

5-2011

# Gender and self-selection among engineering students

Maci Dickson

*University of Arkansas, Fayetteville*

Follow this and additional works at: <http://scholarworks.uark.edu/ineguht>

---

## Recommended Citation

Dickson, Maci, "Gender and self-selection among engineering students" (2011). *Industrial Engineering Undergraduate Honors Theses*. 13.  
<http://scholarworks.uark.edu/ineguht/13>

This Thesis is brought to you for free and open access by the Industrial Engineering at ScholarWorks@UARK. It has been accepted for inclusion in Industrial Engineering Undergraduate Honors Theses by an authorized administrator of ScholarWorks@UARK. For more information, please contact [scholar@uark.edu](mailto:scholar@uark.edu), [ccmiddle@uark.edu](mailto:ccmiddle@uark.edu).

# **Gender and self-selection among Engineering Students**

An undergraduate honors thesis submitted to the  
Department of Industrial Engineering  
University of Arkansas

By

**Maci Lynn Dickson**

**4-18-11**

Mentor: Dr. Justin Chimka

Reader: Dr. Heather Nacthmann

**Key words:** attrition, demographics, engineering, gender, graduation, logistic regression, persistence, retention, and socioeconomics.

# Gender and self-selection among Engineering Students

MACI DICKSON and JUSTIN R CHIMKA\*

Department of Industrial Engineering, University of Arkansas, USA

**Gender and graduation rates of first time engineering college students have been analyzed as a function of academic and demographic variables in order to investigate the hypothesis that an advantage to women with respect to student success might be attributed to their socioeconomic advantages as a student population. The authors present descriptive, graphical, and model-based evidence to support their ideas about gender and self-selection driven by socioeconomic phenomena that leave a disproportionate number of women out of higher education, but create a group of female students more likely than their male counterparts to succeed.**

**Key words:** attrition, demographics, engineering, gender, graduation, logistic regression, persistence, retention, and socioeconomics.

## INTRODUCTION

An increased graduation rate is occurring for females in colleges and universities across the United States. Over the years not only have females caught up to males in graduation rates; some studies show that more females than males are graduating in the present age. The U.S. Department of Education stated that as of 2004 females were

---

\* Corresponding author. E-mail: [jchimka@uark.edu](mailto:jchimka@uark.edu)

awarded 58 percent of all bachelor's degrees in the United States (Buchmann, 2006). Worldwide, more women than men recently graduated in 75 of 98 countries (Church, 2009). The topic of gender equality is popular in modern research. Female graduation rates are climbing, and survival analysis of the same data described in this paper showed, "the significance of standardized math test scores, gender and science ACT scores in explaining variation in student graduation under different conditions (Chimka et al., 2007-2008)."

Many studies have been conducted involving gender differences in education. A main concern is that females are achieving more diplomas than men in the United States. In 2005, females accounted for 61% of graduates and 59% of students (OECD, 2008). Females are even starting to achieve gender parity in performance on standardized math tests (Wente, 2008), but at least one study has showed, "males are more likely than female students to be retained, but less likely to graduate (Wohlgemuth et al., 2006-2007, p. 471)." One interesting way to explain gender differences has been to examine a student's earlier years. A study on public education states, "73% males and 27% females receive discipline referrals in elementary school (Clark et al., 2008, p. 121)." Students who are eager to learn as much as possible are generally not the students getting discipline referrals. Discipline referrals also have consequences such as detention and suspension that inhibit the ability to be present in the classroom for new material.

While females are having great success with a rate of degrees earned, women are still underrepresented in technical fields such as science, math and engineering. There are some popular explanations for this based on differences in topic and content of the

majors. One study on degree performance actually claimed, “on the whole females are better at verbal topics and males at mathematical topics, but even among mathematical topics there are some where females excel (Kornbrot, 1987, p. 516).”

An alternate description is the social impact on females working in technical fields. Women’s rights have come a long way, but there are still some negative views towards females in certain occupations. Hartman and Hartman (2008, p. 253) showed, “relative to men, women in science and engineering programs face difficulties of a largely psycho-cultural nature.” The glass ceiling still exists in some regions and can discourage females from applying themselves in technical fields. A social mindset that females belong in less technical fields can influence female children’s encouragement and effort in math and science even at the elementary school level.

We believe the gender gap between achievements by college students might be explained by socio-economic variables. The difference in graduation rates can be illustrated by a theory of more varied backgrounds in the males who are attending college. Imagine four would-be students: One male-female pair from a higher socioeconomic background, and another male-female pair from a lower socioeconomic background. If three out of the four go to college or university, then it would seem the male and female from the higher socioeconomic background and just the male from the lower socioeconomic background will be college bound.

## **MATERIALS AND METHODS**

The data featured in survival analysis by Chimka et al. (2007-2008), Chimka and Lowe (2008), and Chimka and Wang (2009) also used here recorded University of Oklahoma

(OU) students who started their freshman year in the college of engineering and were admitted directly from high school in 1995. The dataset followed the students for six and one half years, or until graduation or drop out. The data had some constant variables and other time varying covariates that changed throughout the years for the students. The time periods were divided into semester periods with 0 being the first semester at OU. In order to perform binary logistic regression on the data they were collapsed into one row per student; time varying covariates were deleted, and only the constant variables remained. Data were added from the U.S. Census Bureau, and Zillow.com, a free online real estate site accessed in January 2010. The new socioeconomic values correspond with the student's high school zone improvement plan (ZIP) code and contain demographic information about where they came from before attending OU.

We used Minitab software to estimate logistic regression models of gender and graduation. Logistic regression was chosen for the fact that gender and graduation (the responses) are described with a 0-1 (binary) variable. The independent variables used to model gender and graduation divided into two categories: academic, and demographic or socioeconomic. The academic variables used were based on Scholastic Aptitude Test (SAT) and American College Test (ACT) scores. There are two binary variables, SAT and ACT that are 1 if the student took the corresponding exam, and 0 if the student did not. The other variables are actual scores. Since each student did not take both exams a conversion scale was used to transform an ACT score to a SAT score, and vice versa. With the conversions each student had a value for ACT and SAT scores regardless of what test they actually took.

To mitigate problems of inter-correlation among independent variables we computed variance inflation factors (VIF) for each of the independent variables. VIF is equal to  $1 / (1 - R^2)$ , where  $R^2$  is the coefficient of determination associated with a relevant multiple linear regression model. A common cutoff for VIF is ten (Montgomery and Runger, 2011), so we dropped from consideration those independent variables with VIF greater than ten. Here is the final independent variable set:

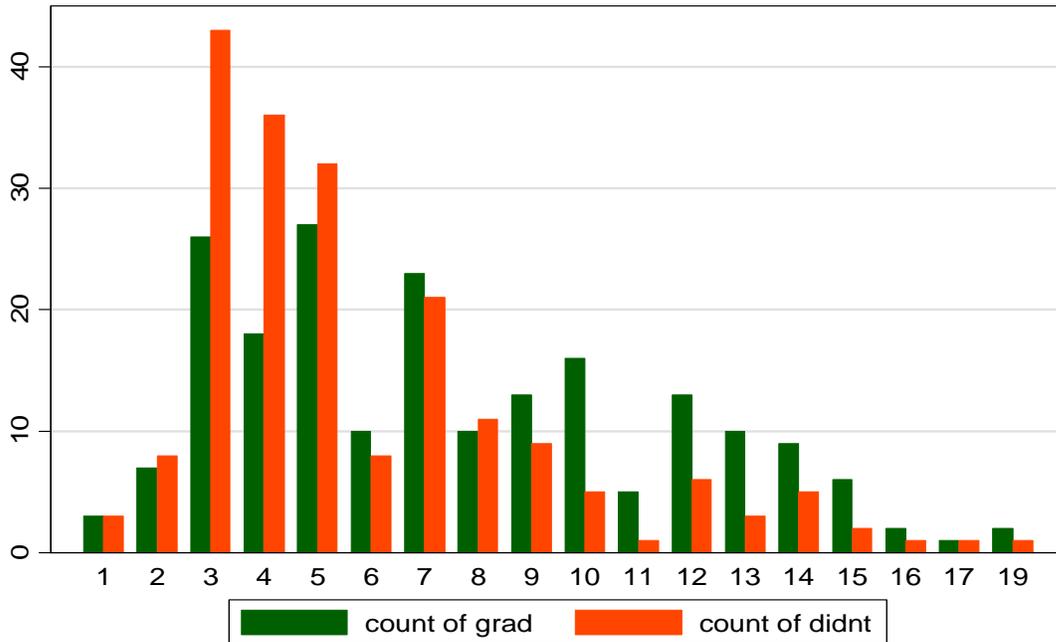
- English ACT score
- English SAT score
- Distance between high school ZIP code and OU
- Gender, home value index of high school ZIP code
- Math ACT score
- Number of Hispanic persons in high school ZIP code
- Number of Native Americans and Alaskan Natives in high school ZIP code
- Number of Pacific Islanders and Native Hawaiians in high school ZIP code
- Number in labor force of high school ZIP code
- Occupied housing units for seasonal or recreational use in high school ZIP code
- Vacant housing units in high school ZIP code
- Whether or not the student is from Oklahoma
- Whether or not the student submitted ACT scores
- Whether or not the student submitted SAT scores

## DESCRIPTIVE RESULTS

The first results we present are descriptive statistics, and graphical representations of the data. They do show a slight difference in the graduation rates of males (50.65%) and females (51.76%) not statistically significant in part perhaps because the sample size for females ( $n = 86$ ) seems absolutely small and is certainly smaller than the sample size for males ( $n = 312$ ). Because engineering is historically a male dominated field we are motivated to explain the difference between graduation rates anyway. Of particular interest is to contrast prevalence of adults with a high school degree in the student's high school ZIP code. Mean and standard deviation of females (mean = 13,096.85, standard deviation = 7371.55) versus males (mean = 12,620.84, standard deviation = 7636.07) suggest that female engineering students are a more advantaged and homogenous group with respect to family and community support for its studies. This would be consistent with our original hypothesis.

Graphical evidence in support of the notion that socioeconomic indicators are important to graduation in general is found in Figures 1 and 2. The histograms in Figure 1 describe percent with a college degree in high school ZIP code for students that did (not) graduate. For the students that did not graduate there appears to be a more dramatically skewed distribution indicating lower education levels in the associated high school ZIP code. This measure among students that do graduate appears to be bimodal and represent a more highly educated high school ZIP code. This graduation result consistent with the descriptive statistics reported above with respect to gender should

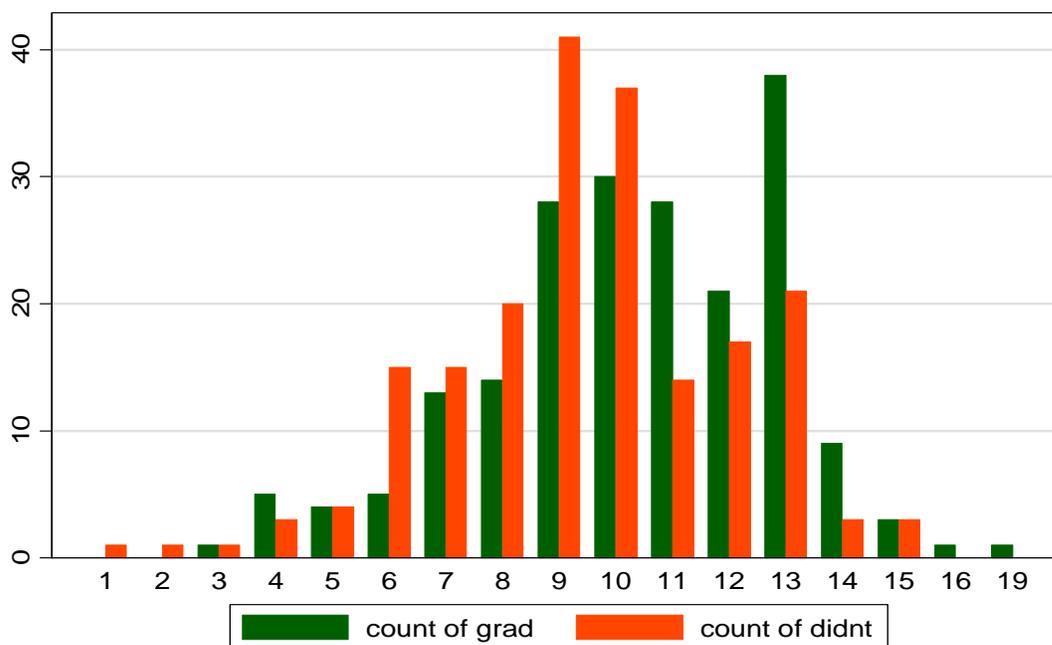
lend even greater support to our hypothesis about gender and self-selection among engineering students. That is females are a more advantaged and homogenous group with respect to socioeconomics that contribute to propensity for academic success.



**Figure 1.** Distribution of percent with a college degree from the high school ZIP code of students that did (not) graduate

Similar histograms of labor force in Figure 2 show two seemingly bimodal groups from the high school ZIP codes of students that did (not) graduate: one more symmetric indicating lesser participation in the labor force, and one more skewed indicating greater participation in the labor force. Among students that did not graduate the more symmetric distribution appears to dominate, where the skewed appears to dominate description of the graduates. This would suggest that engineering graduates are coming from high school ZIP codes with greater employment, another socioeconomic

advantage. If we associate higher education with greater employability, then again we have evidence to support the existence of a more homogenous, better-supported group of female students compared to their male classmates.



**Figure 2.** Distribution of number in labor force from the high school ZIP code of students that did (not) graduate

## DISCUSSION

Significant ( $\alpha = 0.010$ ), theoretically appropriate (logistic) regression models of gender and graduation do not suggest clear alternatives to the hypotheses of gender and self-selection. In a model of graduation as a function of academic-only variables (controlling for gender) the lone significant predictor is math ACT score that indicates the greater the score is the greater is the probability of graduation ( $p = 0.002$ ). In a model of gender as a function of all academic-only variables the lone significant predictor is again MACT,

but the relationship with probability of being female is inverse ( $p = 0.000$ ). This indicates that if female students are succeeding in greater proportion, they are doing so without the benefit of better academic preparation with respect to mathematics. Furthermore in a model of graduation as a function of demographic and socioeconomic variables (controlling for gender and academics) significant predictors include three digit ZIP code total population associated with the student ( $p = 0.024$ ), where the lesser the total population is the greater is the probability of graduation. Once again, in the same model of gender, what is important to graduation is deficient among female students: In a model of gender as a function of demographics and socioeconomic variables (controlling for academics) significant predictors include three digit ZIP code total population associated with the student ( $p = 0.070$ ), but the relationship with probability of being female is inverse, further indication that if female students are unusually successful, it is not because they have some benefit associated with the relatively generic demographic variable about total population. Instead we have observed some evidence of a more homogenous and socioeconomically advantaged female student population which might explain a slight gender gap between achievements by engineering college students.

## **REFERENCES**

Aitken ND (1982). College student performance, satisfaction and retention: Specification and estimation of a structural model. *Journal of Higher Education* 53(1): 32-50.

Buchmann C (2006). The growing female advantage in college completion: The role of family background and academic achievement. *American Sociological Review* 71: 515-541.

Chimka JR, Lowe LH (2008). Interaction and survival analysis of graduation data. *Education Research and Reviews* 3(1): 29-31.

Chimka JR, Reed-Rhoads T, Barker K (2007-2008). Proportional hazards models of graduation. *Journal of College Student Retention* 9 (2): 221-232.

Chimka JR, Wang Q (2009). Accelerated failure time models of graduation. *Educational Research and Reviews* 4(5): 267-271.

Church E (2009). Who's in the know: Women surge, men sink in education's gender gap. *The Globe and Mail* (December 7, 2009).

Clark MA, Lee SM, Goodman W, Yacco S (2008). Examining male underachievement in public education: Action research at a district level. *National Association of Secondary School Principals Bulletin* 92 (2): 111-1132.

Hartman H, Hartman M (2008). How undergraduate engineering students perceive women's (and men's) problems in science, math and engineering. *Sex Roles* 58: 251-265.

Kornbrot DE (1987). Degree performance as a function of discipline studies, parental occupation and gender. *Higher Education* 16(5): 513-534.

Montgomery DC, Runger GC (2011). *Applied Statistics and Probability for Engineers*. John Wiley and Sons, Inc., Hoboken, NJ.

Organisation for Economic Co-operation and Development (OECD) (2008), Higher Education to 2030 Volume 1: Demography.

Wente, M (2008). The gender politics of math scores. The Globe and Mail (July 31, 2008).

Wohlgemuth D, Whalen D, Sullivan J, Nading C, Shelley M, Wang YR (2006-2007). Financial, academic and environmental influences on the retention and graduation of students. *Journal of College Student Retention* 8(4): 457-475.