A Mother's Influence: How a Mother with an Advanced Degree Impacts Various STEM Outcomes

Hannah Amundson*

Department of Economics

University of Colorado Boulder Honors Thesis

Defended April 3rd, 2017

Primary Advisor:

Richard Mansfield, PhD

Department of Economics

Committee Members:

Terra McKinnish, PhD

Department of Economics

Erin Furtak, PhD

School of Education

Abstract

This study analyzes how a mother with an advanced degree directly and indirectly affects her child's participation in a STEM (Science Technology Engineering Math)

*Acknowledgments:** I would like to thank Dr. Richard Mansfield for his enthusaism, support, and guidance throughout this process. I would also like to thank Dr. Terra McKinnish and Dr. Martin Boileau for organizing the honors course. Finally, I would like to thank Dr. Erin Furtak for being willing to serve on my committee.

occupation, with a specific focus on differences between females and males. Indirect effects are determined by examining the effects a mother with an advanced degree has on various adolescent and adult predictors of a STEM occupation, including: math self-efficacy, math abilities, number of hours spent on homework, motivation, future expected career, ACT scores, math ACT scores, number of advanced STEM courses in high school and postsecondary school, high school and postsecondary GPA, Calculus coursework, and college major. This study uses the Education Longitudinal Study of 2002 from the National Center for Education Statistics. This research finds support for previous literature that these STEM predictors actually do predict a major and occupation in STEM. Further, this study finds that a mother with an advanced degree has both direct and indirect impacts on STEM predictors throughout all time periods, most notably in the high school achievement period. Finally, this paper finds that a mother with an advanced degree has both a negative direct and indirect effect on her son's participation of a STEM occupation, but a positive effect on her daughters.

1 Introduction

STEM[1] (Science Technology Engineering Math) majors and STEM occupations are some of the most lucrative majors and occupations (Zafar, 2013). Many STEM occupations are male-dominated though, and there has been a growing concern about the gender gap in this field. For instance, the Bureau of Labor Statistics reports that only 13.6% of architects and engineers were women in 2011 (bur, 2011). Since lucrative fields, such as STEM, are male-dominated, a gender gap in earnings is created in the general population. Due to this, it is important to examine what influences students to choose STEM and non-STEM occupations.

I will examine how a mother with an advanced degree directly and indirectly affects a daughter's likelihood of participation in a STEM occupation compared to a son's. Direct effects include how a STEM occupation is explicitly regressed on maternal attainment of an advanced degree. Indirect effects, on the other hand, will initially look at the impact of maternal attainment of an advanced degree on four different pivotal STEM predictor periods: high school STEM preparedness, high school achievement, between high school and postsecondary school, and postsecondary STEM achievement. STEM occupation will then be regressed on these predictors to see how mother's educational attainment indirectly affects STEM occupation. The data used are from the Educational Longitudinal Study of 2002 from the National Center for Education Statistics.

Previous literature found that the disproportionate female STEM representation is due to pre-college factors and that same-gender educators appear to impact STEM participation (Zafar, 2013; Ceci et al., 2014; Hoffmann and Oreopoulos, 2009; Griffith, 2014). Further studies have found that fathers have an impact on a student's major choice (Leppel et al., 2001). However, there currently exists a gap that examines how same-gender role models at younger ages impact a daughter's participation in a STEM occupation. There is also not research on how same-gender role models indirectly impact STEM occupation participation. Thus, this research will try to fill that gap by examining how mother inputs

[1] In this paper, STEM is defined by the STEM SMART grant definition (sma, 2016). A list of majors included in the STEM SMART grant definition can be found in Appendix A.

specifically impact a woman's occupation choice directly and indirectly in adolescence and adulthood.

The remainder of this paper goes as follows: Section 2 examines the literature surrounding women in STEM and same-gender role models. Section 3 discusses the conceptual framework of the educational production function. Section 4 presents specifications to and necessary assumptions of the educational production function. Section 5 describes the the data used from the Education Longitudinal Study of 2002. Section 6 presents estimated coefficients for these models. Section 7 provides interpretations of the estimated parameters of these inputs, examines potential limitations and robustness checks, and concludes.

2 Literature

Prior literature suggests same-gender role models influence college major choice (Hoffmann and Oreopoulos, 2009; Griffith, 2014). The gender breakdown of students in various fields is influenced by the lack of same-gender role models. Griffith (2014) and Rask and Tiefenthaler (2008) found that female students in male-dominated, quantitative fields are slightly more likely to get a higher grade with a female professor, and, similarly, men in female-dominated fields were more likely to get a higher grade with a male professor. They also found that male students were slightly more likely to get a lower grade with a female professor. Further, Zafar (2013) and Griffith (2010) found that the grades a student gets in a field positively impact the student's continuation in the major. Therefore, the lack of same-gender role models for women in male-dominated fields negatively impacts their grades, which in turn discourages them from further pursuing the major. Thus high ratios of men to women professors in STEM fields are pushing women out of these fields while simultaneously pulling men into them.

While same-gender educators appear to influence students, there are various explanations as to why that is. Hoffmann and Oreopoulos (2009) hypothesized that these role models appear to influence students of the same gender because women know how to relate and teach women better while men know how to relate and teach men better.

Another explanation is that same-gender educators are influential due to differences in perceived gender preferences. Zafar (2013) found that males and females both value majors by workplace outcomes, however they value different workplace outcomes; males value pecuniary outcomes, such as social status of the job, while women value non-pecuniary outcomes, such as reconciling work and family or enjoying work. A student will choose a college major based on whether they think that college major will or will not fit in with their preferences. When women do not see many same-gender educators in a field, beliefs that the field will not balance work and family or that the field is not enjoyable by women are reinforced. On the other hand, when women see female faculty in STEM, these beliefs are challenged.

While it appears that same-gender role models influence students, these role models are more influential at earlier points in life. Hoffmann and Oreopoulos (2009) found that role models were less likely to impact students in higher levels of college than students in lower levels of college. Further, Zafar (2013) claims that the preferences that determine choice of college major appear to already be set by college. Thus he recommends examining how these preferences are set at earlier points in life. This was further supported when Bottia et al. (2015) found that students who attended high schools with more female math and science teachers were more likely to both declare and finish a STEM major in college. Similarly, Ceci et al. (2014) found that pre-college life experiences create the gender disparity in quantitative fields.

Since same-gender role models influence students more at earlier life stages, parents have a significant impact on a child's education attainment (Zafar, 2013; Leppel et al., 2001). Specifically, a child's education experience is impacted by the amount of education a parent obtains. Huesmann et al. (2010) found that the amount of parental education indirectly affects a child's education attainment through adolescent aspirations and educational success. However, there are more direct effects as well. Zafar (2013) found that a student is more likely to choose a major he or she thinks will most gain the approval of his or her parents. Further, Rothstein (1995) found that a mother's educational attainment influences whether her daughter obtains an advanced degree. More research needs to be done as to why a mother's educational attainment influences education attainment.

Perhaps the most closely related work to my own is *The Impact of Parental Occupation and Socioeconomic Status on Choice of College Major* by Leppel et al. (2001). These researchers find that daughters who had mothers or fathers in an executive or professional career were more likely to choose science/engineering. While Leppel, Williams, and Waldauer examined how parental careers influence a college choice of major, my research will examine how parental educational attainment impacts a daughter's choice of major both directly and indirectly.

3 Conceptual Framework

I am attempting to discover the degree to which a mother with an advanced degree impacts her child's occupational choice. Further, I am trying to discover the specific ways in which this advanced degree will impact occupational choice. Perhaps a mother who has an advanced degree helps their child by giving them advice, helping them with job connections, or only because they teach them proper study skills at younger ages. I will use a STEM production function to examine the mechanisms a mother with an advanced degree uses.

An educational production function looks at the relationship between various educational outcomes and a series of inputs (Hanushek, 1986). Similarly, a STEM production function is a production function that looks at STEM specific outcomes. The STEM production function will be similar to an educational production function in the sense that it will mostly measure educational outcomes, however it will be extended in that it will also measure one occupational outcomes.

The STEM production function varies based on time period. In this case, assume there are five different time periods: high school preparedness up through sophomore year, high school achievement by the end of high school, the period between high school and postsecondary school, postsecondary school achievement by the end of postsecondary school, and an occupation period that begins eight years after most of these students graduate high school. While these specific periods are somewhat arbitrary, I chose these periods because they each have a natural collection of measurable outcomes. Then individuals produce a variety of STEM outcomes in these periods that later predict participation in a STEM occupation. Let these period-dependent outcomes be denoted by the following sets:

$$HSP := \text{set of outcomes from high school preparedness period} \tag{1}$$

$$HSA := \text{set of outcomes from high school achievement period} \tag{2}$$

$$BHSPS := \text{set of outcomes from between high school and postsecondary school period} \tag{3}$$

$$PSA := \text{set of outcomes from postsecondary achievement period} \tag{4}$$

$$O := \text{set of outcomes from occupation period} \tag{5}$$

Elements in these sets are outcomes specific to that period.

Additionally, certain outcomes from prior periods become inputs into future outcome production functions. Consider, for instance, the number of advanced STEM courses a student took in high school, which would be an outcome in HSA. Perhaps this outcome serves as a input into the PSA production function, as the increased exposure to these difficult classes in high school made STEM courses in postsecondary school seem easier. Then outcome $psa \in PSA$ production function takes some inputs from the high school achievement outcomes, denoted $HSA_{psa} \subseteq HSP$. In this case, HSA_{psa} is a set of outcome variables from the high school achievement period that serve as inputs for the a specific outcome in the postsecondary school achievement period. Thus a subscript on sets (1) through (5) denotes an outcome in a time period for which a subset of these prior outcomes become inputs.

Similar to the educational outcomes, these STEM outcomes are created from a variety of environmental and innate inputs (Hanushek, 1986; Levin, 1970). Though there are a variety of other factors inputed into the production function, I will specifically focus on inputs from mothers. Conceptually, mothers give their children different tools and skills, such as math help or job advice, throughout the five periods. Some of these tools and skills influence STEM outcomes. These tools and skills given by the mother are represented with the following five sets

$$
\begin{aligned}
M^{HSP} &:= \text{set of tools and skills mother generally provides individual} \\
&\qquad \text{during high school preparedness period} \\
M^{HSA} &:= \text{set of tools and skills mother generally provides individual} \\
&\qquad \text{during high school achievement period} \\
M^{BHSPS} &:= \text{set of tools and skills mother generally provides individual} \\
&\qquad \text{during between high school and postsecondary school period} \\
M^{PSA} &:= \text{set of tools and skills mother generally provides individual} \\
&\qquad \text{during postsecondary school achievement period} \\
M^{O} &:= \text{set of tools and skills mother generally provides individual} \\
&\qquad \text{during occupation period} \qquad (6)
\end{aligned}
$$

Elements in these sets represent a specific tool or skill given in that period. Different mothers likely give these tools in different time periods, however these inputs are put into sets based on when they are generally given.

The STEM production functions are cumulative in the sense that they looks at not only contemporaneous inputs, but also prior inputs (Levin, 1970). These inputs can come into play at later periods even if they did not come into play in the period they were given. For instance, a mother may teach study skills in the high school achievement period, but for whatever reason this skill may not come into play until the postsecondary achievement period. Then M_{psa}^{HSA} denotes the subset of skills given in the high school achievement period that come into play for an outcome $psa \in PSA$. Thus the superscript denotes the time period in which the mother gave the skill, and the subscript denotes the outcome for which the skills came into play.

Previous outcome variables should still be in current STEM production function even if the current STEM production function accounts for previous mother inputs that affected this previous outcome variable. For example, ACT score should be included in a postsecondary achievement production function even if all mother inputs into ACT score were also included. This is because there could be some variability in the previous outcome that mother inputs could not account for. For instance, perhaps a student got lucky on his or her ACT score and correctly guessed on a large sum of questions. Thus their score was not based on mother inputs, but on luck. However, perhaps this higher ACT score positively impacted his or her future STEM outcomes. If ACT score was not accounted for in a future production function, this additional boost from the ACT score could not be accounted for.

Additionally, there are extra inputs into the STEM production function that are not given by the mother. Let these other inputs be represented by the set

$$X := \text{set of all other inputs} \tag{7}$$

Subsets of (7) come into play at different times. For instance, $X_{hsp} \subseteq X$ denotes additional inputs for outcome $hsp \in HSP$. Finally, there is an error term that allows for additional variability within each outcome. A subscript on this error term, ϵ, denotes which outcome the error term corresponds to.

With these inputs in mind, the high school preparedness production function is

$$hsp = f_{hsp}(M_{hsp}^{HSP}, X_{hsp}, \epsilon_{hsp}) \tag{8}$$

for all outcomes $hsp \in HSP$. Since there is not a specific input for childhood inputs, this first production function absorbs any early life differences between individuals, including both innate and environmental factors. The remaining STEM production functions should not be influenced by these early life inputs because the outcomes in the first production function absorbed them. The high school achievement production function is

$$hsa = f_{hsa}(M_{hsa}^{HSP}, M_{hsa}^{HSA}, HSP_{hsa}, X_{hsa}, \epsilon_{hsa}) \tag{9}$$

for all outcomes $hsa \in HSA$. The between high school and postsecondary school production function is

$$bhsps = f_{bhpsps}(M_{bhsps}^{HSP}, M_{bhsps}^{HSA}, M_{bhsps}^{BHSPS}, HSP_{bhsps}, HSA_{bhsps}, X_{bhsps}, \epsilon_{bhsps}) \tag{10}$$

for all outcomes $bhsps \in BHSPS$. The postsecondary achievement production function is

$$psa = f_{psa}(M_{psa}^{HSP}, M_{psa}^{HSA}, M_{psa}^{BHSPS}, M_{psa}^{PSA}, HSP_{psa}, HSA_{psa}, BHSPS_{psa}, X_{psa}, \epsilon_{psa}) \tag{11}$$

for all outcomes $psa \in PSA$. Finally, the occupation production function is

$$o = f_o(M_o^{HSP}, M_o^{HSA}, M_o^{BHSPS}, M_o^{PSA}, M_o^{O}, HSP_o, HSA_o, BHSPS_o, PSA_o, X_o, \epsilon_o) \tag{12}$$

for all outcomes $o \in O$.

4 Empirical Specification

4.1 Defining The Production Function

The production functions produce STEM outcomes throughout the five time periods for individual i. Let $s(i)$ refer to the high school that individual i attended. Assume the production functions are linear[2].

Suppose (6) varies based on a mother's level of education. These production functions will specifically examine how mother inputs vary based on whether or not she has an advanced degree. Then MAD_i represents inputs specific to mothers with advanced degrees. This is a binary variable that takes a value of one when individual i's mother has an advanced degree and zero otherwise. Since gender appears to impact STEM outcomes, gender also needs to be incorporated into the production function (Zafar, 2013). Hence let FEM_i represents a binary variable that has a value of one when the individual i is female, and zero if male. An interaction term between MAD_i and FEM_i allows for differences between males and females in higher educated mothers inputs. Let I_i, F_i, and $S_{S(i)}$ represent a vector of individual, familial, and high school controls dependent on i, respectively. Then (7) is redefined as a vector $X_i := [I_i, F_i, S_{s(i)}]$. Additionally, (7) includes father inputs, such as whether or not he has an advanced degree. This is represented as the binary variable FAD_i, which takes a value of one if the father has an advanced and a value of zero otherwise. An additional interaction term between FAD_i and FEM_i allows

[2]Though the literature does provide some examples of nonlinear education production functions, many education production functions are linear (Hanushek, 1979).

comparisons between mother inputs and father inputs amongst the genders. Let $HSPO_i$, $HSAO_i$, $BHSPSO_i$, and $PSAO_i$ represent a vector of outcomes that correspond to the elements in (1), (2), (3), and (4) respectively.

It is difficult to measure prior and contemporaneous subsets of (6). Attempts to measure mother contemporaneous inputs can result in unintentionally picking up residual components from previous mother inputs, as well as residual components from X_i. To account for this issue, assume prior inputs from mothers only impact prior outcomes for the remaining sections. These prior outcomes then absorb any effects from prior inputs. For example, assume that in (9) HSP_{hsa} fully absorb M_{hsa}^{HSP}. Then MAD_i will only pickup inputs from the current period. Finally, assume X_i fully captures all other inputs into the production functions, so that MAD_i is exogenous and the residual component of X_i is uncorrelated with mother's education.

With equation (8) and these assumptions in mind, the high school preparedness production function is modeled as

$$\begin{aligned} outcome_{ji} = \beta_{0j} + \beta_{Xj} X_i + \beta_{FEMj} FEM_i + \beta_{FADj} FAD_i + \beta_{FADFEMj} FAD_i \times FEM_i \\ + \beta_{MADj} MAD_i + \beta_{MADFEMj} MAD_i \times FEM_i + \epsilon_{ji} \end{aligned} \tag{8A}$$

$\forall$ outcomes $j \in HSP$. The high school achievement production function (9) is modeled as

$$\begin{aligned} outcome_{ki} = \beta_{0k} + \beta_{Xk} X_i + \beta_{FEMk} FEM_i + \beta_{FADk} FAD_i + \beta_{FADFEMk} FAD_i \times FEM_i \\ + \beta_{MADk} MAD_i + \beta_{MADFEMk} MAD_i \times FEM_i + \alpha_{HSPOk} HSPO_{ki} + \epsilon_{ki} \end{aligned} \tag{9A}$$

$\forall$ outcomes $k \in HSA$. The between high school and secondary school production function (10) is modeled as

$$\begin{aligned} outcome_{ri} = \beta_{0r} + \beta_{Xr} X_i + \beta_{FEMr} FEM_i + \beta_{FADr} FAD_i + \beta_{FADFEMr} FAD_i \times FEM_i \\ + \beta_{MADr} MAD_i + \beta_{MADFEMr} MAD_i \times FEM_i + \alpha_{HSPOr} HSPO_{ri} \\ + \alpha_{HSAOr} HSAO_{ri} + \epsilon_{ri} \end{aligned} \tag{10A}$$

$\forall$ outcomes $r \in BHSPS$. The postsecondary achievement production function (11) is modeled as

$$\begin{aligned} outcome_{li} = \beta_{0l} + \beta_{Xl}X_i + \beta_{FEMl}FEM_i + \beta_{FADl}FAD_i + \beta_{FADFEMl}FAD_i \times FEM_i \\ + \beta_{MADl}MAD_i + \beta_{MADFEMl}MAD_i \times FEM_i + \alpha_{HSPOl}HSPO_{li} \\ + \alpha_{HSAOl}HSAO_{li} + \alpha_{BHSPSOl}BHSPSO_{li} + \epsilon_{li} \end{aligned} \tag{11A}$$

$\forall$ outcomes $l \in PSA$. Finally, the occupation production function (12) is:

$$\begin{aligned} outcome_{hi} = \beta_{0h} + \beta_{Xh}X_i + \beta_{FEMh}FEM_i + \beta_{FADh}FAD_i + \beta_{FADFEMh}FAD_i \times FEM_i \\ + \beta_{MADh}MAD_i + \beta_{MADFEMh}MAD_i \times FEM_i + \alpha_{HSPOh}HSPO_{hi} \\ + \alpha_{HSAOh}HSAO_{hi} + \alpha_{BHSPSOh}BHSPSO_{hi} + \alpha_{PSAOh}PSAO_{hi} + \epsilon_{hi} \end{aligned} \tag{12A}$$

$\forall$ outcomes $h \in O$.

Under these assumptions, β_{MADt} and $(\beta_{MADt} + \beta_{MADFEMt})$ for outcome $t \in HSP \cup HSA \cup BHSPS \cup PSA \cup O$ are interpreted as the difference in inputs between mothers with an advanced degree and mothers without an advanced degree for males and females respectively. This analysis will not distinguish between the difference in magnitude of the inputs and the the difference in the number of inputs between more and less educated mothers. For instance, a mother with an advanced degree may positively impact ACT score because she forced her child to go to an ACT prep course (different number of inputs) or because she emphasized the importance of ACT scores more than a mother with less education (difference in magnitude of inputs).

I hypothesize that both β_{MADt} and $(\beta_{MADt} + \beta_{MADFEMt})$ will be greater than zero for all outcomes that positively predict future STEM outcomes, and less than zero for all outcomes that are inversely related to future STEM outcomes. That is, I hypothesize that mother with advanced degree inputs are larger for a son or daughter than mother without advanced degree inputs. Taking into account same-gender role model hypothesis, I further predict that $(\beta_{MADt} + \beta_{MADFEMt})$ will be larger than β_{MADt} for outcomes that positively predict STEM outcomes, and will be smaller for outcomes that are inversely related to STEM outcomes. That is, I hypothesize that mother with advanced degree inputs are larger for daughters than sons. Finally, I hypothesize that $(\beta_{MADt} + \beta_{MADFEMt})$ is larger than

$(\beta_{FADt} + \beta_{FADFEMt})$. In other words, I assume that mother with advanced degree inputs are larger on a daughter than are father with advanced degree inputs.

4.2 Calculating Direct and Indirect Effects

The mother adds direct effects and indirect effects through predictors of that period's success. Define outcome $z \in HSP \cup HSA \cup BHSPS \cup PSA \cup O$. Then a mother with an advanced degree has a β_{MADz} direct effect on men and $(\beta_{MADz} + \beta_{MADFEMz})$ direct effect on women for outcome z.

Let the function $period(x)$ be defined as the period of variable x. For instance, $period(ACTscore) = HSA$. Define the function $prior(y)$ as a union of all periods before period y. For instance, $prior(PSA) = HSP \cup HSA \cup BHSPS$. Then for each vector of prior outcomes p for variable v, the sum of all $\alpha_{pv} p_{vi}$ can be written as

$$\sum_{q \in prior(period(v))} \alpha_{qv} q_i$$

Then a mother with an advanced degree has an indirect effect for variable $v \in HSA \cup BHSPS \cup PSA \cup O$ of magnitude

$$\sum_{q \in prior(period(v))} \alpha_{qv} \beta_{MADq}$$

on her son, and an indirect effect of magnitude

$$\sum_{q \in prior(period(v))} \alpha_{qv} (\beta_{MADq} + \beta_{MADFEMq})$$

on her daughter.

A mother with an advanced degree has a total effect on variable v, which is defined as the sum of the direct and indirect effects for v.

5 Data

5.1 Dataset

The dataset used is the restricted access Educational Longitudinal Study of 2002 (ELS 2002) from the National Center for Education Statistics (els, 2002). The 16,197 individuals from 750 high schools are a nationally representative sample of students in the United States, and were obtained through cluster sampling.

This dataset utilized is panel data. The nature of this research requires access to student inputs at various points in time. Fortunately, ELS 2002 surveys students over a course of ten years, initially surveying students in tenth grade (2002), and then following up two years later (2004), four years later (2006), and ten years later (2012). ELS 2002 includes high school transcripts, postsecondary transcripts, and higher entrance exam data. Furthermore, this research requires access to parental and nonparental inputs. ELS 2002 surveys parents, high school administrators, and high school teachers.

There are a few limitations to this dataset. Ideally this dataset would obtain parental inputs throughout the ten years; unfortunately, it only surveys the parents in the base year. Nevertheless, this dataset has a rich collection of parental inputs. Another problem is that this dataset has a large amount of missing data. Percentages of missing data for primary independent variables, outcome variables, and control variables can be found in Table 1, Table 2, and Appendix B.

5.2 Measuring Time Periods

The framework is based on five different time periods: high school preparedness, high school achievement, between high school and postsecondary school, postsecondary school achievement, and occupation period. This section will define how these periods will be empirically measured with ELS 2002.

High School Preparedness

Previous literature suggests various predictors of a STEM occupation or STEM college major consistent with the high school preparedness period (Moakler and Kim, 2014; McDill et al., 1967; Wang, 2013; Bieri Buschor et al., 2014; Veenstra, 2008). Thus I will measure high school preparedness with five different variables and now define the set (1) as

$$HSP := \{\text{math self-efficacy, math test score, time spent on homework,}$$
$$\text{motivation, prediction of STEM occupation}\} \quad (1A)$$

Math self-efficacy is the belief that one can succeed and accomplish tasks in mathematics (Bandura, 1977). This variable was created by standardizing individual's personal rankings of the following five statements: "Can do excellent jobs on math test," "Can understand difficult math texts," "Can understand difficult math class," "Can do excellent jobs on math assignments," and "Can master math class skills." Higher values of math self-efficacy mean the individual has more self-efficacy.

Math test score was taken from an math test administered by ELS 2002 at the beginning of the study. This score was then estimated to the population as a whole to assess an individual's abilities compared to his or her peers.

Hours spent working on homework is the number of hours a student spent on homework in school or out of school per week. Students who spent more than 21 hours on homework out of school or 26 hours in school were coded as missing.

Motivation is a variable that measures extrinsic motivation (motivation to accomplish goals). It was constructed by standardizing individual's personal rankings of the following three statements: "Studies to get a good grade," "Studies to increase job opportunities," "Studies to ensure financial security." This variable also incorporates awareness of nonpecuniary preferences. Higher values mean the individual has higher extrinsic motivation.

Prediction of STEM occupation is a binary variable that measures whether or not a student predicted having a STEM occupation at age thirty as of their tenth grade year. A value of one means the student predicted having a STEM occupation at age thirty, and a value of zero means the student predicted having a non-STEM occupation at age thirty.

High School Achievement

Prior literature suggests various predictors of a STEM occupation or college major consistent with the high school achievement period (Bottia et al., 2015; Griffith, 2010; Moakler and Kim, 2014; Wang, 2013; Veenstra, 2008). Hence I will measure this period with five different variables, and redefine (2) as the set

$$HSA := \{\text{ACT score, math ACT score, advanced STEM courses, GPA,} \\ \text{Calculus coursework}\} \quad (2A)$$

ACT score is an individual's overall ACT (Advanced College Test) or SAT (Scholastic Achievement Test) score, converted into ACT units. Both of these tests are higher education placement exams. Math ACT score measures the score an individual got on the math section of the ACT or SAT, converted into ACT units.

Advanced STEM courses is a measurement of the number of AP/IB (Advanced Placement/International Baccalaureate) STEM courses the individual took in high school. GPA (grade point average) is high school GPA converted to a four point scale.

Calculus coursework is a binary variable that measures whether a student took any Calculus coursework in high school. A value of one indicates an individual did take at least one Calculus course, while a value of zero indicates an individual did not.

Between High School and Postsecondary School

This period uses two variables to measure the period between high school and postsecondary school. Thus the set (3) is represented as

$$BHSPS := \{\text{time between high school and postsecondary school,} \\ \text{attendance of postsecondary institution}\} \quad (3A)$$

The time between high school and postsecondary school is measured as the number of months an individual took between high school graduation and postsecondary school attendance. This variable accounts for prior literature, which suggests that a higher percentage of younger students major in STEM fields than than older students (Chen, 2009).

Attendance of postsecondary institution is a binary variable. It has a value of one if the student ever attended a postsecondary institution and zero if the student did not. This variable is included because higher education typically is associated with a STEM occupation (Beede et al., 2011).

Postsecondary School Achievement

Literatures suggests various predictors that measure participation in a STEM occupation that are consistent with the postsecondary school achievement time period (Glass et al., 2013; Griffith, 2014; Rask and Tiefenthaler, 2008). I will measure postsecondary school achievement with three variables and define (4) as

$$PSA := \{\text{number of STEM courses, STEM GPA, STEM major}\} \tag{4A}$$

The number of STEM courses is a variable that measures how many STEM courses an individual took in postsecondary school. STEM GPA measures the individual's GPA in all STEM courses at their postsecondary institution, converted to a four point scale. STEM major is a binary variable that indicates whether a student majored in a STEM field at their postsecondary institution. It has a value of one if he or she majored in a STEM field and a value of zero if he or she did not.

Occupation

The occupational period will be measured with one variable, whether or not the individual works in a STEM field. Thus the set (5) is defined as

$$O = \{\text{STEM occupation}\} \tag{5A}$$

STEM occupation is a binary variable that has a value of one if the individual works in a STEM field and a value of zero if the individual does not work in a STEM field as of 2012. If the individual was on track to graduate high school, this is eight years after high school graduation.

5.3 Sample Characteristics

Table 1 displays the means and standard deviations for the primary independent variables of interest. This table shows that 50% of the data are female, 9.2% have a mother with an advanced degree, and 14.2% have a father with an advanced degree.

Table 2 displays the means, standard deviations, minimum values, and maximum values for all outcome variables. Table 3 shows the proportion of each binary variable that is female or male. For instance, it shows that 36% of STEM majors are women.

Males and females have similar averages for most outcomes, however there are a few exceptions. Math self-efficacy in the high school preparedness period displays that the average female has a math self-efficacy score that is 0.111 standard deviations below the mean, while the average male has a math self-efficacy score that 0.182 standard deviations above the mean. However, females on average are .08 standard deviations above males on the motivation scale.

Table 1 shows that 12% of males predict having a STEM occupation at the age of thirty, whereas only 4.2% of females do. Table 3 shows that 27.08% of individuals who predicted having a STEM occupation at 30 were female, while 72.9% were male. Table 1 shows that males tend to preform a little over one point higher than women on the math section of the ACT, though this difference is not seen in the overall ACT score. On the other hand, women's GPA is 0.341 above men's on average. Men tend to take 2.048 more STEM courses than women in postsecondary school, though women's average STEM GPA is 0.154 higher than men's average STEM GPA. Almost a quarter of men major in a STEM field, while only about 10% of women major in a STEM field. Table 3 shows that 36% of STEM majors are women, while 63% are men. Table 1 shows that 9.5% of men and 3% of women work in a STEM occupation at the third follow up. Table 3 shows women make up about 24% of STEM occupations, while men make up 75%. This is consistent with literature (bur, 2011; Beede et al., 2011).

Table 2 shows that an individual who has a mother with an advanced degree has higher means for every outcome variable than an individual who does not have a mother with an advanced degree, except the amount of time taken between high school and postsecondary school. Further, it shows that more individuals with higher educated mothers major and work in STEM fields than individuals with less educated mothers. Hence this table supports the hypothesis that a mother with an advanced degree does impact STEM participation throughout the life course.

Table 2 shows that there is a large amount of missing data. This creates the risk of looking at unintended populations. Assume these variations in missing data do not change the population being looked at across variables. Potential problems with this assumption will be discussed in section 7.

A list of sample statistics for individual, family, and high school controls can be found in Appendix B.

6 Results

Tables 4 through 8 contain regressions split by the five time period. These tables include a mother's direct, indirect, and total effects on a series of variables. These effects are categorized by gender. The mother's direct effect is the amount to which a mother with an advanced degree impacts that specific dependent variable. The mother's indirect effect in each table is the amount a mother with an advanced degree impacts predictors of the variable. It is cumulative in that Table 6's indirect effects are the results of Table 4 and Table 5, and so on. The mother's total effect in each table is a sum of the mother's direct and indirect effect. All tables control for individual, family, and high school variables. Tables that display control coefficients can be found in Appendix C.

Ordinary least square regression is reported for continuous variables, while marginal effects at the means from probit model are reported for binary variables in all regression tables. R-squared means pseudo R-squared for binary variables.

6.1 Statistical Significance

A majority of the reported coefficients in tables 4 through 8 are not statistically significant. That is, there is not sufficient evidence to conclude that the population coefficients are not equal to zero. Specifically, there were almost interaction terms between mother with an advanced degree and female that were statistically significant. Even though these coefficients are imprecisely estimated, they are still the best estimate we have for the population coefficient.

6.2 Direct and Indirect Effects

High School Preparedness

Table 4 contains the regressions of the measurements of high school preparedness on all students in the sample, including math self-efficacy, hours spent on homework per week, math test score, motivation score, and prediction of STEM occupation at age thirty.

Columns (a) and (d) in Table 4 represent the amount to which a one unit increase in a regressor impacts the standard deviation of an individual's math self-efficacy score and motivation score, respectively. Hence a mother with an advanced degree increases her son's math self-efficacy score by 0.032 standard deviations and increases her daughter's math self-efficacy score by 0.082 standard deviations. Further, a mother with an advanced degree increases her son's motivation score by 0.079 standard deviations and decreases her daughter's motivation score by 0.047 standard deviations.

Columns (b) and (c) represent the amount to which a one unit increase in a regressor increase the number of hours spent on homework per week and the math test score, respectively. So a mother with an advanced degree increases her son's hours spent on homework by 0.199 hours and decreases her daughter's by 0.453 hours. Meanwhile, a mother with an advanced degree increases her son's math test score by 1.176 points while only increasing her daughter's by 0.838 points.

Column (e) represent the amount to which a one unit increase in a regressor at the means increases the probability that a student predicts having a STEM occupation at age thirty. This column shows that a mother with an advanced degree decreases her son's probability of predicting a STEM occupation by 0.8% and increases her daughters by 0.6% at the means.

Table 4 shows that a mother with an advanced degree directly impacts a son more than a daughter for every measurement of high school preparedness except math self-efficacy and prediction of STEM occupation at age thirty, which is contrary to same-gender role model hypothesis.

Mothers with an advanced degree have a larger effect than fathers in all cases except hours spent on homework per week and motivation score for women. This is consistent with same-gender role model theories. A father with an advanced degree has a larger impact than mothers for men in all categories except motivation.

High School Achievement

Table 5 contains the regressions of the measurements of high school achievement on all students in the sample, including: the individual's overall ACT score, math ACT score, the number of AP/IB STEM courses taken, GPA, and whether the individual took Calculus.

Columns (a), (b), (c), and (d) in Table 5 represent the amount to which a one unit increase in a regressor increase the dependent variable. A mother with an advanced degree directly increases her son's and her daughter's respective overall ACT score by 0.617 and 0.439 points; math ACT score by 0.487 and 0.336 points; number of AP/IB STEM courses by 0.234 and 0.109 Carnegie Units; and GPA by 0.075 and 0.072. Thus a mother impacts her son more than her daughter in all categories. Column (e) represents the amount to which a one unit increase in a regressor impacts the probability that a student took Calculus in high school at the means. A mother with an advanced degree directly increases the probability that her son took Calculus by 4.9%, while only increasing her daughter's probability by 2.7% at the means. This is contrary to the hypothesis that a mother would increase her daughter's high school achievement more than her son's.

Further, Table 5 shows that all male indirect effects are greater than female indirect effects, and all male total effects are larger than female total effects. This is not consistent with initial same-gender role model hypothesis. For example, this table displays that a mother with an advanced degree indirectly increases her son's ACT score by 0.456 points, and in total increases her son's ACT score by 1.074 points.

Table 5 shows that a father with an advanced degree impacts his daughter more than a mother with an advanced degree in every category except the number of advanced STEM courses taken and Calculus coursework. In both of these cases a father with an advanced degree decreases the number of courses and probability she took Calculus by 0.064 courses and 0.2%, while a mother with an advanced degree increases it by 0.109 courses and 2.7%. Similarly, a father with an advanced degree impacts only impacts his son more than a mother with an advanced degree in ACT score and high school GPA. A mother with an advanced degree increases her son's math ACT score, number of advanced STEM courses, and the probability he took Calculus by 0.617 points, 0.234 courses, and 4.9%, while a father with an advanced degree increases it by 0.879 points, increases it by 0.061 courses, and decreased it by 1.6%. Thus a father has a larger effect on more categories for his daughters than his sons, while a mother has a larger effect on daughters than sons. This is contrary to same-gender role model hypothesis.

Between High School and Postsecondary School

Table 6 contains the regressions of the measurements of the between high school and postsecondary school regression period, which includes both the number of months between high school and postsecondary attendance and whether the individual ever attended a postsecondary institution. Column (b) was regressed on the entire sample, while column (a) was regressed on any student that reported ever attending a postsecondary institution.

Column (a) in Table 6 represents the amount to which a one unit increase in a regressor impacts the number of months between high school and postsecondary school attendance. A mother with an advanced degree directly decreases the number of months her son and daughter, respectively, take between high school and postsecondary school by 0.276 months and 0.554 months. Column (b) represents how much a one unit increase in a regressor increases the probability that an individual attends a postsecondary institution at the means. A mother with an advanced degree directly decreases the probability that her son attends a postsecondary institution by 0.2% at the means, and her daughter by 0.3%. This is inconsistent with same-gender role model hypothesis.

Table 6 shows that a mother with an advanced degree indirectly decreases the number of months between schooling for males by 0.146 months and for women by 0.109 months. In total, a mother with an advanced degree decreases the number of months between schooling for her son by 0.422 months and for her daughter 0.663 months. Further, in total a mother increases the probability a son attends a postsecondary institution by 0.3% at the means, while increasing a daughter's probability by 0.7%. These results support same-gender role model hypothesis.

Column (a) in Table 6 shows that a father with an advanced degree decreases the number of months his daughter takes between schooling by 0.366 months, while a mother only decreases it by 0.089 months for a daughter. This is consistent with same-gender. However, a mother with an advanced degree decreases the number of months her son takes between school by 0.276 months for her son, while a father increases it by 0.187 months.

Postsecondary School Achievement

Table 7 contains regressions for the postsecondary achievement period, which includes the number of STEM courses taken, STEM GPA, and whether a student majored in a STEM field.

Column (a) and (b) in Table 7 represent the extent to which a one unit increase in a regressor increases the number of STEM courses and STEM GPA. A mother with an advanced degree directly decreases the number of STEM courses her son takes by 0.198 courses, and directly decreases his STEM GPA by 0.033 points. On the other hand, she directly increases her daughter's total number of STEM courses by 0.282 courses and STEM GPA by 0.16 points. This supports the same-gender role model hypothesis. Column (c) displays the degree to which a one unit increase in a regressor increases the probability that an individual majored in a STEM field at the means. Shockingly, a mother with an advanced degree directly decreases the probability that her son majors in a STEM field by 0.3% at the means and decreases her daughters by 0.8% at the means. This is contrary to the initial hypothesis.

However Table 7 shows that a mother with an advanced degree has an overall positive total effect on the probability that her son or daughter majors in STEM and does not have an effect on the probability that her daughter majors in STEM. Additionally, a mother with an advanced degree indirectly increases the probability that both her son and daughter majors in STEM by 1.2% at the means.

Table 7 shows that fathers have more of an impact on their sons than mothers do for number of STEM courses taken and STEM GPA. A father with an advanced degree increases a son's total number of STEM courses and STEM GPA by 0.543 courses and

0.083, while a mother with an advanced degree increases it by 0.198 courses and decreases it by 0.033. Contrarily, a father with an advanced degree decreases the probability his son majors in a STEM field by 1.9%, while a mother only decreases it by 0.3%. On the other hand, a mother with an advanced degree increases her daughter's number of STEM courses and STEM GPA by 0.282 courses and 0.16, while a father with an advanced degree decreases it by 1.14 courses and 0.0.082. However, a father with an advanced degree increases his daughter's probability of majoring in a STEM field by 0.8% while a mother decreases her daughter's by 0.8%.

Occupation

Table 8 contains regressions from the final period: the occupation period. Column (a) shows that a mother with an advanced degree directly decreases the probability her son works in STEM by 6.4% at the means, while increasing her daughter's probability of working in STEM by 6% at the means. This supports same-gender role model hypothesis, however it contradicts the hypothesis that a mother with an advanced degree directly increases any child's probability of working in STEM. A mother with an advanced degree also indirectly increases the probability that her daughter majors in STEM by 0.7% at the means. A mother with an advanced degree has a positive total effect on a daughter's probability of majoring in STEM and a negative total effect on a son's. This is consistent with same-gender role model hypothesis, however again it is not consistent with the hypothesis that a mother with an advanced degree positively impacts all of her children's participation in a STEM occupation.

A father with an advanced degree decreases the probability that his son works in a STEM occupation by 1.9%, while a mother decreases it by 4.6%. The father increases the probability his daughter works in a STEM occupation by 3.3%, while a mother increases it by 6%. A mother increases it by twice as much a father for women, which supports a same-gender role model hypothesis. This shows that a mother has a larger impact on her daughters than her sons, and daughters are more impacted by their mothers than their fathers.

7 Discussion

7.1 Limitations and Robustness Checks

I did a series of robustness checks to account for potential limitations. Tables corresponding to these robustness checks can be found in Appendix D.

Constrained Model

One of the obvious limitations to this study is the large amount of missing data. There is a large amount of data missing from each variable, which means each regression in tables 4 through 8 is analyzing different samples. In order to do a sensitivity check, constrained models with less missing data are reported in Appendix D. In general the coefficients in the constrained model appears to be relatively close in magnitude to the original model, with mother has advanced degree being slightly larger in the constrained model. This makes sense as the constrained model does not account for various family inputs, such as whether the parents have rules for GPA. Since the additional parental inputs, such as having rules for GPA, are theoretically positively correlated with mother has advanced degree and positively correlated with STEM predictors, it would make sense that the constrained model is upward biased and thus larger than the original model. Thus the missing data in the original models does not appear to have much of a bias. Additionally, the STEM predictors appear to have similar magnitudes between the original and constrained models.

Nevertheless, the amount of missing data even in the constrained model can create bias in the sample. One of the largest percentages of missing data comes from ACT score data, as all high school students were not required to take ACT or SAT tests at this time. This can specifically bias the coefficients upward, as students from lower income brackets are less likely to take these tests in this data[3].

Additional limitations stem out of the fact that some of the missing data appears to be

[3] 60% of students in the lowest income bracket (0-$25,000) had a missing ACT score, whereas only 17% of students were missing ACT score in the highest income bracket ($200,000 or more).

conditional on parents education[4]. This can result in inaccurately measured coefficients for mothers with advanced degree. Future studies, however, will likely be able to account for this problem since the ACT and SAT are now required in most public high schools.

Field of Education

The most accurate estimates for direct effect of a mother who has an advanced degree is in the high school achievement period (Table 5). One potential explanation of this is that many advanced degrees are within the field of education (Kena et al., 2016). Thus the variable ”mother with an advanced degree” could be unintentionally picking up effects that really have more to do with knowing the inner workings of education. This would create an upward bias in the high school achievement period, and could explain why mother with an advanced degree is statistically significant in the high school achievement period.

A robustness check was conducted in Appendix D.2 to account for this. The data did not include information on what specific advanced degree a mother had, therefore the variable ”mother works in education” was used as a proxy for whether the mother's degree was in education. This is not a perfect proxy, as many people with non-education degrees work in primary or secondary education, and many people with education degrees do not work in primary or secondary education; however, this is the best the data would allow.

The robustness check showed that working in the field of education did not affect statistical significance for males or females who have mothers with advanced degrees. As expected, Table D.2 and D.2 show that both men and women are less affected by a mother with an advanced degree when working in the field of education is taken into account for all categories in the high school achievement period, except number of AP/IB STEM courses taken. Thus Table D.2's omitted variable bias did bias the coefficients for mother has advanced degree slightly upward, however not by much.

[4]For instance, whether or not an individual had a missing ACT score is related to how far a parent expects a student will go in their education. A mother with an advanced degree is significantly less likely to expect her child to not graduate from high school. Therefore a mother with an advanced degree is less likely to have a child with a missing ACT score.

Accounting for Graduate School in STEM

There are still direct inputs that a mother can put in during occupational period (Table 8). However, these inputs are negative for men and positive for women. This could be due to the fact that many of these mothers' children seek advanced degrees themselves, meaning they would not be in a STEM occupation as of the third follow up survey. A robustness check was conducted in Appendix D.3, which included attending a graduate school for STEM as a STEM occupation. In this case, a mother with an advanced degree has less, but still a negative, direct impact on her son's participation in a STEM occupation, and has a positive impact on her daughter's participation.

7.2 Revisiting Assumptions

A mother with an advanced degree directly impacts a child's participation in a STEM occupation for both men and women in each new period. These results imply that a mother with an advanced degree does have an impact outside of the indirect effects. Under the current assumption that all inputs given in prior periods were controlled for, this direct effect is interpreted as a more highly educated mother puts in new inputs at each new period (whether they increase the likelihood of the STEM predictor or not). However, there are likely prior inputs that do not come into play until later. Therefore, it is likely that prior mother inputs are absorbed by the mother with an advanced degree variable. Due to the large array of prior mother inputs, it is unclear whether this would bias upward or downward the mother with an advanced degree coefficient.

Another assumption that was made was that the variable for a mother with an advanced degree was exogenous and thus, uncorrelated with the residual components of the control variables. Realistically, there are other control variables that the data did not have which are likely correlated with a mother's educational attainment. This would bias the coefficients. However, many individual, family, and school controls were used in order to reduce as much of the correlated residual term as possible.

7.3 Concluding Remarks

I examined the direct and indirect effects of a mother's attainment of an advanced degree on a STEM occupation for women and men. I used eleven ordinary least squares linear regressions and five logistic regressions from the Education Longitudinal Study of 2002. A mother with an advanced degree impacts almost all STEM occupation predictors. However, the most accurate coefficients are found in the high school achievement period, even after robustness checks.

Many of the limitations of this analysis attest to the importance of collecting data that include richer parent surveys. Ideally this data would interview parents throughout the entire time period and ask questions about inputs. If this data were available, the assumptions could be loosened, which would allow a slightly more realistic model. Additionally, this data would allow more accurate estimates of the parameters.

Surprisingly, a mother with an advanced degree has a negative direct impact and neutral indirect impact on the probability that her son works in a STEM occupation at the means. However, as expected, a mother with an advanced degree has both an indirect and direct positive impact on her daughter's work in a STEM occupation. Both of these conclusions are similar to Leppel et al. (2001) conclusion that a professional mother decreases the likelihood that her son majors in an engineering field, while she increases the likelihood that her daughters majors in one. Inconsistent with this literature, a father with an advanced degree has a negative impact on his son's participation in a STEM occupation, however has a positive impact on his daughter's participation in one. One policy implication that can arise out of this finding is that "STEM mentors" can be used, not only in early year outcomes, but even as far out as the occupation period.

This dataset did not contain information on preferences or views on various occupations, however as previous researchers have found, preferences and views on occupations appear to play an important role in determining occupation choice (Zafar, 2013). Thus future research should examine potential relationships between younger states of life and the formation of preferences.

This paper reaffirmed various STEM major predictors, including math self-efficacy, anticipated career, ACT score (both the overall score and math score), number of advanced STEM courses taken in high school, high school GPA, and whether the individual took Calculus. However, the research did not show support for direct impacts of math abilities, time spent on homework, motivation, or the time taken in between high school and postsecondary school. That being said, all STEM predictors impacted at least one future STEM outcome. Additionally, many of these STEM predictors, such as STEM GPA in postsecondary school and prediction of STEM occupation, not only impacted the probability of majoring in a STEM field but also the probability of working in one. This emphasizes the importance of early life outcomes, as they influence later life outcomes.

Unfortunately most interaction terms between mother with an advanced degree and female were not statistically significant. Nevertheless, many of the imprecise estimates of this interaction term were consistent with same-gender role models. For instance, mother with advanced degree inputs were larger than fathers in most cases in the high school preparedness period for women. Additionally, fathers with advanced degree inputs for sons were larger than mother inputs in this period. In the postsecondary school achievement period, mother with advanced degree inputs impact daughters more than sons for both STEM GPA and number of STEM courses taken. Most notably, mother with advanced degree inputs are twice as large as father with advanced degree inputs for daughters in occupation period. However, other outcomes, such as taking Calculus do not lend support for same-gender role model theories. Thus it appears that these same-gender effects are specific to certain outcomes. This analysis does not lend insight into what causes these larger mother inputs on daughters, and we cannot conclude that these inputs do or do not include merely seeing someone of the same gender. Future research should examine what causes these differences in inputs between females and males.

8 References

(2002). Education longitudinal study of 2002. *National Center for Education Statistics*.

(2011). Women as a percent of total employed in selected occupations, 2011. *Bureau of Labor Statistics, U.S. Department of Labor*.

(2016). About smart. *Science, Mathematics and Research for Transformation*.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, 84(2):191.

Beede, D. N., Julian, T. A., Langdon, D., McKittrick, G., Khan, B., and Doms, M. E. (2011). Women in stem: A gender gap to innovation.

Bieri Buschor, C., Berweger, S., Keck Frei, A., and Kappler, C. (2014). Majoring in stem—what accounts for women's career decision making? a mixed methods study. *The Journal of Educational Research*, 107(3):167–176.

Bottia, M. C., Stearns, E., Mickelson, R. A., Moller, S., and Valentino, L. (2015). Growing the roots of stem majors: Female math and science high school faculty and the participation of students in stem. *Economics of Education Review*, 45:14–27.

Ceci, S. J., Ginther, D. K., Kahn, S., and Williams, W. M. (2014). Women in academic science a changing landscape. *Psychological Science in the Public Interest*, 15(3):75–141.

Chen, X. (2009). Students who study science, technology, engineering, and mathematics (stem) in postsecondary education. stats in brief. nces 2009-161. *National Center for Education Statistics*.

Glass, J. L., Sassler, S., Levitte, Y., and Michelmore, K. M. (2013). What's so special about stem? a comparison of women's retention in stem and professional occupations. *Social forces; a scientific medium of social study and interpretation*, 92(2):723.

Griffith, A. L. (2010). Persistence of women and minorities in stem field majors: Is it the school that matters? *Economics of Education Review*, 29(6):911–922.

Griffith, A. L. (2014). Faculty gender in the college classroom: Does it matter for achievement and major choice? *Southern Economic Journal*, 81(1):211–231.

Hanushek, E. A. (1979). Conceptual and empirical issues in the estimation of educational production functions. *Journal of human Resources*, pages 351–388.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of economic literature*, 24(3):1141–1177.

Hoffmann, F. and Oreopoulos, P. (2009). A professor like me the influence of instructor gender on college achievement. *Journal of Human Resources*, 44(2):479–494.

Huesmann, L. R., Dubow, E. F., and Boxer, P. (2010). *The effect of religious participation on aggression over one's lifetime and across generations*. na.

Kena, G., Hussar, W., McFarland, J., de Brey, C., Musu-Gillette, L., Wang, X., Zhang, J., Rathbun, A., Wilkinson-Flicker, S., Diliberti, M., et al. (2016). The condition of education 2016. nces 2016-144. *National Center for Education Statistics*.

Leppel, K., Williams, M. L., and Waldauer, C. (2001). The impact of parental occupation and socioeconomic status on choice of college major. *Journal of Family and Economic issues*, 22(4):373–394.

Levin, H. M. (1970). A new model of school effectiveness.

McDill, E. L., Meyers Jr, E. D., and Rigsby, L. C. (1967). Institutional effects on the academic behavior of high school students. *Sociology of Education*, pages 181–199.

Moakler, M. W. and Kim, M. M. (2014). College major choice in stem: Revisiting confidence and demographic factors. *The Career Development Quarterly*, 62(2):128–142.

Rask, K. and Tiefenthaler, J. (2008). The role of grade sensitivity in explaining the gender imbalance in undergraduate economics. *Economics of Education Review*, 27(6):676–687.

Rothstein, D. S. (1995). Do female faculty influence female students' educational and labor market attainments? *ILR Review*, 48(3):515–530.

Veenstra, C. (2008). A research study of student success for engineering majors compared to other stem majors. *Ann Arbor Michigan*, 48109:1340.

Wang, X. (2013). Why students choose stem majors: Motivation, high school learning, and postsecondary context of support. *American Educational Research Journal*, 50(5):1081–1121.

Zafar, B. (2013). College major choice and the gender gap. *Journal of Human Resources*, 48(3):545–595.

9 Tables

Table 1: Summary statistics for primary variables of interest

Variable	Mean	SD	Percent of data missing
Female	0.501	0.500	0%
Mother has advanced degree	0.092	0.289	0%
Father has advanced degree	0.142	0.350	0%

Table 2: Sample characteristics for outcome variables

Variable	Mean					SD	Min	Max	Percent of data missing
	Total	Male	Female	Mother has advanced degree	Mother does not have advanced degree				
High School Preparednes									
Math self-efficacy (standardized)	0.028	0.182	-0.111	0.273	0.002	1.002	-1.831	1.772	35.5%
Math test (compared to all peers in U.S.)	50.71	51.241	50.183	56.267	50.147	9.912	19.38	86.68	1.9%
Hours per week spent on homework	9.797	9.152	10.427	11.452	9.631	7.395	1	45	15.4%
Motivation (standardized)	0.049	0.007	0.087	0.245	0.029	0.998	-1.994	1.579	34.3%
Prediction of STEM occupation at 30	0.08	0.12	0.042	0.093	0.079	0.271	0	1	19.6%
High School Achievement									
ACT score	21.527	21.805	21.285	24.407	21.131	5.06	10	36	41.1%
Math ACT score	21.328	21.987	20.757	24.063	20.952	5.287	11	36	41.4%
Number of advanced STEM courses	0.268	0.278	0.258	0.639	0.231	0.745	0	8	8.6%
GPA	2.574	2.403	2.744	2.985	2.533	0.835	0	4	8.7%
Took Calculus	0.135	0.136	0.134	0.295	0.119	0.342	0	1	8.6%
Between High School and Postsecondary School									
Time between high school and postsecondary school	8.512	8.894	8.18	5.671	8.861	16.77	0	119	25.5%
Attended postsecondary institution	0.868	0.834	0.899	0.964	0.858	0.338	0	1	18.2%
Postsecondary School Achievement									
Number of STEM courses	9.284	10.388	8.34	11.876	8.961	10.57	0	123	28.8%
STEM GPA	2.493	2.409	2.565	2.785	2.454	0.987	0	4	37.2%
Majored in STEM	0.165	0.247	0.104	0.209	0.157	0.371	0	1	57.1%
Occupational									
STEM occupation	0.061	0.095	0.03	0.11	0.055	0.239	0	1	21.3%

Table 3: Proportion of gender per outcome variable

Variable	Female	Male	Total
Predict STEM occupation at age 30	27.089%	72.911%	100%
Took Calculus	49.925%	50.075%	100%
Ever attended college	54.554%	45.446%	100%
STEM major	36.133%	63.867%	100%
STEM occupation	24.966%	75.034%	100%

Table 4: High school preparedness regression for all students

	(a)	(b)	(c)	(d)	(e)
	-0.39***	1.575***	-1.934***	0.014	0.008
Female	(0.032)	(0.208)	(0.212)	(0.032)	(0.008)
Mother has	0.032	0.199	1.176***	0.079	-0.008
advanced degree	(0.069)	(0.436)	(0.45)	(0.068)	(0.009)
Female X mother	0.05	-0.652	-0.338	-0.126	0.014
has advanced	(0.097)	(0.642)	(0.657)	(0.096)	(0.028)
Father has	0.053	1.315***	2.231***	-0.024	0.001
advanced degree	(0.057)	(0.378)	(0.39)	(0.057)	(0.012)
Female X father	0.012	-0.536	-0.726	0.014	-0.003
has advanced	(0.078)	(0.53)	(0.54)	(0.078)	(0.019)
Male direct effects	0.032	0.199	1.176	0.079	-0.008
Female direct effects	0.082	-0.453	0.838	-0.047	0.006
R-squared	0.0764	0.0844	0.361	0.0764	0.0889
Number of observations	4788	6021	6549	4788	5715

Standard errors are in parentheses.

***Significance at 1% level, **Significance at 5% level, *Significance at 1% level.

(a) Math self-efficacy (standardized), (b) hours spent on homework per week, (c) math test score (out of 100), (d) motivation score (standardized), (e) prediction of STEM occupation at age thirty

Table 5: High school achievement regression for all students

	(a)	(b)	(c)	(d)	(e)
Female	0.462***	0.001	0.035	0.32***	0.015
	(0.121)	(0.122)	(0.026)	(0.02)	(0.01)
Mother has advanced degree	0.617***	0.487**	0.234***	0.075*	0.049**
	(0.229)	(0.23)	(0.053)	(0.04)	(0.024)
Female X mother has	-0.179	-0.151	-0.125*	-0.003	-0.022
advanced degree	(0.32)	(0.323)	(0.075)	(0.056)	(0.02)
	0.879***	0.359*	0.061	0.087***	-0.016
Father has advanced degree	(0.195)	(0.197)	(0.044)	(0.034)	(0.013)
Female X father has	-0.135	0.024	-0.125*	0.013	0.014
advanced degree	(0.27)	(0.272)	(0.075)	(0.046)	(0.023)
High School Preparedness					
Math self efficacy	0.138**	0.69***	0.07***	0.061***	0.037***
	(0.06)	(0.06)	(0.013)	(0.01)	(0.005)
Math test score	0.382***	0.414***	0.03***	0.032***	0.014***
	(0.007)	(0.008)	(0.002)	(0.001)	(0.001)
Time spent on homework	0.03***	0.022***	0.007***	0.004***	0.002***
per week	(0.007)	(0.007)	(0.002)	(0.001)	(0.001)
Motivation	0.011	-0.087	0.009	0.081***	0.006
	(0.056)	(0.057)	(0.013)	(0.009)	(0.005)
Expected Career at 30 is	0.316*	0.545***	0.167***	0.009	0.043***
in STEM	(0.18)	(0.181)	(0.04)	(0.03)	(0.018)
Male direct effects	0.617	0.487	0.234	0.075	0.049
Male indirect effects	0.456	0.499	0.037	0.047	0.018
Male total effects	1.074	0.986	0.271	0.122	0.067
Female direct effects	0.439	0.336	0.109	0.072	0.027
Female indirect effects	0.317	0.398	0.027	0.026	0.014
Female total effects	0.756	0.734	0.136	0.098	0.041
R-squared	0.3208	0.6928	0.3208	0.4809	0.3748
Number of Observations	3995	3194	3995	3994	3987

Standard errors are in parentheses.

***Significance at 1% level, **Significance at 5% level, *Significance at 1% level.

(a) ACT score, (b) math ACT score, (c) number of AP/IB STEM courses taken (Carnegie Unit), (d) HS GPA, (e) took Calculus

Table 6: Between high school and postsecondary school regression

	(a)	(b)
Female	-0.067	0.003
	(0.48)	(0.002)
Mother has advanced degree	-0.276	-0.002
	(0.836)	(0.006)
Female X mother has advanced degree	-0.279	-0.002
	(1.173)	(0.01)
Father has advanced degree	0.187	-0.008
	(0.721)	(0.008)
Female X father has advanced degree	-0.553	-
	(0.996)	-
High School Preparedness		
Math self efficacy	0.032	-0.003**
	(0.231)	(0.001)
	0.043	0
Math test score	(0.04)	(0)
Time spent on homework per week	-0.002	0
	(0.027)	(0)
Motivation	-0.204	0.002**
	(0.214)	(0.001)
Expected Career at 30 is in STEM	-0.289	0.003
	(0.669)	(0.002)
High School Achievement		
ACT score	0.019	0
	(0.098)	(0)
Math ACT score	-0.163*	0.001**
	(0.098)	(0.001)
Number of advanced STEM courses	-0.11	0.003
	(0.28)	(0.002)
High school GPA	-1.615***	0.005**
	(0.426)	(0.002)
Took Calculus	0.593	-0.003
	(0.591)	(0.005)
Male direct effects	-0.276	-0.002
Male indirect effects	-0.146	0.002
Male total effects	-0.422	0.003
Female direct effects	-0.554	-0.003
Female indirect effects	-0.109	0.001
Female total effects	-0.663	0.007
R-squared	0.055	0.238
Number of Observations	2,677	2390

Standard errors are in parentheses.

***Significance at 1% level, **Significance at 5% level, *Significance at 1% level.

(a) number of months between high school and postsecondary school attendance, (b) attended postsecondary institution

Table 7: Postsecondary school achievement regression for all students who ever attended a postsecondary institution

	(a)	(b)	(c)
Female	-2.282***	0.043	-0.12***
	(0.499)	(0.038)	(0.021)
Mother has advanced	0.198	-0.033	-0.003
degree	(0.866)	(0.066)	(0.028)
Female X mother has	0.084	0.193**	-0.005
advanced degree	(1.219)	(0.092)	(0.041)
Father has advanced	0.543	0.083	-0.019
degree	(0.751)	(0.057)	(0.023)
Female X father has	-1.683	-0.165**	0.027
advanced degree	(1.033)	(0.079)	(0.041)
High School Preparedness			
Math self efficacy	0.75***	-0.039**	0.03***
	(0.241)	(0.018)	(0.009)
Math test score	-0.017	0.003	0.002
	(0.042)	(0.003)	(0.002)
Time spent on	-0.018	0	0
homework per week	(0.028)	(0.002)	(0.001)
Motivation	-0.38*	0.026	-0.007
	(0.221)	(0.017)	(0.008)
Expected Career at 30	5.012***	-0.066	0.14***
is in STEM	(0.692)	(0.052)	(0.034)
High School Achievement			
ACT score	-0.427***	-0.008	-0.014***
	(0.102)	(0.008)	(0.004)
Math ACT score	0.474***	0.029***	0.014***
	(0.101)	(0.008)	(0.004)
Number of advanced	1.383***	0.008	0.021**
STEM courses	(0.29)	(0.022)	(0.009)
High school GPA	1.91***	0.642***	0.082***
	(0.444)	(0.034)	(0.019)
Took Calculus	1.089*	0.04	0.04*
	(0.612)	(0.046)	(0.022)
Between High School and Postsecondary School			
Months between high school and	-0.076***	0.013***	0
postsecondary school	(0.021)	(0.002)	(0.002)
Male direct effects	0.198	-0.033	-0.003
Male indirect effects	0.408	0.063	0.012
Male total effects	0.605	0.03	0.009
Female direct effects	0.282	0.16	-0.008
Female indirect effects	0.401	0.045	0.012
Female total effects	0.683	0.205	0.004
R-squared	0.1763	0.3348	0.2134
Number of Observations	2575	2462	1927

Standard errors are in parentheses.

***Significance at 1% level, **Significance at 5% level, *Significance at 1% level.

(a) number of STEM courses taken, (b) STEM GPA, (c) majored in STEM

Table 8: Occupation regression for all students

	(a)
Female	-0.066***
	(0.016)
Mother has advanced degree	-0.046***
	(0.012)
Female X mother has advanced degree	0.106**
	(0.067)
Father has advanced degree	-0.019
	(0.015)
Female X father has advanced degree	0.052
	(0.041)
High School Preparedness	
Math self efficacy	0
	(0.007)
Math test score	-0.001
	(0.001)
Time spent on homework per week	0
	(0.001)
Motivation	-0.003
	(0.006)
Expected Career at 30 is in STEM	0.079***
	(0.026)
High School Achievement	
ACT score	-0.002
	(0.003)
Math ACT score	0.005*
	(0.003)
Number of advanced STEM courses	0.006
	(0.007)
High school GPA	0.009
	(0.014)
Took Calculus	-0.012
	(0.014)
Between High School and Postsecondary School	
Months between high school and postsecondary school	-0.003
	(0.002)
Postsecondary School	
Number of STEM courses taken	0.004***
	(0.001)
STEM GPA	0.026***
	(0.01)
Majored in STEM	0.106***
	(0.035)
Male direct effects	-0.064
Male indirect effects	0
Male total effects	-0.046
Female direct effects	0.06
Female indirect effects	0.007
Female total effects	0.067
R-squared	0.3652
Number of Observations	1817

Standard errors are in parentheses.

***Significance at 1% level, **Significance at 5% level, *Significance at 1% level.

(a) works in a STEM occupation

10 Appendices

A Appendix

In this paper, STEM is defined by the STEM SMART grant definition (sma, 2016). SMART grant definition includes the following majors:

- Aeronautical and Astronautical Engineering
- Biosciences
- Chemical Engineering
- Chemistry
- Civil Engineering
- Cognitive, Neural, and Behavioral Sciences
- Computer and Computational Sciences and Computer Engineering
- Electrical Engineering
- Geosciences
- Industrial and Systems Engineering (technical tracks only)
- Information Sciences
- Materials Science and Engineering
- Mathematics
- Mechanical Engineering
- Naval Architecture and Ocean Engineering
- Nuclear Engineering
- Oceanography
- Operations Research (technical tracks only)
- Physics

B Appendix

Individual Controls

Race and ethnicity is represented as a collection of binary variables that hold a value of one if the individual is that race or ethnicity, and zero if not. The races included are American Indian or Alaska Native, Asian, Black, Hispanic, Multiracial, Native Hawaiian or Pacific Islander, and White.

Number of schools attended represents how many schools the student has attended since first grade as of their tenth grade year. It is capped at five times. Held back one or more grades is a binary variable that holds a value of one if the student was ever held back a grade, and zero if not.

Table B.1 (at the end of the section) displays summary statistics for individual controls.

Family Controls

Income in thousands is split into six categories: 0-25, 25-50, 75-100, 100-200, and 200 or more. Each of these binary variables hold the value one if the individual falls in the income category, and zero otherwise.

Discuss report card is a binary variable that holds a value of one if the parent said they frequently or usually discussed report cards with their tenth grader. It holds a value of zero if not. Has rules about Homework, GPA, chores, or TV are binary variables that take a value of one if the family has rules about it, and zero if not. Father and mother is immigrant are binary variables that hold a value of one if the father or mother immigrated to the US, and zero otherwise.

Parent education expectation for child is split into five binary variables: not graduate high school or get GED, graduate high school or get GED, some college, graduate college, and advanced degree. These are binary variables that hold a value of one if the parent responded to the statement ”How far in school you expect your tenth grader will go” with the given education level. These variables hold a value of zero if the parent did not respond with the given education level

Number of dependents represents how many dependents (both under or over eighteen) the family supports. Growth mindset is a binary variable that represents whether a parent agreed/strongly agreed or disagreed/strongly disagreed/didn't know with/about the statement "Most people can learn to be good at math". Attend PTO meetings is a binary variable that holds a value of one if the individual went to a PTO meeting in the ninth grade school year, and zero if not. Number of meals per week with child represents the number of days per week the parent ate at least one meal with the student.

Table B.2 (at the end of the section) displays summary statistics for family controls.

School Controls

Urbanicity is split into three categories: urban, rural, and suburb. These are binary variables that take a value of one if the school is in this type of areas based on the Common Core of Data (CCD) 1999-2000 and the Private School Survey (PSS) 1999-2000. These variables take a value of zero if they are not in this type of area. Public high school is a binary variable that takes on a value of one if the school is a public school based on CCD 1999-2000 and PSS 1999-200. It takes a value of zero if the school is a private school.

School academic climate is a scale of the high school administrators views of the school's academic climate. Higher values mean a more academically-oriented climate. Percent free and reduced lunch represents the percent of the student body that uses free or reduced-price lunch. Years required to graduate for math, science, and computer coursework represents how many years in a subject a student is required to take to graduate high school.

Table B.3 (at the end of the section) displays summary statistics for high school controls.

Table B.1: Sample characteristics for individual controls

Variable	Mean	SD	Min	Max	Percent of data missing
Race/Ethnicity					
American Indian or Alaska Native	0.009	0.093	0	1	0%
Asian	0.098	0.297	0	1	0%
Black	0.134	0.341	0	1	0%
Hispanic	0.15	0.357	0	1	0%
MultiRacial	0.047	0.213	0	1	0%
Native Hawaiian or Pacific Islander*	0.01	0.07	0	1	0%
White	0.558	0.497	0	1	0%
Number of schools attended	1.249	1.472	0	5	24%
Held back one or more grades	0.121	0.327	0	1	23%

*These variables have very few people. These means and standard deviations have been rounded to protect the privacy of these individuals.

Table B.2: Sample characteristics for family controls

Variable	Mean	SD	Min	Max	Percent of data missing
Income (in thousands)					
0-25	0.21	0.407	0	1	0%
25-50	0.304	0.46	0	1	0%
50-75	0.205	0.404	0	1	0%
75-100	0.134	0.341	0	1	0%
100-200	0.112	0.315	0	1	0%
200 or more	0.036	0.186	0	1	0%
Discusses report card	0.963	0.188	0	1	24%
Has rules about					
Homework	0.925	0.263	0	1	24%
GPA	0.82	0.384	0	1	24%
Chores	0.877	0.329	0	1	24%
TV	0.653	0.476	0	1	24%
Father is immigrant	0.212	0.409	0	1	19%
Mother is immigrant	0.213	0.409	0	1	18%
Parent education expectations for child					
Not graduate high school/get GED or didn't know*	0.01	0.06	0	1	24%
Graduate high school or get GED	0.069	0.254	0.000	1	24%
Some college	0.157	0.363	0.000	1	24%
Graduate college	0.443	0.497	0.000	1	24%
Advanced Degree	0.328	0.470	0.000	1	24%
Number of dependents	2.674	1.377	0.000	8	23%
Growth mindset	0.833	0.373	0.000	1	23%
Attend PTO meetings	0.366	0.482	0.000	1	24%
Number of meals per week with child	5.371	1.860	0.000	7	18%

*These variables have very few people. These means and standard deviations have been rounded to protect the privacy of these individuals.

Table B.3: Sample characteristics for high school controls

Variable	Mean	SD	Min	Max	Percent of data missing
Urbanicity					
Urban	0.339	0.473	0	1	0%
Rural	0.182	0.386	0	1	0%
Suburb	0.479	0.5	0	1	0%
Public high school	0.788	0.49	0	1	0%
Region					
Northeast	0.183	0.387	0	1	0%
South	0.363	0.481	0	1	0%
West	0.205	0.404	0	1	0%
Midwest	0.249	0.432	0	1	0%
School academic climate	0.008	0.16	-0.63	0.266	18%
Percent free or reduced lunch	26.832	25.154	0	100	30%
Years required to graduate in					
Math	2.905	0.624	1	4	28%
Science	2.734	0.611	1	4	29%
Computer	0.591	0.586	0	3	31%

C Appendix

Table C.1 is the full regression for the high school preparedness period. This table corresponds to Table 4.

Table C.2 is the full regression for the high school achievement period. This table corresponds to Table 5.

Table C.3 is the full regression for the between high school and postsecondary school period. This table corresponds to Table 6.

Table C.4 is the full regression for the postsecondary school achievement period. This table corresponds to Table 7.

Table C.5 is the full regression for the occupation period. This table corresponds to Table 8.

Table C.1: High school preparedness regression for all students

	(a)	(b)	(c)	(d)	(e)
School controls					
Urbanicity (compared to suburbs)					
Urban	0.038	0.197	0.197	0.045	0.008
	(0.036)	(0.233)	(0.237)	(0.035)	(0.008)
Rural	-0.022	0.032	-0.118	-0.009	-0.008
	(0.039)	(0.256)	(0.261)	(0.039)	(0.009)
Public high school	0.056	-0.831***	0.36	-0.089**	0.026***
	(0.042)	(0.283)	(0.29)	(0.042)	(0.009)
Region (compared to Midwest)					
Northeast	-0.085**	-2.086***	-0.078	-0.003	0.008
	(0.043)	(0.29)	(0.297)	(0.043)	(0.011)
South	0.093**	-1.235***	0.063	0.087**	-0.005
	(0.042)	(0.275)	(0.28)	(0.042)	(0.009)
West	-0.054	0.835***	-0.392	-0.021	-0.009
	(0.045)	(0.307)	(0.309)	(0.045)	(0.01)
School academic climate	-0.28***	0.832	4.12***	-0.248**	-0.031
	(0.102)	(0.672)	(0.677)	(0.101)	(0.023)
Percent free and reduced lunch	-0.002***	-0.014***	-0.032***	0.001	0
	(0.001)	(0.005)	(0.005)	(0.001)	(0)
Years of coursework required to graduate					
Math	0.072*	-0.132	-0.326	-0.02	0.016*
	(0.037)	(0.244)	(0.25)	(0.037)	(0.008)
Science	-0.014	-0.141	0.178	0.041	-0.011
	(0.038)	(0.254)	(0.258)	(0.038)	(0.009)
Computer	0.017	0.164	0.042	-0.008	-0.004
	(0.026)	(0.168)	(0.172)	(0.026)	(0.006)
Family controls					
Income (compared to 50-75K)					
0-25K	-0.127**	-0.465	-2.811***	-0.144***	-0.023**
	(0.053)	(0.332)	(0.334)	(0.052)	(0.01)
25-50K	-0.072*	-0.391	-0.701***	-0.056	-0.005
	(0.039)	(0.262)	(0.267)	(0.039)	(0.009)
75-100K	-0.003	-0.663**	0.4	0.036	-0.011
	(0.046)	(0.308)	(0.315)	(0.045)	(0.01)
100-200K	0.008	-0.265	1.346***	0.09*	-0.01
	(0.049)	(0.331)	(0.34)	(0.048)	(0.01)
200K+	0.054	1.665***	1.705***	0.074	-0.036**
	(0.074)	(0.505)	(0.516)	(0.074)	(0.012)
Discuss report card	-0.192**	-0.293	-0.475	0.01	-0.043*
	(0.088)	(0.574)	(0.576)	(0.089)	(0.027)
Has rules for					
Homework	-0.086	-0.101	-0.47	0.012	0
	(0.062)	(0.409)	(0.413)	(0.061)	(0.014)
GPA	-0.068*	-0.982***	-0.58**	-0.002	-0.044***
	(0.041)	(0.269)	(0.274)	(0.041)	(0.012)
Chores	-0.06	-0.514*	-0.815***	-0.019	0.006
	(0.046)	(0.3)	(0.306)	(0.045)	(0.01)
TV	0.016	0.212	-0.367*	0.019	0.01
	(0.032)	(0.211)	(0.214)	(0.031)	(0.007)
Father is immigrant	0.112*	0.265	-0.885**	0.232***	-0.016
	(0.065)	(0.418)	(0.423)	(0.064)	(0.013)
Mother is immigrant	-0.076	0.12	-0.825**	-0.014	0.016
	(0.062)	(0.409)	(0.411)	(0.062)	(0.016)
Parent expects (compared to parent expects no high school graduation)					
High school graduation	-1.332*	1.986	-4.027	-1.56**	0.904***
	(0.682)	(4.175)	(4.423)	(0.685)	(0.029)
Some college	-1.175*	2.443	-2.066	-1.356**	0.932***
	(0.68)	(4.159)	(4.408)	(0.683)	(0.023)
Graduate college	-0.938	3.78	2.721	-1.092	0.721***
	(0.679)	(4.155)	(4.405)	(0.682)	(0.046)
Advanced degree	-0.61	4.684	5.966	-0.83	0.877***
	(0.679)	(4.156)	(4.406)	(0.682)	(0.035)
Number of dependents	0.016	-0.052	-0.14*	0	-0.006**
	(0.011)	(0.073)	(0.074)	(0.011)	(0.003)
Math growth mindset	0.059	0.068	0.271	0.023	-0.006
	(0.038)	(0.246)	(0.25)	(0.037)	(0.009)
Parent attend PTO meetings	-0.019	0.258	-0.983***	0.005	0
	(0.03)	(0.199)	(0.202)	(0.03)	(0.007)
Number of meals per week	0.01	-0.001	-0.015	0.023***	0.002
	(0.008)	(0.053)	(0.054)	(0.008)	(0.002)
Individual controls					
Race/Ethnicity (compared to white)					
American Indian or Alaska Native	0.03	-0.578	-3.092***	-0.453**	0.006
	(0.215)	(1.174)	(1.147)	(0.211)	(0.041)
Asian	0.055	0.234	2.203***	0.064	0.027
	(0.075)	(0.495)	(0.501)	(0.075)	(0.021)
Black	-0.048	-1.325***	-6.018***	0.039	0.006
	(0.059)	(0.363)	(0.365)	(0.059)	(0.014)
Hispanic	-0.083	-0.801**	-3.889***	-0.009	-0.013
	(0.057)	(0.358)	(0.363)	(0.057)	(0.012)
Multiracial	-0.056	-0.09	-1.756***	-0.005	-0.001
	(0.072)	(0.479)	(0.488)	(0.071)	(0.016)
Pacific Islander	-0.031	-2.164	-0.926	0.179	0.094
	(0.341)	(1.94)	(1.925)	(0.324)	(0.106)
Number of schools attended	0.003	-0.052	-0.135*	-0.014	0.002
	(0.011)	(0.069)	(0.07)	(0.01)	(0.002)
Held back	-0.081	-0.69*	-4.42***	-0.036	-0.028**
	(0.06)	(0.356)	(0.353)	(0.059)	(0.01)
Constant	1.268*	9.793**	56.525***	0.964	
	(0.695)	(4.267)	(4.513)	(0.698)	
R-squared	0.0764	0.0844	0.361	0.0764	0.0889
Number of observations	4788	6021	6549	4788	5715

Standard errors are in parentheses.

***Significance at 1% level, **Significance at 5% level, *Significance at 1% level.

(a) Math self-efficacy (standardized), (b) hours spent on homework per week, (c) math test score (out of 100), (d) motivation score (standardized), (e) prediction of STEM occupation at age thirty

Table C.2: High school achievement regression for all students

	(a)	(b)	(c)	(d)	(e)
School controls					
Urbanicity (compared to suburbs)					
Urban	0.129	0.161	0.053*	-0.019	-0.01
	(0.127)	(0.128)	(0.028)	(0.021)	(0.01)
Rural	0.112	0.018	-0.114***	0.033	-0.01
	(0.145)	(0.146)	(0.03)	(0.023)	(0.012)
Public high school	-0.272*	-0.107	0.201***	-0.043*	-0.026**
	(0.148)	(0.149)	(0.033)	(0.025)	(0.014)
Region (compared to Midwest)					
Northeast	-0.481***	-0.401***	0.092***	-0.194***	0.014
	(0.154)	(0.156)	(0.034)	(0.026)	(0.014)
South	-0.132	-0.261*	0.141***	-0.078***	0.011
	(0.151)	(0.152)	(0.033)	(0.025)	(0.013)
West	-0.103	-0.043	-0.03	0.042	-0.014
	(0.169)	(0.17)	(0.036)	(0.027)	(0.013)
School academic climate	0.428	0.519	0.199**	-0.055	-0.054*
	(0.375)	(0.377)	(0.081)	(0.061)	(0.031)
Percent free and reduced lunch	-0.012***	-0.014***	0	0.001***	0*
	(0.003)	(0.003)	(0.001)	(0)	(0)
Years of coursework required to					
Math	0.129	0.27**	-0.02	-0.001	0.017
	(0.131)	(0.132)	(0.029)	(0.022)	(0.011)
Science	-0.069	-0.273**	-0.024	0.009	0.001
	(0.135)	(0.136)	(0.03)	(0.023)	(0.012)
Computer	0.035	-0.008	-0.008	0.024	0.009
	(0.095)	(0.095)	(0.021)	(0.016)	(0.008)
Family controls					
Income (compared to 50-75K)					
0-25K	-0.65***	-0.572***	-0.073*	-0.104***	-0.033*
	(0.215)	(0.217)	(0.042)	(0.032)	(0.015)
25-50K	-0.233	-0.305**	-0.042	-0.066***	-0.019
	(0.143)	(0.144)	(0.031)	(0.023)	(0.012)
75-100K	0.11	0.195	0.019	-0.043	0.025*
	(0.159)	(0.16)	(0.036)	(0.027)	(0.015)
100-200K	0.266	0.281*	0.005	-0.022	0.017
	(0.166)	(0.167)	(0.038)	(0.029)	(0.015)
200K+	0.611**	0.773***	0.21***	-0.06	0.061***
	(0.249)	(0.25)	(0.059)	(0.044)	(0.029)
Discuss report card	0.093	-0.148	0.016	-0.069	-0.063*
	(0.329)	(0.331)	(0.071)	(0.053)	(0.039)
Has rules for					
Homework	-0.047	-0.18	0.047	-0.019	0.011
	(0.225)	(0.227)	(0.05)	(0.037)	(0.017)
GPA	-0.62***	-0.397***	-0.069**	-0.078***	-0.034***
	(0.149)	(0.15)	(0.032)	(0.025)	(0.015)
Chores	-0.154	0.008	-0.156***	-0.016	-0.012
	(0.162)	(0.163)	(0.036)	(0.027)	(0.014)
TV	0.141	0.103	-0.017	-0.037*	0.003
	(0.115)	(0.116)	(0.025)	(0.019)	(0.01)
Father is immigrant	-0.361	-0.182	0.023	-0.115***	-0.024
	(0.239)	(0.242)	(0.051)	(0.038)	(0.017)
Mother is immigrant	0.359	0.461**	0.021	0.049	0.036*
	(0.233)	(0.235)	(0.049)	(0.037)	(0.023)
Parent expects (compared to parent expects no high school graduation)					
High school graduation	-5.476*	-5.515*	0.851	-1.225**	0.942***
	(2.914)	(2.934)	(0.702)	(0.529)	(0.005)
Some college	-6.325**	-5.976**	0.747	-1.154**	0.978***
	(2.879)	(2.9)	(0.7)	(0.528)	(0.005)
Graduate college	-5.737**	-5.805**	0.704	-0.954*	0.915***
	(2.873)	(2.894)	(0.699)	(0.527)	(0.039)
Advanced degree	-4.908*	-5.227*	0.919	-0.805	0.986***
	(2.872)	(2.892)	(0.699)	(0.527)	(0.011)
Number of dependents	-0.105**	-0.063	0.003	0.001	0.003
	(0.043)	(0.044)	(0.009)	(0.007)	(0.004)
Math growth mindset	-0.045	-0.012	-0.003	0.046**	0.027**
	(0.135)	(0.137)	(0.03)	(0.022)	(0.01)
Parent attend PTO meetings	-0.065	-0.074	0.031	0.002	0.001
	(0.109)	(0.109)	(0.024)	(0.018)	(0.009)
Number of meals per week	0.021	-0.005	0.023***	0.011**	0.006**
	(0.03)	(0.03)	(0.006)	(0.005)	(0.003)
Constant	6.976**	5.376*	-2.308***	2.064***	
	(2.964)	(2.985)	(0.716)	(0.54)	
R-squared	0.3208	0.6928	0.3208	0.4809	0.3748
Number of observations	3995	3194	3995	3994	3987

Standard errors are in parentheses.

***Significance at 1% level, **Significance at 5% level, *Significance at 1% level.

(a) ACT score, (b) math ACT score, (c) number of AP/IB STEM courses taken (Carnegie Unit), (d) HS GPA, (e) took Calculus

Table C.3: Between high school and postsecondary school regression for all students

	(a)	(b)
School controls		
Urbanicity (compared to suburbs)		
Urban	-0.199	0.002
	(0.475)	(0.002)
Rural	-0.137	-0.006**
	(0.556)	(0.004)
Public high school	1.394**	-0.006*
	(0.557)	(0.003)
Region (compared to Midwest)		
Northeast	-0.573	-0.002
	(0.588)	(0.004)
South	0.081	-0.005
	(0.572)	(0.004)
West	-0.281	-0.003
	(0.628)	(0.004)
School academic climate	-2.543*	0.003
	(1.411)	(0.007)
Percent free and reduced lunch	-0.013	0
	(0.012)	(0)
Years of coursework required to graduate		
Math	0.387	-0.003
	(0.494)	(0.003)
Science	-0.124	0.004*
	(0.511)	(0.003)
Computer	-0.495	0.001
	(0.357)	(0.002)
Family controls		
Income (compared to 50-75K)		
0-25K	1.714**	-0.004
	(0.826)	(0.005)
25-50K	0.619	-0.004
	(0.543)	(0.003)
75-100K	-1.132*	0.002
	(0.593)	(0.003)
100-200K	-0.632	0.007*
	(0.613)	(0.002)
200K+	-0.552	-
	(0.937)	-
Discuss report card	-1.073	-0.001
	(1.23)	(0.005)
Has rules for		
Homework	0.233	0.001
	(0.846)	(0.005)
GPA	-0.613	0.003
	(0.557)	(0.004)
Chores	0.391	-0.002
	(0.606)	(0.002)
TV	0.209	-0.001
	(0.435)	(0.002)
Father is immigrant	0.282	0.003
	(0.897)	(0.003)
Mother is immigrant	-1.947**	-0.006
	(0.883)	(0.008)
Parent expects (compared to parent expects no high school graduation)		
High school graduation	-	-
	-	-
Some college	-0.313	0.004
	(2.229)	(0.002)
Graduate college	-2.912	0.015***
	(2.132)	(0.009)
Advanced degree	-2.972	0.014***
	(2.151)	(0.007)
Number of dependents	-0.08	-0.001
	(0.164)	(0.001)
Math growth mindset	0.545	0
	(0.512)	(0.002)
Parent attend PTO meetings	0.448	0.004**
	(0.41)	(0.002)
Number of meals per week	-0.005	0
	(0.116)	(0.001)
Individual controls		
Race/Ethnicity (compared to white)		
American Indian or Alaska Native	-3.298	-
	(3.73)	-
Asian	0.66	0.005
	(1.033)	(0.002)
Black	-0.566	0.004
	(0.836)	(0.002)
Hispanic	1.811**	0.003
	(0.869)	(0.002)
Multiracial	0.57	0.002
	(1.045)	(0.003)
Pacific Islander	0.532	-
	(4.995)	-
Number of schools attended	0.157	0.001
	(0.15)	(0.001)
Held back	-1.269	0
	(1.081)	(0.004)
Constant	12.923***	
	(3.35)	
R-squared	0.055	0.238
Number of observations	2677	2390

Standard errors are in parentheses.

***Significance at 1% level, **Significance at 5% level, *Significance at 1% level.

(a) number of months between high school and postsecondary school attendance, (b) attended postsecondary institution

Table C.4: Postsecondary school achievement regression for all students who attended college

	(a)	(b)	(c)
School controls			
Urbanicity (compared to suburbs)			
Urban	0.037	0.005	0.025
	(0.494)	(0.038)	(0.02)
Rural	1.216**	-0.098**	0.038
	(0.577)	(0.044)	(0.025)
Public high school	0.919	0.033	0.027
	(0.58)	(0.044)	(0.021)
Region (compared to Midwest)			
Northeast	0.051	0.082*	0.041
	(0.614)	(0.047)	(0.027)
South	1.47**	-0.018	0.048**
	(0.591)	(0.045)	(0.025)
West	0.264	-0.062	-0.024
	(0.654)	(0.051)	(0.023)
School academic climate	0.082	-0.099	-0.052
	(1.465)	(0.113)	(0.056)
Percent free and reduced lunch	-0.015	-0.002**	0
	(0.012)	(0.001)	(0)
Years of coursework required to graduate			
Math	0.917*	-0.031	0.029
	(0.511)	(0.039)	(0.019)
Science	-0.711	0.03	-0.034*
	(0.529)	(0.041)	(0.02)
Computer	-0.157	-0.041	-0.016
	(0.371)	(0.029)	(0.015)
Family controls			
Income (compared to 50-75K)			
0-25K	-0.087	0.071	0.015
	(0.856)	(0.066)	(0.039)
25-50K	0.175	0.039	0.009
	(0.564)	(0.043)	(0.022)
75-100K	-0.69	-0.022	-0.016
	(0.617)	(0.047)	(0.022)
100-200K	-0.684	-0.008	-0.022
	(0.636)	(0.048)	(0.021)
200K+	-0.299	0.05	0.012
	(0.975)	(0.075)	(0.036)
Discuss report card	-0.1	0.008	-0.019
	(1.274)	(0.098)	(0.054)
Has rules for			
Homework	-0.082	-0.016	-0.034
	(0.872)	(0.067)	(0.036)
GPA	0.609	-0.133***	0.003
	(0.575)	(0.044)	(0.021)
Chores	1.563**	0.027	0.061***
	(0.629)	(0.048)	(0.017)
TV	-0.556	-0.079**	-0.037**
	(0.451)	(0.035)	(0.018)
Father is immigrant	0.817	0.005	0.042
	(0.931)	(0.072)	(0.042)
Mother is immigrant	0.855	-0.04	-0.018
	(0.92)	(0.071)	(0.034)
Parent expects (compared to parent expects no high school graduation)			
High school graduation	0.851	-0.322*	-
	(2.28)	(0.183)	-
Some college	-	-	-0.022
	-	-	(0.117)
Graduate college	2.248***	-0.027	0.018
	(0.846)	(0.067)	(0.129)
Advanced degree	2.867***	-0.021	0.044
	(0.902)	(0.071)	(0.132)
Number of dependents	0.167	-0.003	0
	(0.171)	(0.013)	(0.007)
Math growth mindset	0.314	-0.023	0
	(0.532)	(0.041)	(0.021)
Parent attend PTO meetings	-0.551	0.008	-0.013
	(0.426)	(0.033)	(0.016)
Number of meals per week	-0.095	-0.012	0.007
	(0.121)	(0.009)	(0.005)
Individual controls			
Race/Ethnicity (compared to white)			
American Indian or Alaska Native	-0.581	-0.616**	-
	(3.797)	(0.284)	-
Asian	0.24	0.004	-0.006
	(1.067)	(0.082)	(0.037)
Black	1.665*	-0.149**	0.063
	(0.871)	(0.067)	(0.047)
Hispanic	0.313	-0.036	0.001
	(0.907)	(0.07)	(0.038)
Multiracial	-0.818	-0.11	0.011
	(1.101)	(0.085)	(0.048)
Pacific Islander	-2.088	0.791*	0.006
	(5.88)	(0.441)	(0.006)
Number of schools attended	-0.196	-0.008	-0.02
	(0.156)	(0.012)	(0.044)
Held back	-0.111	0.04	-0.12***
	(1.133)	(0.089)	(0.021)
Constant	0.069	0.277	
	(2.833)	(0.223)	
R-squared	0.1763	0.3348	0.2134
Number of observations	2575	2462	1927

Standard errors are in parentheses.

***Significance at 1% level, **Significance at 5% level, *Significance at 1% level.

(a) number of STEM courses taken, (b) STEM GPA, (c) majored in STEM

Table C.5: Occupation regression for all students

	(a)
School controls	
Urbanicity (compared to suburbs)	
Urban	-0.014
	(0.013)
Rural	-0.014
	(0.014)
Public high school	0.024
	(0.014)
Region (compared to Midwest)	
Northeast	0.015
	(0.019)
South	0.007
	(0.017)
West	0.013
	(0.02)
School academic climate	-0.03
	(0.04)
Percent free and reduced lunch	0
	(0)
Years of coursework required to graduate	
Math	0.008
	(0.014)
Science	-0.025*
	(0.014)
Computer	0.005
	(0.01)
Family controls	
Income (compared to 50-75K)	
0-25K	-0.016
	(0.021)
25-50K	-0.011
	(0.014)
75-100K	-0.003
	(0.016)
100-200K	-0.006
	(0.016)
200K+	0.029
	(0.031)
Discuss report card	0.011
	(0.033)
Has rules for	
Homework	-0.048*
	(0.033)
GPA	0.022
	(0.013)
Chores	0.011
	(0.015)
TV	-0.016
	(0.013)
Father is immigrant	-0.023
	(0.019)
Mother is immigrant	0.016
	(0.03)
Parent expects (compared to parent expects no high school graduation)	
High school graduation	0.375**
	(0.227)
Some college	-
	-
Graduate college	0.015
	(0.029)
Advanced degree	0.009
	(0.031)
Number of dependents	-0.003
	(0.005)
Math growth mindset	-0.02
	(0.017)
Parent attend PTO meetings	0.008
	(0.012)
Number of meals per week	0.007**
	(0.003)
Individual controls	
Race/Ethnicity (compared to white)	
American Indian or Alaska Native	-
	-
Asian	0.023
	(0.035)
Black	0.024
	(0.033)
Hispanic	0.023
	(0.033)
Multiracial	0.046
	(0.046)
Pacific Islander	-
	-
Number of schools attended	-0.004
	(0.004)
Held back	-0.027
	(0.023)
R-squared	0.3652
Number of observations	1817

Standard errors are in parentheses.

***Significance at 1% level, **Significance at 5% level, *Significance at 1% level.

(a) works in STEM occupation

D Appendix

D.1 Constrained model

Tables D.1.1 through Tables D.1.2 are constrained models of Tables 4 through 8.

Table D.1.1: Constrainted high school preparedness regression for all students

	(a)	(b)	(c)	(d)	(e)
School controls					
Urbanicity (compared to suburbs)					
Urban	0.029	0.075	-0.465***	0.04*	0.009
	(0.023)	(0.146)	(0.162)	(0.023)	(0.005)
Rural	-0.024	0	-0.154	0.005	-0.004
	(0.027)	(0.173)	(0.196)	(0.027)	(0.006)
Public high school	-0.063**	-1.587***	-2.336***	-0.14***	0.009
	(0.026)	(0.166)	(0.188)	(0.026)	(0.006)
Region (compared to Midwest)					
Northeast	0.004	-1.532***	0.732***	0.03	0.01
	(0.029)	(0.189)	(0.214)	(0.029)	(0.007)
South	0.091***	-1.302***	0.477***	0.096***	0.007
	(0.025)	(0.161)	(0.182)	(0.025)	(0.006)
West	-0.037	0.603***	0.255	0.018	0.001
	(0.03)	(0.195)	(0.218)	(0.03)	(0.007)
Family controls					
Income (compared to 50-75K)					
0-25K	-0.195***	-0.817***	-4.13***	-0.145***	-0.026***
	(0.032)	(0.199)	(0.222)	(0.031)	(0.006)
25-50K	-0.08***	-0.49***	-1.62***	-0.062**	-0.005
	(0.027)	(0.177)	(0.199)	(0.027)	(0.006)
75-100K	0.04	-0.146	0.959***	0.057*	-0.002
	(0.033)	(0.215)	(0.243)	(0.033)	(0.007)
100-200K	0.066*	0.341	2.256***	0.109***	-0.009
	(0.035)	(0.232)	(0.264)	(0.035)	(0.008)
200K+	0.105*	1.537***	2.561***	0.095*	-0.029**
	(0.054)	(0.362)	(0.408)	(0.054)	(0.01)
Individual controls					
Race/Ethnicity (compared to					
American Indian or	0.052	-0.975	-4.833***	-0.018	0.017
Alaska Native	(0.111)	(0.686)	(0.756)	(0.112)	(0.028)
Asian	0.235***	1.167***	1.967***	0.299***	0.03***
	(0.036)	(0.233)	(0.257)	(0.036)	(0.01)
Black	0.06*	-1.165***	-6.406***	0.112***	0.006
	(0.034)	(0.203)	(0.224)	(0.034)	(0.008)
Hispanic	0.018	-0.893***	-5.142***	0.14***	-0.006
	(0.032)	(0.198)	(0.219)	(0.032)	(0.007)
Multiracial	-0.028	-0.283	-1.973***	-0.005	-0.003
	(0.045)	(0.294)	(0.331)	(0.045)	(0.01)
Pacific Islander	0.189	-0.801	-3.24***	0.134	0.008
	(0.156)	(0.984)	(1.102)	(0.157)	(0.036)
Variables of interest					
Female	-0.278***	1.398***	-0.803***	0.099***	-0.076***
	(0.021)	(0.135)	(0.151)	(0.021)	(0.005)
Mother has advanced degree	0.144***	0.711**	2.151***	0.126**	0.021*
	(0.051)	(0.317)	(0.356)	(0.05)	(0.012)
Father has advanced degree	0.167***	1.47***	3.144***	0.096**	-0.002
	(0.043)	(0.268)	(0.304)	(0.042)	(0.009)
FEMALE X Mother has	-0.036	-0.2	-0.243	-0.059	-0.022
advanced degree	(0.069)	(0.45)	(0.505)	(0.069)	(0.013)
FEMALE X Father has	-0.054	-0.37	-0.703*	-0.041	0.007
advanced degree	(0.057)	(0.37)	(0.418)	(0.057)	(0.015)
Constant	0.172***	11.128***	54.651***	-0.009	
	(0.036)	(0.232)	(0.264)	(0.036)	
R-squared	0.0475	0.0591	0.234	0.0273	0.045
Number of observations	10,443.00	13,699.00	15,892.00	10,647.00	13,022.00

Standard errors are in parentheses.

***Significance at 1% level, **Significance at 5% level, *Significance at 1% level.

(a) Math self-efficacy (standardized), (b) hours spent on homework per week, (c) math test score (out of 100), (d) motivation score (standardized), (e) prediction of STEM occupation at age thirty

Table D.1.2: Constrained high school achievement regression for all students

	(a)	(b)	(c)	(d)	(e)
School controls					
Urbanicity (compared to suburbs)					
Urban	-0.009	-0.009	0.016	-0.017	-0.006
	(0.091)	(0.092)	(0.018)	(0.016)	(0.006)
Rural	0.049	-0.198*	-0.064***	0.044**	-0.004
	(0.112)	(0.113)	(0.021)	(0.019)	(0.007)
Public high school	-0.703***	-0.436***	0.126***	-0.096***	-0.019***
	(0.094)	(0.096)	(0.02)	(0.018)	(0.007)
Region (compared to Midwest)					
Northeast	-0.227**	-0.219*	0.065***	-0.182***	0.014*
	(0.112)	(0.114)	(0.022)	(0.02)	(0.008)
South	0.006	-0.244**	0.109***	-0.069***	0.015**
	(0.099)	(0.1)	(0.02)	(0.018)	(0.007)
West	0.05	0.072	0.024	0.001	-0.01
	(0.125)	(0.126)	(0.023)	(0.021)	(0.007)
Family controls					
Income (compared to 50-75K)					
0-25K	-0.598***	-0.431***	-0.022	-0.126***	-0.022**
	(0.142)	(0.144)	(0.025)	(0.022)	(0.008)
25-50K	-0.259**	-0.289***	-0.026	-0.061***	-0.009
	(0.109)	(0.111)	(0.021)	(0.019)	(0.007)
75-100K	0.249**	0.293**	0.007	-0.023	0.011
	(0.123)	(0.124)	(0.025)	(0.023)	(0.009)
100-200K	0.457***	0.422***	0.047*	-0.002	0.007
	(0.127)	(0.129)	(0.027)	(0.024)	(0.009)
200K+	0.763***	0.925***	0.165***	-0.036	0.01
	(0.193)	(0.195)	(0.042)	(0.038)	(0.013)
Individual controls					
Race/Ethnicity (compared to white)					
American Indian or Alaska Native	-1.1*	-1.259*	0.052	-0.417***	0.03
	(0.626)	(0.648)	(0.09)	(0.081)	(0.055)
Asian	-0.027	1.081***	0.544***	0.062**	0.108***
	(0.141)	(0.143)	(0.028)	(0.025)	(0.016)
Black	-0.586***	-0.267*	0.042	-0.187***	0.019
	(0.147)	(0.149)	(0.027)	(0.025)	(0.013)
Hispanic	-0.579***	-0.371**	0.089***	-0.157***	0.009
	(0.147)	(0.148)	(0.025)	(0.023)	(0.011)
Multiracial	-0.052	0.003	0.066*	-0.098***	0.008
	(0.188)	(0.191)	(0.035)	(0.032)	(0.013)
Pacific Islander	-0.339	0.687	0.009	-0.116	-0.025
	(0.677)	(0.684)	(0.116)	(0.105)	(0.036)
Variables of interest					
Female	0.463***	-0.04	0.046***	0.396***	0.032***
	(0.09)	(0.091)	(0.017)	(0.015)	(0.006)
Mother has advanced degree	0.543***	0.446**	0.193***	0.059*	0.028**
	(0.181)	(0.183)	(0.038)	(0.034)	(0.014)
Father has advanced degree	1.09***	0.653***	0.119***	0.139***	0.014
	(0.154)	(0.156)	(0.032)	(0.029)	(0.01)
FEMALE X Mother has	0.216	0.089	-0.074	0.006	-0.011
advanced degre	(0.247)	(0.25)	(0.052)	(0.047)	(0.013)
FEMALE X Father has advanced degree	-0.355*	-0.169	-0.013	-0.034	0.002
	(0.205)	(0.208)	(0.043)	(0.039)	(0.013)
High school preparedness					
Math self-efficacy	0.129***	0.652***	0.074***	0.056***	0.029***
	(0.045)	(0.046)	(0.009)	(0.008)	(0.003)
Math test score	0.4***	0.418***	0.026***	0.039***	0.011***
	(0.005)	(0.005)	(0.001)	(0.001)	(0)
Hours spent on homework per	0.026***	0.023***	0.008***	0.005***	0.001***
week	(0.005)	(0.005)	(0.001)	(0.001)	(0)
Motivation	0.068	-0.087**	0.024***	0.114***	0.008***
	(0.042)	(0.043)	(0.008)	(0.007)	(0.003)
Prediction of STEM occupation at age thirty	0.408***	0.525***	0.168***	0.039	0.045***
	(0.136)	(0.137)	(0.027)	(0.025)	(0.011)
Constant	-0.136	-1.402***	-1.438***	0.536***	
	(0.318)	(0.321)	(0.057)	(0.052)	
R-squared	0.6452	0.6779	0.2636	0.4425	0.3532
Number of observations	5885	5864	8327	8321	8327

Standard errors are in parentheses.

***Significance at 1% level, **Significance at 5% level, *Significance at 1% level.

(a) ACT score, (b) math ACT score, (c) number of AP/IB STEM courses taken (Carnegie Unit), (d) HS GPA, (e) took Calculus

Table D.1.3: Constrained between high school and postsecondary school regression for all students

	(a)	(b)
School controls		
Urbanicity (compared to suburbs)		
Urban	-0.134	0.001
	(0.356)	(0.002)
Rural	0.07	-0.005*
	(0.443)	(0.003)
Public high school	1.697***	-0.009***
	(0.368)	(0.002)
Region (compared to Midwest)		
Northeast	-0.908**	0.003
	(0.444)	(0.002)
South	0.123	0.001
	(0.388)	(0.002)
West	-0.453	0.004
	(0.483)	(0.002)
Family controls		
Income (compared to 50-75K)		
0-25K	0.517	-0.009**
	(0.573)	(0.005)
25-50K	0.586	-0.006**
	(0.428)	(0.003)
75-100K	-1.084**	0.003
	(0.472)	(0.003)
100-200K	-0.679	0.006*
	(0.487)	(0.002)
200K+	-0.718	-
	(0.749)	-
Individual controls		
Race/Ethnicity (compared to		
American Indian or	-1.763	-
Alaska Native	(2.637)	-
Asian	-0.77	0.007**
	(0.568)	(0.002)
Black	-0.275	0.007***
	(0.588)	(0.002)
Hispanic	0.463	0.002
	(0.584)	(0.003)
Multiracial	-0.368	-0.002
	(0.766)	(0.005)
Pacific Islander	4.811*	-
	(2.655)	-
Variables of interest		
Female	-0.455	0.004*
	(0.372)	(0.002)
Mother has advanced degree	0.241	0.001
	(0.685)	(0.005)
Father has advanced degree	-0.243	-0.002
	(0.593)	(0.005)
FEMALE X Mother has	-0.51	0.004
advanced degre	(0.936)	(0.006)
FEMALE X Father has	0.075	0.005
advanced degree	(0.785)	(0.003)
High school preparedness		
Math self-efficacy	0.176	-0.002**
	(0.18)	(0.001)
Math test score	0.015	0
	(0.031)	(0)
Hours spent on homework per	0.026	0
	(0.021)	(0)
Motivation	-0.519***	0.003**
	(0.166)	(0.001)
Prediction of STEM occupation	-0.434	0.003
	(0.527)	(0.003)
High school achievement		
ACT score	-0.043	0
	(0.077)	(0)
Math ACT score	-0.078	0.001
	(0.076)	(0.001)
Advanced STEM courses	-0.199	0.005*
	(0.217)	(0.003)
GPA	-2.126***	0.008***
	(0.32)	(0.002)
Took Calculus	0.452	-0.007
	(0.475)	(0.006)
Constant	12.622***	0***
	(1.4)	(0)
R-squared	0.0485	0.1971
Number of observations	4675	4627

Standard errors are in parentheses.

***Significance at 1% level, **Significance at 5% level, *Significance at 1% level.

(a) number of months between high school and postsecondary school attendance, (b) attended postsecondary institution

Table D.1.5: Constrained occupation regression for all students

	(a)
School controls	
Urbanicity (compared to suburbs)	
Urban	-0.016*
	(0.009)
Rural	-0.026**
	(0.009)
Public high school	0.013
	(0.009)
Region (compared to Midwest)	
Northeast	0.017
	(0.013)
South	0
	(0.01)
West	0.009
	(0.014)
Family controls	
Income (compared to 50-75K)	
0-25K	-0.039***
	(0.009)
25-50K	-0.018*
	(0.01)
75-100K	-0.018
	(0.01)
100-200K	-0.02*
	(0.01)
200K+	0.008
	(0.018)
Individual controls	
Race/Ethnicity (compared to white)	
American Indian or Alaska Native	-
	-
Asian	0.001
	(0.013)
Black	0.012
	(0.015)
Hispanic	-0.002
	(0.017)
Multiracial	0.002
	(0.022)
Pacific Islander	-0.02
	(0.049)
Variables of interest	
Female	-0.059***
	(0.012)
Mother has advanced degree	-0.009
	(0.013)
Father has advanced degree	-0.017
	(0.011)
FEMALE X Mother has advanced degre	0.044
	(0.033)
FEMALE X Father has advanced degree	0.012
	(0.022)
High school preparedness	
Math self-efficacy	0.005
	(0.005)
Math test score	0
	(0.001)
Hours spent on homework per week	0.001
	(0.001)
Motivation	-0.005
	(0.004)
Prediction of STEM occupation at age	0.062***
	(0.018)
High school achievement	
ACT score	0.001
	(0.002)
Math ACT score	0.002
	(0.002)
Advanced STEM courses	-0.002
	(0.005)
GPA	0.017*
	(0.01)
Took Calculus	-0.002
	(0.011)
Between high school and postsecondary school	
Months between high school and postsecondary school	-0.002
	(0.001)
Postsecondary school achievement	
Number of STEM courses taken	0.003***
	(0)
STEM GPA	0.019***
	(0.007)
Majored in STEM	0.134***
	(0.028)
R-squared	0.3268
Number of observations	3096

Standard errors are in parentheses.

***Significance at 1% level, **Significance at 5% level, *Significance at 1% level.

(a) works in STEM occupation

Table D.1.5: Constrained occupation regression for all students

	(a)
School controls	
Urbanicity (compared to suburbs)	
Urban	-0.016*
	(0.009)
Rural	-0.026**
	(0.009)
Public high school	0.013
	(0.009)
Region (compared to Midwest)	
Northeast	0.017
	(0.013)
South	0
	(0.01)
West	0.009
	(0.014)
Family controls	
Income (compared to 50-75K)	
0-25K	-0.039***
	(0.009)
25-50K	-0.018*
	(0.01)
75-100K	-0.018
	(0.01)
100-200K	-0.02*
	(0.01)
200K+	0.008
	(0.018)
Individual controls	
Race/Ethnicity (compared to white)	
American Indian or Alaska Native	-
	-
Asian	0.001
	(0.013)
Black	0.012
	(0.015)
Hispanic	-0.002
	(0.017)
Multiracial	0.002
	(0.022)
Pacific Islander	-0.02
	(0.049)
Variables of interest	
Female	-0.059***
	(0.012)
Mother has advanced degree	-0.009
	(0.013)
Father has advanced degree	-0.017
	(0.011)
FEMALE X Mother has advanced degre	0.044
	(0.033)
FEMALE X Father has advanced degree	0.012
	(0.022)
High school preparedness	
Math self-efficacy	0.005
	(0.005)
Math test score	0
	(0.001)
Hours spent on homework per week	0.001
	(0.001)
Motivation	-0.005
	(0.004)
Prediction of STEM occupation at age	0.062***
	(0.018)
High school achievement	
ACT score	0.001
	(0.002)
Math ACT score	0.002
	(0.002)
Advanced STEM courses	-0.002
	(0.005)
GPA	0.017*
	(0.01)
Took Calculus	-0.002
	(0.011)
Between high school and postsecondary school	
Months between high school and postsecondary school	-0.002
	(0.001)
Postsecondary school achievement	
Number of STEM courses taken	0.003***
	(0)
STEM GPA	0.019***
	(0.007)
Majored in STEM	0.134***
	(0.028)
R-squared	0.3268
Number of observations	3096

Standard errors are in parentheses.

***Significance at 1% level, **Significance at 5% level, *Significance at 1% level.

(a) works in STEM occupation

D.2 Field of Education

Table D.2.1 includes a proxy if a mother's advanced degree is in education.

D.2.1: High school achievement regression for all students

	(1a)	(1b)	(1c)	(1d)	(1e)
Variables of interest					
Female	0.463***	0.001	0.034	0.321***	0.015
	(0.122)	(0.122)	(0.026)	(0.02)	(0.01)
Mother has advanced degree	0.597***	0.475**	0.244***	0.061	0.045**
	(0.232)	(0.233)	(0.053)	(0.04)	(0.023)
Father has advanced degree	0.875***	0.356*	0.062	0.084**	-0.016
	(0.196)	(0.197)	(0.045)	(0.034)	(0.013)
FEMALE X Mother has advanced degre	-0.183	-0.153	-0.125*	-0.004	-0.022
	(0.32)	(0.323)	(0.075)	(0.056)	(0.02)
FEMALE X Father has advanced degree	-0.132	0.026	-0.055	0.013	0.014
	(0.27)	(0.272)	(0.062)	(0.046)	(0.023)
High school preparedness					
Math self-efficacy	0.139**	0.691***	0.07***	0.062***	0.037***
	(0.06)	(0.06)	(0.013)	(0.01)	(0.005)
Math test score	0.382***	0.414***	0.03***	0.032***	0.014***
	(0.007)	(0.008)	(0.002)	(0.001)	(0.001)
Hours spent on homework per week	0.03***	0.022***	0.007***	0.004***	0.002***
	(0.007)	(0.007)	(0.002)	(0.001)	(0.001)
Motivation	0.011	-0.087	0.009	0.081***	0.005
	(0.056)	(0.057)	(0.013)	(0.009)	(0.005)
Prediction of STEM occupation at age thirty	0.319*	0.547***	0.166***	0.01	0.043***
	(0.18)	(0.181)	(0.04)	(0.03)	(0.018)
Mother works in education	0.099	0.058	-0.043	0.062**	0
	(0.174)	(0.175)	(0.04)	(0.03)	0.375
R-squared	0.321	0.6928	0.321	0.4815	0.375
Number of observations	3995	3194	3995	3994	3987

Standard errors are in parentheses.

***Significance at 1% level, **Significance at 5% level, *Significance at 1% level.

(1a) Math self-efficacy (standardized), (1b) hours spent on homework per week, (1c) math test score (out of 100), (1d) motivation score (standardized), (1e) prediction of STEM occupation at age thirty

D.3 Accounting for Graduate School in STEM

Table D.3.1 redefines a career in STEM to include attending STEM graduate school.

Table D.3.1: Occupation Regression for All Students

	(a)
Female	-0.067***
	(0.017)
Mother has advanced degree	-0.049***
	(0.013)
Female X mother has advanced degree	0.092*
	(0.063)
Father has advanced degree	-0.019
	(0.016)
Female X father has advanced degree	0.056
	(0.043)
High School Preparedness	
Math self efficacy	-0.002
	(0.007)
Math test score	-0.001
	(0.001)
Time spent on homework per week	0
	(0.001)
Motivation	-0.002
	(0.006)
Expected Career at 30 is in STEM	0.08***
	(0.027)
High School Achievement	
ACT score	-0.002
	(0.003)
Math ACT score	0.005*
	(0.003)
Number of advanced STEM courses	0.01
	(0.007)
High school GPA	0.012
	(0.015)
Took Calculus	-0.011
	(0.015)
Between High School and Postsecondary School	
Months between high school and postsecondary school	-0.002
	(0.002)
Postsecondary School Achievement	
Number of STEM courses taken	0.004***
	(0.001)
STEM GPA	0.03***
	(0.01)
Majored in STEM	0.125***
	(0.038)
R-squared	0.3759
Number of Observations	1819

Standard errors are in parentheses.

***Significance at 1% level, **Significance at 5% level, *Significance at 1% level.

(5a) works in a STEM occupation, which includes STEM graduate school