

A Large-Scale Quantitative Study of Women in Computer Science at Stanford University

Katie Redmond
Computer Science Department
Stanford University
Stanford, CA 94305, USA
kredmond@cs.stanford.edu

Sarah Evans
Computer Science Department
Stanford University
Stanford, CA 94305, USA
sevans1@stanford.edu

Mehran Sahami
Computer Science Department
Stanford University
Stanford, CA 94305, USA
sahami@cs.stanford.edu

ABSTRACT

In this paper, we analyze gender dynamics in the undergraduate Computer Science program at Stanford University through a quantitative analysis of 7209 academic transcripts and 536 survey responses. We examine previously studied effects as well as present new findings. We also introduce Fisher's Noncentral Hypergeometric Distribution as a model for estimating the impact of program changes on underrepresented populations and explain why it is a more robust measure than changes in the *percentage* of minority participants.

Categories and Subject Descriptors

K.3.2 [Computers and Education]: Computer and Information Science Education – *Computer science education*

General Terms

Management, Measurement, Human Factors.

Keywords

Gender diversity, women in computing science.

1. INTRODUCTION

Despite the awareness of the need to increase participation by women in computing, the National Center for Women and Information Technology reports that only 18% of 2010 Computer and Information Sciences undergraduate degree recipients were female [16]. The Computer Research Association reports that less than 12% of Bachelor's degrees in Computer Science were awarded to women at North American research universities in 2011 [22]. While there has been much qualitative analysis about what drives women's relationship with computing, there are few large-scale quantitative studies that offer actionable results.

In this paper, we examine a number of factors related to women's participation in Computer Science through a quantitative analysis of 7209 academic transcripts and 536 survey responses from students at Stanford University. We examine previously studied effects as well as present new findings.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference '10, Month 1–2, 2010, City, State, Country.

Copyright 2010 ACM 1-58113-000-0/00/0010 ...\$15.00.

As part of our study, we also examine the impact of a recent CS curriculum change at Stanford on gender dynamics in the major. We also present Fisher's Noncentral Hypergeometric Distribution [6] as an effective model for gauging the impact of a curriculum change on female participation and show that simply measuring growth in the percentage of women in a program can be a flawed indicator when program changes lead to a total increase in the number of participants.

Our study examines students at Stanford University, where the Computer Science department is housed within the School of Engineering. Students have until the end of their sophomore year to declare a major. Stanford has a set of introductory programming/systems courses numbered CS106A, CS106B and CS107. CS106A and CS106B correspond to CS1 and CS2, respectively, with the former being taught in Java and the latter in C++. Additionally, an accelerated course, CS106X, is offered as an alternative to the CS106A/B sequence for students with previous computing background. CS107 is the first systems course that CS majors are required to take. It is taught in C, and the main emphasis of the course is on understanding low-level topics (such as memory management and compilation) as opposed to the mechanics of programming.

While CS106A (and to some extent CS106B) are required for a variety of majors and are actually taken by a large percentage of the entire undergraduate population, CS107 is only required for students majoring in Computer Science or a few other highly-related majors (Symbolic Systems, Math and Computational Sciences, and the software concentration in Electrical Engineering). CS107 is commonly regarded by students as a "weeder" class, a critical juncture for students to decide whether they wish to continue on with a major in CS or not. As mentioned previously, there was a significant curriculum change made in the CS program during the 2008/09 school year that created a track (i.e., concentration area) structure in the major [18], and also resulted in a broad revision of the CS107 curriculum (although the curricula for CS106A/B/X remained unchanged). Thus, we give CS107 special consideration in our study to both understand its impact on CS declarations as well as see if the curriculum revision had any impact on participation by women.

2. RELATED WORK

The level of gender diversity in computing has far reaching consequences. On a purely economic level, the projected significant shortfall in computing graduates [5] could be better addressed by broadening participation by women in the field. More socially, the lack of women in computing enables sexism and perpetuates stereotype threat [3, 13, 19] by corroborating the

misconception that the small number of women in computing is indicative of a lack of belonging or ability. It has also been suggested that lack of gender diversity may potentially inhibit the diversity of ideas generated in the field [7]. As a result, understanding the factors that contribute to a woman’s choice whether or not to pursue computing has been studied from different perspectives for more than two decades [2, 8, 14, 20].

Two similar studies have been conducted at Stanford in the past decade. The first [11] analyzed women’s attitudes and participation in Stanford’s introductory curriculum in 2003. This study concluded that while CS107 did not seem to filter more women than men out of the CS pipeline, the women studied reported lower self-confidence and comfort with computers than their male counterparts. The study focused primarily on a qualitative analysis of women’s confidence in CS classes in the context of gendered communication and self-presentation. Although this study was performed before the curriculum change we analyze here, it provides a crucial context for our understanding of CS107 and the way such “gatekeeper” classes are experienced by each gender.

The second Stanford study [9] provided an ethnographic analysis of the importance of family influence and interactions with instructors as a means for encouraging women in CS, reinforcing the importance of mentoring previously posited as a crucial factor in helping retain women in computing majors [4]. The study also raised broader social issues, such as the degree to which women self-identified as “engineers” and the perceived relevance of one’s major toward future career paths— themes also studied qualitatively by Margolis, Fisher, and Miller [15] that we revisit more quantitatively in this study.

The finding that women self-report lower skills with computing technology upon entering college has been documented in other settings as well [1, 10, 17], and is related to fact that men generally have more experience than women in computing prior to entering college [1, 15, 17]. This observed difference between men and women in prior computing experience sets an important foundation for the study we conduct here. Indeed, we not only provide further quantitative validation of this phenomenon in our setting, but delve further into its implications with respect to gender differences in course performance, finding significant downstream impacts. Moreover, prior familiarity with computing also influences the point in a student’s academic career when CS courses are taken, and we find clear gender differences along these lines. This point ultimately becomes a crucial feature in the dynamic of whether students choose to pursue a major in Computer Science.

3. DATA GATHERING

To analyze gender differences related to academic performance as well as issues related to confidence, prior computing experience and views of computing, our study gathered two types of information: academic transcripts and survey data.

3.1 Academic Transcripts

We obtained partial academic transcripts for all students at Stanford from 1995 to 2012 who had taken either CS106B or CS106X. We chose to include all students who had taken CS106B/X (equivalent to CS2) as this population includes those students who had shown a level of interest in computing beyond CS106A (CS1). The anonymous partial transcripts include students’ grades for the courses CS106A/B/X and CS107 (if

taken) as well as the students’ major and major declaration date. This data was used to analyze gender differences in the choice of major and CS course performance. This data set includes transcripts for 7209 students, 4281 of whom also took CS107.

3.2 Surveys

To gather information related to confidence, prior computing experience, and views of computing, we distributed an online survey to all students who had taken CS106A/B/X or CS105 (Stanford’s CS0 course) during the 2011-12 year, as well as to all current Computer Science and Symbolic Systems majors. By surveying students in CS105 (i.e., CS0) as well as Symbolic Systems (a non-CS, but computing-related major in the school of Humanities and Sciences) we could get a broader sampling of students across majors and programs. Survey questions included both free response questions (which were independently coded by the first two authors until consistency was achieved) and Likert-scale rating questions on a range of issues addressing self-perception and views of one’s major. Students were incentivized to complete the anonymous survey by being entered into a random drawing for one of eight \$50 gift cards from Amazon.com. The survey was distributed to approximately 2500 students and we received 536 responses.

4. A NEW STATISTICAL MODEL

As we report below, for many comparisons of gender differences with respect to some variable or attribute, classical statistical models (e.g., t-tests, Pearson correlation, Cramer’s V measure of association, etc.) are useful and appropriate, given the quantity of data available. However, when reporting increases in participation by women in computing, the traditionally used measure of the *percentage* of women in a particular population (e.g., in a specific class, majoring in CS, etc.) can be (surprisingly) problematic when the overall size of that population is also in flux. For example, consider a college that has (as many do) a total population that is roughly 50% women, but a population of CS majors which is only 18% women. As the number of CS majors as a whole grows (hypothetically approaching the population of the whole college), the *percentage* of women majoring in CS must also grow to approach the total college percentage of women (50%). Thus, it would appear that CS was becoming more attractive to women (relative to men), even if the true dynamic was simply that CS had become more attractive to *both* genders. The increase in the percentage of women is simply a statistical artifact of growth in the CS major population, not an indication that the major has actually become more attractive to women (relative to men) than it was before.

To provide a more robust statistical model that better measures the likelihood of women choosing to pursue a major in CS relative to men, we suggest the use of Fisher’s Noncentral Hypergeometric (FNCH) distribution [6]. FNCH is a generalization of the Hypergeometric distribution where sampling probabilities (of black and white balls in an urn) are *unequally* weighted. The probability mass function of the FNCH distribution is given by:

$$P(X = i) = \frac{\binom{m_1}{i} \binom{m_2}{n-i} w^i}{\sum_{j=x_{\min}}^{x_{\max}} \binom{m_1}{j} \binom{m_2}{n-j} w^j},$$

where m_1 = number of white balls in the urn, m_2 = number of black balls in the urn, n = number of balls drawn (simultaneously) from the urn, $x_{\min} = \max(0, n - m_2)$, $x_{\max} = \min(n, m_1)$, and w = relative weight of drawing a white ball as opposed to a black ball. The variable X denotes the number of white balls drawn from the urn (after n draws).

To explain the analogy with gender composition in CS, consider an urn (college) which contains a particular number of black balls (men) and white balls (women). We then choose as many balls (simultaneously) from the urn as the number of CS majors, where the color of the chosen balls reflects the gender composition in CS. If the numbers of black and white balls in the urn were the same to begin with and either color was *equally* likely to be drawn ($w = 1$), then our sample representing CS majors would have a maximum likelihood outcome of containing the same number of black and white balls (men and women). However, if the white balls were to be *weighted* so as to be less likely to be drawn than black balls, then our sample would likely contain a higher proportion of black balls (men), as we see in real-world CS enrollments. Note that if there are the same number of black and white balls in the urn, and all the balls are drawn, we would still produce the a 50/50 outcome regardless of the weighting used.

To reiterate, the weight of the white balls in the FNCH model reflects the likelihood of a woman choosing to major in CS relative to a man. Given the other parameters (m_1, m_2, n) we can obtain a maximum likelihood estimate for w using numerical optimization, allowing us to measure the weighting factor in different populations. By focusing on this weighting factor (instead of percentages) we are able to more accurately measure the impact of changes aimed at making CS more attractive to women even in the face of changes in overall enrollment levels in CS. The underlying FNCH model dynamics are not distorted by the size of the sample of balls taken from the urn in the same way that a simple percentage measurement would be (as it would be forced to approach the population mean). This is especially important in accurately comparing women's participation in computing over time as overall enrollment levels fluctuate, which has certainly been the case in recent years.

5. RESULTS

5.1 Role Models and Social Factors

One of the main goals of the study was to provide quantitative evidence to support or refute issues that are often mentioned in relation to gender and computer science. In this vein, a frequent hypothesis is that women's decision to pursue computing is affected by the lack of female professors and role models [4, 9]. While we did not investigate the effect of industry role models, we found there was no significant impact of a professors' gender on women's propensity to take a class with him/her (t-test based on 2885 classes taught by men and 347 classes taught by women, $p = 0.2$). While this result is impacted by the fact that some courses are required rather than elective (and may only be taught by faculty of one gender), it still provides evidence that female students did not seem to seek out courses with female instructors.

On the other hand, we found that having parental support, especially maternal support, was a greater influence for women in computing than their male counterparts. The Cramer's V association between maternal/paternal support and sureness of one's career aspirations was measured stratifying by gender (male/female) and major (CS/non-CS). The results are below.

Association with sureness of career	Female CS	Male CS	Female non-CS	Male non-CS
Maternal support	0.26	0.18	0.18	0.15
Paternal support	0.24	0.15	0.18	0.16

Similarly, we found parental support (either maternal or paternal) to be highly associated with women calling themselves "hardcore" about CS. And while we found a similar high associate among men, the correlation is more pronounced for women:

Association with being "hardcore"	Female CS	Male CS
Maternal support	0.33	0.30
Paternal support	0.25	0.21

These results show that that parental support has a significant impact on female students' attitudes toward their academic and career paths, especially with regard to computing.

Another social factor that has been posited for why women may not pursue CS is the solitary nature of computing and lack of interaction with others. We found conflicting evidence regarding this claim. On a free-response survey question asking students for potential "cons" of majoring in CS, 14.7% of female CS majors (N=34) and 15.5% of female non-computing related majors (N=71) listed the solitariness of CS as a "con." However, only 11.9% of male CS majors (N=84) and 13% of male non-computing related majors (N=83) did so. Interestingly, women, regardless of whether they major in CS or not, seem to find CS a more solitary discipline (in a negative sense) than their male counterparts.

Paradoxically, based on transcript data we did not find evidence that women were more likely to take courses involving group work as a potential means to avoid a solitary working environment. Fitting weight parameters in the FNCH model across classes involving group projects (N=444) and those involving individual work (N=5680), we did not find that women were statistically any more likely to take CS classes involving group projects than CS classes involving only individual work ($p = 0.3$). From this finding we posit that simply developing a CS curriculum including more group projects may not necessarily help address women's view of solitary working conditions in the field unless there is a commensurate compelling reason for women to take such courses.

Another common belief regarding women in computing is that feeling like a "minority" in the field may deter women from considering or continuing on in CS. In our survey data, 84% of female CS majors (N=50) self-reported feeling like a gender minority in a free-response question asking if they identified as any form of minority in their major. This number was 52% for women survey respondents overall (N=229). However, feeling like a minority did not appear to be correlated with students' self-reported grades in CS106A or self-reported confidence asking questions of a CS professor. Since the survey and transcript data are both anonymous, we unfortunately cannot cross-correlate answers across them to validate actual grades in CS106A or grades in other CS courses. Nevertheless, we note that the high proportion of women who feel like a minority in CS creates greater potential for stereotype threat, which has been observed in other settings [13].

5.2 Confidence and Enjoyment

Next, we analyzed issues related to women's confidence with and enjoyment of factors related to computing. First, we examined women's self-reported confidence in their mathematical abilities. Female CS majors' (N=51) rating of confidence in their math abilities was statistically indistinguishable from that of their female non-CS (N=180) counterparts ($p = 0.7$). Thus, confidence in one's math abilities did not appear to be an important factor in women's choice to pursue CS. Nevertheless, in the overall population, men's (N=299) self-reported confidence in their math abilities was higher than women's (N=231) at a highly statistically significant level ($p < 0.001$).

Perhaps more importantly, we also observed a statistically significant ($p < 0.001$) gender discrepancy with respect to confidence asking questions in CS classes, as males (N=294) rated themselves more confident asking questions than female (N=226) respondents did. Such gender differences are important for instructors to be aware of in considering classroom dynamics.

Moving from confidence to enjoyment, we wanted to better understand how factors related to the enjoyment of CS might reveal gender differences. Looking specifically at the introductory programming course CS106A, survey results showed that men (N=204) self-reported enjoying this course significantly more than women (N=164) overall ($p = 0.02$). Restricting the data to only CS majors, this gender difference is no longer statistically significant ($p = 0.2$). This provides quantitative evidence for the unsurprising conclusion that the enjoyment of CS106A is an important factor in choosing to continue on in computing, but that the level of enjoyment is not gender balanced. It is important to consider how such a gender discrepancy can be decreased when designing introductory CS courses as it indeed has significant downstream impact (i.e., choice of major).

One possible cause for the discrepancy in enjoyment of CS106A is that we also found a gender difference in the enjoyment of problem solving, a more general factor we hypothesized would be strongly related to computing. Indeed, male survey respondents (N=299) reported enjoying problem solving more than females (N=230) by a wide margin ($p < 0.001$). And enjoyment of problem solving was indeed correlated with enjoying CS106A (Cramer's V value = 0.291). Digging deeper, we found that the enjoyment of problem solving is also linked to the likelihood of majoring in CS, and we found significantly ($p = 0.01$) higher enjoyment of problem solving among women who major in CS (N=51) versus those who do not (N=179).

5.3 Prior Experience and Grades

Corroborating previous findings [1, 15, 17], we found higher rates of self-reported CS experience prior to college for males (N=298) than females (N=232):

Prior CS experience	All students	CS majors	Non-CS majors
Female	42.4%	45.1%	41.7%
Male	66.3%	75.0%	59.0%

Indeed, the difference in prior CS experience between all men and women was highly statistically significant ($p < 0.001$). So was the difference between prior CS experience for CS major men and women ($p < 0.001$) and for non-CS major men and women ($p < 0.002$). Of note, while some difference in the level of prior CS

experience does exist between female CS majors and non-CS majors, this difference is much less pronounced than among men. This would seem to indicate that women's prior experiences with CS before coming to college were not as compelling a driver of their collegiate major choices as they were for men.

We hypothesized that having prior CS experience would improve students' performance in the introductory programming class (CS106A) and may also be one of the contributing factors to the gender-based confidence gap in the course, discussed in the previous section. Indeed, we found a relatively high (Cramer's V value = 0.293) correlation between having prior CS experience and students' grades in CS106A. Indeed, comparing the grades for students in the class with and without prior CS experience revealed a clear statistically significant difference ($p < 0.001$).

This finding led to a larger-scale comparison of course performance between men and women using transcript data. We examined the mean course grades (GPA) for men and women in our introductory series of programming/systems courses (CS106A, CS106B, CS106X, and CS107) from 1995 to 2012. We found statistically significant differences in grades by gender in every course examined (N is total number of students in all offerings of the course over the period examined):

Course grades	CS106A		CS106B		CS106X		CS107	
	N	GPA	N	GPA	N	GPA	N	GPA
Female	1367	3.63	1330	3.30	269	3.22	870	3.22
Male	3467	3.68	3590	3.41	1538	3.51	3408	3.33
p-value	0.005		< 0.001		< 0.001		< 0.001	

Interestingly, our results agree with smaller-scale studies conducted at other institutions [21], but contradict the conclusions of a previous study conducted by Irani a decade prior at our own institution [12]. We believe the difference stems from the fact that the previous study was based on only one year's worth of data. While it found that men did receive slightly higher grades than women in CS107, it did not include a large enough sample to detect statistical significance in this difference.

It is also important to point out that while we detect a clear difference in grades for men and women in these courses, we do not have a clear explanation as to *why*. Indeed, many factors shown to have gender differences, such as experience with CS before college, enjoyment of problem solving, or confidence asking questions may all be contributing factors to the difference in grades. We believe that this is a rich area for further study.

We did want to assess the potential impact of grades on whether women choose to continue on in CS beyond the introductory courses. Since CS107 is informally considered a "weeder" class by students, we examined women's grade differences between CS106B and CS107 (using transcript data) to see if students who chose not to major in CS after taking CS107 experienced a more significant drop in grade from CS106B than students who did major in CS. Interestingly, we found no significant difference in grade drop between women who became CS majors and those who did not ($p = 0.6$). We also found no significant difference in grade drop between women and men ($p = 0.6$). These results suggest that performance differences from the introductory classes do not affect women's choice of major after CS107 any more than men's choice, in alignment with the previous work of Irani that came to a similar conclusion using a different analysis.

5.4 Dynamics of Choosing a Major

Examining the theme of choice of major, our survey asked students how sure they were of their major upon entering college. We found that among CS majors, men ($N=143$) were more sure of what their major would be upon entering college than were women ($N=51$) at a statistically significant level ($p = 0.001$). This difference is likely related to the differing rates of experience with CS prior to college between men and women, further punctuating the importance of early CS exposure to attract women to the field.

Unfortunately, we found women (perhaps exacerbated by less CS familiarity prior to college) typically take their first college CS class later than men—a pattern seen in both survey responses and transcript data. Figure 1 shows the quarter/year in which men and women took CS106A (in the survey data, $N_{\text{men}}=185$, $N_{\text{women}}=149$).

The idea of encouraging women to take a CS class early in their college years takes on even greater importance in light of the fact that 25.4% of female non-CS majors taking the survey ($N=71$) reported that they had started taking CS courses too late in their academic career, a factor that was cited by only 8.8% of female respondents who did major in CS ($N=34$). Even more strikingly, 61% of female survey respondents said they would have considered a CS major more strongly if they had taken CS106A earlier. Indeed, based on these findings, we believe that encouraging women to take a CS course as early as possible during their undergraduate careers is one of the most critical factors in promoting the number of women CS majors.

We also asked students to rate how relevant they believed their choice of major was to their future careers. Overall, men ($N=299$) viewed their major as significantly more relevant to their future career ($p = 0.05$) than women ($N=231$). Restricting to just CS majors data ($N_{\text{men}}=144$, $N_{\text{women}}=51$), this pattern was still present, though with a lesser degree of statistical significance ($p = 0.1$). This finding aligns with previous qualitative work [14] indicating that women look more at the broader social impacts of computing rather than focusing primarily on the technology itself, reflecting that the technology they learn as a result of their major is just a facet of what they will do in their careers after graduation.

5.5 Curriculum Changes

Finally, we want to understand the gender impact of the recent curriculum revision made in our CS program. While locally, we are interested in whether the curriculum change resulted in more gender diversity, the broader research question is one of how to robustly evaluate the impact of curricular changes on gender diversity, especially in light of overall fluctuating enrollments. Hence, we identified the FNCH model, and used it extensively in evaluating our curriculum revision.

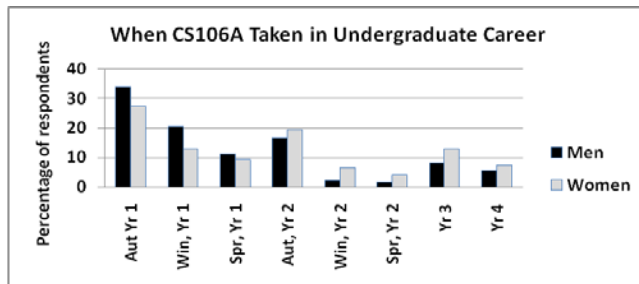


Figure 1. Percentage of survey respondents who took CS106A during various quarters/years of their undergraduate career. Aut, Win, and Spr refer to Autumn, Winter and Spring quarters of the year (1 = Freshman, 4 = Senior).

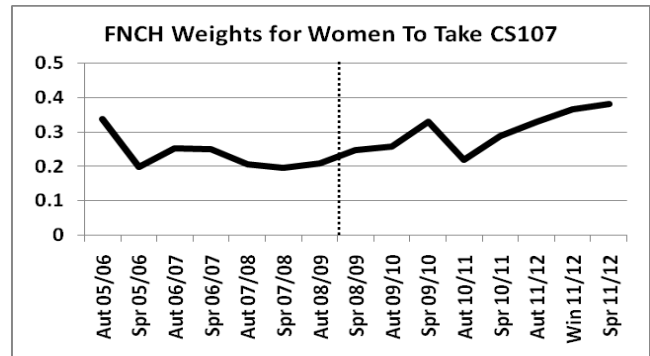


Figure 2. FNCH weights for women to take CS107. The dashed line delineates when the course content was revised.

We began by focusing on CS107, since—as part of the curriculum revision—the content of CS107 was also significantly revised. We examined the relative propensity for women (vs. men) from the entire campus population to take CS107 based on weight estimation in the FNCH model. We looked at a symmetric (7 year total) period before and after the curriculum revision. The results are given in Figure 2.

We find that the average weight after the curriculum revision ($w = 0.30$) is notably higher than the average weight before ($w = 0.23$). Using a (non-parametric) Mann-Whitney test, we find this difference in weights to be statistically significant ($p = 0.02$), indicating a higher propensity for women to take CS107 after the curriculum revision.

To be precise, we noted one possible confounding factor of this analysis, which is that all offerings of CS107 examined before the curriculum revision were taught by the same (male) instructor (Prof. A) and all but one offering of the CS107 after the curriculum revision were taught by the same (female) instructor (Prof. B). Thus, we wanted to see if the difference observed above was potentially due to the instructor as opposed to the course content. Luckily, we found that these two instructors had both taught several offerings of CS106B and CS106X in the past (during which time the content for those courses remained stable), so we examined those classes to see if they exhibited a similar difference in mean FNCH weights across instructors:

FNCH weights	CS106B	CS106X
Prof. A	0.322	0.209
Prof. B	0.341	0.202

Here we found no statistically significant difference between Prof. A and Prof. B in the FNCH weights for either of the other courses they teach, leading us to believe that the difference observed with respect to CS107 is not due to the instructor, but rather due to the revision in the course content.

Moving from CS107 to majoring in CS, we considered the same 7 year population of CS107 students and calculated the weights in the FNCH model based on the number of men and women who eventually became CS majors. We found that prior to the curriculum revision, the mean relative weight for women vs. men to major in CS was $w = 0.66$, and after the curriculum revision it was 0.80, an encouraging result showing that women are more likely to major in CS after completing the revised CS107 course.

Our final analysis involved looking at the revised curriculum as a whole to determine if it has positive impact on female enrollment.

Again, the FNCH model was employed using CS enrollment data over the past 14 years. We grouped the data by students' year of graduation. Thus most students who were juniors (and had likely declared their major previously) when the curriculum revision went into effect would be in the graduating class of 2009/10. Students in later graduating classes would likely have declared their major after the new curriculum was in effect, so we consider the class of 2009/10 an approximate delineating point between the old and new curriculum. (Note that students who have declared their major, but not yet graduated are projected to be on a 4-year program, common at Stanford. So, students who were freshman in 2010/11 would be reported in the (projected) graduating class of 2013/14.) The weights estimated using the FNCH model are reported in Figure 3. A Mann-Whitney test finds the difference in weights between the old and new curriculum to be statistically significant ($p < 0.05$), indicating a higher propensity for women to major in CS after the curriculum revision.

6. CONCLUSIONS AND FUTURE WORK

The results reported here provide large-scale quantitative evidence for several gender-related issues in computing, both in relation to previously observed phenomena as well as newly discovered ones. We highlight three of our more significant conclusions.

- At the college level, we found that one of the biggest obstacles between women and a CS major was starting CS classes too late in their undergraduate education. We believe that providing strong encouragement for women to take a CS course early in their academic career is a key factor to promoting greater gender diversity in the field.
- Corroborating previous studies, we believe CS educators should be aware of the differences in confidence and previous background in computing that men and women bring with them into courses. These differences can impact interactions with instructors (such as asking questions or seeking help) and potentially overall performance in courses. Providing strong encouragement and mentorship can potentially help to address some of these issues.
- Curricular revision has the potential to have a real impact on gender diversity, as we have witnessed with our own curriculum change. Trying to identify the aspects of our new curriculum that appear most responsible for greater female participation is part of our future work. We conjecture that the program's track structure provides more options and a broader context for impactful work in computing.

Of course, much work still remains to be done, including building better structures for mentorship and community to help increase women's confidence and comfort in CS classes. Our study also highlights the importance of gender dynamics in computing before students even start college, especially the need to expose more women to computing early and engender parental support for their pursuits. It may also be instructive to examine in more detail the factors that inhibit women from taking CS classes earlier in the academic careers. We are hopeful that with continued study we may make further progress toward reaching equitable gender representation in computing.

7. ACKNOWLEDGMENTS

We thank Paddy McGowan and Claire Stager for providing data that was essential to our study. This work was supported by a gift from the Holtzschue/Schloss family.

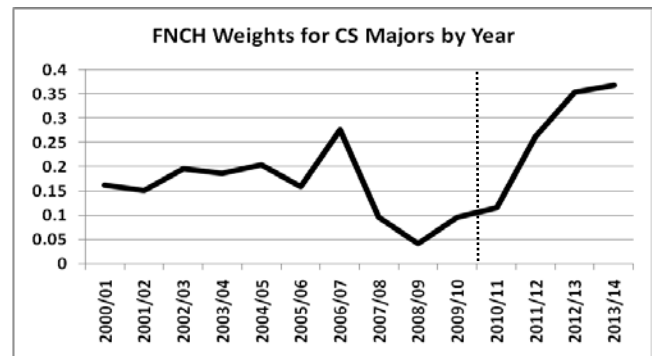


Figure 3. The FNCH weights for CS majors, grouped by year of graduating. Graduating class beyond 2011/12 are projected based on a 4-year program. The dashed line indicates the approximate change point from the old to the new curriculum.

8. REFERENCES

- [1] Beyer, S., Rynes, K., Perrault, J., Hay, K., and Haller, S. Gender differences in computer science students. In *Proc. of SIGCSE '03*.
- [2] Camp, T. The Incredible Shrinking Pipeline, *Communications of the ACM (CACM)*, Volume 40 Issue 10, Oct. 1997.
- [3] Cheryan, S., Plaut, V. C., Davies, P., and Steele, C. M. Ambient belonging: How stereotypical environments impact gender participation in computer science. *Journal of Personality and Social Psychology*, 97, 1045-1060.
- [4] Cohoon, J.M. Recruiting and retaining women in undergraduate computing majors. *SIGCSE Bull.*, v.34 n.2, June 2002.
- [5] Computer Science Teachers' Association, United States' Annual STEM Job Openings vs College Graduates through 2018, *Voice Newsletter*, v.6, n.5, 2010.
- [6] Fisher's Noncentral Hypergeometric Distribution. *Wikipedia*. Wikimedia Foundation, 20 July 2012. http://en.wikipedia.org/wiki/Fisher's_noncentral_hypergeometric_distribution.
- [7] Gendered Innovations in Science, Health & Medicine, and Engineering. <http://genderedinnovations.stanford.edu/>
- [8] Güter, D., and Camp, T. An ACM-W Literature Review on Women in Computing, *SIGCSE Bull.*, v.34 n.2, June 2002.
- [9] Heddleston, K. Women in Computer Science: An Ethnographic Study of Women in Computer Science at Stanford University. Undergraduate Honors Thesis. Stanford University, 2010.
- [10] Hoffman, M.E. and Vance, D.R. Gender difference trends in computer literacy of first-year students, *SIGCSE Bull.*, v39, n1. 2007
- [11] Irani, L. A Different Voice: Women Exploring Stanford Computer Science. Undergraduate Honors Thesis. Stanford University, 2003.
- [12] Irani, L. Understanding gender and confidence in CS course culture. In *Proceedings of SIGCSE '04*, 195-199.
- [13] Kumar, A.N. A Study of Stereotype Threat in Computer Science. In *Proceedings of ITiCSE '12*, 273-278.
- [14] Margolis, J. and Fisher, A. *Unlocking the Clubhouse: Women in Computing*. MIT Press, Cambridge, MA. 2002.
- [15] Margolis, J., Fisher, A., and Miller, F. The Anatomy of Interest: Women in Undergraduate Computer Science. *Women's Studies Quarterly*, 28(1/2).
- [16] National Center for Women & IT, By the Numbers, 2012. www.ncwit.org/sites/default/files/legacy/pdf/BytheNumbers09.pdf
- [17] Ogan, C., Robinson, J.C., Ahuja, M., and Herring, S.C. Gender differences among students in computer science and applied information technology. In W. Aspray & J. McGrath Cohoon (Eds.), *Women and Information Technology: Research on the Reasons for Under-Representation*, MIT Press, 2006.
- [18] Sahami, M., Aiken, A., and Zelenski, J. Expanding the frontiers of computer science: designing a curriculum to reflect a diverse field. *Proc. of SIGCSE '10*.
- [19] Spencer, S.J., Steele, C.M., and Quinn, D.M. Stereotype threat and women's math performance, *Journal of Experimental Social Psychology*, 35, 1999, 4-28.
- [20] Spertus, E., *Why Are There So Few Female Computer Scientists?*, MIT Artificial Intelligence Laboratory Technical Report 1315, 1991.
- [21] Vilner, T., and Zur, E. Once she makes it, she is there: gender differences in computer science study. In *Proc. of ITiCSE'06*.
- [22] Zweben, S., and Bizot, B. 2010-2011 Taulbee Survey. *Computing Research News*, May 2012.