argue for a causal pathway in gender socialization from older to younger siblings.

The overall aim of this study is to focus on tertiary educational choices and analyze to what extent younger siblings' educational choices in two-child families are shaped by their older siblings' educational choice, and more specifically the impact of the older sibling's gender. Previous research focusing on the underrepresentation of women in STEM programs has analyzed either the role of schools (Mann and DiPrete 2013; Legewie and DiPrete 2014), the effect of extra support provided to women during introductory STEM courses (e.g., Herrmann et al., 2016), or the association between the older sibling's gender and the younger sibling's probability of choosing a STEM program (Gabay-Egozi et al., 2016). This study contributes to existing research by first analyzing to what extent a younger sibling's probability of acquiring a STEM degree is influenced by having an older sibling with a STEM degree. Second, the impact of the older singling's gender on their younger singling's educational choice is analyzed, a perspective thus far not considered. With administrative registers covering the entire Swedish population, this study focuses on Swedish two-child families, which constitute the largest family-size group in Sweden (Statistics Sweden 2017). This specification also facilitates both the theoretical and empirical understanding of sibling interaction.

Sweden is, for several reasons, a particularly interesting case study for educational choices, as parents' financial resources play a much smaller role than in many other countries. Swedish society is characterized by a relatively compressed income distribution and generous social transfers. More importantly, the educational system has no tuition fees at any levels, all students receive student aid, and housing opportunities are available through student associations (Högskoleverket 2012). Thus, findings on the transition to tertiary education are less likely to be influenced by parents' economic constraints than in other societies in which tuition fees and living costs are, to a large extent, financed by the student's parents.

2. Contextual factors and STEM degrees in Sweden over time

The Swedish educational system is divided into three levels: (i) nine years of compulsory schooling, (ii) three years of secondary school, and (iii) tertiary education. Both primary and secondary education in Sweden have no tuition fees, and tertiary education is free for Swedish citizens (Hallden, 2008). Students in tertiary education are also eligible for government financial aid in the form of study grants (Högskoleverket 2012). The structure of Sweden's educational system along with student aid undoubtedly reduces the importance of family financial resources in the transition to tertiary education than in other contexts.

Increasing the number of women enrolled in STEM programs has been a goal in Sweden for almost 40 years. In 1973, the Swedish National Agency for Education was directed to stimulate the number of female applicants to natural science and technical studies (Lindahl 2003:24). The reason for promoting female participation in science and engineering studies is not only to increase women's earning potential, but also to meet the labor market demand for workers with STEM profiles.

Fig. 1 illustrates the proportion of women who acquired a STEM degree in one of 15 different subjects across birth year. It is evident that the majority of those receiving a STEM degree in Sweden has been and continues to be men, but also that the proportion of women in all 15 STEM subjects increases steadily over time. Although STEM fields continue to be male-dominated on the whole, it is worth noting that chemistry and biotechnology engineering (both at bachelor and master levels) have developed more female graduates than the other subjects. Overall, Fig. 1 underscores the gender segregation of STEM degrees in Sweden.

3. Understanding gender segregation in the field of education

One explanation for gender segregation in the STEM fields has focused on the idea that math-intensive degrees are less attractive to women, as they do not conform to traditional gender roles (e.g., Bradley 2000; Ceci et al., 2010). In a cross-national analysis, Charles and Bradley (2009) argue that gender essentialism (i.e., beliefs that men and women are interested in and suited for different kinds of study and work) is persistent and remains strong in contemporary societies. This persistence in beliefs regarding educational and occupational opportunities could explain gender differences in the decision to pursue a degree in a STEM field of study.

Another explanation is based on gender differences in values and preferences. Hakim (2002) argues that on average, women have a preference for a particular lifestyle that is either work-centered, home-centered, or one that combines paid work and time with family. This preference in turn could drive gender differences in a particular field of study. However, more recent research seems to indicate that work-related values, such as career, only play a minor role in explaining the gender gap in the fields of engineering and natural science (Frehill 1997; Mann and DiPrete 2013).

3.1. Family constellation

Yet another avenue for understanding gender segregation in educational fields is family constellation. Previous research on individuals' educational outcomes has analyzed the relationship between sibling configuration, including sibship size (Åslund and Grönqvist, 2010), birth order (Härkönen, 2013), birth spacing (Buckles and Munnich, 2012), and sex composition (Amin, 2009), and educational outcomes in terms of years of schooling, educational levels, and math or language scores at the primary/secondary level (see also Steelman et al., 2002 for a literature overview). In a recent study, Barclay et al. (2017) studied the relationship between birth order and university programs in Sweden. The authors find that first-born children are more likely to graduate from tertiary medical or engineering programs than are subsequent siblings.

The two most prominent explanations for sibling configuration effects are the *dilution hypothesis* and the *confluence model*. According to the former, an increased number of children in the family leads to a dilution of the available resources (i.e., money, time, attention) parents can invest in their children (Blake 1981). The latter instead postulates that family size and birth order influence the intellectual environment within a household (Zajonc and Markus 1975; Zajonc 1976). The arrival of the second child thus impairs the intellectual environment for the first-born child, as the parents now have two children to consider. Moreover, older siblings may benefit more from teaching their younger sibling(s) than the younger sibling(s) benefits from learning.

However, neither of these two explanations consider siblings' gender. Given that gender can be understood as not only an inborn quality, but also as a social and cognitive construction reproduced in everyday interaction (Bussey and Bandura, 1999), it is plausible to assume that sibling gender could also influence the parent/child interaction. For instance, there is evidence of differential paternal time investment in children from U.S. data: Fathers invest more time in their sons than their daughters (Price 2008), and mothers tend to spend more time alone with their daughters than their sons (Mammen, 2011).

Recent literature has focused on how younger siblings' choices can be shaped by the downstream social influence of older siblings. Joensen and Nielsen (2018) focus on whether older siblings' educational choices influence their younger siblings when it comes to upper secondary mathematics and science courses. Gabay-Egozi et al. (2016) analyze the topic further, looking into whether older siblings' gender has a socialization effect on younger siblings' decision for a STEM program in the US. Their findings indicate that both younger girls and boys are more likely to pursue a STEM degree if they have an older brother. However, they do not condition this on whether the older sibling has a STEM degree, and therefore cannot give evidence to whether they act as a role model.

3.2. Sibling influence

Older siblings often serve as role models for their younger siblings. Theoretical explanations for this include the 1) *social cognitive career theory* and 2) the *revised sex minority hypothesis*, which are here related to sibling gender configuration.

According to the social cognitive career theory (Lent et al. 1994, 2002; Lent 2013), a further development of Bandura's (1986) social cognitive theory, individuals form their interests and make educational choices based on internal (i.e., self-efficacy, outcome expectations, goal setting) as well as external (i.e., gender, social support, social constraints) factors. The focus of this theory is on the personal constructions that individuals place on events related to career decision-making (Lent et al., 2002; Lent 2013). As the topic at hand focuses on the downstream social influence of older siblings, the external support mechanisms present in individuals' environments, where support is argued to help shape individuals' interests and beliefs about themselves, are highlighted here. For instance, social support (e.g., emotional, financial) for engaging in specific behaviors creates opportunities and exposes individuals to specific tasks and role models (Lent et al., 2000). As an example, to illustrate the cognitive process, younger siblings are more likely to observe and take into account the experiences of their older sibling(s) before deciding on how to behave. For instance, if the older sibling has chosen a science program at upper secondary school and been successful, the younger sibling might be more inclined to follow in the same path. Moreover, the older sibling can function as a role model by introducing their younger sibling to specific activities and providing support for their continuing participation. The following *gender-neutral hypothesis* results from these thoughts:

H.1. Individuals with an **older sibling** with a STEM degree are more likely to acquire a STEM degree than individuals with an older sibling *without* a STEM degree

Furthermore, the cognitive process also implies that if children routinely experience certain types of activities to be performed by women, while other activities are performed by men, then it is not surprising that these type of activities are gender-typed as "feminine" and "masculine" (Lent et al., 2000, 2013).[1] For example, a young girl with strong mathematical skills may not choose to major in engineering due to her perception of engineering as an inappropriate occupation for women. Younger siblings may then be more likely to imitate their older siblings if they are gender-similar, further reinforcing such behaviors. Following the social cognitive career theory reasoning about role models as support systems, individuals who have an older sibling with a STEM degree are expected to be more likely to acquire a STEM degree than those who have an older sibling without a STEM degree. Although the imitation of role models' actions is not necessarily gender typical, as discussed above, several factors stimulate individuals to reproduce the action or behavior that their nearest surrounding social environment consider gender-appropriate. As a result, younger siblings in same-sex sibling pairs in which the older sibling has a STEM degree are expected to be more likely to acquire a STEM degree themselves, compared to younger siblings in same-sex sibling pairs in which the older sibling does not have a STEM degree. The following

[1] Empirical evidence from time-use data on housework in Swedish two-child families. Evertsson (2006:431) writes: "*… the notion of housework as a gendered activity is transferred from parents – and the surrounding society – to children.*"

hypotheses have thus been labeled *role-modeling hypotheses*:

H.2a. Females with an **older sister** with a STEM degree are more likely to acquire a STEM degree than females with an older sister ***without*** a STEM degree

H.2b. Males with an **older brother** with a STEM degree are more likely to acquire a STEM degree than males with an older brother ***without*** a STEM degree

The sex composition of sibling pairs becomes a strong socialization factor in such a way that it paves the path for siblings' educational trajectory (Conley 2000). The *revised sex minority hypothesis* suggests that same-sex siblings create a more competitive environment, which in turn influences educational choices, while opposite-sex siblings establish a more sociable environment (Conley 2000). Sibling competition in same-sex sibling pairs is therefore more likely to lead to siblings with similar educational choices, an assumption supported by a Danish study that found that the younger sibling, in particular in a boy-boy constellation, was more likely to choose math and natural science courses in secondary school if the older sibling had done the same (Joensen and Nielsen, 2018). Studies using U.S. data, such as Buser et al. (2014), estimates that around 20 percent of the gender difference in the choice of an academic math and natural science high school track is explained by gender differences in competition. However, sibling competition in same-sex sibling pairs does not necessarily lead to similar siblings, as suggested above. Some theoretical arguments suggest that siblings might actually niche away from each other in order to ease the competition between them (Sulloway 2007). One reason for a younger sibling to niche away from older siblings is to gain parental resources, representing not only a reaction to an older sibling's behavior, but also a reaction to the parents' behavior. It is, in other words, reasonable to assume that due to seniority, the first-born sibling will have an advantage in almost all aspects, including educational tracks, hobbies, etc. If the second-born sibling does not niche away and instead engages in similar activities, they may risk settling into "second place", thereby receiving less parental appreciation and praise. By niching, younger siblings avoid competition and parents are more likely to judge them as equally good or successful, rewarding both siblings with the same amount of parental attention.

If both competiveness and support systems impact educational decision-making, having an older sibling role model who has demonstrated that it is possible to achieve academic success in a male-dominated area (such as a STEM field) might increase the likelihood that the younger sibling will follow in their footsteps. The first-born sibling's gender might then exacerbate the sibling influence of having an older sibling with a STEM degree, especially if siblings have the same sex. Following the *revised sex hypothesis* formulated by Conley (2000), which emphasizes competition as a mechanism, and the *social cognitive career theory* (Lent et al., 2000), which emphasizes the support system from role models as a mechanism, it is therefore more likely that younger sisters are more influenced by having an older sister than an older brother with a STEM degree, and vice versa for boys. The following hypotheses from these notions have thus been labeled *sibling competition/support hypotheses*:

H.3a. Girls with an **older sister** with a STEM degree are more likely than girls with an ***older brother*** with a STEM degree to acquire a STEM degree themselves.

H.3b. Boys with an **older brother** with a STEM degree are more likely than boys with an ***older sister*** with a STEM degree to acquire a STEM degree themselves.

4. Data and research strategy

Using Swedish administrative registers covering the entire population, this study examines whether and to what extent older siblings shape their younger sibling's educational choices. In Swedish register data, children are linked to their parents by a multigenerational identifier, making it possible to identify sibling groups. This identifier is managed and maintained by the Statistics Sweden, and updated annually and part of the registry system for the Total Population Register (SCB, 2016). Here, siblings are defined as having the same mother and father (i.e., all children born in Sweden between 1960 and 1980 with ID for both parents: n = 2,141,988 observations). Twin siblings are excluded (n = 2,106,215 observations), and sibling constellations are restricted to two children (i.e., only two-child families: n = 945,153 observations). The analytical population in this study can be described as an open window, meaning the second-born sibling must be born between 1960 and 1980, but the first-born sibling can be born outside the window (i.e., before 1960).

Educational information for all family members is extracted from the Statistics Sweden's Integrated Database for Labour Market Studies (LISA), which covers the years 1990–2012. The analytical population includes two-child families in which both siblings have finished at least one semester at the tertiary level or obtained a tertiary degree. The most recent educational information is used for the following analyses. Finally, as the focus of this study is to analyze the influence of older siblings on younger, the analytical population is further restricted to sibling pairs in which the older sibling either started or finished their tertiary degree first (n = 243,808; of which 121,904 are second-born siblings).

4.1. The variables

The dependent variable here measures whether the second-born sibling has completed at least one semester in a STEM program at the tertiary level.[2] All analyses were also conducted with only those siblings who had finished their STEM degree (n = 162,280; of

[2] The reason for focusing on second-born siblings who have completed at least one semester in a STEM program, is because there has historically been a strong labor demand for STEM graduates in Sweden, especially in the field of engineering, resulting in a common tradition of offering STEM students positions even before receiving their degree.

which 81,640 are second-born siblings).[3] This produces similar results and can be seen in section *sensitivity test* below. Educational attainment is classified according to the Swedish national educational classification, SUN,[4] and the identified STEM subjects are reported in Table A1 in the Appendix.

The first independent variable ***older sibling STEM*** distinguishes whether the older sibling has attended/completed a STEM program (1 = yes, 0 = no). Having an ***older sister*** (1 = older sister; 0 = older brother) and having an ***older brother*** (1 = older brother; 0 = older sister) are dummy variables. The interactions ***older sibling STEM * older sister*** and ***older sibling STEM * older brother,*** distinguish whether the second-born sibling has an older sister (brother) with a STEM background.

As noted above, some of the association between the two siblings' educational choices is likely to be confounded by other individual, sibling, and parental characteristics. Previous research has indicated that a student's intellectual ability (Davies and Guppy 1997), having an absent parent (Downey 1994; Buchmann and DiPrete 2006), parental occupation (Leppel et al., 2001), having at least one parent working in a STEM field (Dabney et al., 2013), and parental income (Seehuus and Reisel 2017) influences gender differences in educational decision-making. This analysis therefore also controls for the following characteristics that could influence siblings' choice of pursuing a STEM degree:

Year of birth refers to the birth year of the younger sibling, measured as year dummies, to control for changes over time between 1960 and 1980. ***Birth spacing intervals*** are included to capture the influence of sibling age differences on the relationship between siblings and parental investment in each child (Zajonc 1976; Shahbazian 2020). Birth spacing intervals are divided into eight dummy variables with the following intervals: 1) less than 24 months, 2) 24–35 months, 3) 36–47 months, 4) 48–59 months, 5) 60–71 months, 6) 72–83 months, 7) 84–95 months, and 8) more than 95 months. ***Living with both parents from ages 8–15*** is included to capture the length of time the siblings have been interacting and have had a similar environment (i.e., living with both parents).[5] This measure is divided into nine different dummy variables with the following time spans: 1) did not live with both parents from ages 8–15, 2) lived with both parents up to age 8, 3) up to age 9, 4) up to age 10, 5) up to age 11, 6) up to age 12, 7) up to age 13, 8) up to 14, and 9) up to age 15. ***GPA rank*** is a measure used to capture intellectual ability. The variable ranks pupils GPA (from 1 to 100), at upper secondary level, at each examination year relative to all pupils in Sweden that graduated the same year. By ranking the GPA, any potential difference in grading between cohorts are adjusted. ***Adopted*** is a dummy variable indicating whether one of the two siblings have been adopted (1 = at least one sibling adopted; 0 = no sibling adopted). ***Biological two-child family*** is a dummy variable indicating whether the two siblings are biological siblings (i.e., have the same parents) and that both of their parents have no previous (or subsequent) children from other relationships (1 = biological siblings (with no stepsiblings); 0 = at least one stepsibling). ***Born in Sweden*** is a dummy variable indicating whether the second-born sibling was born in Sweden (1 = yes; 0 = no).

The first set of parental indicators included in this study is whether the parents have a STEM educational background. ***Father STEM*** and ***mother STEM*** are dummy variables indicating whether the father and mother have a STEM background (1 = yes; 0 = no). ***Father's education level*** and ***mother's education level*** are divided into nine dummy variables, indicating the father's/mother's educational level: 1) primary education, 2) upper secondary education, 3) post upper secondary education (less than two years), 4) post upper secondary education (two years), 5) post upper secondary education (three years), 6) post upper secondary education (five years), 7) post upper secondary education (more than five years), 8) PhD degree, and 9) missing value. ***Father's occupation*** and ***mother's occupation*** were measured when the second-born sibling was approximately 15 years old, and is classified according to the three-digit Nordic Occupational Classification (NYK). The measure is divided into 294 and 277 dummy variables, respectively. ***Household disposable income rank*** is constructed by combining both parents' income for each sibling from the age of 8–16 (see Hjalmarsson et al. (2015)). ***Residential region*** indicates in which region (in Swedish: län) the second-born sibling lived at age 16, in order to control for different demands for STEM graduates based on the local labor market conditions.

4.2. Analytical strategy

The following analyses are conducted with linear probability models (LPM). This approach is similar to ordinary least squares (OLS) models, but run on a binary outcome. The coefficients are therefore interpreted as percentage point increases in the probability of the outcome. LPM models are preferred over logistic regression because estimates are less influenced by unobserved heterogeneity and omitted variable, can be more readily compared across models, and intuitively interpreted (see Mood 2010). All models were also estimated via logistic regression, and the results remain unchanged (see Section 5.4 below).

The gender neutral hypothesis, namely whether younger siblings are more likely to acquire a STEM degree if an older sibling has already obtained a STEM degree can be tested by estimating the following model as Eq(1):

$$STEM_i = \alpha + \beta_1(older\ sibling\ STEM_i) + \boldsymbol{\beta}(all\ controls_i) + \varepsilon_i$$

[3] There are no missing observations in the dependent variable.

[4] *https://www.scb.se/dokumentation/klassifikationer-och-standarder/svensk-utbildningsnomenklatur-sun/*.

[5] Due to data restrictions, the age frame of reference begins at age 8. The *fastighetsnr* [in English: residential building number] is available only from 1968 onward, at the time the first sibling in the sample was 8 years old. Please note that having the same *fastighetsnr* does not necessarily mean that all family members live together. It is theoretically possible that the parents (or one of the siblings) live in different apartments in the same building [*fastighetsnr*]. However, the probability that two parents who have children together live in the same residential building (i.e., have the same *fastighetsnr*) but in different apartments, from age 8 of their oldest child, until age 16 of their youngest child, can be assumed to be extremely low.

In order to test the role modeling and sibling competition/support hypotheses, either a three-way interaction model (interaction: younger siblings' gender, older siblings' gender, and older sibling STEM) or a two-way interaction model (population split by the gender of the second-born sibling) can be performed. For these analyses, the latter was chosen and is estimated as Eq(2) for second-born girls:

$$STEM_i = \alpha + \beta_1(older\ sibling\ STEM_i) + \beta_2(older\ sister_i) + \beta_3(older\ sibling\ STEM_i * older\ sister_i) + \boldsymbol{\beta}(all\ controls_i) + \varepsilon_i$$

and as Eq(3) for second-born boys

$$STEM_i = \alpha + \beta_1(older\ sibling\ STEM_i) + \beta_2(older\ brother_i) + \beta_3(older\ sibling\ STEM_i * older\ brother_i) + \boldsymbol{\beta}(all\ controls_i) + \varepsilon_i$$

for which i is an index for the second-born sibling and *STEM* indicates whether the second-born sibling i has chosen a STEM program. *Older sibling STEM* indicates the first-born sibling's STEM background, *older sister* indicates whether the first-born sibling is a girl, *older brother* whether the first-born sibling is a boy, and the interaction between *older sibling STEM* and *older sister* or *older brother*. Table 1 shows the interpretation of the coefficients of both interaction models. In Panel A (second-born girls), having an older brother without a STEM background corresponds to the reference category, having an older brother with a STEM background corresponds to β_1, having an older sister without a STEM background corresponds to β_2, and having an older sister with a STEM background corresponds to $\beta_1 + \beta_2 + \beta_3$.

The second set of hypotheses (H.2a & H.2b) tests the role-modeling hypothesis (i.e., having an older same-sex sibling with a STEM background). H.2a tests whether second-born girls are more likely to pursue a STEM degree if their older sister had also studied a STEM program compared to if their older sister had studied something else (i.e., whether $\beta_1 + \beta_2 + \beta_3 > \beta_2$ in Eq (2)). H.2b tests the opposite: whether second-born boys are more likely to pursue a STEM degree if their older brother had also studied a STEM program compared to if their older brother had studied something else ($\beta_1 + \beta_2 + \beta_3 > \beta_2$ in Eq (3)).

The third set of hypotheses (H.3a & H.3b) tests the competition/support hypothesis of same-sex siblings (i.e., having an older same-sex sibling with a STEM background compared to an opposite-sex sibling with a STEM background). H.3a tests whether having an older sister with a STEM background compared to an older brother with a STEM background increases second-born girls' likelihood to pursue a STEM degree (i.e., whether $\beta_1 + \beta_2 + \beta_3 > \beta_1$ in Eq (2)). H.3b tests the opposite: whether having an older brother with a STEM background compared to an older sister with a STEM background increases second-born boys' likelihood to pursue a STEM degree ($\beta_1 + \beta_2 + \beta_3 > \beta_1$ in Eq (3)).

4.3. Descriptive statistics

Table 2 lists descriptive statistics for the interaction variables used in this analysis, separately for women and men. The analysis sample consists of 121,904 second-born siblings, of which 53% are women. Individuals with a STEM degree total 80,268, of which 75% are men. Slightly more first-born siblings have a STEM background compared to second-born siblings. In addition, 73% (n = 12,325) of second-born siblings have an older brother with a STEM background compared to 27% (n = 4488) who have an older sister with a STEM background. These descriptive statistics clearly show a strong gender bias in STEM degrees in the sample. Regardless of how the sample is divided, the proportion of women with a STEM background remains under 1/3 (fluctuating between 24% and 32%), while the proportion of men never dips below 2/3 (fluctuating between 68% and 76%). Cell sizes are sufficient to perform interaction models. Descriptive statistics for most control variables are presented in Table A2 in the Appendix.

5. Results

5.1. Gender-neutral hypothesis

Whether individuals with an older sibling who has a STEM background are more likely to acquire a STEM degree themselves compared to individuals with an older sibling with no STEM background is tested in H.1 – the gender-neutral hypothesis. Table 3 lists the likelihood of a second-born sibling choosing to pursue a STEM degree after the stepwise inclusion of: i) whether the second-born has an older sibling who attended a STEM program (Model 1) and ii) all control variables (Model 2). As previously described, these analyses are based on population data from Swedish official records, which provide a satisfactory estimate precision. Significance levels will not be further elaborated upon here.

The results presented in Table 3 show that second-born siblings who have an older sibling with a STEM background are 13 percentage points (effect size ≈ 41%)[6] more likely to pursue a STEM degree themselves compared to second-born siblings whose older sibling has a non-STEM background. This result is robust for the inclusion of the control variables in Model 2 (b = 0.12, 12 percentage points). As a result, H.1 cannot be rejected.

5.2. Role modeling hypotheses

In a second step, the role modeling hypotheses (H.2a & H.2b) are tested to evaluate whether younger siblings are more likely to acquire a STEM degree if their older sibling has a STEM degree compared to another degree. Table 3 shows the likelihood of second-

[6] This effect size is calculated by dividing the *point estimate* with the *mean value of the outcome* (the proportion of **second born siblings** with STEM) and multiplying with 100: (0.13/0.32) * 100 ≈ 41%.

Table 1
Interpretation of coefficients according to Eq(2) and Eq(3).

Panel A: Second-born girls according to Eq (2)	
Having an older Brother without STEM	*Ref.*
Having an older Brother with STEM	β_1
Having an older Sister without STEM	β_2
Having an older Sister with STEM	$\beta_1 + \beta_2 + \beta_3$
Panel B: Second-born boys according to Eq (3)	
Having an older Sister without STEM	*Ref.*
Having an older Sister with STEM	β_1
Having an older Brother without STEM	β_2
Having an older Brother with STEM	$\beta_1 + \beta_2 + \beta_3$

Note: Table 1 provides the interpretation of the coefficients of the interaction models according to Eq(2) (Panel A) and Eq(3) (Panel B).

Table 2
Descriptives (the variables in focus).

	Total	Women	Men
# of 2nd born with tertiary education	121,904	64,422	57,482
	(100%)	(53%)	(47%)
# of 1st born with STEM	40,944	9752	31,192
	(100%)	(24%)	(76%)
# of 2nd born with STEM	39,324	10,631	28,693
	(100%)	(27%)	(73%)
# of 1st born with a younger sibling with STEM	16,183	4369	11,813
	(100%)	(27%)	(73%)
# of 2nd born with an older sibling with STEM	16,813	5066	11,747
	(100%)	(30%)	(70%)
# of 2nd born with an older brother with STEM	12,325	3617	8708
	(100%)	(30%)	(70%)
# of 2nd born with an older sister with STEM	4488	1449	3039
	(100%)	(32%)	(68%)

Note: this table was compiled by the author.

Table 3
LPM regression coefficients for the association between sibling sex configuration and STEM programs for second-born girls and boys.

Dependent variables	Second-born		Second-born girl				Second-born boys			
	M. 1	M. 2	M. 3	M. 4	M. 5	M. 6	M. 7	M. 8	M. 9	M. 10
Older SIBLING STEM	0.13**	0.12**	0.10**	0.12**	0.10**	0.09**	0.17**	0.20**	0.16**	0.14**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Having an older Sister				0.04**	0.02**	0.02**		–	–	–
				(0.00)	(0.00)	(0.00)		–	–	–
* *Older SIBLING STEM*					0.06**	0.05**			–	–
					(0.01)	(0.01)			–	–
Having an older Brother				–	–	–		−0.08**	−0.11**	−0.10**
				–	–	–		(0.00)	(0.01)	(0.01)
* *Older SIBLING STEM*					–	–			0.08**	0.07**
					–	–			(0.01)	(0.01)
Controls	No	Yes	No	No	No	Yes	No	No	No	Yes
Constant	0.28**	0.27**	0.13**	0.11**	0.12**	0.09**	0.44**	0.47**	0.48**	0.44**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.03)
Observations	121,904	121,904	64,422	64,422	64,422	64,422	57,482	57,482	57,482	57,482
R-squared	0.020	0.039	0.017	0.019	0.020	0.045	0.025	0.031	0.032	0.066

Note: Robust standard errors in parentheses, $^{**}p < 0.01$, $^{*}p < 0.05$. The follow controls are included, year of birth dummies, birth spacing interval dummies, GPA rank, adopted sibling, biological two-child families, living with both parents, born in Sweden, mother having a STEM degree, father having a STEM degree, mothers educational level, fathers educational level, mothers occupation around age 15, fathers occupation around age 15, household $ rank (ages 8–15) and residential region.

born girls/boys to pursue a STEM degree after the stepwise inclusion of: i) whether the second-born sibling has an older sibling with a STEM background (Model 3/7), ii) the older singling's gender (Model 4/8), iii) the interaction of the two previous variables (Model 5/9), and iv) the full model with all control variables (Model 6/10).

The results from Model 3 in Table 3 show that second-born girls who have an older sibling with a STEM background are more likely to choose to pursue a STEM degree themselves compared to second-born girls whose older sibling has not studied a STEM subject ($b = 0.10$, 10 percentage points). Model 4 in Table 3 shows that having an older sister increases the likelihood for second-born girls to pursue a STEM degree ($b = 0.04$, 4 percentage points), and Model 5 shows that the additional association of having an older sister with a STEM background is positive for second-born girls' probability to pursue a STEM degree themselves ($b = 0.06$). When all individual, sibling, and family characteristics, such as parental education, occupation, and whether they have a STEM degree, are added in Model 6, the results for second-born girls remain stable.

Model 7 in Table 3 shows that second-born boys who have an older sibling with a STEM background are more likely to pursue a STEM program themselves. However, Model 8 indicates that having an older brother, compared to an older sister, decreases the likelihood for second-born boys to pursue a STEM degree. Model 9 shows that the additional association of having an older brother with a STEM background is positively associated with second-born boys' probability to pursue a STEM degree.

However, the purpose of implementing an interaction model is to analyze the combined association of the main and interaction effect(s). In doing so, the role model hypotheses can be tested: whether older siblings with a STEM background act as role models for their younger siblings' educational decisions, compared to older siblings who studied something else. H.2a tests whether having an older sister with a STEM background increases the likelihood of second-born girls to pursue a STEM degree themselves, compared to an older sister with a different educational background. Table 4 presents support for this hypothesis, based on results from the interaction models in Table 3 above. Second born-girls who have an older sister with a STEM background are approximately 13 percentage points (effect size ≈ 80%)[7] more likely than second-born girls whose older sister has a different educational background to choose to pursue a STEM program themselves.

In order to test the corresponding hypothesis for second-born boys, the combined association net of individual, sibling, and parental characteristics is tested according to M10 in Table 3. The result presented in Table 4 indicates support for H.2b; second-born boys with an older brother with a STEM background are approximately 20 percentage points (effect size ≈ 41%)[8] more likely than second-born boys whose older brother has a different educational background to pursue a STEM degree. Thus far, the results indicate that both second-born girls and boys in same-sex sibling pairs are more likely to pursue a STEM degree if their older sibling has a STEM background, net of individual, sibling, and parental characteristics.

5.3. Sibling competition and/or support hypotheses

As discussed previously, STEM fields are male-dominated. The second set of hypotheses focuses on sibling competition/support; more specifically, whether second-born siblings become similar or niche away from their older siblings with a STEM background, depending on the sex of the older sibling.

H3.a tests the likelihood of second-born girls to pursue a STEM degree, depending on whether they have an older sister or brother with a STEM background. Table 4 provides support for this hypothesis, and therefore it cannot be rejected. This finding indicates that having a same-sex sibling increases second-born girls' likelihood of choosing to pursue a STEM program by almost 7 percentage points (effect size ≈ 41%)[9]. Thus, second-born girls are somewhat more affected by their sisters' educational path than by their brothers'. The corresponding finding for second-born boys indicates only a marginal negative effect: Second-born boys whose older brother has a STEM background are somewhat less likely to choose to pursue a STEM degree themselves, 3 percentage points (effect size ≈ - 6%)[10], compared to boys with an older sister who has a STEM degree.

These results indicate that second-born girls niche towards their older sister with a STEM background, but the same finding is not found between second-born boys and their older brothers.

5.4. Sensitivity tests

As discussed above, all those who completed at least one full semester of a STEM program are defined as having a STEM educational background. The main reason for this wide definition is the strong demand for STEM skills in the labor market, resulting in many STEM students (especially engineering students) finding employment before obtaining their final degree. The question emerges as to whether the result changes if only siblings who successfully attained a degree in a STEM field are included in the analysis sample. Table 5 presents the results of the two sets of hypotheses (i.e., role modeling and sibling competition/support) based on results from the

[7] This effect size is calculated by dividing the *point estimates* with the *mean value of the outcome* (the proportion of **second born girls** with STEM) and multiplying with 100: (0.133/0.165) * 100 ≈ 80%.

[8] This effect size is calculated by dividing the *point estimates* with the *mean value of the outcome* (the proportion of **second born boys** with STEM) and multiplying with 100: (0.204/0.499) * 100 ≈ 41%.

[9] This effect size is calculated by dividing the *point estimates* with the *mean value of the outcome* (the proportion of **second born girls** with STEM) and multiplying with 100: (0.068/0.165) * 100 ≈ 41%.

[10] This effect size is calculated by dividing the *point estimates* with the *mean value of the outcome* (the proportion of **second born boys** with STEM) and multiplying with 100: (−0.029/0.499) * 100 ≈ −6%.

Table 4
Tests for role modelling and sibling competition hypotheses.

	Coefficients	Estimates	Std. E.	t	P > |t|	99% Conf. Interval	
Panel A: Role modelling hypotheses							
H.2a	$\beta_1 + \beta_2 + \beta_3 > \beta_2$	0.133	0.007	19.23	0.000	0.116	0.152
H.2b	$\beta_1 + \beta_2 + \beta_3 > \beta_2$	0.204	0.006	33.73	0.000	0.188	0.219
Panel B: Sibling competition/support hypotheses							
H.3a	$\beta_1 + \beta_2 + \beta_3 > \beta_1$	0.068	0.007	9.37	0.000	0.051	0.088
H.3b	$\beta_1 + \beta_2 + \beta_3 > \beta_1$	−0.029	0.008	−3.54	0.000	−0.050	−0.008

Note: Table 4 shows the linear combination of coefficients according to the study's hypotheses, based on results from M(6) and M(10) in Table 3. The coefficients have the following interpretation.
Ref = having an older sister (brother).
β_1 = having an older sister (brother) who has a STEM degree.
β_2 = having an older brother (sister).
$\beta_1 + \beta_2 + \beta_3$ = having an older brother (sister) who has a STEM degree.

interaction model in Table A3 in the Appendix.

These results indicate that this sample restriction does not alter the findings reported with the wider sample definition. Thus, second-born siblings are still more likely to follow in their older sibling's footsteps. Moreover, second-born girls tend to niche toward older sister with a STEM degree, and although the results show that second-born boys tend to niche away from their older brothers, the magnitude of the result remains small. Furthermore, it should be noted that the magnitude of the combined association has increased compared to the wider sample, indicating that when the sample definition is restricted to having completed a STEM degree, the magnitude of results increases.

The reported findings do not appear to be an effect of regression technique choice. As Table 6 indicates, a logistic regression approach produces similar results to those discussed here (see also Table A4 in the Appendix).

6. Discussion

The aim of this study was to analyze the extent to which tertiary educational choices are associated by sibling gender configuration in Swedish two-child families. Three sets of hypotheses were tested: i) whether younger siblings are more likely to pursue a STEM degree if their older siblings have a STEM background compared to an older sibling with a different educational background, ii) whether having a same-sex sibling with or without a STEM background is more likely to shape second-born siblings' choice to pursue a STEM degree themselves, and iii) whether second-born siblings are more influenced by their same or opposite-sex sibling's STEM educational path. Swedish population register data is used to analyze these three sets of hypotheses. As the Swedish educational system has no tuition fees at any levels, all students receive student aid in the form of grants, and student housing options exist, the results of this analysis are less likely to be influenced by the family's economic constraints.

The presented findings indicate that second-born siblings are more likely to pursue a STEM degree if their older siblings have a STEM background as compared to an older sibling with a different educational background, net of parental characteristics such as level of education, occupation, and STEM background. Both second-born girls and boys have a higher likelihood of following in their older siblings' STEM footsteps; however, as STEM fields remain male-dominated, there is reason to investigate whether the gender of the older sibling shapes younger siblings' decision to pursue a STEM degree. Girls are more likely to pursue a STEM degree themselves if

Table 5
Sensitivity tests for role modelling and sibling competition hypotheses (only completed degree).

	Coefficients	Estimates	Std. E.	t	P > |t|	99% Conf. Interval	
Panel A: Role modelling hypotheses							
H.2a	$\beta_1 + \beta_2 + \beta_3 > \beta_2$	0.145	0.009	17.31	0.000	0.124	0.168
H.2b	$\beta_1 + \beta_2 + \beta_3 > \beta_2$	0.215	0.007	28.51	0.000	0.195	0.234
Panel B: Sibling competition/support hypotheses							
H.3a	$\beta_1 + \beta_2 + \beta_3 > \beta_1$	0.074	0.009	8.42	0.000	0.053	0.098
H.3b	$\beta_1 + \beta_2 + \beta_3 > \beta_1$	−0.027	0.010	−2.81	0.005	−0.052	−0.002

Note: Table 5 shows the linear combination of coefficients according to the study's hypotheses, based on results from M4 and M8 in Table A3 in the Appendix. The coefficients have the following interpretation.
Ref = having an older sister (brother).
β_1 = having an older sister (brother) who has a STEM degree.
β_2 = having an older brother (sister).
$\beta_1 + \beta_2 + \beta_3$ = having an older brother (sister) who has a STEM degree.

Table 6
Sensitivity tests for role modelling and sibling competition hypotheses (Logistic regression).

	Coefficients	Estimates	Std. E.	z	P > |z|	99% Conf. Interval	
Panel A: Role modelling hypotheses							
H.2a	$\beta_1 + \beta_2 + \beta_3 > \beta_2$	2.31	0.058	20.00	0.000	2.160	2.459
H.2b	$\beta_1 + \beta_2 + \beta_3 > \beta_2$	2.35	0.061	32,90	0.000	2.202	2.518
Panel B: Sibling competition/support hypotheses							
H.3a	$\beta_1 + \beta_2 + \beta_3 > \beta_1$	1.390	0.038	14.91	0.000	1.292	1.489
H.3b	$\beta_1 + \beta_2 + \beta_3 > \beta_1$	0.880	0.032	−3.37	0.000	0.802	0.965

Note: Table 6 shows the coefficients according to the study's hypotheses, based on results from M4 and M8 in Table A4 in the Appendix. The coefficients have the following interpretation.
Ref = having an older sister (brother).
β_1 = having an older sister (brother) who has a STEM degree.
β_2 = having an older brother (sister).
$\beta_1 + \beta_2 + \beta_3$ = having an older brother (sister) who has a STEM degree.

their older sister has a STEM background compared to girls who have an older brother. The corresponding result for second-born boys is not found. Instead, having an older brother with a STEM background decreases the probability for second-born boys to pursue a STEM degree themselves compared to having an older brother without a STEM background. However, it should be noted that the magnitude of this result is quite small and may in turn indicate that for younger boys, the association of role models is not gendered.

This study finds, for instance, that second-born girls who have an older sibling with STEM have 13 percentage points higher probability to select a STEM field at tertiary level, compared to second-born girls whose older siblings have not studies STEM field. By calculating the effect size for this group, which corresponds to 80%, we can conclude that the association is large.

The results of this analysis indicate that the family, more specifically siblings, shape individuals' educational choices with regard to a preference for STEM fields. Recent research has shown that gender segregation in STEM fields is related to mathematics performance only to a very small extent (Mann and DiPrete 2013). This finding suggests that social environment factors shape how children and adolescents develop interest and abilities associated with STEM fields. If there is a societal desire to provide both boys and girls with the same opportunity in terms of labor market performance, policy makers should focus on how to counteract the transmission of educational gender norms within the family. One specific example would be for schools to expose students to female role-models occupying gender atypical occupations, i.e. mathematicians or engineers, in conjunction with career day.

Admittedly, register data lacks information on attitudes concerning tertiary educational choices, complicating a test of the exact mechanism behind the reported findings, for instance whether older siblings act as role models. However, following social cognitive career theory and revised sex minority hypothesis, sibling role modeling and sibling competition might be expected as reasonable explanations. For both girls and boys, the results presented here are in line with a role modeling hypothesis. The findings suggest that same-sex sibling constellations are positively associated with second-born girls' decision to follow their older sisters' STEM path. One mechanism might be that having the same sex as one's older sister creates an environment that motivates and encourages competition amongst sisters, therefore influencing second-born girls' educational choices (Conley 2000). Another possible mechanism is the support that an older sister provides, encouraging younger girls to overcome gender-based challenges and normalizes feelings of not belonging in a gender atypical educational program (Lent et al., 2000). Some research indicates that support systems function as important mechanisms for women, such as a female role model during university studies (Herrmann et al., 2016).

This same finding was not evident for second-born boys, however, who tend to niche away from their older brothers (compared to sisters) with STEM backgrounds. One explanation might be that in the presence of such strong societal influence for men to pursue a STEM degree, having a same- or opposite-sex sibling does not exert any additional impact on a second-born boy, while women seem to be in need of an extra "push" (in the form of an older female sibling role model) to pursue a STEM degree.

It is interesting to note that the results for second-born siblings' choice of degree is net of parental characteristics, indicating that the relevance of the association between sibling influence and understanding their educational choices, alongside the influence parents exert. An interesting avenue for future research would be to investigate the exact mechanism behind the influence of older siblings on women's choice of STEM educational fields. Although competition between siblings is proposed as a mechanism for educational outcomes (Conley 2000), support and encouragement from older siblings to pursue a gender atypical degree could also be a possible mechanism (Lent et al., 2000). To test this, however, prospective information on sibling relationships during childhood and adolescence are necessary. Furthermore, since social norms are not entirely static, it is critical for future research to document and analyze the importance of older siblings' examples for their younger siblings' educational choices.

Funding

This work was supported by the Swedish Research Council (VR) (Dnr: 2008:7499), and the Swedish Research Council for Health, Working Life and Welfare (FORTE) (Dnr: 2013-1119).

Acknowledgement

I have benefited greatly from comments on previous versions by Magnus Nermo, Erik Bihagen, Katarina Boye, Robert Erikson, Anette Fasang, Linus Andersson, Juho Härkönen, Anders Stenberg, Mohammad H. Sephavand, Iman Dagdar, discussion with Martin Kolk and Limor Gabay-Egozi, and comment by three anonymous reviewers. I would like to also thank the seminar participants at the Swedish Institute for Social Research, September 2016, and the Summer Meeting of Research Committee 28 on Social Stratification and Mobility (RC28), New York, August 2017.

Appendix A

Table A1
STEM degrees, according to Statistic Sweden's code name. *SUN2000grupper*.

		STEM	
		Not completed degree	Completed degree
44P	Finished at least 1 term in a program within science, mathematics and data (no degree)	x	
45D	Programming and system informatics	x	x
45N	Degrees in natural science at university	x	x
45X	Other Degrees in science, mathematics and data	x	x
54P	Finished at least 1 term in a program within engineering (no degree)	x	
55C	Engineering: Road and water, building, etc. (≈4.5 years)	x	x
55D	Engineering: Machine, vehicle/craft, industrial, etc. (≈4.5 years)	x	x
55E	Engineering: Electrical, technical physics and data (≈4.5 years)	x	x
55F	Engineering: Chemical, biotechnology, etc. (≈4.5 years)	x	x
55G	Engineering: Others (≈4.5 years)	x	x
55H	Engineering: Road and water, building (≈3 years)	x	x
55I	Engineering: Machine, vehicle/craft, industrial, etc. (≈3 years)	x	x
55J	Engineering: Electrical, technical physics and data (≈3 years)	x	x
55K	Engineering: Chemical and biotechnology (≈3 years)	x	x
55L	Engineering: Others (≈3 years)	x	x

Table A2
Descriptives for some of the control variables used in the analyses

Variable name	%	N	%	N	%	N
	First-born child		Second-born child		ALL	
Female	52%	63,146	53%	64,422	52%	127,568
GPA (rank: 1–100)	65	114,230	64	113,884	65	228,114
Adopted	0%	205	0%	226	0%	431
Born in Sweden	98%	119,009	98%	119,794	98%	238,803
Biological two-child family	86%	105,194	86%	105,194	86%	210,388
Living with both parents through age 15	74%	89,944	79%	96,556	76%	186,500
Living with both parents through age 14	4%	5478	2%	2496	3%	7974
Living with both parents through age 13	4%	4744	2%	1900	3%	6644
Living with both parents through age 12	3%	4158	1%	1787	2%	5945
Living with both parents through age 11	3%	3276	1%	1625	2%	4901
Living with both parents through age 10	2%	2777	1%	1545	2%	4322
Living with both parents through age 9	1%	1647	1%	1581	1%	3228
Living with both parents through age 8	1%	1743	1%	1582	1%	3325
Did not live with both parents ages 8-15	7%	8137	11%	12,832	9%	20,969
Birth spacing < 24 months	13%	15,704	13%	15,704	13%	31,408
Birth spacing 24–35 months	29%	34,860	29%	34,860	29%	69,720
Birth spacing 36–47 months	26%	31,990	26%	31,990	26%	63,980
Birth spacing 48–59 months	15%	18,506	15%	18,506	15%	37,012
Birth spacing 60–71 months	8%	9382	8%	9382	8%	18,764
Birth spacing 72–83 months	4%	5076	4%	5076	4%	10,152
Birth spacing 84–95 months	2%	2824	2%	2824	2%	5648
Birth spacing > 95 months	3%	3562	3%	3562	3%	7124
Household income rank ages 8-15	67	115,327	69	118,112	68	233,439
Mother STEM education	1%	1746	1%	1746	1%	3492
Father STEM education	11%	13,924	11%	13,924	11%	27,848
Mother education:						
Primary edu	17%	20,341	17%	20,341	17%	40,682
Upper secondary edu	39%	47,251	39%	47,251	39%	94,502
Post upper secondary edu (<2 years)	3%	3302	3%	3302	3%	6604
Post upper secondary edu (2 years)	14%	16,799	14%	16,799	14%	33,598

(*continued on next page*)

Table A2 (*continued*)

Variable name	%		N	%	N	%	N
First-born child		Second-born child		ALL			
Post upper secondary edu (3 years)	15%		18,722	15%	18,722	15%	37,444
Post upper secondary edu (5 years)	8%		9943	8%	9943	8%	19,886
Post upper secondary edu (>5 years)	2%		2132	2%	2132	2%	4264
Having a PhD degree	1%		1279	1%	1279	1%	2558
Missing value	2%		2135	2%	2135	2%	4270
Father education:							
Primary edu	19%		22,872	19%	22,872	19%	45,744
Upper secondary edu	38%		46,393	38%	46,393	38%	92,786
Post upper secondary edu (<2 years)	5%		5606	5%	5606	5%	11,212
Post upper secondary edu (2 years)	8%		9981	8%	9981	8%	19,962
Post upper secondary edu (3 years)	10%		12,438	10%	12,438	10%	24,876
Post upper secondary edu (5 years)	10%		12,242	10%	12,242	10%	24,484
Post upper secondary edu (>5 years)	2%		2735	2%	2735	2%	5470
Having a PhD degree	3%		3883	3%	3883	3%	7766
Missing value	5%		5754	5%	5754	5%	11,508

Note: this table was compiled by the author.

Table A3
LPM regression coefficients for the association between sibling sex configuration and STEM programs for second-born girls and boys. Only completed degree.

Dependent variables	Second-born girl				Second-born boys			
	M. 1	M. 2	M. 3	M. 4	M. 5	M. 6	M. 7	M. 8
Older SIBLING STEM	0.11**	0.13**	0.11**	0.09**	0.18**	0.21**	0.16**	0.14**
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Having an older Sister		0.04**	0.02**	0.02**		–	–	–
		(0.00)	(0.00)	(0.00)		–	–	–
** Older SIBLING STEM*			0.07**	0.05**			–	–
			(0.01)	(0.01)			–	–
Having an older Brother		–	–	–		−0.08**	−0.11**	−0.10**
		–	–	–		(0.01)	(0.01)	(0.01)
** Older SIBLING STEM*			–	–			0.08**	0.07**
			–	–			(0.01)	(0.01)
Controls	No	No	No	Yes	No	No	No	Yes
Constant	0.13**	0.10**	0.12**	0.09**	0.46**	0.49**	0.50**	0.44**
	(0.00)	(0.00)	(0.00)	(0.03)	(0.00)	(0.00)	(0.00)	(0.05)
Observations	43,904	43,904	43,904	43,904	37,736	37,736	37,736	37,736
R-squared	0.020	0.022	0.024	0.055	0.028	0.034	0.035	0.074

Note:**$p < 0.01$, *$p < 0.05$. The follow controls are included: year of birth dummies, birth spacing interval dummies, GPA rank, adopted sibling, biological two-child families, living with both parents, born in Sweden, mother having STEM degree, father having STEM degree, mothers educational level, fathers educational level, mothers occupation around 15, fathers occupation around 15, household $ rank (ages 8–15) & residential region.

Table A4
Logistic regression (odds-ratio) coefficients for the association between sibling sex configuration and STEM programs for second-born girls and boys.

Dependent variables	Second-born girl				Second-born boys			
	M. 1	M. 2	M. 3	M. 4	M. 5	M. 6	M. 7	M. 8
Older SIBLING STEM	2.03**	2.29**	1.83**	1.75**	1.98**	2.30**	1.88**	1.78**
	(0.04)	(0.06)	(0.06)	(0.06)	(0.04)	(0.05)	(0.06)	(0.06)
Having an older Sister		1.33**	1.15**	1.14**		–	–	–
		(0.03)	(0.02)	(0.02)		–	–	–
** Older SIBLING STEM*			1.38**	1.26**			–	–
			(0.04)	(0.04)			–	–
Having an older Brother		–	–	–		0.71**	0.65**	0.66**
		–	–	–		(0.01)	(0.01)	(0.01)
** Older SIBLING STEM*			–	–			1.37**	1.33**
			–	–			(0.06)	(0.06)
Controls	No	No	No	Yes	No	No	No	Yes
Constant	0.15**	0.12**	0.16**	0.15**	0.80**	0.89**	0.92**	0.70**
	(0.00)	(0.00)	(0.00)	(0.02)	(0.01)	(0.01)	(0.01)	(0.10)
Observations	64,422	64,422	64,422	64,422	57,482	57,482	57,482	57,482
Pseudo R2	0.018	0.021	0.022	0.045	0.017	0.023	0.024	0.057

Note: **$p < 0.01$, *$p < 0.05$. The follow controls are included: year of birth dummies, birth spacing interval dummies, GPA rank, adopted sibling, biological two-child families, living with both parents, born in Sweden, mother having STEM degree, father having STEM degree, mothers educational level, fathers educational level, mothers occupation around 15, fathers occupation around 15, household $ rank (ages 8–15) & residential region.

References

Alon, Sigal, Gelbgiser, Dafna, 2011. The female advantage in college academic achievements and horizontal sex segregation. Soc. Sci. Res. 40 (1), 107–119.
Amin, Vikesh, 2009. Sibling sex composition and educational outcomes: a review of theory and evidence for the UK. Lab. Travail 23 (1), 67–96.
Åslund, Olof, Grönqvist, Hans, 2010. Family size and child outcomes: is there really no trade-off? Lab. Econ. 17 (1), 130–139.
Bandura, Albert, 1986. Social Foundations of Thought and Action: A Social Cognitive Theory. Prentice-Hall, Inc, Englewood Cliffs, NJ, US.
Barclay, Kieron, Hällsten, Martin, Myrskylä, Mikko, 2017. Birth order and college major in Sweden. Soc. Forces 96 (2), 629–660.
Barone, Carlo, 2011. Some things never change: gender segregation in higher education across eight nations and three decades. Sociol. Educ. 84 (2), 157–176.
Blake, Judith, 1981. Family size and the quality of children. Demography 18 (4), 421–442.
Boli, John, Allen, Mary Lou, Payne, Adrienne, 1985. High-ability women and men in undergraduate mathematics and chemistry courses. Am. Educ. Res. J. 22 (4), 605–626.
Bradley, Karen, 2000. The incorporation of women into higher education: paradoxical outcomes? Sociol. Educ. 1–18.
Brim, Orville G., 1958. Family structure and sex role learning by children: a further analysis of Helen Koch's data. Sociometry 21 (1), 1–16.
Buchmann, Claudia, DiPrete, Thomas A., 2006. The growing female advantage in college completion: the role of family background and academic achievement. Am. Socio. Rev. 71 (4), 515–541.
Buckles, Kasey S., Munnich, Elizabeth L., 2012. Birth spacing and sibling outcomes. J. Hum. Resour. 47 (3), 613–642.
Buser, Thomas, Niederle, Muriel, Oosterbeek, Hessel, 2014. Gender, competitiveness, and career choices. Q. J. Econ. 129 (3), 1409–1447.
Bussey, K., Bandura, A., 1999. Social cognitive theory of gender development and differentiation. Psychol. Rev. 106 (4), 676.
Caprile, M., Palmén, R., Sanz, P., Dente, G., 2015. Encouraging STEM Studies for the Labour Market. Directorate General for Internal Policies, European Union.
Ceci, Stephen J., Williams, Wendy M., 2010. Sex differences in math-intensive fields. Curr. Dir. Psychol. Sci. 19 (5), 275–279.
Charles, Maria, Bradley, Karen, 2002. Equal but separate? A cross-national study of sex segregation in higher education. Am. Socio. Rev. 573–599.
Charles, Maria, Bradley, Karen, 2009. Indulging our gendered selves? Sex segregation by field of study in 44 countries. Am. J. Sociol. 114 (4), 924–976.
Conley, Dalton, 2000. Sibship sex composition: effects on educational attainment. Soc. Sci. Res. 29 (3), 441–457.
Dabney, Katherine P., Chakraverty, Devasmita, Tai, Robert H., 2013. The association of family influence and initial interest in science. Sci. Educ. 97 (3), 395–409.
Davies, Scott, Guppy, Neil, 1997. Fields of study, college selectivity, and student inequalities in higher education. Soc. Forces 75 (4), 1417–1438.
Downey, Douglas B., 1994. The school performance of children from single-mother and single-father families: economic or interpersonal deprivation? J. Fam. Issues 15 (1), 129–147.
Fouad, Nadya, A., Hackett, Gail, Smith, Philip L., Kantamneni, Neeta, Fitzpatrick, Mary, Haag, Susan, Spencer, Dee, 2010. Barriers and supports for continuing in mathematics and science: gender and educational level differences. J. Vocat. Behav. 77 (3), 361–373.
Frehill, Lisa M., 1997. Education and occupational sex segregation: the decision to major in engineeringxs. Socio. Q. 38 (2), 225–249.
Gabay-Egozi, L., Grieger, L., Nitsche, N., 2016. Does the Likelihood that a Younger Sibling Will Pursue a STEM Major Depend on the Gender of Other Siblings? *Manuscript presented at American Sociological Association 2016*.
Hakim, Catherine, 2002. Lifestyle preferences as determinants of women's differentiated labor market careers. Work & Occup. 29 (4), 428–459.
Halldén, Karin, 2008. The Swedish educational system and classifying education using the ISCED-97. In: Schneider, Silke (Ed.), The International Standard Classification of Education (ISCED-97). An Evaluation of Content and Criterion Validity in 15 European Countries. MZES, Mannheim, pp. 253–267.
Halpern, Diane F., Camilla P Benbow, David C Geary, C Gur, Ruben, Hyde, Janet Shibley, Morton, Ann Gernsbacher, 2007. The science of sex differences in science and mathematics. Psychol. Sci. Publ. Interest 8 (1), 1–51.
Härkönen, Juho, 2013. Birth order effects on educational attainment and educational transitions in West Germany. European Sociological Review, p. jct027.
Herrmann, Sarah D., Adelman, Robert Mark, E Bodford, Jessica, Graudejus, Oliver, Morris, A Okun, Kwan, Virginia SY., 2016. The effects of a female role model on academic performance and persistence of women in STEM courses. Basic Appl. Soc. Psychol. 38 (5), 258–268.
Hjalmarsson, Simon, Mood, Carina, Shahbazian, Roujman, 2015. Kodbok För Inkomstvariabler 1968-89 (I*n English: Codebook for Income Variables 1968-89*). The Swedish Institute for Social Research.
Högskoleverket, 2012. Swedish Universities and University Colleges: Short Version of Annual Report 2012. Högskoleverket, Stockholm.
Joensen, Juanna Schrøter, Nielsen, Helena Skyt, 2018. Spillovers in education choice. J. Publ. Econ. 157, 158–183.
Legewie, Joscha, DiPrete, Thomas A., 2014. The high school environment and the gender gap in science and engineering. Sociol. Educ. 87 (4), 259–280.
Lent, Robert W., 2013. "Career-life preparedness: revisiting career planning and adjustment in the new workplace. Career Dev. Q. 61 (1), 2–14.
Lent, Robert W., Brown, Steven D., Hackett, Gail, 1994. Toward a unifying social cognitive theory of career and academic interest, choice, and performance. J. Vocat. Behav. 45 (1), 79–122.
Lent, Robert W., Brown, Steven D., Hackett, Gail, 2000. Contextual supports and barriers to career choice: a social cognitive analysis. J. Counsel. Psychol. 47 (1), 36.
Lent, Robert W., Brown, Steven D., Hackett, Gail, 2002. Social Cognitive Career Theory, vol. 4. Career choice and development, pp. 255–311.
Leppel, Karen, Williams, Mary L., Waldauer, Charles, 2001. The impact of parental occupation and socioeconomic status on choice of college major. J. Fam. Econ. Issues 22 (4), 373–394.
Lindahl, Britt, 2003. Lust att lära naturvetenskap och teknik? En longitudinell studie om vägen till gymnasiet (Göteborg Studies in Educational Sciences, 196.) Göteborg, Acta Universitatis Gothoburgensis.
Mammen, Kristin, 2011. Fathers' time investments in children: do sons get more? J. Popul. Econ. 24 (3), 839–871.
Mann, Allison, DiPrete, Thomas A., 2013. Trends in gender segregation in the choice of science and engineering majors. Soc. Sci. Res. 42 (6), 1519–1541.
Mood, Carina, 2010. Logistic regression: why we cannot do what we think we can do, and what we can do about it. Eur. Socio Rev. 26 (1), 67–82.
Price, Joseph, 2008. Parent-child quality time does birth order matter? J. Hum. Resour. 43 (1), 240–265.
Seehuus, Sara, Reisel, Liza, 2017. Betydningen av sosial bakgrunn for kjønnsdeling i høyere utdanning. Tidsskr. Samfunnsforskning 58, 284–310, 03.
Shahbazian, Roujman, 2020. The consequence of birth spacing for first- and second-born siblings' long-term income rank: a restrictive two-child family approach. J. Family Res. 32 (1), 72–104.
Statistics Sweden, 2017. Statistical Database, Access 7 April 2018. http://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START_LE_LE0102_LE0102J/LE0102T19/?rxid=ffbf3169-abe6-4288-9733-6a32d86a62e1.
Steelman, Lala Carr, Powell, Brian, Werum, Regina, Carter, Scott, 2002. Reconsidering the effects of sibling configuration: recent advances and challenges. Annu. Rev. Sociol. 28 (1), 243–269.
Sulloway, Frank J., 2007. Birth order." Family Relationships: an Evolutionary Perspective, pp. 162–182.
Webber, Douglas A., 2014. The lifetime earnings premia of different majors: correcting for selection based on cognitive, noncognitive, and unobserved factors. Lab. Econ. 28, 14–23.
Zajonc, Robert B., 1976. Family configuration and intelligence. Science 192 (4236), 227–236.
Zajonc, Robert B., Markus, Gregory B., 1975. Birth order and intellectual development. Psychol. Rev. 82 (1), 74.