2017

A New Frontier: But for Whom? An Analysis of the Micro-Computer and Women's Declining Participation in Computer Science

Eliana Keinan
Claremont McKenna College

Claremont McKenna College

A New Frontier: But for Whom? An Analysis of the Micro-Computer and Women's Declining Participation in Computer Science

SUBMITTED TO

Professor Manfred Keil

AND

Professor Serkan Ozbeklik

AND

Dean Peter Uvin

BY

Eliana Keinan

for
Senior Thesis
Fall 2016
December 5, 2016

Abstract

Though women's participation in science, technology, engineering, and mathematics (STEM) fields has greatly increased over the past 60 years, women's participation in computer science peaked in the 1980s. The paper searches for key motivators for women entering computer science at the peak in order to isolate factors for the subsequent steep decline. A major finding of the paper is that having a computer at home is (weakly) statistically significant as a determinant for female students choosing to pursue computer science. This relationship is insignificant for students in other STEM and non-STEM fields. A final section of the paper examines employment in computing. There is some support to suggest that early exposure to computing is correlated with individuals, both male and female, subsequently using a computer at work.

Acknowledgements

I would first like to thank my readers, Professors Keil and Ozbeklik, for their ongoing mentorship and guidance throughout my research and study of Economics. Thank you for continually pushing me to think more critically and dive deeper in my analysis.

I would also like to thank my friends for constantly motivating me to learn more and for all their support and happiness over the past four years. I am very grateful to be a part of such an incredibly inspiring, intellectually curious, and warm community at Claremont McKenna College.

Finally, without the endless support and love of my family, this would not have been possible. I am so lucky and thankful for them always pushing me to try new things and be a better version of myself. In particular, thank you to my mother, who majored in computer science in the 1980s, for instilling in me her love of science and math and showing me the importance of hard work and always following your passions.

Table of Contents

I. Introduction

In a Los Angeles Times article published on November 28, 2016, journalist Sonali Kohli recounts the journey of FemSTEM, a robotics team comprised of eight Muslim girls aged 10 to 14 from Chino Hills, California that just won Best Overall Performance at the First Lego League robotics competition. These young females joined the competition because they "wanted the world to know that Muslim girls can be strong, intelligent, successful career women, whose parents and community support them," particularly in light of the recent political climate (Kohli 2016). As Kohli notes, this accomplishment is meaningful for both the Muslim community and young women who have not always had equal access to the world of technology.

Since the 1960s, women's participation in the labor market and pursuit of college degrees has increased significantly. In fact, women now earn more college degrees than their male counterparts. According to the National Center for Education Statistics (NCES), 57 percent of bachelor's degrees and 63 percent of master's degrees were obtained by women in 2010 (U.S. Department of Education 2012). Though in many fields, particularly in psychology and other social sciences, this percentage has been high since the 1980s, in Science, Technology, Engineering, and Mathematics (STEM) fields, females are still significantly underrepresented. In 2009, only 27 percent of the 9.2 million workers in the United States with a STEM degree were women (U.S. Department of Commerce 2011). Though the total percentage of STEM degrees awarded to female graduates has increased over the past fifty years, still only 26 percent of bachelor degrees are awarded in

STEM fields[1]. This participation gender gap is also reflected in the STEM workforce percentage – only 24 percent workforce in these fields was comprised of women in 2009, a percentage that only increased overall by three percentage points between 1989 and 2009 (U.S. Department of Commerce 2011).

Disproportions in participation rates in science and technology have been relevant in recent policy discussions, including at the federal level. This has resulted in several initiatives such as the *Investing in Innovation* (I3) Fund, which is focused on increasing women's interest and performance in these fields starting at a young age. In a speech in February 2013, President Obama said:

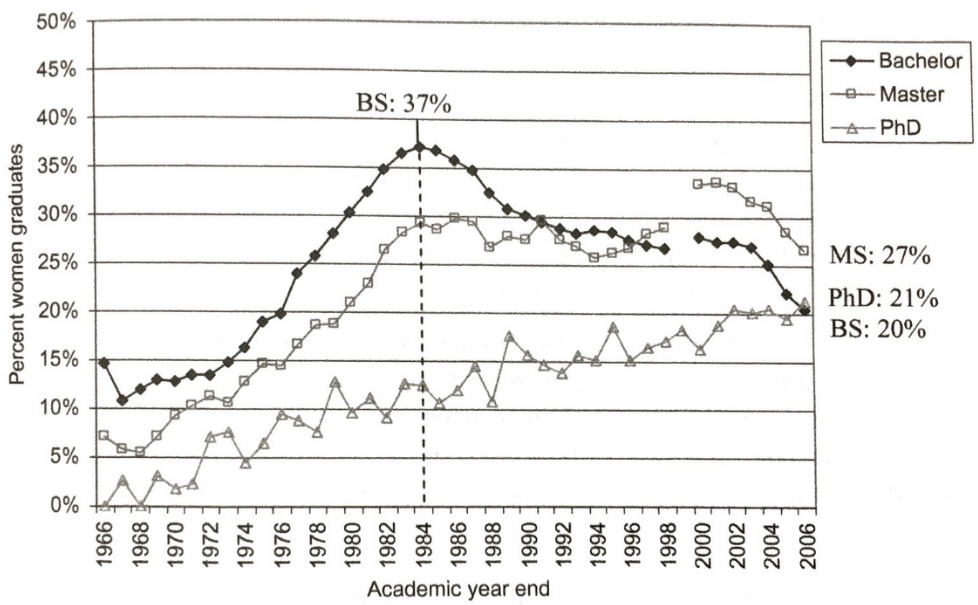
> "One of the things that I really strongly believe in is that we need to have more girls interested in math, science, and engineering. We've got half the population that is way underrepresented in those fields and that means that we've got a whole bunch of talent…not being encouraged the way they need to" (The White House 2013).

Though such efforts have potentially been successful across some STEM fields, computer science and computing have actually experienced a decline in female participation in both undergraduate education and the workforce. In fact, according to surveys collected by the National Science Foundation (NSF), computer science is the only STEM field to have seen a decrease in the percentage of bachelor's degrees awarded to women since 2002 (NSF 2008). Interestingly, the peak of women's participation was in the mid-1980s, where 37 percent of bachelor's degrees in computing were awarded to women and 38 percent of United States white-collar computing employees were female (Misa 2010). However, according to the NSF, women now make up only one seventh, or approximately 15 percent, of all undergraduate computing students. Thus, the current

[1] This excludes biology, which has traditionally had a significant number of women in the field.

problem is twofold: an extremely low current participation rate of women in computing, and a downward trend since the mid-1980s, both in the workforce and education. Currently, only 0.4 percent of first-year college students list computer science as a probable major entering college, according to the Taulbee Survey of top-ranked North American Computer Science and Engineering programs; this number was approximately ten times higher in the early 1980s. Figure 1 displays this high growth rates and eventual peak in women's participation rates in computer science as the field. In 1967, 24 out of the 222 (11 percent) computer science majors in the U.S. were women. This increased to 12,066 out of the 43,435 (37 percent) at the peak in 1984, before declining to only 14,406 out of 57,405 (20 percent) in 2006.

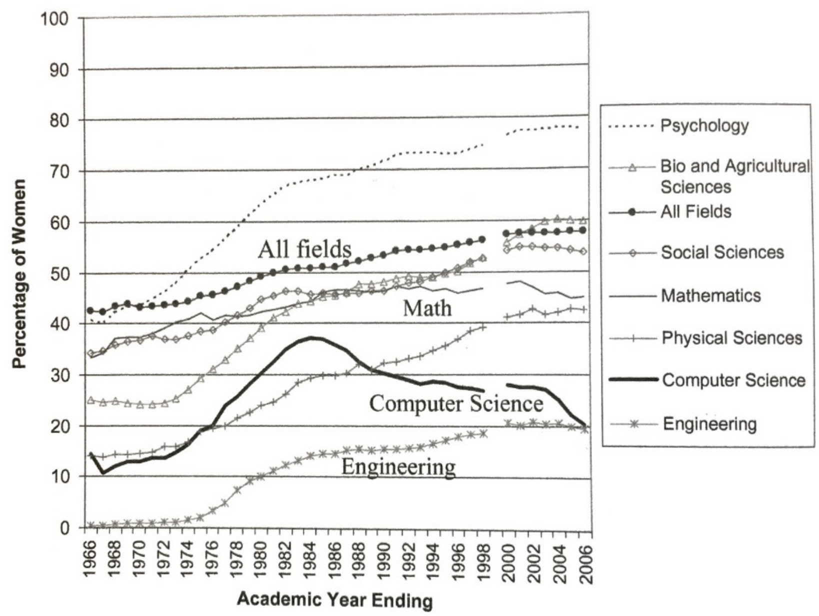
Figure 1: Proportions of Women receiving Bachelor, Master, and PhD. Degrees in Computer Science in the U.S.



Source: Misa 2010, 29

Thus, although the women's movement encouraging young women to pursue any undergraduate major can help explain the significant growth in female participation in computer science, the drastic downfall is unique amongst STEM fields. This is shown in Figure 2, where, until 1984, the growth in women earning computer science degrees was far more rapid than any other STEM field; however, it is the only field to experience a subsequent decline.

Figure 2: Proportions of Bachelor's Degrees Awarded to Women in the U.S. for Various Disciplines



Source: Misa 2010, 30

This is somewhat puzzling since women actually played a significant role in the early development stages of the now male-dominated field of computing. In the 1940s and 50s, women thought of digital computing as exciting and fun; it allowed them to be proud of their work and explore completely new territory (Abbate 2012, 11). Elsie Shutt, hired by Raytheon, an aerospace and defense manufacturing company, in 1953 who went on to start her own, all-female freelance programming company, reflecting on computer science today, said, "It really amazed me that these *men* were programmers, because I thought it was women's work!" (Abbate 2012, 1).

The field of computing largely began and greatly developed during the Second World War in an effort to use analysis to outsmart the opponent. During wartime, the supply of male labor was extremely scarce, and thus the United States Army began staffing hundreds of women for computing roles. The word "computer" actually originated in this period to refer to people – typically women – who manually did calculations at mechanical desks (Abbate 2012, 12-13).

One important example of female involvement is the development of the Electronic Numerical Integrator and Computer (ENIAC) at the University of Pennsylvania in the early 1940s. At the time, the U.S. was particularly interested in the potential for machines to enhance ballistic calculations, such as using firing tables to map out how gunners could aim weapons to better hit targets at specific ranges. These computations were done by either the Moore School's Differential Analyzer or the two hundred female computers trained in mathematics using desk calculators[2]. Thus, in 1943, the Army recruited a team of six women to program the ENIAC, a machine that would allow for significantly faster calculations. Similar to most technical innovations at the time, the physical engineering and production of the machine was built by men, but the machine was programmed solely by this team of women. Though the ENIAC was not completed until November 1945 and thus arrived too late to help the war effort, it played a major role in shaping the exponentially growing industry of computing and demonstrated how women were able to

[2] We see a similar trend in Great Britain during WWII, where two-thirds of the individuals on Alan Turing's code-breaking team at the Government Code and Cypher School were in fact women (Heath 2015).

enter a completely unknown field and make significant contributions (Abbate 2012, 12-13).

Yet, quickly we see a shift from women's high involvement in the cutting edge of technological development, such as Elizabeth "Jake" Feinler in the 1970s defining the top-level internet domain names still used today, to computing being seen as a field dominated by white males – often referred to as "brogramming" (Misa 2010). In the mid-1980s, with the innovation of the microcomputer, suddenly the computer became a gadget for younger boys. Similar to most other previous mechanical toys, they tended to be advertised specifically to one gender – males. The culture of the "geek" was formed, evident especially in popular culture (NPR 2014). The 1980s were infiltrated with movies such as *Revenge of the Nerds* and *Weird Science* that followed groups of socially awkward, male protagonists extremely skilled at computing. If women were featured at all, they were in highly-sexualized, objectifying contexts (NPR 2014). All of this framed computer science as a hostile field for women. Levy (1984, Ch. 4) states "You knew that horribly inefficient and wasteful things like women, they burn too many cycles, occupy too much memory space." We see this culture reinforced even today, such as in the hit HBO television show *Silicon Valley*, which tells the story of five introverted, male programmers who try to make it as entrepreneurs in the Bay Area.

In this paper, I aim to explore the factors that contribute to the creation of this culture and stereotype in the attempt to understand this shift. In particular, it examines the role the invention of the personal computer played in the growth of the gender gap and solidification of cultural norms associated with the field. Using the High School & Beyond (HS&B) Longitudinal Study conducted by the National Center for Education Statistics

(NCES) starting in 1980, I study how high school experiences such as having a computer at home and using a computer terminal in school are correlated with male and female students' decisions to study computer science, other STEM fields, or non-STEM fields in college.

I find that having a computer at home was a statistically significant factor positively correlated with female students choosing to pursue computer science undergraduate degrees in the mid-1980s but an insignificant factor for their male counterparts and for all students in both other STEM fields and non-STEM fields. I then explore how this relationship compares to employment in computing fields and find that early exposure to computing, such as having a computer at home and using a computer terminal in high school, is correlated with individuals, both male and female, then using computers at work a few years later.

This paper proceeds as follows: the next section reviews relevant literature. This is followed by a discussion of the data set used and the two models I create to analyze undergraduate major choice. I then look at the computing workforce and computer usage on the job. A final section concludes the paper.

II. Literature Review

There exists extensive literature attempting to explain the underrepresentation of women in STEM fields. Hunt (2016) finds that a key reason for women leaving science and engineering professions is their relative dissatisfaction with wages and the lack of promotion opportunities. Hewlett *et al.* (2008) argue that the combination of hostile workplaces, the isolating feeling of being the only woman on a team, and the difficulty of managing intensive work hours and family responsibilities are some of the key factors in the "brain drain" of women exiting the STEM workforce at a higher rate than their male colleagues.

Oguzoglu and Ozbeklik (2016) find that women are more likely to pursue undergraduate degrees in STEM fields if their fathers had STEM occupations, unless they have a brother. This study therefore attempts to explain the intergenerational impact on gender roles associated with majors in science and engineering, namely that they are primarily "male" fields. The negative stereotypes established by parents and teachers about gender-related differences in math abilities shape young girls' math abilities and compromise their interests in STEM fields (Shapiro and Williams 2011). Carrell *et al*. (2010) find that professor gender has a significant impact on female students' performance in science and math classes, as well as their desire to pursue and eventual success in earning STEM undergraduate degrees, particularly for female students who are already strong in math and science. Yet, they find professor gender composition has minimal impact on female students in other, non-STEM fields and on male students regardless of field. That said, Cheryan *et al*. (2011) find that only non-stereotypical role models in STEM positively influenced women's success beliefs.

There have been some efforts to explain the decline in female participation percentages in computer science; though overall the subject is relatively overlooked, as computing tends to be lumped under the larger category of STEM. Margolis and Fisher (2002) try to explain the decline through collecting first-hand accounts of computer science students and find that “most [women] did not have the same experience of falling in love at an early age [with computing] that many boys did” (Margolis and Fisher 2002, 18). Instead, of the girl gamers and computer club members, most did not come to the subject until high school after first being interested in math and science and excited by problem-solving (18).

One prominent theory for this discrepancy between genders in their early exposure to computing is that computers were seen much more as a “male toy” kept in the boys’ room in a household. Margolis and Fisher (2002) find that families were more likely to buy personal computers for their children in the 1990s if they were male than female, even when individuals in both genders were equally interested. They hypothesize this helped contribute to computers and computing being seen as a male hobby. They find that more than half of the male students interviewed had a computer in their bedrooms – whether it was their own or the family computer. Schofield (1995) examines students who spent time in their school’s computer lab during their lunch period in the mid-1980s and finds every boy claimed to have a computer at home, while not a single girl did. Giacquinta, Bauer, and Levin (1993) find the majority of families with computers kept them in locations primarily accessible to the males in the household. Similarly, Goldstein (1994) find that fathers spend more time playing with their sons than their daughters, which results in them also spending more time teaching their sons about computing as well. All of this early exposure to computing results in a significant experience gap in both formal and informal computing between interested students of different genders (Margolis and Fisher 2002, 40).

Further theories include the idea of computer science being seen as a “boys club.” In particular, Turkle (1984) argues that successful computer scientists must enjoy risk taking and the unknown, yet there is an expectation that boys should be bold and daring while girls are taught to be attentive and careful. This is reinforced by a 1999 report by the Fairfax County Human Relations Advisory Committee which finds that female students tended to sign up for more basic word processing classes in high school while 94 percent of students in the artificial intelligence class and 77 percent in the business computer programming class were male (AAUW 2000). Further, the American Association of University Women (AAUW) Commission finds that, in general, assignments in programming classes focused on more male-dominated interests, such as sports, which likely deterred and isolated female students (AAUW 2000).

When looking at key motivators for pursuing an undergraduate degree in computing, Margolis and Fisher (2002) find that women’s reasons primarily included enjoying computing, the value versatility of computing, computing’s relation to their interests in math and science, and the employment security in its career path. Alternatively, the male decision making primarily focused on studying computer science in college, simply a natural extension of their longtime passion for computing. Thus, they support AAUW (2000)’s conclusions in arguing that high school is a critical time for sparking girls’ interest in computer science.

In this paper, I attempt to further explain the reason for this decline in women pursuing computer science degrees. This paper adds to the literature because, though there are several qualitative explanations for the phenomenon, there is little empirical evidence regarding this trend. Thus, I specifically look at how certain high school experiences in the 1980s, the critical period where we see this shift from rapid growth to a significant decline in women's participation in the field of computer science, were correlated with young men and women deciding to pursue computer science in college.

III. Data

The data set used in this analysis is the High School & Beyond (HS&B) Longitudinal Study conducted by the National Center for Education Statistics (NCES) for its National Longitudinal Studies Program. It was designed to provide nationally representative data in the United States by randomly selecting high schools of differing types across the country. This data is comprised of two cohorts of students – the senior class of 1980 and the sophomore class of 1980. The data comes from a base year survey administered in 1980 and four follow-up surveys in 1982, 1984, 1986, and 1992 (though the senior cohort did not participate in the fourth survey in 1992) along with supplemental information obtained such as related transcripts. The data includes information collected from a Student Questionnaire, a School Questionnaire, a Teacher Comment Checklist, and a Parent Questionnaire, all gathered throughout the duration of the survey years. In this paper, I analyze only the sophomore cohort data because it contains more extensive variables, especially relating to computing.

This sophomore cohort portion of the data set is helpful for my analysis because it provides detailed student demographic information (e.g., gender, race, and family income). Moreover, it has detailed information regarding their college experiences including, but not limited to, their bachelor's major and credits obtained in specific fields (e.g., computer science and computer-related subjects). It also includes information on their home experiences in high school, such as if the student had a microcomputer at home in 1982, if the student used a computer terminal senior year of high school, and planned college degree while still in high school. It also has detailed occupational information (e.g., occupation field) upon undergraduate graduation. Finally, this data set is valuable because the time period in which the surveys were administered is crucial for examining why the gender participation percentages for women peaked in the mid-1980s. Using high school student data from these critical peak years can help illustrate key characteristics correlated with participation in computer science degrees and jobs. From these correlations, I can then infer the impact on larger trends of the population of students in the United States.

A key limitation in the data is nonresponse error. Of the 1,120 schools selected in the original sample, only 72 percent of the schools (811 schools) chose to participate in the survey series and thus an additional 204 schools were selected to be in the replacement sample. That said, completion rates at the participating schools were quite high, with an 88 percent participation rate in the base-year, and an 86 percent participation rate for the sophomores in the fourth follow-up. The nonresponse error becomes more significant when looking at several variables at once because data points have missing values for specific questions, thus greatly reducing the sample size in the analysis. I also restrict my undergraduate major analysis to only examine students who attended a 4-year college or university, further reducing the sample size. Taking into account only observations without missing information for each variable studied, my final sample includes only 2,982 observations, clearly a significant limitation.

III. A. Key Variable Definitions:

The key variables used in this analysis are undergraduate degree, gender, microcomputer at home in 1982, computer terminals used in high school, and in computer programming senior year of high school.

For undergraduate degree, I create two indicator variables for undergraduate college major. First, I create an indicator variable for computing majors equal to 1 if the student majored in any computer science or computing-related undergraduate degree and 0 otherwise. Second, I create an indicator variable for STEM majors (including computing) equal to 1 if the student majored in Biological Sciences, Computer Information Science, Engineering, Health Occupations, Health Sciences, Mathematics, and Physical Science and 0 otherwise.

I create an indicator variable for female equal to 1 if the student identifies as female and 0 if the student identifies as male. I also create an indicator variable for microcomputer equal to 1 if the student claimed to have a microcomputer in their household in 1982, regardless of it being their own or a different family member's, and 0 otherwise. Similarly, I create an indicator variable equal to 1 if the student used computer terminals in high school and 0 otherwise. Finally, I create an indicator variable equal to 1 if the student was involved in computer programming during their senior year of high school and 0 otherwise.

In this reduced dataset, 51.2 percent of the observations are females compared to the 48.8 percent male. In fact, women make up 51.1 percent of the computer science majors in the data set, 49.0 percent of other STEM majors, and 52.0 percent of all non-STEM majors (see Table 1 and Table 2).

Table 1: Breakdown of Undergraduate Major Choice

Undergraduate Major	**Number of Observations**	**Percentage Female (Male)**
Computer Science	139	51.1% (48.9%)
STEM Non-Computer Science	785	49.0% (50.0%)
Non-STEM	2,058	52.0% (47.0%)

Table 2: Breakdown of Undergraduate Major Choice by Gender

	Men		Women	
Undergraduate Major	**Number of Observations**	**Percentage of Observations in Major**	**Number of Observations**	**Percentage of Observations in Major**
Computer Science	68	4.7%	71	4.7%
STEM Non-Computer Science	400	27.5%	385	25.2%
Non-STEM	987	67.8%	1,071	70.1%
Total	1,455	100%	1,527	100%

III. B. Additional Control Variables

I create an indicator variable for race called *nonwhite* equal to 1 for students who identify as Hispanic or Spanish, American Indian, Asian, Pacific Islander, Black, or Other, and 0 for students who identify as White.

I include several additional variables to control for differing backgrounds that could affect a student's likelihood to pursue computing and STEM fields. I control for family income by creating two dummy variables to represent middle-income and high-income families. I create four dummy variables to describe the highest level of education the student's father has obtained. I create analogous dummy variables for the mother's level of education. I create two dummy variables to describe the amount of influence the student's father has on the student's post-high school plans, specifically large influence and medium influence. I create analogous dummy variables for the mother's influence on post-high school plans.

I also control for the individual's quantitative ability by including sophomore year standardized math part 2 test scores. This variable is also likely correlated with general enjoyment in mathematics.

IV. Methodology and Regression Analysis:

I aim to analyze how individuals' experiences by gender are correlated with the students pursuing STEM, particularly computer science and computing, undergraduate majors. I estimate two models to investigate this: a linear probability model and a multinomial logit model. I discuss each in turn.

IV. A. Linear Probability Model

I first run a basic linear probability model[3] to analyze the determinants of undergraduate major choice. Specifically, I estimate a model of the following form:

$$major_i = \alpha + \beta_1 female + \beta_2 compAtHome + \beta_3 female*compAtHome + \beta_4 X + i.schlid + \epsilon_i,$$

where *major* can be either computer science major or STEM major, *female* is an indicator variable equal to 1 if the student identifies as female and 0 otherwise, *compAtHome* is an indicator variable equal to 1 if the student has a microcomputer at home and 0 otherwise, and *female*compAtHome* is the interaction of the two and is included to determine if the effect of having a computer at home differs between genders.[4] X is a vector of additional variables including computer terminal used in high school, in computing junior and senior year, math test scores, race, family income, parents' education levels, and parents'

[3] An alternative to using the Linear Probability Model (LPM) is to estimate the relationship through a logit/probit model. While it is easier and more intuitive to interpret the coefficients of the LPM model, this is not a sufficient justification. Instead, it turned out that the linear model is a good approximation for the logit model for this data set, evident in that the statistically significant coefficients closely mirror those in the results of the linear model I describe in the next section. Additionally, testing the extremes of the linear model demonstrates that the estimations are still primarily within 0 and 1, and thus this model is appropriate to use.

[4] I chose to exclude interaction variables between other explanatory variables of interest (namely, *sy11c2* and *fy9l*) because when included they yield insignificant results, thus suggesting their effects are consistent for both male and female students.

influence on post-high school plans. *i.schlid* is a vector of school fixed effects included to account for differences in school types. Finally, ε is an error term with the usual properties.

I run two separate specifications for major choice. I first run the regression where the major specification is computer science. I next run the regression where the major specification is STEM major in order to test the robustness of the results and examine if they are true for all majors or just computer science majors. The next section discusses the results, shown in Table 3.

IV. B. Linear Regression Results

In this section I will discuss the various outcomes from my estimation. Table 3 looks at the linear probability for the given major options. Columns 1 - 3 display the results of three regressions for the indicator variable, *CSmajor*. Column 1 lists the outcomes for a regression just of my primary variable of interest - *compAtHome*. Column 2 and 3 then introduce the other key explanatory variables studied, as well as the additional controls. Columns 4 - 6 display analogous results for the indicator variable *STEMmajor*.

Table 3: Linear Regression for CS and STEM Majors

	CS MAJOR			STEM MAJOR		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
female	-0.00599	-0.00668	-0.00735	-0.0287	0.0230	0.0290
	(0.00800)	(0.0102)	(0.0104)	(0.0177)	(0.0237)	(0.0243)
compAtHome	-0.0203	-0.00836	-0.0105	-0.0263	-0.0365	-0.0313
	(0.0152)	(0.0195)	(0.0200)	(0.0406)	(0.0502)	(0.0511)
female*compAtHome	0.0786**	0.0680*	0.0656*	0.0742	0.0891	0.0707
	(0.0343)	(0.0381)	(0.0384)	(0.0639)	(0.0798)	(0.0811)
ybmth2sd		0.000370	0.000258		-0.000901	-0.00123
		(0.000508)	(0.000517)		(0.00110)	(0.00111)
sy11c2		0.0265*	0.0223		0.0219	0.0160
		(0.0140)	(0.0142)		(0.0318)	(0.0322)
fy91		-0.0116	-0.00626		0.00939	0.0187
		(0.0117)	(0.0116)		(0.0290)	(0.0295)
highincome		-0.00894	-0.0124		0.0262	0.0205
		(0.0195)	(0.0201)		(0.0367)	(0.0379)
midincome		-0.0143	-0.0166		0.00932	0.00316
		(0.0186)	(0.0193)		(0.0351)	(0.0365)
nonwhite		0.000786	-0.000849		0.0342	0.0307
		(0.0118)	(0.0119)		(0.0285)	(0.0289)
father_HSeduc		0.00601	0.00942		0.0385	0.0158
		(0.0185)	(0.0205)		(0.0400)	(0.0438)
father_UGeduc		-0.0256	-0.0212		-0.00155	-0.0179
		(0.0209)	(0.0237)		(0.0483)	(0.0524)
father_MDeduc		-0.0214	-0.0186		-0.0275	-0.0583
		(0.0241)	(0.0271)		(0.0543)	(0.0588)
father_DDeduc		-0.00783	-0.00413		0.0375	0.00370
		(0.0239)	(0.0261)		(0.0580)	(0.0621)
mother_HSeduc		0.00986	0.0113		-0.0126	-0.0151
		(0.0120)	(0.0122)		(0.0284)	(0.0291)
mother_UGeduc		0.00865	0.00881		0.00537	0.00532
		(0.0163)	(0.0164)		(0.0398)	(0.0406)
mother_MDeduc		0.0259	0.0267		-0.0124	-0.00542
		(0.0221)	(0.0225)		(0.0519)	(0.0533)
mother_DDeduc		0.0650	0.0654		-0.0147	0.0103
		(0.0423)	(0.0429)		(0.0868)	(0.0881)
father_largeinfluence			0.00248			0.0400
			(0.0205)			(0.0402)
father_medinfluence			-0.00867			0.0138
			(0.0178)			(0.0369)
mother_largeinfluence			-0.000143			-0.0391
			(0.0200)			(0.0445)
mother_medinfluence			-0.00842			-0.0212
			(0.0185)			(0.0427)
Control for school fixed effects?	No	Yes	Yes	No	Yes	Yes
Constant	0.0487***	0.00237	0.0145	0.324***	-0.0457	-0.0112
	(0.00594)	(0.0372)	(0.0424)	(0.0129)	(0.0807)	(0.0920)
Observations	2,982	2,982	2,899	2,982	2,982	2,899
R-squared	0.003	0.294	0.304	0.001	0.278	0.293

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The regression results for computer science majors indicate that the relationship between pursuing a computer science degree and having a computer at home is significant at the 10 percent level if the student is female in comparison to male, as seen in Column 2 and 3. This coefficient on the interaction variable *female*compAtHome* says that women relative to men are approximately 7 percentage points more likely to be pursuing a computer science degree if they have a computer at home in high school. Column 2 indicates this relationship is slightly stronger when we do not include parents' influence on post-high school plans. Looking at the *compAtHome* variable, which is statistically insignificant, we can see this same relationship regarding having a computer at home does not exist for men in computing.

Column 2 indicates the variable for using a computer terminal in high school is significant at the 10 percent level. This means that both male and female students are approximately 2.7 percent more likely to pursue a computer science degree if they have this high school computing experience.

Analyzing the results for STEM majors, we see that none of the variables included are statistically significant, as was expected in the hypothesis. This is also true for non-STEM majors, which will simply be the inverse of the coefficients in Columns 4-6. These results therefore suggest that having a computer at home only is significantly correlated for female in comparison to male students choosing to study computer science.

IV. C. Multinomial Logit Model

I next run multinomial logit regressions in order to analyze the probability of an individual choosing to pursue a computer science degree in comparison to their other major options. I choose to use a multinomial logit model rather than other logit models because the data set only includes case-specific variables yet still includes data from multiple alternatives (i.e. major choice). The multinomial logit model specifies that:

$$p_{ij} = \frac{\exp\left(x_i' \beta_j\right)}{\sum_{l=1}^{m} \exp\left(x_i' \beta_l\right)}$$

where x_i are the case specific regressors (i.e. the same variables included in the linear regression above) and p_{ij} refers to the probability that an individual will choose a particular major. In order to run this regression, I first create a variable *major* that equals 1 if the student is a computer science major, 2 if the student is a non-computer science STEM major, and 3 if the student is a non-STEM major.

The multinomial logit model relies on the assumption of independence of irrelevant alternatives (IIA), which states that presenting an additional alternative (i.e. an additional major choice) does not change the relative probabilities of individuals selecting between existing alternatives. This means that error terms of different choice equations are independent. This seems like a reasonable assumption especially when comparing non-STEM majors to computer science and other STEM majors because they require very different expertise. This assumption becomes less clear when considering computer science in comparison to other STEM fields because they require similar basic skills, though still are considered fairly distinct types of work. Though this is a potential limitation in using this model for the data set, the data set lacks the necessary information to use a nested logit or conditional logit model, thus making the multinomial logit model a good choice for my analysis.

I run two identical regressions except one where non-STEM major is the base outcome and the other where non-computer science STEM major is the base outcome. This allows me to directly understand the probability of going into computer science given the regressors in comparison to the individual going into other STEM fields or a non-STEM field altogether. Positive coefficients in this model can be interpreted to mean that as the given regressor increases, the probability of picking the chosen alternative in comparison to the base outcome increases. The inverse is true for negative coefficients.

IV. D. Multinomial Logit Regression Results

Table 4 displays the results of the regression, both when non-STEM major is the base outcome and when non-computer science STEM major is the base outcome. Column 1 and 3 contain the coefficients in the multinomial logit model for students choosing computer science over the base outcome. Table 5 then lists these coefficient estimates transformed into relative-risk ratios, i.e. the relative odds of picking the given alternative instead of the base outcome.

Table 4: Multinomial Logit Regression

	Base: Non-STEM		Base: STEM, Non-CS	
VARIABLES	CS	STEM Non-CS	CS	Non-STEM
female	-0.179	-0.110	-0.0688	0.110
	(0.188)	(0.0889)	(0.197)	(0.0889)
compAtHome	-0.504	-0.101	-0.403	0.101
	(0.529)	(0.202)	(0.546)	(0.202)
female*compAtHome	1.482**	0.0923	1.390**	-0.0923
	(0.645)	(0.320)	(0.680)	(0.320)
ybmth2sd	0.00543	-0.00574	0.0112	0.00574
	(0.00916)	(0.00437)	(0.00962)	(0.00437)
sy11c2	0.284	0.0641	0.219	-0.0641
	(0.240)	(0.119)	(0.253)	(0.119)
fy91	0.00473	0.0280	-0.0233	-0.0280
	(0.235)	(0.114)	(0.247)	(0.114)
highincome	0.0322	0.374**	-0.342	-0.374**
	(0.292)	(0.155)	(0.312)	(0.155)
midincome	-0.0459	0.312**	-0.358	-0.312**
	(0.280)	(0.150)	(0.300)	(0.150)
nonwhite	0.0576	0.0486	0.00903	-0.0486
	(0.196)	(0.0947)	(0.206)	(0.0947)
father_HSeduc	-0.0770	0.223	-0.300	-0.223
	(0.300)	(0.171)	(0.325)	(0.171)
father_UGeduc	-1.013**	0.146	-1.158**	-0.146
	(0.432)	(0.200)	(0.455)	(0.200)
father_MDeduc	-0.507	0.0703	-0.578	-0.0703
	(0.450)	(0.229)	(0.479)	(0.229)
father_DDeduc	-0.768	0.274	-1.042*	-0.274
	(0.514)	(0.239)	(0.540)	(0.239)
mother_HSeduc	-0.0280	-0.0403	0.0122	0.0403
	(0.233)	(0.114)	(0.245)	(0.114)
mother_UGeduc	-0.157	0.0314	-0.189	-0.0314
	(0.383)	(0.163)	(0.398)	(0.163)
mother_MDeduc	0.177	0.0156	0.162	-0.0156
	(0.439)	(0.213)	(0.461)	(0.213)
mother_DDeduc	0.634	-0.245	0.879	0.245
	(0.600)	(0.361)	(0.653)	(0.361)
Constant	-2.748***	-1.123***	-1.625**	1.123***
	(0.667)	(0.335)	(0.706)	(0.335)
Observations	2,982	2,982	2,982	2,982

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

In Table 4, Columns 1 and 3 indicate that for female students, having a computer at home has a positive impact on their decision to enter computer science in comparison to non-STEM fields as well as other STEM fields. This result is statistically significant at the 5 percent level. To understand this impact, we can look at relative-risk ratios for the variable *female*compAtHome* for computer science in comparison to the other alternatives. Table 5 lists the relative-risk ratio for the interaction variable *female*compAtHome* for computer science in comparison to the two other major options. These values indicate that a female student who has a computer at home is approximately 4.4 times more likely to pursue a computer science degree rather than a non-STEM degree because they have a computer. Similarly, a female student who has a computer at home is 4.0 times more likely to pursue a computer science degree rather than an alternative STEM degree because they have a computer. These relative-risk ratios were obtained by taking the coefficients in Table 4 and raising the value *e* to the power of the coefficient, i.e. $e^{\beta_{jr}}$.

Column 1 and 3 in Table 4 also suggest that increases in father's education from not having a high school degree to having one is correlated with both male and female students being less likely to choose to major in computer science over either of the two alternative major choices. This influence is the same for both genders, as is evident by statistically insignificant coefficients on interaction variables created between female and parent education.

Table 5: Relative-Risk Ratios for Statistically Significant Coefficients

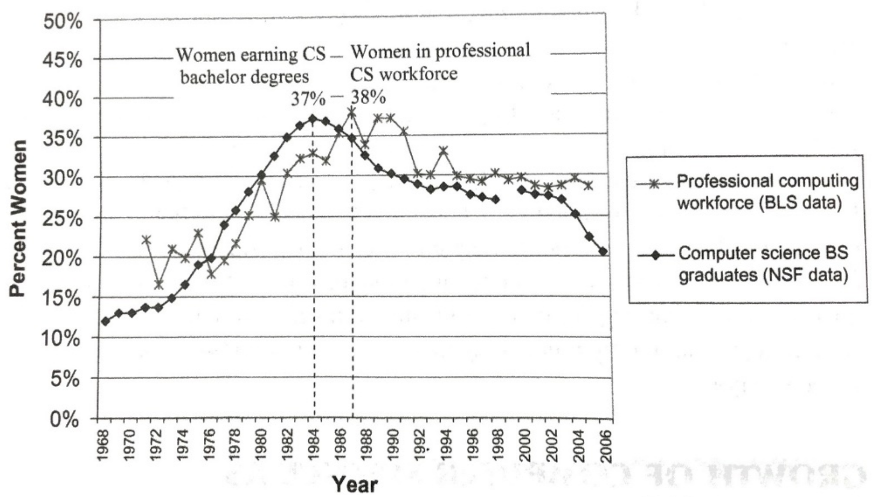
Variable	Base Outcome Major	Relative-Risk Ratio for Computer Science
*female*compAtHome*	Non-STEM	4.4030 (2.840)
	Non-CS STEM	4.0148 (2.7314)
father_UGeduc	Non-STEM	0.3632 (0.1568)
	Non-CS STEM	0.3140 (0.1428)
father_DDeduc	Non-CS STEM	0.3528 (0.1905)

Thus, the basic linear regression model and the multinomial logit model both suggest the same result: having a computer at home for female students is correlated with the students pursuing undergraduate computer science degrees instead of other degrees.

V. Employment

According to the NSF, employment trends follow similar trends to undergraduate education for computer science. This is shown in Figure 3, where female participation in the professional computing workforce closely mirrors trends in computer science Bachelor's degrees earned.

Figure 3: Proportion of Professional Computing Workforce and Computer Science BS Graduates that are Women in the U.S.



Source: Misa 2010, 33

Therefore, I attempt to confirm my findings about women in computer science by looking to see if similar relationships exist for undergraduate education and employment for the HS&B data set. In order to examine this, I first generate an indicator variable *CSjob* equal to 1 if the individual reported their first, second, third, or fourth job to be either a computer programmer, a computer systems analyst, or a computer specialist and a 0 otherwise. However, this generated variable has only 22 individuals reported as having a computing occupation. This is too small of a sample size to generate meaningful analysis. An interesting observation though is that only 3 of these 22 individuals (14 percent) are women, compared to the 51 percent of computer science majors that are women.

That said, there are several variables asking about computer usage on the job, in education, and recreationally. I use the job-related computer usage variables to create a joint variable *compAtWork* that can roughly estimate employment in the fields of computing and computer science in the 1980s. Though greatly limited in describing the computer science workforce, the variable *compAtWork* is understandably correlated with *CSjob* at the 1 percent level.

I run a similar linear regression to the model used for undergraduate major, but where the dependent variable is now *compAtWork*. Table 6 indicates that individuals with a computer at home are approximately 6.6 percent more likely to also use a computer at work, significant at the 10 percent level. However, no difference is seen across genders. The coefficient of *sy11c2*, significant at the 1 percent level, suggests that individuals who used a computer terminal in high school are approximately 5.8 percent more likely to use a computer at work than those who did not. This result is also the same for both genders because the interaction variable created between *female* and *sy11c2*, is not statistically significant. The coefficient of *fy9l*, significant at the 5 percent level, suggests that individuals are also more likely to use a computer at work if they were involved with computing junior and senior year of high school. Like with *sy11c2*, when interacted with *female*, it is insignificant, thus indicating the impact is consistent across genders.

Though these results suggest using a computer at work is correlated equally for men and women for the variables for having a computer at home, using computer terminals in high school and being involved in computer junior and senior year of high school, they do still support the claim that high school experiences, both in and out of the classroom, have a significant impact on later work and study in the field.

Once again, this data analysis is incredibly limited in describing relationships between the chosen variables and field of employment. This is especially evident in the fact that only 25 percent of computer science majors also later reported using a computer on the job compared to the 23 percent of non-computer science majors who also reported using a computer at work. This potentially also explains the fact that 65 percent of the individuals who reported using one of the above computers at work were women (646 females and 342 males).

Table 6: Linear Regression for *compAtWork*

VARIABLES	(1) Model 1
female	0.133***
	(0.0128)
compAtHome	0.0656**
	(0.0314)
female*compAtHome	-0.0528
	(0.0516)
ybmth2sd	0.00242***
	(0.000668)
sy11c2	0.0576***
	(0.0197)
fy91	0.0400**
	(0.0185)
highincome	0.0357*
	(0.0204)
midincome	0.0235
	(0.0191)
nonwhite	0.0360**
	(0.0143)
father_HSeduc	-0.0311
	(0.0238)
father_UGeduc	-0.0275
	(0.0297)
father_MDeduc	-0.0544
	(0.0345)
father_DDeduc	-0.0944***
	(0.0359)
mother_HSeduc	0.0207
	(0.0159)
mother_UGeduc	-0.0161
	(0.0251)
mother_MDeduc	-0.0236
	(0.0334)
mother_DDeduc	0.0669
	(0.0559)
Constant	-0.0486
	(0.0469)
Observations	4,380
R-squared	0.036

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

VI. Conclusion

This paper was written in an attempt to understand key factors in high school students, and more specifically women, deciding to pursue computer science undergraduate degrees in the 1980s. The variable determined to be the most significant in female students choosing to study computer science was having a computer at home (i.e. *compAtHome*). Female students in comparison to male students with a computer at home were shown to be more likely to choose computer science over both non-STEM fields and other STEM fields than students who did not have a computer at home. No such relationship was found to be significant for male computer science students or for both men and women in STEM fields in general.

These results are consistent with previous literature about key motivators for individuals pursuing computer science degrees. In particular, the results suggest that early exposure to computing in the 1980s was a significant factor correlated with young girls choosing to enter computer science. Having a computer at home in the 1980s is largely a proxy for having early exposure to computing because computers were just becoming popular and were not yet a common household item like they are today. This fits well into the larger narrative because, as computers were increasingly seen as a toy given to boys and advertised accordingly especially in the world of gaming, it is understandable that women's participation began declining. Then, as "brogramming" was continually reinforced in the 1980s and 1990s, further barriers to entry into the world of computing were established and solidified, even though access to computers became easier for members of both genders.

These results can also suggest the importance of family environment in students, and particularly female students, pursuing specific undergraduate degrees. After all, having a computer at home can be seen as a proxy for parental interest in computing and technology, especially because the time period I study is right before the computer revolution truly takes off. Thus, technologically curious parents were probably more likely to purchase a computer for their household, as well as pass down their interest in computing to their children. This is consistent with previous literature written about STEM fields in general, as previously described.

One important limitation of the data is the small sample size. Though the original data set has 14,825 data points, once all of the variables are included, this number is reduced drastically due to missing values for each variable. Additionally, I only look at a cross-section and thus can only infer later trends based on the results. Therefore, this study leaves several paths for future research.

First, my findings can be confirmed with a larger data set, particularly for understanding how the computer science workforce trends compare to the those in education. Specifically, it would be interesting to look at key reasons why women drop out of the computing workforce. Bertrand *et. al* (2010) find that though women and men start their careers with similar earnings, the wage discrepancy increases greatly by the time individuals get their MBAs, particularly due to shorter work hours for females. Thus, looking at how in computer science, female workers can potentially stay more engaged while at home than in other fields could yield interesting results.

Second, it would be interesting to examine how trends differ over time rather than looking at a specific cross-section. This data set was collected right before the huge boom in the computer revolution and thus describes motivations for women entering computing before computers became much more easily accessible and widespread. Examining how the role of the computer in students choosing to enter computing changed, particularly in the 1990s, could potentially be very telling of additional reasons for the decline in female participation in computing.

Third, examining students' experiences during college could further explain the "boys club" hypothesis for declining female participation in computer science. Colleges such as Harvey Mudd College (HMC) and Carnegie Mellon University (CMU) have spearheaded the recent push towards gender equality in computer science education. HMC, which boasts its high gender equality of faculty across all fields, just this past year had more female than male computer science graduates (Staley 2016). CMU credits its similar gender makeup of computer science majors to its "Women @ SCS [School of Computer Science]" organization comprised of both faculty and students that works on establishing a welcoming environment for both genders (Spice 2016). Thus, it would be interesting to analyze factors such as gender composition of professors in computing classes and additional resources available to students to see if there are correlations with female students choosing to switch out of computer science during their undergraduate careers.

Finally, looking at how female participation in computing compares in other countries could be valuable in understanding motivating factors and how these potentially differ across varying cultures. Further, analyzing countries such as China could reveal insights about intergenerational effects in computer science. The one-child policy, for example, could increase female participation in computer science, as parents focus all their attention on transferring skills to their single child rather than only male siblings, a trend seen with STEM fields described previously.

The results of my study suggest that in order to reverse this downward trend in female participation, it is likely important to focus on early exposure to computing. Though in today's world personal computers are very widespread and accessible, it is still important to introduce children of both genders, but in particular young girls, to the deeper world of computing. Whether this involves encouraging them to form robotics teams like FemSTEM, develop their own applications from scratch, or even enter gaming just like so many of their male counterparts, increased familiarity with computers from early on will likely inspire young girls to develop that same lifelong passion for computer science that found in so many male students.

VII. Bibliography

Abbate, J. 2012. *Recoding Gender: Women's Changing Participation in Computing.* Cambridge: MIT Press.

American Association of University Women Educational Foundation. 2000. "Tech-Savvy: Educating Girls in the New Computer Age."

Bertrand, M., C. Goldin, and L. Katz. 2010. "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors." *American Economic Journal: Applied Economics 2.*

Carrell, S.E., and J.E. West. 2010. "Sex and Science: How Professor Gender Perpetuates the Gender Gap." *Quarterly Journal of Economics 125* 1101-1144.

Cheryan, S., J.O. Siy, M. Vichayapai, B. Drury, and Saenam Kim. 2011. "Do Female and Male Role Models Who Embody STEM Stereotypes Hinder Women's Anticipated Success in STEM?" *Social and Psychological and Personality Science.*

Giacquinta, J.B., J. Bauer, and J. Levin. 1993. *Beyond Technology's Promise: An Examination of Children's Educational Computing at Home.* New York: Cambridge University Press.

Goldstein, J.H. 1994. *Toys, Plays, and Child Development.* Cambridge: Cambridge University Press.

Hewlett, S.A., C. Buck Luce, L.J. Servon, P. Shiller, E. Sosnovich, and K. Sumberg. 2008. "The Athena Factor: Reversing the Brain Drain in Science, Engineering and Technology." *Harvard Business Review Research Report* (Harvard Business Publishing).

Heath, N. 2015. *Hacking the Nazis: The secret story of the women who broke Hitler's codes.* March 27. www.techrepublic.com.

Hunt, J. 2016. “Why do Women Leave Science and Engineering?” *ILR Review*. 199-225.

Kohli, S. 2016. "This Muslim teen has her own way to protest the election -- winning robotics competitions." *Los Angeles Times,* November 28.

Levy, S. 1984. "Chapter 4, Greenblatt and Gosper." In *Hackers: Heroes of the Computer Revolution*. Garden City, NY: Anchor Press/Doubleday.

Margolis, J, and A Fisher. 2002. *Unlocking the Clubhouse: Women in Computing.* Cambridge: The MIT Press.

Misa, T. 2010. *Gender Codes: Why Women are Leaving Computing.* Hoboken, New Jersey: John Wiley & Sons, Inc.

National Center for Education Statistics. 2012. “The Digest of Education Statistics 2011” NCES 2012-001, U.S. Department of Education.

National Public Radio (NPR). 2014. "Episode 576: When Women Stopped Coding." http://www.npr.org/templates/transcript/transcript.php?storyId=487069271.

National Science Foundation, Division of Science Resources Statistics. 2008. *Science and Engineering Degrees: 1966-2006.* Detailed Statistical Tables NSF 08-321. Arlington, VA: National Science Foundation. www.nsf.gov/statistics/nsf08321/.

Oguzoglu, U., and S. Ozbeklik. 2016. "Like Father, Like Daughter (Unless There is a Son): Sibling Sex Composition and Women's STEM Major Choice in College." *IZA Discussion Paper Series.*

Schofield, J. 1995. *Computers and Classroom Culture.* Cambridge: Cambridge University Press.

Shapiro, J.R. and A.M. Williams. 2012. “The Role of Stereotype Threats in Undermining Girls’ and Women’s Performance and Interest in STEM Fields.” *Sex Roles*. 175-183.

Spice, B. 2016. "Women Are Almost Half of Carnegie Mellon's Incoming Computer Science Undergraduates." *Carnegie Mellon School of Computer Science.* September 12. http://www.cs.cmu.edu/news.

Staley, O. 2016. “Harvey Mudd College took on gender bias and now more than half its computer science majors are women.” *Quartz*. August 22. http://qz.com.

The White House, Office of Science and Technology Policy. 2013. “Fact Sheet: Women and Girls in Science, Technology, Engineering, and Math (STEM),” Retrieved from: https://www.whitehouse.gov/administration/eop/ostp/women.

Turkle, S. 1984. *The Second Self: Computers and the Human Spirit.* New York: Simon and Schuster.

U.S. Department of Education, National Center for Education Statistics. 2012. *The Condition of Education 2012 (NCES 2012-045), Table A-47-2.*

U.S. Department of Commerce. 2011. “Women in STEM: A Gender Gap to Innovation,” Executive Summary, Washington, DC.

VIII. Appendix

Table A: Summary Statistics for Key Variables

Name	**Description**	**Min**	**Max**	**Mean**	**Standard Deviation**
CSmajor	Is a computer science major	0	1	0.0466	0.2108
STEMmajor	Is a STEM major, including CS	0	1	0.3098	0.4625
female	Is female	0	1	0.5121	0.4999
compAtHome	Has a computer at home in 1982	0	1	0.0805	0.2721
sy11c2	Uses computer terminals in H.S.	0	1	0.2062	0.4047
fy9l	In computer programming, junior and senior year of H.S.	0	1	0.2357	0.4245

Table B: Summary Statistics for Additional Control Variables

Name	**Description**	**Min**	**Max**	**Mean**	**Standard Deviation**
nonwhite	Identifies as a race other than white	0	1	0.3206	0.4668
highincome	Family income $30,000 or more	0	1	0.4816	0.4997
midincome	Family income $15,000 - $29,000	0	1	0.3960	0.4892
father_HSeduc	Father has a high school degree	0	1	0.6026	0.4894
father_UGeduc	Father has an undergraduate college degree	0	1	0.1640	0.3703
father_MDeduc	Father has a master’s degree	0	1	0.0818	0.2741

father_DDeduc	Father has a doctorate degree	0	1	0.0694	0.2542
mother_HSeduc	Mother has a high school degree	0	1	0.6060	0.4887
mother_UGeduc	Mother has an undergraduate college degree	0	1	0.1315	0.3380
mother_MDeduc	Mother has a master's degree	0	1	0.0550	0.2280
mother_DDeduc	Mother has a doctorate degree	0	1	0.0178	0.1321
ybmth2sd	Sophomore math part 2 standardized test score	28.60	77.55	54.32	10.30
father_largeinfluence	Father has a large influence on post-high school plans	0	1	0.4403	0.4965
father_medinfluence	Father has a medium influence on post-high school plans	0	1	0.3949	0.4889
mother_largeinfluence	Mother has a large influence on post-high school plans	0	1	0.5051	0.5001
mother_medinfluence	Mother has a medium influence on post-high school plans	0	1	0.3986	0.4897

Table C: Summary Statistics for Additional Variables for Employment

Name	**Description**	**Min**	**Max**	**Mean**	**Standard Deviation**
CSjob	Has a job in computing	0	1	0.0053	0.0725
compAtWork	Uses a computer on the job	0	1	0.2262	0.4184