MAX PLANCK INSTITUTE
FOR DEMOGRAPHIC RESEARCH

In My Brother's Footstep or Shadow? Siblings' Compositional Characteristics and Gender Differences in STEM Major

Limor Gabay-Egozi ᛁ limor.gabay-egozi@biu.ac.il
Lloyd Grieger ᛁ lloyd.grieger@gmail.com
Natalie Nitsche ᛁ nitsche@demogr.mpg.de

In My Brother's Footstep or Shadow? Siblings' Compositional Characteristics and Gender Differences in STEM Major

Limor Gabay-Egozi, Lloyd Grieger, and Natalie Nitsche

Limor Gabay-Egozi
Department of Sociology and Anthropology
Bar Ilan University, Ramat Gan, Israel
Tel: +972-3-5318964   Mobile: +972-54-2754145
Email address: limor.gabay-egozi@biu.ac.il
(Corresponding author)

Lloyd Grieger
University of Hawai'i at Mānoa
Sinclair Library, Room 504
2425 Campus Road, Honolulu, HI 96822
Email address: lloyd.grieger@gmail.com

Natalie Nitsche
Max Planck Institute for Demographic Research (MPIDR)
Konrad-Zuse-Str. 1
18057 Rostock, Germany
Tel: +49 381 2081-257
Email address: nitsche@demogr.mpg.de

Authors Affiliation and Research Interests

LIMOR GABAY-EGOZI is assistant professor in the Department of Sociology and Anthropology at Bar-Ilan University, Israel. Her research involves social and educational stratification and inequality, focusing primarily on rational choice in education, life-course dynamics, gender, class, and ethnic inequality, and how sib-group structure influences educational attainment. Her recent work appeared in Social Science Research, European Sociological Review, Socius, Acta Sociologica, the Journal of Education Policy, and the Journal of Law and Society.

LLOYD GRIEGER is a data analyst in the HAWAI'I P-20 Partnerships for Education led by the Executive Office on Early Learning, the Hawai'i State Department of Education and the University of Hawai'i System. His research interests include social demography, poverty, affluence and inequality, sociology of the family, social stratification, social welfare policy, young adulthood, and quantitative methods. His work has appeared in Demography, Journal of Children and Poverty, Journal of Economic and Social Measurement, Focus, and Development Southern Africa.

NATALIE NITSCHE is deputy head of the Laboratory of Fertility and Well-Being at the Max Planck Institute for Demographic Research. She is a social demographer with research interests in family formation, fertility, work-divisions in families, and gender inequality. Her research has been published in journals including Demography, the European Sociological Review, and Advances in Life Course Research.

Acknowledgements

In My Brother's Footstep or Shadow? Siblings' Compositional Characteristics and Gender Differences in STEM Major

Abstract

While significant associations between siblings' compositional characteristics and years of schooling are well documented, little is known on whether siblings' features may be relevant for substantive educational preferences. Using the NLSY-79 Mothers and Children Files (N=1545; 57% females; 22% STEM-major), LPM analysis revealed that siblings' compositional characteristics – sibship size, gender composition, and the sex and math ability of the older sibling - matter for a STEM-major preference in college, but for young women only. Net of other compositional features, young females were more likely to prefer a STEM major if they were raised in smaller sib-groups, in male sib-group dominance, and with a high math achieving older sister rather than a highly math achieving older brother. These results correspond with the resource dilution approach, illuminate the role of a normative male-role sib-group climate, and suggest an interplay between role modeling and same-sex competitive stimulation in shaping gendered educational outcomes. For young men, STEM major preference is mostly driven by their own math ability. Our findings underscore the importance of siblings in creating gendered educational preferences and later life trajectories, and that social dynamics between siblings differ for males and females.

Keywords: *STEM, Sibling compositional characteristics, Older sibling, Role modeling, Same-sex competitive stimulation, A normative male-role sib-group climate.*

Introduction

Sociologists, psychologists, and economists have long studied the link between sibling group (hence forth: sib-group) compositional characteristics and individual educational attainment. Sib-group size, birth ordering, birth spacing, and sex composition have received extensive attention as potential sources of (dis)advantage for education outcomes such as years of completed schooling or grade point average (see Steelman et al. 2002 for a review). Most of this research, however, has focused on examining whether the resource dilution hypothesis is predictive of time spent in the educational system (Anastasi 1956; Black et al. 2005; Blake 1981, 1986, 1989; Downey 1995, 2001; Powell and Steelman 1990, 1993; Hertwig et al. 2002). Indeed, net of other features of the sib-group (i.e., birth ordering, spacing, and sex composition), having more siblings can be a disadvantage since siblings dilute parental resources (Blake 1981, Coleman 1988, Conley 2004). However, parents are not the only resourceful and influential family members – developmental psychology demonstrates that siblings also assert influence on each other while growing up, in particular regarding social and attitudinal learning (McHale et al. 2001; Whiteman et al. 2009). Therefore, in addition to the perspective of resource dilution, we draw from social learning aspects embedded in role modeling perspectives (Brim 1958; Dunn 1983) and the same-sex competitive stimulation approach (Conley 2000) to derive hypotheses on how sib-group compositional characteristics may be linked to young adults' decision to later pursue a STEM major - science, technology, engineering and math - in college. Research on whether and how sib-group features and dynamics may be relevant for gender differences in college major preferences is rare.

Theoretical Frameworks

Historically, gender differences in college attendance and completion in the United States have favored males over females but have narrowed, disappeared or even reversed in some sub-populations (DiPrete and Buchmann 2013). Despite women surpassing men in tertiary education attendance and completion, women today are still less likely than men to major in STEM fields, even as they have achieved parity in the fields of medicine, law and chemistry (Alon and Gelbgiser 2010; Charles and Bradley 2009; National Science Foundation 2013). As a result, despite making up nearly half of the U.S. workforce and half of the college-educated workforce, women hold less than 25 percent of STEM-type jobs (Beede et al. 2011). This horizontal gender segregation contributes to persisting income inequality between men and women because STEM occupations tend to be higher paying than the helping and service-type careers in which female college graduates are overrepresented (Beede et al. 2011; Charles and Grusky 2004). Recent work suggests that the conventional narratives explaining gender inequality in college major, such as relative academic performance, gendered occupational aspirations, work-family attitudes and general life goals (Eccles 2011a, 2011b; Hakim 2002; Jacobs 1989, 1995; Jonsson 1999), explain only small portions of the gender gap in the choice of a STEM major (Mann and DiPrete 2013; Morgan et al. 2013; Riegle-Crumb et al. 2012).

Students' family background and characteristics are known to influence interest in and motivation towards a STEM career (Torche 2011). It has been further argued that the differential filtration of females into STEM careers is an outcome of gender role socialization during childhood which is mediated by the family. Parents expose and provide their children insights and lessons about gender roles, encourage their sons and daughters to perform and engage in specific behaviors, activities, and tasks, as well as offer emotional and financial support for engaging in mathematical and technical activities (see Fouad et al. 2010 for a review, Bussey and Bandura 1999; Lytton and Romney 1991). These social experiences create bridges and barriers to developing an interest in STEM; experiences which differ for boys and girls (e.g., Crouter at al. 1995, 2007; McHale at al. 2003). As females grow up, their participation in STEM-related activities substantially decreases. By the time girls are in second grade, boys outperform them in science (yet not in math) and girls already begin to exhibit the belief that math and science are for boys (Legewie and DiPrete 2014). From this standpoint, the gender gap in STEM majors is no surprise as parents are more likely to support, expect from, and encourage their sons rather than daughters to be interested in and to pursue a STEM activity (Fouad et al. 2010; Author A and—— 2015).

Parents, nonetheless, are not the only influential family members – siblings are another socialization source. Even though siblings share genes and often a similar environment by living at the same home during their childhood, sibling outcomes are greatly diverse (Björklund and Jäntti 2012). This sibling divergence is also due to the fact that each individual, depending on sibling compositional characteristics, experiences the social environment within the family differently. Different theoretical approaches have been formulated to link sib-group compositional characteristics to individual educational attainment. We present a three-fold theoretical framework to postulate on how sib-group compositional characteristics may be linked to young females' and males' decision to later pursue a STEM major in college.

Resource dilution hypothesis: Sib-group size

Most researchers articulate this link between siblings' compositional characteristics and educational attainment via a resource dilution hypothesis (Anastasi 1956; Black et al. 2005; Blake 1981; Downey 1995, 2001; Powell and Steelman 1990, 1993; Hertwig et al. 2002). Siblings from smaller families and older siblings tend to hold an advantage over siblings from larger families and younger counterparts in access to parental financial resources (De Haan 2010), time investments (Price 2008), energy, monitoring and supervision efforts (Hotz and Pantano 2015). This resource differentiation between siblings appears to result in lower educational attainment for children from larger families and for later-born siblings (Black et al. 2005 for Norway; Björklund and Jäntti 2012 for Sweden; Booth and Kee 2005 for England; Powell and Steelman 1990, 1993; Price 2010 for the US). Moreover, longer spacing between siblings allows for greater parental investment in older children, contributing to the positive link between spacing and academic outcomes (Bu 2014; Buckles and Munnich 2012; Galbraith 1982; Kidwell 1981; Powell and Steelman 1990, 1993, 1995; Price 2010). Still, with wide spacing a last-born child might benefit more as older children leave the household or through the expansion of time inputs as both parents and older siblings spend time with the last-born child (Behrman and Taubman 1986; Hanushek 1992). Recent literature, however, indicates that birth spacing has no significant implication on educational attainment (Grätz 2018; Barclay and Kolk 2017).

The only recent study we know that tests the link between sibling compositional characteristics and college major, using Swedish register data, found that first-borns compared with later-born siblings are more likely to study college majors with greater expected earnings and greater expected occupational prestige (Barclay et al. 2017). Specifically, first-borns were more likely to study engineering, medicine, and life sciences in university and to graduate, but there were no statistically significant differences by birth order in applications and graduation from physics, math and statistics, or computing. However, since the Swedish education system provides free tertiary education, these results are not mainly driven by the exhaustion of family financial resources. Hence, this link could be more pronounced in a context where tuition fees are high as in the US. Following the resource dilution argument, *we expect that sib-group size negatively affects the likelihood of selecting a STEM-major in college (net of birth ordering and spacing), implying that individuals from larger sib-groups are less likely to choose a STEM major in college compared with those from smaller families* (Hypothesis 1).

Role modeling perspective: Older sibling math ability

In addition to the dynamics of parental resource distribution, siblings can have direct impacts on one another's educational development by serving as social partners and role models during everyday interactions (McHale et al. 2001; Whiteman et al. 2009). Siblings often spend time with one another, playing, bonding, arguing and quarrelling, this strong social attachment can produce a downstream positive role modeling, in which younger siblings imitate older siblings' constructive social behaviors and learn social skills and conflict resolution strategies. Studies of social learning show that older siblings are more influential role models than younger siblings (Brim 1958), and Dunn (1983) reported that, by the second year, second-born children imitate many behaviors of their first-born sibling. Older siblings usually possess the characteristics of effective socialization agents such as status, nurturance,

and similarity, and hence are being particularly influential models for their younger siblings (Bandura 1977). Interactions with older siblings, moreover, assist in developing children's social-cognitive skills such as emotional understanding, negotiation, persuasion and problem solving (Brown et al. 1996; Dunn 2007; Howe et al. 2002). This asymmetric effect between older and younger siblings is also affected by the organization of schooling by age (Hanushek 1992; Oettinger 2000). Older siblings are known to contribute to individuals' academic engagement during adolescence (Bouchey et al. 2010), offer social support and information transmission, and provide a longitudinal perspective of how educational journey looks like (Davies 2019). Social learning via sibling modeling leads to our expectation that a younger sibling learns from and imitates the characteristics of an older sibling. As such, we expect that older siblings with high interest/ability in science and math will encourage younger siblings to cultivate the same set of skills.

Alternatively, individuals may choose to pursue interests and abilities that are explicitly different from their older sibling. Competition for parental investment amongst children, Sulloway (1996) argues, causes siblings to adapt their behavior and develop a personality that would allow them to occupy particular roles within the family. In the scramble for parental investment from the very beginning, later-born children are forced to become more creative, original and follow risks in order to attract that investment. Following Sulloway's work, negative role modeling may take place, with younger siblings selecting different niches in the family by developing distinct qualities. Therefore, we hypothesize that *younger siblings with an older math-skilled sibling are more likely to opt for a STEM major in college* (Hypothesis 2a); or, alternatively, that *having a math-skilled older sibling is positively associated with a younger sibling's preference to opt for a non-STEM major in college* (Hypothesis 2b).

Competition hypothesis: Sib-group sex composition

Previous research on the relationship between sib-group sex composition and educational attainment focuses mainly on the amount of educational attainment and yields mixed results (Butcher and Case 1994; Conley 2000; Kaestner1997; Kuo and Hauser 1997; Powell and Steelman 1989, 1990). Many of the studies focus on explaining differences in educational attainment between family types (i.e. at the family-level of analysis), which, as ——and Author A (2013) point out, assumes that the effects of sib-group compositional characteristics are the same for women and men. In the few studies we know of that explicitly test how sib-group sex composition relates to educational attainment, none accounted for the possibility that group-level sex composition is experienced differently by girls and boys, nor did any offer theoretical explanations for their results (Bu 2014; Butcher and Case 1994; Kaestner 1997; Kuo and Hauser 1997; Powell and Steelman 1989, 1990).

One exception is Conley's (2000) finding that kids raised with siblings of the opposite sex have a lower level of educational attainment than their same-sex peers with same-sex siblings. Conley (2000) suggests that it is relatively disadvantageous to have siblings of the opposite sex, as same-sex siblings stimulate a competitive environment, which pushes children to perform better. This propensity for siblings of the same sex to compete with one another in similar tasks leads to higher outcomes

compared to individuals who lack a same-sex sibling.[1] Focusing on advanced math and science courses in high school, Joensen and Nielsen (2018) have shown that older siblings causally affect younger sibling's educational choices. This siblings' spillover effect is the strongest among closely spaced male siblings. The authors suggest that competition is likely one of the driving forces behind younger siblings conforming to their older siblings' choices. Following Conley's (2000) revised sex minority hypothesis we expect that *a girl with an older sister (and with more sisters) is more likely to opt for a STEM major in college than a girl who was raised with an older brother (and with more brothers).* Similarly, *a boy with an older brother (and with more brothers) is more likely to opt for a STEM field than a boy with an older sister (and with more sisters)* (Hypothesis 3).

Indeed, most research assumes positive asymmetrical associations in which older brothers and sisters promote and evoke similarity on the part of their younger siblings, yet recently a growing body of work recognizes more complicated differentiation dynamics (e.g. Whiteman et al. 2014). Alternatively, the positive framing that Conley attached to the competition among same-sex siblings may not be uniformly positive. It is possible for competition between siblings to be non-collaborative and even adversarial, potentially leading to less positive or to opposite outcomes, as implied by Sulloway's work (1996). As such, in choosing different niches and developing distinct personal qualities, youth are thought to protect themselves from siblings' rivalry and resentment and in turn receive their share of parental love and attention (Feinberg and Hetherington 2000; Schachter et al. 1976). Accounting for siblings' sex composition, these differentiation processes were found to be more prevalent when siblings were the same gender (Schachter et al. 1976, 1978; Whiteman et al. 2014). We continue our analysis while acknowledging this possibility, and we refer to this further in the discussion section.

Intersected hypotheses: Role modeling via same-sex competitive stimulation

The sib-group is a dynamic social institution and it is possible that the theoretical approaches and hypotheses we listed above may operate simultaneously and are not necessarily mutually exclusive. Social psychology research suggests, for example, that the impact of older sibling role modeling is enhanced with same-sex siblings (Brim 1958; Koch 1960). Accordingly, both boys and girls by age 3 with same-sex older siblings acted more 'sex-typed' than children with older opposite-sex siblings (McHale et al. 2003; McHale et al. 2012). Hence, namely if, and if so, we hypothesize that *the effect of a math skilled older sibling on a younger sibling's likelihood to opt for a STEM major in college is more pronounced in same-sex pairs* (Hypothesis 4). Younger siblings of an opposite sex older sibling with STEM or math affinity may respond with competitive specialization in other fields to find their own niche, implying negative role modeling. Nevertheless, it is well known that math and sciences are competitive disciplines that appeal more to males who are also known to be attracted to competition,

---

[1] Conley's research is a response to Rosenberg's (1965) sex minority hypothesis in which it is argued (without direct empirical support) that parents will have a greater attachment to children who are a minority in the sib-group with respect to sex, as they enjoy a special status in contrast to children in a same-sex sib-group. In other words, Rosenberg's theory predicts that girls with only brothers have better educational outcomes than girls who were raised with sisters since the former occupy a position of privilege.

while females tend to avoid mixed-sex competition and do worse in high-stakes mixed-sex competition domains (Niederle and Vesterlund 2007, 2010). Hence, it might be that the positive effect of role modeling on preferring STEM fields is present mostly among brothers.

The Current Study

At the core of our study is the question of how and to what extent sib-group compositional characteristics are associated with gender differences in STEM majoring later in life. Since we are interested in the spillover effect of sib-group features, we focus our attention on college major preference of respondents who are non-firstborn siblings ('focal' respondents). First-born children spend at least some part of the beginning of their life without other siblings – and some never have other siblings – while non-firstborn children, by contrast, will have older siblings as role models from birth on (Zajonc and Markus 1975). Theoretically, a younger sibling is born into a family after rules and roles have already been established with the first-born and are more likely to be influenced by the roles already adopted by other siblings. Sibling socialization, hence, is at least partially a top-down process in which older siblings serve as role models for younger siblings to look up to and imitate (Whiteman et al. 2014). We focus on the relationship between the focal child and the sibling born immediately before the focal child (henceforth, older sibling) in terms of sex-composition, spacing, and math ability while controlling for the presence of other siblings.[2] Math ability during adolescence is a predictor for later STEM preference, and we use math ability of the older sibling as an indicator for role modeling in this respect. To recap, testing whether and how sibling compositional characteristics have a lifelong intergenerational effect on younger siblings we suggest four hypotheses.

H1: Sib-group size negatively affects the likelihood of selecting a STEM-major in college (net of birth ordering and spacing), implying that individuals from larger sib-groups are less likely to choose a STEM major in college compared with those from smaller families.

H2: Younger siblings with an older math-skilled sibling are more likely to opt for a STEM major in college; or, alternatively, that having a math-skilled older sibling is positively associated with a younger sibling's preference to opt for a non-STEM major in college.

H3: A girl with an older sister (and with more sisters) is more likely to opt for a STEM major in college than a girl who was raised with an older brother (and with more brothers). Similarly, a boy with an older brother (and with more brothers) is more likely to opt for a STEM field than a boy with an older sister (and with more sisters).

H4: The effect of a math skilled older sibling on a younger sibling's likelihood to opt for a STEM major in college is more pronounced in same-sex pairs.

Method

Participants

To assess the probability that a younger sibling will pursue a STEM major in college, we use data from the U.S. *Child and Young Adult* (CY-NLSY79) module of the *National Longitudinal Survey of Youth*

---

[2] To anticipate our analytic strategy, we also tested our models using parallel measurements referring to the relationship between the focal child and the first-born sibling, i.e., the eldest sibling in a sib-group, instead of the sibling born immediately before the focal child. We show and discuss the results for both measurement approaches, which provide the same results.

1979 cohort (NLSY79). The NLSY79 is a nationally representative sample consisting of 12,686 individuals between the ages of 14-22 in 1979. These individuals were re-interviewed annually until 1994 and biennially thereafter. In follow-up waves, NLSY79 female sample members with children were asked to complete a module for their own children. This multigenerational design, which is not available in younger cohorts of the NLSY, allows for the collection of comprehensive information about all the siblings in a family throughout their childhoods as well as later life outcomes including academic histories and longitudinal information on parental and family background, family structure and household composition (Center for Human Resource Research 2001). The CY-NLSY79 module, administered annually to the children of all original NLSY79 female members, represents approximately 70%-75% of all children that will be born to a typical cohort of American women (Center for Human Resource Research 2009). We use these CY-NLSY79 modules (1986-2012) as the basis target sample for our analysis with initial number of 11,512 respondents.

Appendix A presents our sample selection strategy, showing how we limit the sample based on sib-group size, birth ordering, and birth cohort. Because we focus our attention on non-firstborn siblings, first, only-child families (1,176 children out of 11,512) were dropped, and first-born siblings were also excluded, (3,754 out of the remaining 10,336). Of the remaining 6,582 non-firstborn respondents who had at least one sibling, we excluded respondents who were too young to have graduated from college at the most recent wave, limiting our analytic sample to 4,498 children born prior to 1991 who are not the first-born in their family. Next, among the remaining 4,498 respondents, the sample is further reduced to the 3,130 respondents who attended college. All further sample reductions were due to listwise deletion. As can be seen in Table 1, in this stage most sample attrition was attributable to missing values for college attendance information (1,243 of the remaining cases), and mathematic ability during early adolescence for the older sibling (207 respondents were eliminated from our sample) and the respondent (additional 111 dropped). The listwise deletion yields an analytic sample of 1,545 non-firstborn children – 888 females and 657 males, originating from 1,202 households. Of those, 634 respondents were co-residential siblings (in 291 households).

Measures

College STEM-field preference

Our main dependent variable measures whether or not a respondent's primary college major was in a STEM-field. Following other studies, we grouped programs of science, technology, engineering, mathematics and doctoral-track medicine as STEM majors, while a course of study focusing on the humanities, social sciences, and clinical or health sciences (non-doctoral-track) were considered non-STEM majors (Morgan et al. 2013). To be considered a STEM-major, respondents' first declared field of study for their Bachelor's degree must have been in one of the identified STEM fields. All other non-STEM fields or an undeclared major were considered non-STEM.[3] Utilizing the first declared major enabled us to eliminate college-related contextual factors that may be associated with persistence

---

[3] Specifically, 895 sample members specified a non-STEM field and 307 did not declare a major. Both groups were collapsed into the non-STEM preference.

within major.[4] Among our 1,545 sample members who attended college, 343 of them (22%) declared majors in STEM fields.

Next we present six key independent variables accounting for sibling group compositional characteristics and older sibling math ability.

Sib-group size, order and spacing

First, we considered three measurements for sib-group compositional characteristics related to resource dilution: *sib-group size*, which measured the number of children, *birth order* measured as a dummy variable indicating second-born (coded 1) compared to third born and above order (coded 0)[5], and *birth spacing*, which measured the age distance (months) between the focal child and the sibling born immediately before him/her.[6]

Sib-group sex composition measurements

To test the same-sex competition stimulation hypothesis, we constructed two measurements of sib-group sex composition: the sex of the older sibling born just before the respondent (male coded as 1, female as 0) and the proportion of boys in a sib-group.

Older sibling's math ability

Finally, to test the role modeling perspective we included a measure of the older sibling's math ability, which was based on scores from the math subtest of the Peabody Individual Achievement Test (PIAT).[7] The PIAT is a widely used measure of academic achievement for children aged 5-15 and is generally considered to be highly reliable (Center for Human Resource Research 2009). This assessment is composed of multiple-choice questions that increase in difficulty, ranging from identifying numbers to trigonometry and was administered in each survey wave for children aged 5-15. Our measure of math ability relied on siblings' standardized test scores when they were aged 10-14. We used the most recent standardized score assessment available within that age bracket. Controlling for math ability for other siblings other than older sibling born immediately before the respondent substantially diminished our sample size, thus we decided to exclude these indicators from our analysis.

Control variables

---

[4] We aim to explain STEM major preference. About 28 percent of bachelor's degree students and 20 percent of associate's degree students entered a STEM field (Chen 2013), yet more than half of those who declared STEM majors while entering college left these fields before graduation (Chen 2009; Higher Education Research Institute 2010). Type of institution, STEM course-taking, relative performance in STEM were found to increase the likelihood of switching majors to non-STEM fields (Chen 2013). To avoid these possible effects of STEM degree completion, we choose first declared major.

[5] We tested for alternative measurement of birth order as a continuous variable and got consistent results (not shown).

[6] We tested for alternative measurement of birth spacing (continuous and as dummy variable) expressing the age distance between the eldest sibling and the youngest child, which yielded similar results (not shown). We preferred to use the selected measurement since our hypotheses refer to older sibling born immediately before the respondent.

[7] We also tested the effect of older siblings STEM major (not shown), which yielded no significant results net of older sibling's math ability, sib-group, individual and family characteristics.

Controls included various measures of individual and parental/household characteristics. We constructed measures for each sample member's birth cohort, race/ethnicity, and math ability similar to older sibling measurement. Information about parents measured when the focal child was aged 6 to 10 was used to construct controls for mother's education, number of parents in the household, and family income (logged and adjusted to constant dollars). We also constructed a measure of mother's occupation capturing her employment status *and* occupational gender-orientation when the focal child was between the ages 6-14 using mothers' self-reported work history data. We classified mother's work status information into five categories: male-type occupation, female-type occupation, unemployed, out of labor, and missing. The classification of occupations to male or female type was based on Weeden's (2004) classification scheme, which relies on saturated scale values of female and male participation in U.S. occupations during 1970 and 1990 (Weeden 2004, appendix table A5.1). Weekly employment information was obtained from each NLSY79 mother, yielding up to 416 observations per mother. The modal value across all 416 records was used to determine the mother's overall employment status and occupation when she was in the labor force. Controlling for mother's occupation is important in order to differentiate between parental influences and the effects of sib-group compositional characteristics on our outcome of interest.

Analytic Strategy

First we provide descriptive statistics. Then, to model the dependent dummy outcome, college major preference - STEM rather than non-STEM, we gave priority to linear probability model (LPM) over nonlinear models (probit or logistic regression) because LPM allows direct comparisons of coefficients across models and groups (Mood 2010).[8] We estimated four nested models predicting a focal sibling preference for a STEM-major vs. non-STEM major as a function of sib-group compositional characteristics. Following our research hypotheses, we first considered the sibling characteristics related to resource dilution, accounting for sib-group size, birth order, and birth spacing, net of respondent's sex, his/her math ability, and all other individual and family background characteristics (Model I, testing H1). Next, we focused on sibling characteristics based on Conley's (2000) theory on same-sex competitive stimulation, adding older sibling sex and the proportion of boys in a sib-group (Model II, testing H3). Accounting for the math ability of the older sibling in Model III we tested our hypotheses related to the role-modeling perspective (testing H2a/H2b). In Model IV, our final model, we tested for the possibility that the role modeling effect is more likely to occur among same- or opposite- sex siblings by adding an interaction term for older sibling's math ability and older sibling sex (testing H4). To explore the association of sib-group compositional characteristics with second-and later born females' and males' preference for a college STEM major, all sib-group configurations in all four models were interacted with the sex of the respondent. To facilitate comparison of the effects of the continuous independent variables within the models, we standardized them to a mean of zero and a unit standard deviation. Dummy variables retained their original metric. All models were estimated

---

[8] We get similar results with nonlinear models, see Appendix B.

using the appropriate weights as provided by the NLSY79 with standard errors clustered on household ID in order to account for dependence between co-residential siblings.

Robustness Checks

We conducted a variety of robustness checks. First, we assessed our analytical strategy by testing nonlinear models (probit or logistic regression), these yield the same conclusions (results are presented in Appendix B). Next, we evaluated the robustness of our findings on role modeling, testing whether downstream influences remain present when replacing the measurements from the *older sibling born immediately before* the respondent (math ability, sex, birth spacing) with measures from the *first-born sibling*. Results from these models are very similar, shown in the last two columns in Appendix B. Further we considered several potential biases due to sample selection (models not shown, but available upon request). Here, we excluded children from families with 5 or more children, as commonly done in the literature, to eliminate potential outliers, because families with many children may present different dynamics. Results did not change. Then, we assessed potential biases due to sample selection stemming from a) selection into college attendance, and b) selective panel attrition (missing values on education outcome in young adulthood). To this end, we created a dependent variable with four outcomes (STEM major, non-STEM major, no college attendance, missing college education information) and re-estimated all models using multinomial logit models. Substantive results remained the same. Furthermore, sib-group compositional characteristics and math ability did not significantly predict non-college attendance or panel attrition (models not shown, available upon request).

Results

Descriptive Analysis

Table 1 contains descriptive statistics for dependent and independent variables for our respondents separately by sex. Expectedly, the results show that males were more likely than females to prefer STEM fields as their major in college (0.30 vs. 0.17, p<0.001) and show higher math ability in young adolescence (103.4 vs. 100, p<0.001). The distribution of sib-group compositional characteristics at large followed the same patterns for both female and male respondents. About 35%-41% respondents were raised in 3-child families, about 30% came from 2-child families, 19% came from 4-child families, and the remaining 12%-16% were raised in families with five children or more. The majority (66%-67%) were second born siblings, 23%-26% were third-born siblings and 11%-7% were the fourth child or higher birth order. Finally, female and male respondents had an equal likelihood of having a brother or a sister as their older sibling born immediately before them. Exceeds to the gender similarity patterns are the proportion of boys in a sib-group and older sibling's math ability. Specifically, males were more likely to be raised in a sib-group with higher proportion of boys than females (0.67 vs. 0.35, p<0.001) and their older sibling had on average a higher math score (101.3 vs. 99.49, p<0.05). Also of note is that males were more likely to be raised in two-parent family than females (0.74 vs. 0.66, p<0.01).

[Table 1 here]

Table 2 provide Pearson correlations between siblings' compositional characteristics and STEM Major preference in college. For male respondents, results show positive correlation between preference for STEM major and the sex (male) of the older sibling (r =0.09, p<0.05). Among female, positive relation was observed between preference for STEM major and older sibling's math ability (r =0.104, p<0.01). Preferring STEM field in college was negatively associated with sib-group size among females (r =-0.07, p<0.05), but not among males. Also, and as expected, respondent's math ability is positively associated with young females' and males' preference for STEM field in college (r =0.13, p<0.001).

[Table 2 here]

LMP Analysis

Table 3 contains the standardized estimates from four LPMs, each predicting respondents' preference for a STEM versus a non-STEM major by siblings' compositional characteristics, net of socio-demographic and family background (for full models see Appendix C). In our sample, 30% of male respondents prefer STEM fields as college majors, compared with only 17% among females (Table 1). Controlling for socio-demographic and family background, yet without accounting for siblings' configuration or math ability, male respondents on average were 10% more likely to choose a STEM major in college (not shown). Starting with Model I in Table 3 the results show that higher math achievers were more likely than low math achievers to choose STEM fields and this inclination is similar among female and male respondents (b(se)=0.057 (0.018), p<0.01, with an insignificant interaction term of b(se)=0.008 (0.028), p>0.05). With each additional standard deviation point in a respondent's math ability score, his/her probability to prefer a STEM major increases by 5.7%. That is, increases of about 6.5-13 points in a respondent's math ability score (respondents' math ability (se) =13, Table 1) pushes his/her chances of pursuing a STEM major by 3%-6%. Corresponding to the resource dilution argument (H1), sib-group size is negatively associated with preferring STEM fields in college, equally for male and female respondents (b(se) =-0.038 (0.015), p<0.05, with an insignificant interaction term b(se)=0.040 (0.031), p>0.05). With each additional sibling in a sib-group, a respondent's preference for a STEM major decreases by about 4%. Both birth order and spacing are negatively associated with preferring STEM fields as a college major. Females and males who were the second born in a sib-group (the majority of our sample, Table 1) were 8% less likely to prefer a STEM major compared to higher birth order respondents (b(se)=-0.080 (0.041), p<0.05, with an insignificant interaction term b(se) =0.108 (0.069), p>0.05). Birth spacing is negatively associated with preferring STEM majors, each 13 months of birth spacing (birth spacing(se) =26.79, Table 1) reduces the probability of STEM majoring on average by 1%, yet the effect did not reach the statistical significance threshold (b(se)=-0.018 (0.013), p>0.05, with an interaction term b(se)=0.017 (0.026), p>0.05).

Model II addresses the role of same-sex competition stimulation in curricular preference, testing whether males and females were more likely to pursue a STEM degree if their older sibling was of the same sex, and when they came from a family with a higher proportion of same-sex siblings. The

results suggest that having an older brother, rather than a sister, increases the preference for a STEM major by 12% among male respondents, whereas for female respondents having an older brother decreases their chances of preferring a STEM major by about 5%, yet this effect for females did not reach acceptable statistical significance levels (b(se) =-0.047 (0.046), p>0.05 and an interaction term b(se)=0.165 (0.083), p<0.05). The effect of having a higher proportion of brothers in a sib-group is positively associated with preferring STEM fields for both sexes (b(se)=0.056 (0.028), p<0.05 and an insignificant interaction term b(se)=-0.051 (0.051), p>0.05). In other words, with each additional 15%-30% of brothers in a sib-group (proportion of brothers in a sib-group(se) =0.28, Table 1), both female and male respondents would increase their preference for STEM fields by about 3%-6%. Putting things differently, being raised in a sib-group consisting of 66% brothers is associated with a 6% increase in STEM majoring changes compared with those who were raised in a sib-group with 33% brothers. This implies that our findings are at odds with the same-sex siblings' competition stimulation hypothesis suggested by Conley (2000) (our hypothesis H3). For both males and females, having a male-climate sib-group (proportion of brothers) increases the chances of preferring STEM fields, whereas for males, being raised with an older male sibling further enhances their chances of pursuing a STEM college major. That said, however, the effect of older sibling's sex, although retaining its size, loses statistical significance once we control for older sibling's math ability in Model III (b(se)=-0.047 (0.046), p>0.05 and an insignificant interaction term b(se)=0.162 (0.083), p>0.05). This suggests that the association of the sex of older sibling may be channeling via his/her math ability, and is supported by the results of our final Model (Model IV).

[Table 3 here]

In Model III we considered the role modeling perspective by examining the likelihood that males and females with a math-skilled older sibling opted for a STEM major in college. Despite a positive coefficient, an older sibling's math ability showed no statistically significant association with the disposition to prefer a STEM major (b(se) = 0.015 (0.017), p>0.05) when not interacted with the sex of the older sibling, providing no support for H2a and H2b. However, the estimates from Model IV (our final model) indicate that the effect of older sibling's math ability on a younger sibling is differential by the sex of the respondent and of his/her older sibling. For female respondents, an older sister's math ability was positively associated with preferring a STEM major, increasing her young sister's preference of STEM by 2.5%-5% with each 7-14 points in her older sister math score (b(se)=0.050 (0.023), p<0.05; older sibling's math ability(se) =13.77, Table 1). Contrarily, having an older brother with high math ability is negatively associated with his younger sister's likelihood of preferring a STEM major (b(se)=0.050 (0.023), p<0.05 with the interaction term b(se)=-0.067 (0.031), p<0.05). An increase of 7-14 points in her older brother's math score, resulted in female respondents being 1%-2% less likely to prefer STEM fields in college. Among male respondents we find a similar negative pattern for opposite-sex sibling pairs; an older sister with high math ability reduces male respondents' preference for STEM major by 2.5%-5% with each additional 7-14 points in her math score (b(se)=0.050 (0.023), p<0.05, and the interaction term, b(se)=-0.102 (0.039), p<0.01). Math-skilled older brothers, however, show a marginal negative effect on male respondents' college major preference (b(se)=0.101 (0.051), p<0.05, calculation of the effect: [(-0.067+0.101)+(0.050+-0.102)]=-

0.018). With each additional 7-14 points in his older brother's math score, a male respondent's preference for STEM major is reduced slightly by about 1%-2%.

Despite being statistically significant, the magnitude of the effects relating older brother's math score and respondent's sex are relatively small. We further explore and illustrate the results of model IV by presenting predicted probabilities of STEM majoring for selected sib-group compositional characteristics, holding covariates at their mean values (Figure 1). The bottom panes in Figure 1 show predicted probabilities of preferring a STEM major conditional on older siblings' math ability for female respondents (left figure) and male respondents (right figure). Two findings come to the fore. First, variation in the math ability of an older brother or sister within itself is not predicting relevant differences in STEM majoring for male or female respondents. Second, the mix of the siblings' sexes is important in how the older sibling's high (but not low) math ability predicts her or his younger sibling's STEM major preference. For girls, an older sister at the 75th percentile of the math score distribution significantly boosts her STEM majoring probability by about 10%, compared with her having a high math achieving brother. This difference is statistically significant, as tested via an additional model which categorized the continuous measurement of older sibling's math ability into three math score dummy variables – low (0-33%), medium (34-66%), and high (66-100%) (models not shown but available upon request). Female respondents with a high math achieving older sister were significantly more likely to pursue a STEM career than those with a highly math achieving older brother (p=0.0149). Among male respondents, although the difference between a high math achieving older sister versus brother shows a parallel same-sex dyad advantage as in the case of females, it did not reach statistical significance (p>0.05). We conclude that for females, the sex of an older sibling with high math ability is important for her STEM-major choice later in life, whereas this does not apply to males. The upper panes of Figure 1 depict the negative effect of sib-group size and the positive effect of the proportion of brothers on STEM major preference, however, again, effects are larger among female respondents, and confidence intervals of the predictions for males overlap for both measures. Hence, both the resource dilution argument (H1) and the STEM preference enhancing effect of a male-climate subgroup are confirmed, but only for female respondents.

[Figure 1 here]

We estimated our models separately for focal males and females (see Appendix D). Results reinforce that sib-group size, its gender composition, and the interaction of sex and math ability of one's older sibling are significant predictors of STEM majoring among female respondents, whereas for male respondents, STEM major preference is rather driven by own math ability. Our pooled models (Table 3) are more parsimonious compared with the separate models (Appendix D), because they do not interact the control variables with respondent's sex. It is therefore possible low statistical confidence in the sib-group configuration measures among male respondents is due to sample size limitations. Finally, as recall additional analyses are presented in Appendix B for our final model IV to serve as robustness check. First, we show that nonlinear models (probit or logistic regression) yield the same conclusions. Then we present results from analysis in which we exchanged characteristics from

the sibling born immediately before the respondent with those from the first-born sibling. This alternative analysis bolsters our findings.

Discussion

Much is known on the association between sib-group compositional characteristics and years of schooling, however, whether sib-group features may be relevant for substantive educational preferences has not yet been examined. Our results show that sib-group compositional characteristics matter for STEM field preference in college, but for young women only. First, net of other compositional features, women were less likely to prefer a STEM major if they were raised in larger sib-groups. This finding provides descriptive evidence for the resource dilution hypothesis, which so far has been confirmed by research for years of schooling, but had not been investigated with respect to major choice, in other words a 'qualitative' educational outcome. Perhaps excelling in math and other STEM subjects, still gender atypical subjects for girls, requires more support from parents or other adults in the girls' private realm, as girls may receive this support less readily in school than boys (Sansone 2019). This may be one possible explanation for why large sib-group size has a detrimental effect on STEM major preference of women, but not men.

Second, we find that the proportion of boys in a sib-group is positively associated with preferring a STEM career, again, mainly for female respondents. Male sib-group dominance, as indicated by a higher proportion of brothers in a family, may suggest the formation of a normative male-role sib-group climate, as suggested by Powell and Steelman (1990). This could affect the boys' sister's STEM interest in several ways, for instance via repeated exposure to and familiarization with their more 'male' interests and activities. A family with more male children may thus more readily contain children that are socialized to male role expectations, regardless of their individual sex.

Third, and quite interestingly, females who were raised with a high math achieving older sister during young adolescence were significantly more likely to pursue a STEM major in college compared with a highly math achieving older brother, net of all other sib-group configurations. For male respondents, results were much weaker, which we interpret as suggesting that males did not gain (nor lose) much from having same-sex or opposite-sex math achieving older sibling. These findings indicate that role modeling and the influence of an older sibling on channeling younger siblings' educational STEM preferences is stronger and mostly relevant for females. This is, again, intuitive, given that choosing a STEM major is still atypical for women, and may require extra exposure or support during childhood and adolescence. A math-skilled older sister may create a competitive yet positive environment that pushes her younger sister to excel in similar domains; alternatively, increased cooperation and support may take place among sisters who both excel in gender atypical school subjects. Of course, this could also be an outcome of parental features. For instance, parents with strong gender egalitarian values may invest strongly in their daughters' math and science education in order to provide them with equal life chances. In this scenario, parental gender ideology would affect both, the older sister's high math score and the younger daughter's STEM major preference. Indeed, gender equality has increased during the last several decades, and there has been a corresponding increase in adults' endorsement of gender-egalitarian attitudes (Blakemore and Hill 2008; Marks et al. 2009). Our

analyses control for a variety of socio-economic factors of the siblings' family of origin, measured during childhood, including mother's employment status, her occupation and its gendered context. While this eliminates the effect of some parental influences, specifically of mother's occupation role modeling and knowledge, future research should investigate if parental gender ideology may mediate the relationship between older sister's math ability and younger sister's STEM major preference, by directly modeling parental gender ideology, and not indirectly deriving it from mother's occupation. Yet another explanation is that a girl's older math-skilled brother pushes her to seek alternative and contrasting pathways for her own career, perhaps engaging in contrasting gendered behavior. Such an interpretation would be in line with Conley's sex-minority competition hypothesis (2000) or could be understood as an opposite gendered role modeling process (Sulloway 1996; Whiteman et al. 2014). Either way, our findings suggest that future efforts to understand persisting gender differences in educational and career preferences and choices should look deeper into the role modeling effects of siblings in families, and how much they may be mediated by parental features or are independent thereof. Overall, our findings propose that the mechanisms underlying the association between sib-group compositional characteristics and college major account for long-term differences in college curricular preferences.

How Could Sibling Group Composition Matter? What Are the Mechanisms?

One of the main limitations of this study is that it does not directly test the mechanisms driving these patterns. Despite this limitation, a descriptive study such as ours is an important first step in developing a deeper understanding of how gendered family environments may matter for educational and other outcomes. Our models are not designed to test mechanisms; however, we can compare the observed population-level patterns to what we might expect using the theoretical approaches we discussed earlier as our guide: sibling role modeling, sibling competition, and parental resource allocation. The first two theoretical approaches explicitly involve social interactions among siblings while the third involves the social interactions between parents and siblings.

Our findings put forward that the theoretical approaches are not mutually exclusive, as our results for sib-group sex composition strongly conform to the expectations of role-modeling *and* competitive stimulation. Hence, either a hybrid of these theories is operating, additional mechanisms are at play, or sibling group characteristics are mediators for certain theoretical mechanisms. We also observe additional associations beyond the hypotheses we derived from existing theories. For instance, higher likelihoods of STEM majoring among younger sisters with older highly math skilled sisters may not only indicate positive role modeling, but could also imply a particularly pronounced cooperation and mutual support among siblings when they both excel or are interested in gender atypical domains. The same pattern may be present among boys when gender atypical outcomes are predicted. Since we only look at STEM majoring, such a potential pattern necessarily remains undetected by our study, and remains to be investigated by future research.

The patterns we observe raise interesting questions about the underlying mechanisms and an examination using more detailed qualitative information about parental investments and interactions among siblings beyond our set of measures is needed. What are the direct and indirect impacts of

parents, siblings and the family environment on a child's eventual preference of college major? One reasonable pathway could occur if parents prepare their home environment and adapt parenting practices differently depending on the gender of their first child, or the majority of their children, for instance especially present if child number one and two are of the same sex. The ways in which this could be meaningful for subsequent children are myriad: gendered toys, clothes, child-geared equipment like furniture or sports equipment, games, family activity routines, parenting styles established in the family and interactions with same-sex friends of the older child all form a 'parenting infrastructure' that is established after the first child is born, and this could have downstream spillover effects that set a 'gendered tone' for subsequent children. This 'gendered tone' logic is not exclusive to the first-born sibling, but can also operate via the overall sib-group sex ratio, as suggested by Powell and Steelman's (1990) normative climate hypothesis and also displayed in our findings. Hence, having more male siblings in a household regardless of the sex of the first-born child may make male role expectations more common, which could impact all children regardless of their sex. Again, future work must build on what we know about sibling compositional characteristics by combining this knowledge with more qualitative measures of the home environment and the gendered nature of siblings' dynamics and interactions.

Some may argue that our analyses exclude several key variables that are known to relate to gender differentiation in preference and choice of college major. Traditional theories connect the gender segregation in field of study to males' relative advantage compared to females in math and verbal abilities (Jonsson 1999), general life goals, gender roles attitudes (Jacobs 1989, 1995), expectations about work–family compatibility (Eccles 2011a, 2011b; Hakim 2002), and gendered desires for extrinsic or intrinsic occupational preferences (Marini et al. 1996). All these measurements are absent from our analysis to maintain statistical power, and some for lack of available data. Though our results lend preliminary support for some theories over others, future work is needed in order to determine the mechanisms by which older siblings impact younger siblings and how parental influences may shape these dynamics. Still, our results suggest that a key aspect of the gender educational stratification is overlooked without considering a sib-group compositional characteristics and that social dynamics between siblings differ for boys and girls.

Conclusion

Though women reached parity with men in terms of college attendance, fewer women choose STEM majors. We examine whether later life preference of STEM major in college depends on sib-group composition. Siblings' compositional characteristics have received extensive attention in predicting gender differences in educational attainment measured 'quantitatively', namely focused on the time spending in the educational system. However, little is known on the linkage between sib-group configurations and gendered outcomes with respect to more 'qualitative' educational outcomes in the same level of education, namely college major. Testing the tendency for younger siblings' educational preferences to be influenced by sib-group compositional features and the sex and math ability of the older sibling allows us to understand the possible downstream influences of older siblings while also accounting for the presence of other siblings. Going beyond years of schooling and point grade

average, our findings suggest that a sib-group compositional characteristics influences second-and later born females' and males' decision to later pursue a STEM major in college, though in different ways. Females who were raised in a smaller sib-group, in a male sib-group dominance, and with high-achieving older sister are more likely to prefer STEM major in their later educational career. STEM preference among males, our results suggest, is mostly driven by their own math ability, whereas sib-group characteristics revealed as insignificant. Our research advocates that the theoretical approaches we offered are not mutually exclusive, as our results conform an interplay between role modeling and same-sex competitive stimulation accompanied by resource dilution effect. Future research should open the 'black-box' of parental investments and interactions among siblings to reveal the underlying mechanisms behind these patterns starting with those we put forward in our discussion.

Bibliography

Alon, S., & Gelbgiser, D. (2010). The female advantage in college academic achievements and horizontal sex segregation. *Social Science Research*, 40(1), 107–119.

Anastasi, A. (1956). Intelligence and family size. *Psychological Bulletin,* 53(3), 187–209.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.

Barclay, K., & Kolk, M. (2017). The long-term cognitive and socioeconomic consequences of birth intervals: a within-family sibling comparison using Swedish register data. *Demography*, 54, 459–484.

Barclay, K., Hällsten, M., & Myrskylä, M. (2017). Birth order and college major in Sweden. *Social Forces,* 96(2), 629–660.

Beede, D. N., Julian, T. A., Langdon, D., McKittrick, G., Khan, B., & Doms, M. E. (2011). Women in STEM: A gender gap to innovation." Issue Brief #04-11. Economics and Statistics Administration. U.S. Department of Commerce. http://ssrn.com/abstract=1964782.

Behrman, Jere, and Paul Taubman. 1986. "Birth Order, Schooling, and Earnings." *Journal of Labor Economics* 4 (3): S121-45.

Björklund, A., & Jäntti, M. (2012). How important is family background for labor-economic outcomes? *Labour Economics,* 19 (4), 465–474.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2005). The more the merrier? The effect of family size and birth order on children's education. *The Quarterly Journal of Economics*, 120(2), 669–700.

Blake, J. 1981. "Family size and the quality of children. *Demography*, 18(4), 421–442.

———. (1986). Number of siblings, family background, and the process of educational attainment. *Biodemography and Social Biology,* 33(1–2), 5–21.

———. (1989). *Family size and achievement*. Berkeley and Los Angeles, California: University of California Press.

Blakemore, J. E. O., & Hill, C. A. (2008). The child gender socialization scale: A measure to compare traditional and feminist parents. *Sex Roles,* 58(3–4), 192–207. https://doi.org/10.1007/s11199-007-9333-y

Booth, A. L. & Kee, H. J. (2005). Birth order matters: The effect of family size and birth order on educational attainment." IZA DP No. 1713.

Bouchey, H. A., Shoulberg, E. K., Jodl, K. M., & Eccles, J. S. (2010). Longitudinal links between older sibling features and younger siblings' academic adjustment during early adolescence." *Journal of Educational Psychology*, 102(1), 197–211.

Brim, O. G. (1958). Family structure and sex role learning by children: A further analysis of Helen Koch's data. *Sociometry,* 21(1), 1–16.

Brown, J. R., Donelan-McCall, N., & Dunn, J. (1996). Why talk about mental states? The significance of children's conversations with friends, siblings, and mothers." *Child Development*, 67(3), 836–849.

Bu, F. (2014). Sibling configurations, educational aspiration and attainment." ISER Working Paper Series. University of Essex: Institute for Social and Economic Research.

Buckles, K. S., & Munnich, E. L. (2012). Birth spacing and sibling outcomes. *Journal of Human Resources*, 47(3), 613–642.

Bussey, K., & Bandura, A. (1999). Social cognitive theory of gender development and differentiation." *Psychological Review,* 106(4), 676–713.

Butcher, K. F., & Case, A. (1994). The effect of sibling sex composition on women's education and earnings." *The Quarterly Journal of Economics,* 109(3), 531–563.

Center for Human Resource Research. (2001). *NLSY79 Users Guide: A Guide to the 1979–2000 National Longitudinal Survey of Youth Data*. Columbus, Ohio: The Ohio State University.

———. (2009). *NLSY79 Child & Young Adult: Data Users Guide*. Columbus, Ohio: The Ohio State University.

Charles, M., & Bradley, K. (2009). Indulging our gendered selves? Sex segregation by field of study in 44 countries." *American Journal of Sociology,* 114(4), 924–976.

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

Chen, X. (2009). "Students who study science, technology, engineering, and mathematics (STEM) in postsecondary education (NCES 2009-161)." Washington, DC.: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

———. (2013). "STEM attrition: college students' paths into and out of STEM fields (NCES 2014-001)." National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94, 95–120.

Conley, D. (2000). Sibship sex composition: Effects on educational attainment. *Social Science Research*, 29, 441–457.

———. (2004). *The pecking order: Which siblings succeed and why*. New York, NY: Pantheon.

Crouter, A. C., Manke, B. A., & McHale, S. M. (1995). The family context of gender intensification in early adolescence. *Child Development,* 66, 317–329.

Crouter, A. C., Whiteman, S. D., McHale, S M., & Osgood, D. W. (2007). Development of gender attitude traditionality across middle childhood and adolescence. *Child Development*, 78, 911–926.

Davies, K. (2019). 'Sticky' proximities: Sibling relationships and education. *The Sociological Review,* 67(1), 210–225

De Haan, M. (2010). Birth order, family size and educational attainment. *Economics of Education Review*, 29(4), 576–588.

DiPrete, T. A., & Buchmann, C. (2013). *The rise of women: The female advantage in education and what it means for American schooling*. New York: Russell Sage Foundation Press.

Downey, D. B. (1995). When bigger is not better: Family size, parental resources, and children's educational performance." *American Sociological Review,* 60(5), 746–761.

———. (2001). Number of siblings and intellectual development: The resource dilution explanation. *American Psychologist,* 56(6–7), 497–504.

Dunn, J. (1983). Sibling relationships in early childhood. *Child Development*, 54(4), 787–811.

———. (2007). Siblings and socialization. In *Handbook of Socialization: Theory and Research.*, edited by J.E. Grusec and P.D. Hastings. New York: Guilford Press.

Eccles, J. S. (2011a). Gendered educational and occupational choices: applying the Eccles et al. model of achievement-related choices. *International Journal of Behavioral Development,* 35(3), 195–201.

———. (2011b). Understanding women's achievement choices looking back and looking forward. *Psychology of Women Quarterly*, 35(3), 510–516.

Feinberg, M. E., & Hetherington, E. M. (2000). Sibling differentiation in adolescence: Implications for behavioral genetic theory. *Child Development*, 71, 1512–1524.

Fouad, N. A., Hackett, G., Smith, P. L., Kantamneni, N., Fitzpatrick, M., Haag, S., & Spencer, D. (2010). Barriers and supports for continuing in mathematics and science: Gender and educational level differences. *Journal of Vocational Behavior,* 77(3), 361–373.

Galbraith, R. (1982). Sibling spacing and intellectual development: A closer look at the confluence models." *Developmental Psychology,* 18(2), 151–173.

Grätz, M. (2018). Competition in the family: Inequality between siblings and the intergenerational transmission of educational advantage." *Sociological Science,* 5, 246-269.

Hakim, C. (2002). Lifestyle preferences as determinants of women's differentiated labor market careers." *Work and Occupations* 29(4), 428–459.

Hanushek, E. A. 1992. The trade-off between child quantity and quality. *Journal of Political Economy*, 100(1), 84-117.

Hertwig, R., Davis, J. N. and Sulloway, F. J. (2002). Parental investment: How an equity motive can produce inequality. *Psychological Bulletin*, 128(5), 728–745.

Higher Education Research Institute. (2010). Degrees of Success: Bachelor's Degree Completion Rates Among Initial STEM Majors. Los Angeles: Author.

Hotz, J. V., and Pantano, J. (2015). Strategic parenting, birth order, and school performance. *Journal of Population Economics,* 28(4), 911–936.

Howe, N., Rinaldi, C. M., Jennings, M., & Petrakos, H. (2002). 'No! The lambs can stay out because they got cozies': Constructive and destructive sibling conflict, pretend play, and social understanding. *Child Development,* 73(5), 1460–1473.

Jacobs, J. A. (1989). *Revolving doors: Sex segregation and women's careers*. Stanford, CA: Stanford University Press.

Jacobs, J. A. (1995). "Gender and academic specialties: Trends among recipients of college degrees in the 1980s." *Sociology of Education*, 68(2), 81-98.

Joensen, J. S. & Nielsen, H. S. (2018). Spillovers in education choice. *Journal of Public Economics*, 157, 158-183.

Jonsson, J. O. (1999). Explaining sex differences in educational choice an assessment of a rational choice model. *European Sociological Review*, 15(4), 391–404.

Kaestner, R. (1997). Are brothers really better? Sibling sex composition and educational achievement revisited. *The Journal of Human Resources,* 32(2), 250–284.

Kidwell, J. S. (1981). Number of siblings, sibling spacing, sex, and birth order: Their effects on perceived parent-adolescent relationships. *Journal of Marriage and Family,* 43(2), 315–332.

Koch, H. L. (1960). *The relation of certain formal attributes of siblings to attitudes held toward each other and toward their parents.* Monographs of the Society for Research in Child Development 25, 78.

Kuo, H. D., & Hauser, R. M. (1997). How does size of sibship matter? family configuration and family effects on educational attainment." *Social Science Research,* 26(1), 69–94.

Legewie, J., & DiPrete, T. A. (2014). Pathways to science and engineering bachelor's degrees for men and women. *Sociological Science,* 1, 41–48.

Lytton, H., & Romney, D. M. (1991). Parents' differential socialization of boys and girls: A meta-analysis. *Psychological Bulletin,* 109(2), 267–296.

Mann, A., & DiPrete, T. A. (2013). Trends in gender segregation in the choice of science and engineering majors. *Social Science Research,* 42, 1519–1541.

Marini, M. M., Fan, P., Finley, E., & Beutel, A. M. (1996). Gender and job values. *Sociology of Education,* 69(1), 49–65.

Marks, J., Chun Bun, L., & McHale S. M. (2009). Family patterns of gender role attitudes. *Sex Roles,* 61(3–4), 221–234.

McHale, S. M., Crouter, A. C., & Whiteman, S. D. (2003). The family contexts of gender development in childhood and adolescence. *Social Development,* 12(1), 125–148.

McHale, S. M., Updegraff, K. A., Helms-Erikson, H., & Crouter, A. C. (2001). Sibling influences on gender development in middle childhood and early adolescence: A longitudinal study. *Developmental Psychology*, 37(1), 115.

McHale, S. M., Updegraff, K. A., & Whiteman, S. D. (2012). Sibling relationships and influences in childhood and adolescence. *Journal of Marriage and the Family,* 74(5), 913–930.

Mood, C. (2010). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European Sociological Review,* 26(1), 67–82.

Morgan, S. L., Gelbgiser, D., & Weeden, K. A. (2013). Feeding the pipeline: Gender, occupational plans, and college major selection. *Social Science Research*, 42(4), 989–1005.

National Science Foundation. (2013). Women, minorities, and persons with disabilities in science and engineering: 2013. Special Report NSF 13-304." Arlington, VA. http://www.nsf.gov/statistics/wmpd/.

Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics,* 122(3), 1067–1101.

Niederle, M., & Vesterlund, L. (2010). Explaining the gender gap in Math test scores: the role of competition. *Journal of Economic Perspectives,* 24(2), 129–144.

Oettinger, G. S. (2000). Sibling similarity in high school graduation outcomes: Causal interdependency or unobserved heterogeneity? *Southern Economic Journal,* 66(3), 631–648.

Powell, B., & Steelman, L. C. (1989). The liability of having brothers: Paying for college and the sex composition of the family. *Sociology of Education*, 62(2), 134–147.

———. (1990). Beyond sibship size: Sibling density, sex composition, and educational outcomes. *Social Forces,* 69(1), 181–206.

———. (1993). The educational benefits of being spaced out: sibship density and educational progress. *American Sociological Review,* 58(3), 367–381.

———. (1995). Feeling the pinch: Child spacing and constraints on parental economic investments in children. *Social Forces,* 73(4), 1465–1486.

Price, J. (2008). Parent-child quality time does birth order matter? *Journal of Human Resources*, 43(1), 240–265.

———. (2010). The effect of parental time investments: evidence from natural within-family variation. *Working Paper*. http://byuresearch.org/home/downloads/price_parental_time_2010.pdf.

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

Rosenberg, M. (1965). *Society and the adolescent self image*. Princeton, NJ: Princeton University Press.

Sansone, D. (2019). Teacher characteristics, student beliefs, and the gender gap in STEM fields. *Educational Evaluation and Policy Analysis*, 41(2), 127–144.

Schachter, F. F., Gilutz, G., Shore, E., & Adler, M. (1978). Sibling deidentification judged by mothers: Cross-validation and developmental studies. *Child Development*, 49(2), 543–546.

Schachter, F. F., Shore, E., Feldman-Rotman, S., Marquis, R. E., & Campbell, S. (1976). Sibling deidentification. *Developmental Psychology*, 12(5), 418–427.

Sulloway, F. J. (1996). *Born to rebel: Birth order, family dynamics, and creative lives*. London: Little, Brown and Company.

Steelman, L. C., Powell, B., Werum, R., & Carter, S. (2002). Reconsidering the effects of sibling configuration: Recent advances and challenges. *Annual Review of Sociology*, 28(1), 243–269.

Torche, F. (2011). Is a college degree still the great equalizer? Intergenerational mobility across levels of schooling in the United States. *American Journal of Sociology,* 117(3), 763–807.

Weeden, K. A. (2004). Profiles of change: Sex segregation in the United States 1910-2000. In *Occupational ghettos*, edited by M. Charles & D. B. Grusky. Stanford CA: Stanford University Press.

Whiteman, S. D., Becerra, J. M., & Killoren, S. E. (2009). Mechanisms of sibling socialization in normative family development. *New Directions for Child and Adolescent Development*, 126, 29-43.

Whiteman, S. D., Jensen, A. C., & Maggs, J. L. (2014). Similarities and differences in adolescent siblings' alcohol-related attitudes, use, and delinquency: Evidence for convergent and divergent influence processes. *Journal of Youth and Adolescence,* 43, 687–697.

Zajonc, R. B., & Markus, G. B. (1975). Birth order and intellectual development. *Psychological Review,* 82(1), 74–88.

Tables and Figures

Table 1: Means (standard deviations) and proportions of respondents' educational choice, sib-group compositional characteristics, and other individual and family background characteristics by gender

	Range	All		Female Respondents		Male Respondents		Gender difference
		mean	std.	mean	std.	mean	std.	Δ p-value
College STEM field	0, 1	0.22	0.42	0.17	0.37	0.30	0.46	p<0.001***
Respondent's math ability	65 – 135	101.5	13.00	100	12.55	103.4	13.35	p<0.001***
Sib-group characteristics								
Sib-group size	2 – 11	3.29	1.29	3.33	1.34	3.24	1.22	p=0.165
2-child	0, 1	0.29	0.46	0.30	0.46	0.28	0.45	p=0.443
3-child	0, 1	0.37	0.48	0.35	0.48	0.41	0.49	p<0.05*
4-child	0, 1	0.19	0.39	0.19	0.39	0.19	0.39	p=0.850
5-child and above	0, 1	0.15	0.35	0.16	0.37	0.12	0.33	p<0.05*
Birth ordering	2 - 8	2.46	0.77	2.49	0.82	2.42	0.69	p=0.068
2nd born	0, 1	0.66	0.47	0.66	0.47	0.67	0.47	p=0.641
3rd born	0, 1	0.25	0.43	0.23	0.42	0.26	0.44	p=0.157
4th born and above	0, 1	0.09	0.29	0.11	0.31	0.07	0.25	p<0.01**
Birth spacing (months)	0-191	39.84	26.79	39.95	27.61	39.68	25.64	p=0.846
Sibling just above is a male	0, 1	0.51	0.50	0.52	0.50	0.49	0.50	p=0.183
% boys in a sib-group	0 – 1	0.48	0.28	0.35	0.23	0.67	0.24	p<0.001***
Sibling just above math ability	65 - 135	100.3	13.77	99.49	13.86	101.3	13.60	p<0.05*
Cohort								
Before 1975	0, 1	0.002	0.05	0.001	0.06	0.005	0.07	p=0.189
1975-1979	0, 1	0.07	0.25	0.07	0.26	0.06	0.24	p=0.407
1980-1984	0, 1	0.33	0.47	0.36	0.48	0.30	0.46	p<0.05*
1985-1990	0, 1	0.60	0.49	0.57	0.50	0.63	0.48	p<0.05*
Race								
White non-Hispanic	0, 1	0.42	0.49	0.40	0.49	0.44	0.50	p=0.114
Hispanic	0, 1	0.22	0.41	0.21	0.40	0.24	0.43	p= 0.123
Black	0, 1	0.36	0.48	0.39	0.49	0.32	0.47	p<0.01*
Two-parent family	0, 1	0.69	0.46	0.66	0.47	0.74	0.44	p<0.01**
Family income (logged)	0 - 11.68	8.28	0.94	8.25	0.83	8.33	1.08	p=115
Mother's education								
Less than HS	0, 1	0.08	0.27	0.09	0.29	0.07	0.26	p= 0.154
HS grad	0, 1	0.43	0.50	0.45	0.50	0.40	0.49	p<0.05*
Some college	0, 1	0.31	0.46	0.29	0.46	0.33	0.47	p= 0.161
College	0, 1	0.18	0.38	0.16	0.37	0.20	0.40	p<0.05*
Mother's occupation								
Male-type occupation	0, 1	0.02	0.14	0.02	0.15	0.01	0.12	p= 0.207
Female-type occupation	0, 1	0.07	0.26	0.08	0.27	0.07	0.26	p= 0.516
Unemployed	0, 1	0.11	0.31	0.12	0.32	0.10	0.30	p= 0.245
Out of labor force	0, 1	0.63	0.48	0.63	0.48	0.63	0.48	p= 0.910
Missing	0, 1	0.17	0.38	0.15	0.36	0.19	0.40	p< 0.05*
Ns		1,545		888		657		
%		1.00		0.57		0.43		

Asterisks indicate statistically significant gender differences, with the following p-values:
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Table 2: Pearson correlations between college major preferences and siblings' compositional characteristics for female (below diagonal values) and male respondents (above diagonal values)

	1	2	3	4	5	6	7	8
1. Respondent's college STEM-field	--	0.130***	-0.005	0.017	-0.032	0.057	0.090*	-0.015
2. Respondent's math ability	0.134***	--	-0.085*	0.090*	-0.028	-0.018	0.017	0.371***
3. Sib-group size	-0.067*	-0.157***	--	-0.437***	-0.188***	-0.180***	0.014	-0.092*
4. 2nd born (rather than 3rd+)	-0.032	0.108**	-0.462***	--	0.067	0.113**	0.001	0.092*
5. Birth spacing (months)	-0.026	-0.007	-0.179***	0.105***	--	0.036	-0.010	-0.042
6. % boys in a sib-group	0.035	-0.010	0.130***	-0.053	-0.037	--	0.728***	0.098*
7. Sibling just above is a male	0.033	0.015	-0.092**	0.040	0.048	0.715***	--	0.049
8. Sibling just above math ability	0.104**	0.366***	-0.145***	0.138***	-0.031	0.030	0.059~	--

Above the diagonal values for male respondents (N=657); underneath values for female respondents (N=888)
Asterisks indicate statistically significant gender differences, with the following p-values:
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ~p=0.078.

Table 3: LPM standardized coefficients predicting college STEM major preference by sibling compositional characteristics, older sibling measurements refer to sibling born just above the respondent (N=1,545)

Preferring STEM vis-à-vis non-STEM major	Model I	Interaction X R is Male	Model II	Interaction X R is Male	Model III	Interaction X R is Male	Model IV	Interaction X R is Male
Male	0.028		-0.080		-0.082		-0.079	
	(0.052)		(0.065)		(0.064)		(0.062)	
Respondent' math ability	0.057**	0.008	0.056**	0.005	0.050**	0.023	0.048*	0.026
	(0.018)	(0.028)	(0.018)	(0.028)	(0.019)	(0.032)	(0.019)	(0.033)
Resource dilution hypothesis								
Sib-group size	-0.038*	0.040	-0.045**	0.045	-0.045**	0.046	-0.049**	0.049
	(0.015)	(0.031)	(0.015)	(0.032)	(0.015)	(0.032)	(0.015)	(0.032)
Birth order (2nd vis-à-vis 3rd and above)	-0.080*	0.108	-0.082*	0.104	-0.084*	0.114	-0.088*	0.115
	(0.041)	(0.069)	(0.040)	(0.068)	(0.041)	(0.069)	(0.041)	(0.068)
Spacing b/t focal child and older sibling	-0.018	0.017	-0.017	0.015	-0.017	0.012	-0.017	0.012
	(0.013)	(0.026)	(0.013)	(0.025)	(0.013)	(0.025)	(0.013)	(0.025)
Same-sex sibling competitive stimulation hypothesis								
Older sibling is male			-0.047	0.165*	-0.047	0.162	-0.040	0.144
			(0.046)	(0.083)	(0.046)	(0.083)	(0.044)	(0.083)
% of boys in a sib-group			0.056*	-0.051	0.055*	-0.044	0.058*	-0.045
			(0.028)	(0.051)	(0.027)	(0.051)	(0.027)	(0.051)
Role modeling perspective								
Older sibling's math ability					0.015	-0.050	0.050*	-0.102**
					(0.017)	(0.030)	(0.023)	(0.039)
Role modeling via same-sex sibling competition								
Older sibling's math ability X Older sibling is male							-0.067*	0.101*
							(0.031)	(0.051)
Constant	0.194		0.193		0.179		0.174	
	(0.135)		(0.145)		(0.152)		(0.144)	
F	3.20***		3.43***		3.29***		3.30***	
Adjusted R^2	0.052		0.065		0.068		0.072	
AIC	1671.4		1658.5		1657.6		1654.4	
BIC	1799.6		1808.1		1817.9		1825.4	

All models are adjusted for socio-demographic and socio-economic background. Standard errors in parentheses. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Figure 1: Marginal predictions of main sib-group compositional characteristics on preferring a STEM major (based on LPM Table 4 Model IV)
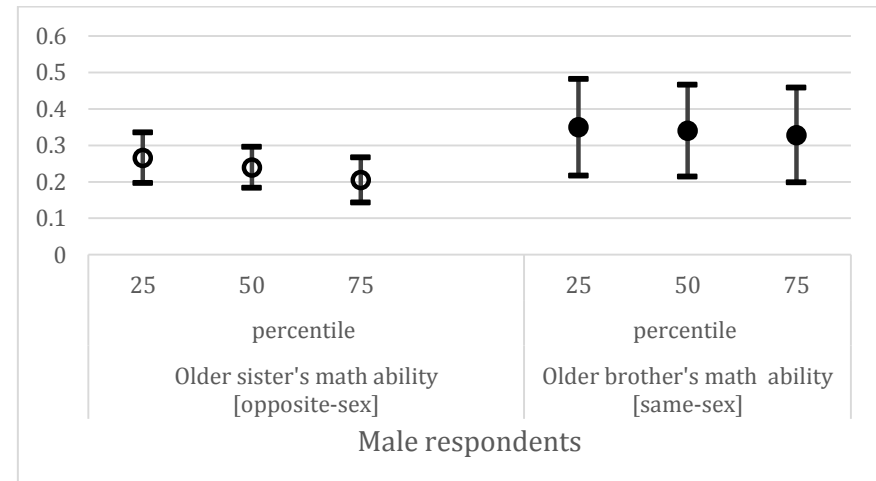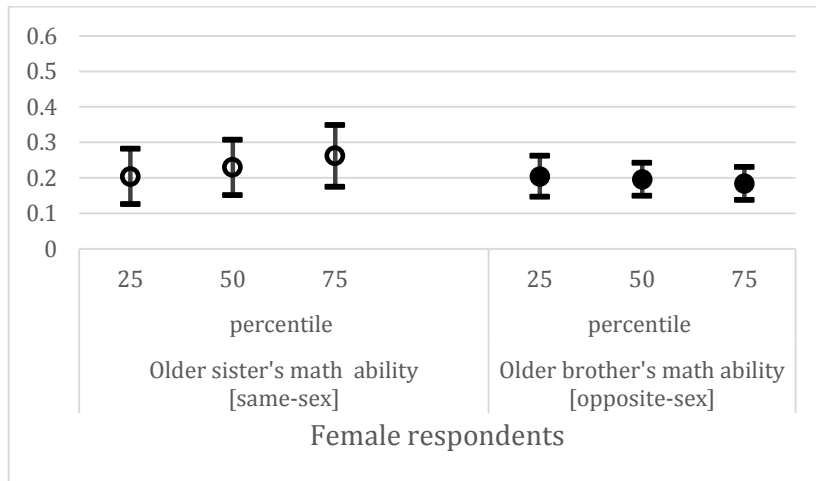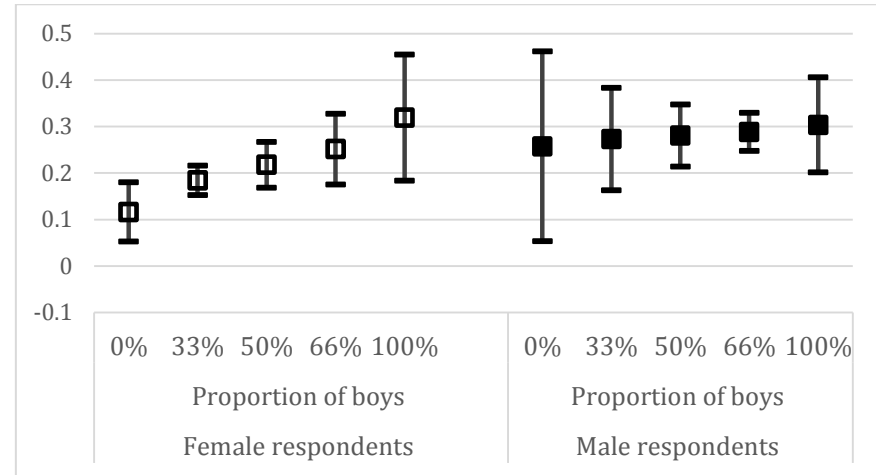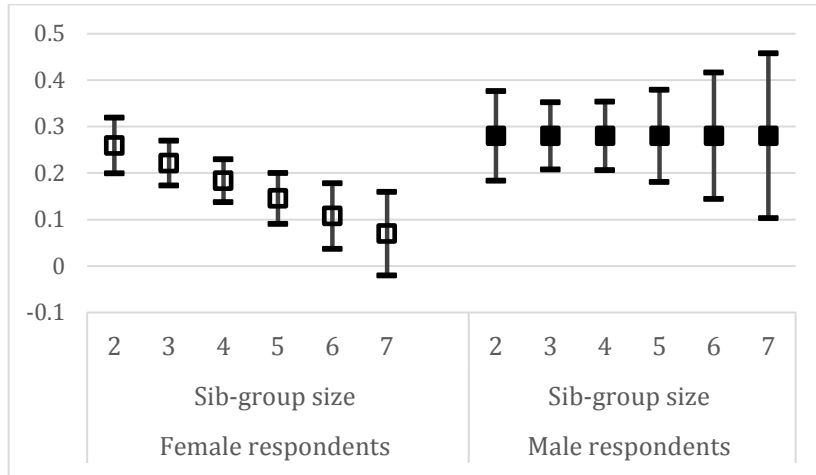
Appendix A: Target sampling attritions information

	Attritions	Ending with target sample of
1st stage: *Sample structural strategy*	Initial sample consists of 11,512 cases	
Excluding…		
those with no siblings	1,176	10,336
firstborn siblings	3,754	6,582
those who were born after 1990	2,084	4,498
those who did not attend college	1,368	3,130
2nd stage: *Listwise deletion*	Target sample consists of 3,130 cases	
Missing information on…		
college attendance	1,243	1,887
birth Spacing	1	1,886
ability math of sibling just above	207	1,679
ability math of the respondent	111	1,568
family income during childhood	23	**1,545**

Appendix B: Standardized coefficients of Probit and logistic regression predicting college STEM major preference by sibling compositional characteristics, and LPM with older sibling measurements refer to the eldest sibling (firstborn), (Model IV, full model, N=1,545)

Preferring STEM vis-à-vis non-STEM major	Testing for modeling specification [As in Table 4, older sibling measurements refer to the sibling born just above the respondent]				Testing alternative measurement of older sibling	
	Probit		**Logit**		**LPM, Firstborn sibling**	
	Model IV	Interaction X R is Male	Model IV	Interaction X R is Male	Model IV	Interaction X R is Male
Male	-0.247		-0.417		-0.0853	
	(0.213)		(0.364)		(0.0725)	
Respondent's math ability	0.193*	0.0224	0.343*	0.0288	0.0528**	0.0205
	(0.0779)	(0.113)	(0.139)	(0.196)	(0.0192)	(0.0310)
Resource dilution hypothesis						
Sib-group size	-0.214**	0.214	-0.406**	0.407	-0.0518**	0.0594
	(0.0765)	(0.116)	(0.143)	(0.210)	(0.0164)	(0.0326)
Birth order (2nd vis-à-vis 3rd and above)	-0.343*	0.434	-0.644*	0.796	-0.0911*	0.0915
	(0.158)	(0.232)	(0.291)	(0.408)	(0.0426)	(0.0808)
Spacing b/t focal child and older sibling	-0.0782	0.0588	-0.130	0.110	-0.00213	-0.00438
	(0.0605)	(0.0906)	(0.109)	(0.157)	(0.0165)	(0.0315)
Same-sex sibling competitive stimulation hypothesis						
Older sibling is male	-0.129	0.444	-0.189	0.692	-0.0981*	0.0826
	(0.172)	(0.271)	(0.310)	(0.473)	(0.0490)	(0.0795)
% of boys in a sib-group	0.236*	-0.206	0.418*	-0.345		-0.0324
	(0.107)	(0.168)	(0.192)	(0.292)	(0.0293)	(0.0507)
Role modeling perspective						
Older sibling's math ability	0.216*	-0.376*	0.424*	-0.706**	0.0483*	-0.115***
	(0.102)	(0.150)	(0.194)	(0.268)	(0.0220)	(0.0321)
Role modeling via same-sex sibling competition						
Older sibling's math ability X Older sibling is male	-0.292*	0.393*	-0.547*	0.726*	-0.0768**	0.168***
	(0.125)	(0.176)	(0.232)	(0.310)	(0.0290)	(0.0476)
Constant	-1.026		-1.677		0.260	
	(0.653)		(1.177)		(0.151)	
Wald chi square / F	81.94***		78.27***		3.47***	
Pseudo R^2 / Adjusted R^2	0.069		0.070		0.075	

All models are adjusted for socio-demographic and socio-economic background. Standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Appendix C: Table 3 full models - LPM standardized coefficients predicting college STEM major preference by sibling compositional characteristics, older sibling measurements refer to sibling born just above the respondent (N=1,545)

Preferring STEM vis-à-vis non-STEM major	Model I	Interaction X R is Male	Model II	Interaction X R is Male	Model III	Interaction X R is Male	Model IV	Interaction X R is Male
Male	0.028		-0.080		-0.082		-0.079	
	(0.052)		(0.065)		(0.064)		(0.062)	
Respondent's math ability	0.057**	0.008	0.056**	0.005	0.050**	0.023	0.048*	0.026
	(0.018)	(0.028)	(0.018)	(0.028)	(0.019)	(0.032)	(0.019)	(0.033)
Resource dilution hypothesis								
Sib-group size	-0.038*	0.040	-0.045**	0.045	-0.045**	0.046	-0.049**	0.049
	(0.015)	(0.031)	(0.015)	(0.032)	(0.015)	(0.032)	(0.015)	(0.032)
Birth order (2nd vis-à-vis 3rd and above)	-0.080*	0.108	-0.082*	0.104	-0.084*	0.114	-0.088*	0.115
	(0.041)	(0.069)	(0.040)	(0.068)	(0.041)	(0.069)	(0.041)	(0.068)
Spacing b/t focal child and older sibling	-0.018	0.017	-0.017	0.015	-0.017	0.012	-0.017	0.012
	(0.013)	(0.026)	(0.013)	(0.025)	(0.013)	(0.025)	(0.013)	(0.025)
Same-sex sibling competitive stimulation hypothesis								
Older sibling is male			-0.047	0.165*	-0.047	0.162	-0.040	0.144
			(0.046)	(0.083)	(0.046)	(0.083)	(0.044)	(0.083)
% of boys in a sib-group			0.056*	-0.051	0.055*	-0.044	0.058*	-0.045
			(0.028)	(0.051)	(0.027)	(0.051)	(0.027)	(0.051)
Role modeling perspective								
Older sibling's math ability					0.015	-0.050	0.050*	-0.102**
					(0.017)	(0.030)	(0.023)	(0.039)
Role modeling via same-sex sibling competition							-0.067*	0.101*
Older sibling's math ability X Older sibling is male							(0.031)	(0.051)
Family background								
Birth cohort (ref. category, before 1975)								
1975-1979	-0.009		0.043		0.053		0.062	
	(0.130)		(0.136)		(0.143)		(0.136)	
1980-1984	0.066		0.116		0.125		0.133	
	(0.127)		(0.133)		(0.140)		(0.132)	
1985-1990	0.016		0.067		0.080		0.089	
	(0.127)		(0.134)		(0.140)		(0.133)	
Race (ref. category, Hispanic)								
Black	-0.006		-0.007		-0.005		-0.005	
	(0.033)		(0.033)		(0.033)		(0.033)	
White non-Hispanic	-0.024		-0.025		-0.024		-0.023	

| | | (0.032) | | (0.032) | | (0.032) | | (0.032) | |

Appendix C (continue)

	Model I	Interaction X R is Male	Model II	Interaction X R is Male	Model III	Interaction X R is Male	Model IV	Interaction X R is Male
Two-parent family	0.009		0.012		0.012		0.012	
	(0.030)		(0.030)		(0.030)		(0.030)	
Family income (logged)	0.004		0.007		0.008		0.008	
	(0.014)		(0.014)		(0.014)		(0.014	
Mother's education (ref. category, HS)								
Less than HS	0.0653		0.0583		0.0587		0.0537	
	(0.0613)		(0.0626)		(0.0628)		(0.0626)	
Post-Secondary	-0.0181		-0.0129		-0.0115		-0.0123	
	(0.0306)		(0.0308)		(0.0310)		(0.0311)	
College (16+)	0.0305		0.0308		0.0317		0.0285	
	(0.0392)		(0.0386)		(0.0388)		(0.0385)	
Mother's occupation (ref. category, Out of labor force)								
Male-type occupation	-0.038		-0.041		-0.041		-0.040	
	(0.066)		(0.064)		(0.066)		(0.063)	
Female-type occupation	-0.012		-0.017		-0.021		-0.026	
	(0.045)		(0.044)		(0.043)		(0.043)	
Unemployed	0.031		0.024		0.0259		0.023	
	(0.042)		(0.041)		(0.041)		(0.041)	
Missing	0.033		0.033		0.0345		0.035	
	(0.036)		(0.036)		(0.036)		(0.036)	
Constant	0.194		0.193		0.179		0.174	
	(0.135)		(0.145)		(0.152)		(0.144)	
F	3.20***		3.43***		3.29***		3.30***	
Adjusted R^2	0.052		0.065		0.068		0.072	
AIC	1671.4		1658.5		1657.6		1654.4	
BIC	1799.6		1808.1		1817.9		1825.4	

All models are adjusted for socio-demographic and socio-economic background. Standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Appendix D: LPM standardized coefficients predicting college STEM major preference by
sibling compositional characteristics separately for female and male respondents, older
sibling measurements refer to sibling born just above the respondent (Model IV, full
model, N=1,545)

Preferring STEM vis-à-vis non-STEM major	Females	Males
Respondent's math ability	0.050*	0.073**
	(0.020)	(0.025)
Resource dilution hypothesis		
Sib-group size	-0.054***	0.007
	(0.016)	(0.031)
Birth order (2nd vis-à-vis 3rd and above)	-0.091*	0.033
	(0.042)	(0.055)
Spacing b/t focal child and older sibling	-0.015	-0.004
	(0.013)	(0.023)
Same-sex sibling competitive stimulation hypothesis		
Older sibling is male	-0.040	0.099
	(0.044)	(0.070)
% of boys in a sib-group	0.057*	0.014
	(0.027)	(0.043)
Role modeling perspective		
Older sibling's math ability	0.053*	-0.052
	(0.024)	(0.031)
Role modeling via same-sex sibling competition		
Older sibling's math ability X Older sibling is male	-0.070*	0.032
	(0.031)	(0.041)
Constant	0.182	0.085
	(0.097)	(0.197)
Observations	888	657
AIC	800.5	838.9
BIC	905.9	942.1

All models are adjusted for socio-demographic and socio-economic background. Standard errors in
parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$