

Model-based Analysis of Diversity in Higher Education

Maryam Alsadat Andalib

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State
University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Industrial and Systems Engineering

Navid Ghaffarzadegan, Chair

Joshua D. Hawley

Niyousha Hosseinichimeh

Konstantinos P. Triantis

Eileen M. Van Aken

May 03, 2018

Blacksburg, Virginia

Keywords: Diversity, Gender, Race, Higher Education, System Dynamics, Statistical
Analysis, Data Envelopment Analysis

Model-based Analysis of Diversity in Higher Education

Maryam Alsadat Andalib

ABSTRACT (academic)

U.S. higher education is an example of a large multi-organizational system within the service sector. Its performance regarding workforce development can be analyzed through the lens of industrial and systems engineering. In this three-essay dissertation, we seek the answer to the following question: How can the U.S. higher education system achieve an equal representation of female and minority members in its student and faculty populations? In essay 1, we model the education pipeline with a focus on the system's gender composition from k-12 to graduate school. We use a system dynamics approach to present a systems view of the mechanisms that affect the dynamics of higher education, replicate historical enrollment data, and forecast future trends of higher education's gender composition. Our results indicate that, in the next two decades, women will be the majority of advanced degree holders. In essay 2, we look at the support mechanisms for new-parent, tenure-track faculty in universities with a specific focus on tenure-clock extension policies. We construct a unique data set to answer questions around the effectiveness of removing the stigma connected with automatic tenure-clock policies. Our results show that such policies are successful in removing the stigma and that, overall, faculty members that have newborns and are employed by universities that adopt auto-TCE policies stay one year longer in their positions than other faculty members. In addition, although faculty employed at universities that adopt such policies are generally more satisfied with their jobs, there is no statistically significant effect of auto TCE policies on the chances of

obtaining tenure. In essay 3, we focus on the effectiveness of training underrepresented minorities (e.g., African Americans and Hispanics) in U.S. higher education institutions using a Data Envelopment Analysis approach. Our results indicate that graduation rates, average GPAs, and post-graduate salaries of minority students are higher in selective universities and those located in more diverse towns/cities. Furthermore, the graduation rate of minority students in private universities and those with affirmative action programs is higher than in other institutions. Overall, this dissertation provides new insights into improving diversity within the science workforce at different organizational levels by using industrial and systems engineering and management sciences methods.

Model-based Analysis of Diversity in Higher Education

Maryam Alsadat Andalib

ABSTRACT (general audience)

One of the goals of higher education institutions is to increase diversity within student and faculty bodies. Equal inclusion of all individuals in students and faculty populations is important to society in several ways. First, providing an equal chance for individuals' higher education and employment, regardless of demographic characteristics, is a cornerstone of any democratic society. Second, improving educational system diversity leads to higher educational achievements, as overall diversity of U.S. universities is a key indicator of global excellence. Despite improvement over the last decades, we still do not see an equitable distribution of women and racial minorities in such populations. The disparities in minority representation are even greater at higher levels of education and academic employment, such as graduate school and tenure-track positions. In this dissertation, our focus is on the trends, processes, and performance of the U.S. higher education system as it relates to diversity. We apply innovative industrial, systems engineering, and management sciences methods to the subject of diversity in the higher education context. The goal is to investigate answers to the following question: How can the U.S. higher education system achieve equal representation of female and minority groups in its student and faculty populations? The results of this dissertation could be used to train policy makers at institution and state levels on the ways of transforming universities into better places for females and minority groups. In particular, the system dynamics model could be used as a flight simulator in performing policy tests for educational workshops. Moreover, the outcomes could inform individuals and policy makers about the barriers doctorate holders face in following a successful academic path. Finally, this dissertation could be used in system dynamics and Data Envelopment Analysis classes as both case study and teaching materials.

Acknowledgment

I would like to express my sincere gratitude to my advisor Dr. Navid Ghaffarzadegan for his patience and continuous support. He has set an example of excellence as a researcher, mentor, instructor and role model. Without his guidance and persistent help this dissertation would not have been possible.

I would also like to thank my committee members, Dr. Konstantinos Triantis, Dr. Joshua Hawley, Dr. Eileen Van Aken, and Dr. Niyousha Hosseinichimeh for their invaluable feedbacks. I am truly appreciative of all that you have done for me over the years.

I would also like to thank my parents for their unconditional love and encouragement. To my mom, Dr. Roshanak Sedaqat, thank you for being my role model and inspiration. I always miss you. To my dad, Reza Andalib, thank you for being my best friend, and mentor and always encouraging me to seek excellence in every aspect of my life.

I am also grateful for the chance of meeting and working with my wonderful coauthors and friends Dr. Richard Larson, Julie Maurer, Dr. Hyungjo Hur, Dr. Arash Baghaei Lakeh, and Sarah Mostafavai. Thank you.

Finally, to my beloved husband, Alireza Sedighi, for his continued and unfailing love, support and understanding during my pursuit of Ph.D. degree that made the completion of this dissertation possible.

Table of Contents

Chapter 1. Introduction	1
References	6
Chapter 2. Simulation of the Leaky Pipeline: Gender Diversity in US K-Graduate Education	7
Abstract	7
Introduction	8
Background	9
Method	13
<i>Procedure</i>	13
<i>Data</i>	14
Modeling	15
<i>The pipeline(s)</i>	15
<i>Recruiting Students</i>	16
<i>Reinforcing Students</i>	18
<i>International Students</i>	20
<i>The Entire Model</i>	21
<i>Base Run Simulation</i>	22
<i>Gender Parity Trends Prediction</i>	25
Discussion	28
Acknowledgement	30
References	31
Chapter 3. Happy with a baby and a declined tenure: Effects of tenure-clock extension policies on career outcomes of new parent faculty members.....	33
Abstract	33
Introduction	35

<i>Automatic Tenure-Clock Extension Policy</i>	41
Method	43
<i>Institution-level Dataset</i>	43
<i>Individual-level Dataset</i>	44
<i>Model</i>	46
Results	47
<i>Which institutions adopt automatic tenure-clock extension policy?</i>	47
<i>Effect of automatic TCE on career outcomes</i>	51
Discussion	59
Acknowledgements	64
References	64
Chapter 4. What Factors Do Make a University More Successful in Diversity Programs?	66
Abstract	66
Introduction	67
Theoretical background	72
<i>Institution’s Environmental Factors</i>	77
<i>Institutional Factors</i>	79
Method	80
<i>Data Envelopment Analysis in education</i>	80
<i>Model</i>	83
<i>Input, output and environmental variables</i>	84
<i>Data</i>	88
Results	90
<i>Outlier Analysis</i>	90
<i>DEA Analysis</i>	91
<i>Hypotheses testing</i>	97
Discussion	100

<i>Implications</i>	103
References.....	105
Chapter 5. Conclusions	108
Contributions.....	112
Broader impact.....	114
Future avenues of research.....	114

List of Figures

Figure 1.1: Racial and gender composition of the US science workforce for different fields within behavioral and social sciences, engineering, and biological sciences subfields in 2015.	2
Figure 1.2: The topic, structure, methods, and data source of this dissertation.	4
Figure 2.1: Gender diversity in the US science and engineering workforce by the highest degree. From National Science Foundation, Scientists and Engineers Statistical Data System (SESTAT).	11
Figure 2.2: Levels of the US Education Pipeline for female population.	16
Figure 2.3: Recruiting feedback structure.	17
Figure 2.4: Reinforcing feedback structure.	19
Figure 2.5: Feedback structure of education pipeline model.	22
Figure 2.6: Female pre-high school enrollment.	23
Figure 2.7: Male pre-high school enrollment.	23
Figure 2.8: Female high school enrollment.	23
Figure 2.9: Male high school enrollment.	23
Figure 2.10: Female college enrollment over time.	25
Figure 2.11: Male college enrollment over time.	25
Figure 2.12: Female graduate school enrollment.	25
Figure 2.13: Male graduate school enrollment.	25
Figure 2.14: Male and female population by highest level of educational attainment from 1970 to 2035.	26
Figure 2.15: Female to male population ratio by highest level of educational attainment from 1970 to 2035.	28
Figure 3.1: Ratio of female doctorate recipients from 2006 to 2016 by field of study.	36
Figure 3.2: Ratio of women among doctorate recipients, new-hire, and all tenure or tenure-track faculty- averaged over 2006-2016.	37
Figure 3.3: Variations in tenure-clock extension policies in the US universities.	42
Figure 3.4: Change in tenure status and employment from 2008 to 2015 according to individual-level dataset.	46
Figure 3.5: Number of automatic TCE policy adopter universities over time- policy adoption. Note: AAUP: American Association of University Professors.	48

Figure 3.6: Selected characteristics of universities by level of tenure-clock extension (TCE) policies.	50
Figure 3.7: Average tenure-track duration by newborn and auto TCE policy.	54
Figure 3.8: Chance of getting tenure (compared to leaving tenure-track position) by having newborn and auto TCE policy.....	56
Figure 3.9: Average job satisfaction by having newborn and auto TCE policy.	59
Figure 4.1: Racial and gender compositions of United States universities, 2015.	70
Figure 4.2: Theory of minority students' underrepresentation in higher education.	76
Figure 4.3: Transformation Process.....	88
Figure 4.4: Distribution of efficiency scores by DEA model.	94
Figure 4.5: DEA efficiency scores distribution on the US map.....	96

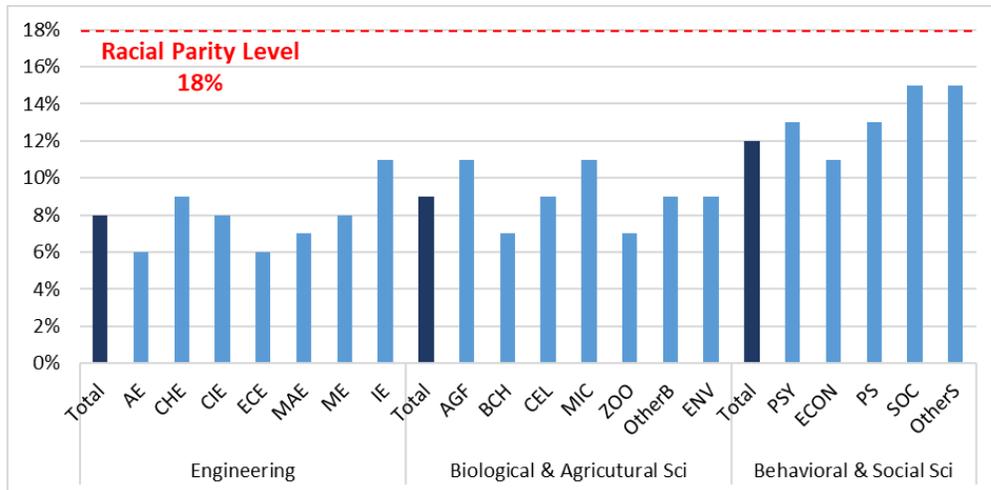
List of Tables

Table 3.1: Descriptive statistics of individual-level dataset and t-test between gender groups.	45
Table 3.2: Results of multivariable linear regression for tenure-track duration.	53
Table 3.3: Binomial logistic regression results for getting tenure	55
Table 3.4: Multivariable linear regression for job satisfaction.	58
Table 4.1: Input, output and environmental variables of data envelopment analysis.	88
Table 4.2: Descriptive statistics table, 2008. (Number of units=156)	90
Table 4.3: Input and output variables of DEA models.	93
Table 4.4: Descriptive statistics of DEA efficiency scores.	94
Table 4.5: Ivy League universities.	97
Table 4.6- The results of the bootstrapped second-stage regression on the environmental factors.	99

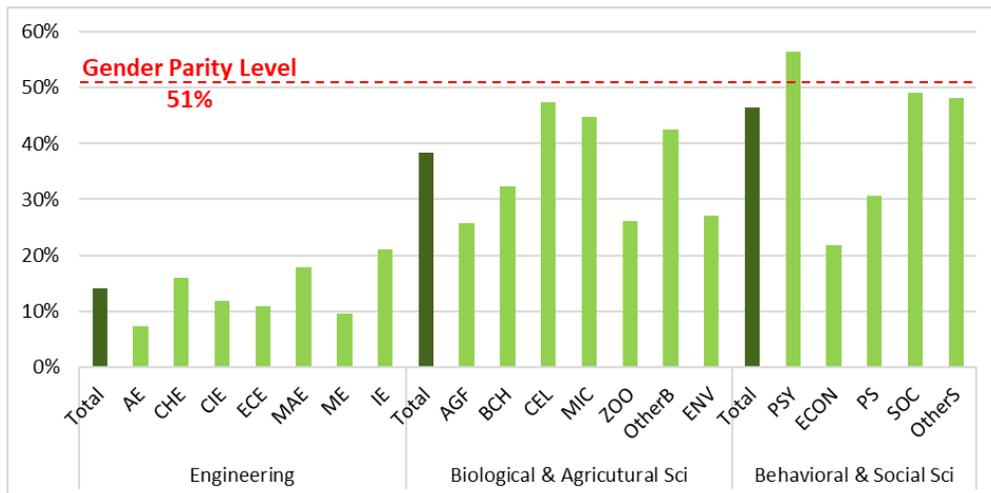
Chapter 1. Introduction

Since the time when there were no female and racial minority students attending universities, United States higher education institutions have made great strides in their inclusion of female and racial minority students. In addition to providing equal educational opportunities to everyone regardless of their demographic characteristics, such improved diversity is also expected to increase overall achievement in the entire education system [1]. Equal inclusion of individuals in the economic and political life of their country is a valuable societal aspect of a diverse education system [2].

However, despite these efforts, we still remain far from equal distribution of educational opportunities in our society. Figure 1.1 shows the 2013 racial and gender composition of doctorate recipients by field of study. Figure 1.1.a shows the percentage of non-white, non-Asian minorities in the academic fields of three major groups—engineering, biological and agricultural sciences, and behavioral and social sciences. The “racial parity level” line shows that the percentage of non-Asian racial minorities in the U.S. population is 18%. As shown, the percentage of non-Asian racial minorities in all scientific fields is less than the same percentage in the overall population. In other words, minorities are underrepresented in the science workforce. Figure 1.1.b depicts the percentage of female students in these fields. Females are 51% of the U.S. population, but in many fields, especially engineering, they comprise less than 20% of the workforce. Looking at the dynamics of the gender and racial gaps over time also demonstrates that such disparities are persistent and even growing in some areas such as medical schools and Science, Technology, Engineering and Mathematics (STEM) fields [3, 4]. Furthermore, women and racial minorities are underrepresented among professionals [5] as well as in tenure-track academic positions [6].



(a) Percentage of non-Asian racial minorities in the U.S. science workforce by fields.



(b) Percentage of women in the U.S. science workforce by fields.

Figure 1.1: Racial and gender composition of the US science workforce for different fields within behavioral and social sciences, engineering, and biological sciences subfields in 2015¹.

-
- ¹ Engineering fields: AE (Aerospace Engineering), CHE (Chemical Engineering), CIE (Civil Engineering), ECE (Electrical and Computer Engineering), MAE (Material Engineering), ME (Mechanical Engineering), IE (Industrial Engineering).
 - Biological and Agricultural Sciences: AGF (Agricultural and Food Sciences), BCH (Biochemistry and Biophysics), CEL (Cell and Molecular Biology), MIC (Microbiology), ZOO (Zoology), ENV (Environmental Life Sciences), OtherB (Other Biological and Agricultural Sciences).
 - Behavioral and Social Sciences: PSY (Psychology), ECON (Economics), PS (Political Sciences), SOC (Sociology), OtherS (Other Social Sciences).

Data Source: The author's estimation based on the Survey of Doctorate Recipients (SDR) 2015, National Science Foundation (NSF). Note: the dash lines, referred to as the parity level, which show the percentage of the specific groups in the U.S. population.

In this dissertation, we study the gender and racial composition of the science workforce and attempt to answer the question: How can an equal representation of female and minority groups in student and faculty populations be achieved in the U.S. higher education? We address this question by looking at various data sources of U.S. higher education at individual, institutional, and national levels.

Figure 1.2 summarizes our approach and the structure of this dissertation. In a three-essay format we utilize three modeling techniques: specifically, system dynamics, statistical analysis, and data envelopment analysis, as described in the next paragraphs. The first objective of this dissertation is to present a systems perspective of the U.S. education pipeline and predict the gender composition of the educated workforce. The second objective is to investigate the effects of work-life balance programs, specifically the automatic extension of the tenure-clock for new parent faculty members and the effect on their career outcomes. Finally, the third objective is to test the effects of several institutional and state-level factors on educational and career outcomes of racial minorities and identify the best practices among U.S. institutions.

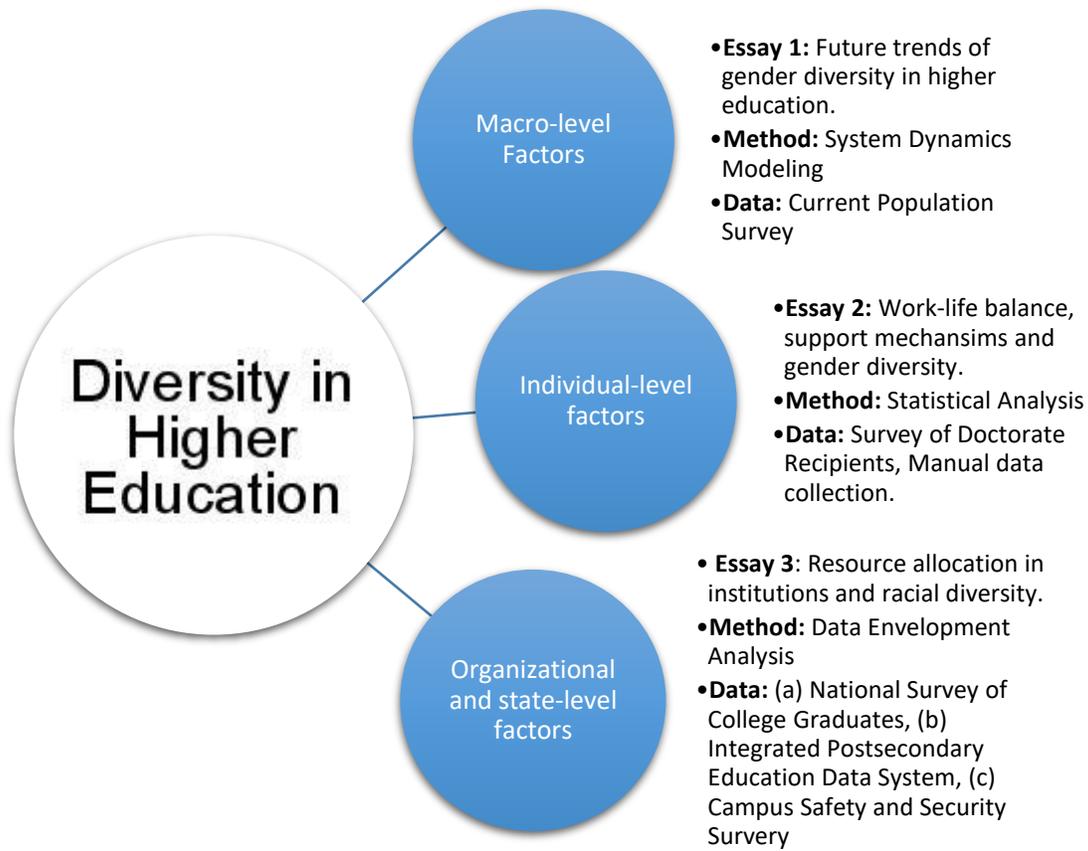


Figure 1.2: The topic, structure, methods, and data source of this dissertation.

Essay 1 discusses a system dynamics approach to present a systems view of the U.S. education pipeline and predict the U.S. workforce gender composition by highest degree attainment level. The education pipeline starts at kindergarten and models the flow of male and female students through different educational levels (i.e., college and graduate school). The model captures several positive and negative factors that affect individuals' decisions to continue their education at a higher level or start a job with their highest degree. The effects of role models and parents are examples of phenomena that help to reduce the attrition rate of university students. The model is calibrated using historical enrollment data from 1980 to 2015 and forecasts future gender-composition trends in regards to the educated workforce to 2035.

Essay 2 focuses on work-life balance programs in higher education institutions and the effect of automatic tenure-clock extensions due to childbirth or adoption on the career outcomes of pre-tenure faculty. Career outcomes are defined as the duration of tenure track, likelihood of promotion to a tenured position, job satisfaction, salary, and chances of getting research funding. I collect an institutional-level data set of the top 250 U.S. universities (according to U.S. News & World Report rankings) to investigate the characteristics of policy adopter universities. The information is gathered through university websites and guidelines, as well as direct contact with provosts of many of these universities. Furthermore, I connect 2008, 2010, 2013, and 2015 surveys of doctorate recipients (SDR) from the National Science Foundation to my constructed data to evaluate career outcomes over seven years at the individual level. Finally, I test the effect of automatic tenure-track extension policy on the abovementioned career outcomes.

Essay 3 looks at the educational and employment outcomes of undergraduate racial minority students at the institutional level and attempts to discern why some universities are more successful than others in terms of improving diversity. I compare two contrasting theoretical explanations for institutional success in bringing more diversity into higher education. Success is defined as a higher graduation rate, higher GPA, and racial minorities' first post-graduate job salaries. I use data envelopment analysis (DEA), a non-parametric method, to compare universities in terms of training and educating minority students (e.g., Hispanic and African-American students). The input measures include number of minority undergraduate and graduate students enrolled, number of minority faculty, and financial support provided to the students. Output measures are related to the educational outcomes of minority students (e.g., graduation rate and GPA) or employment outcomes (e.g., first post-graduate job salary). Number of hate incidents on campus is also considered an undesirable output of the model. I construct my own data set for this analysis by

collecting and integrating different data sources at the institutional level that include the National Survey of College Graduates (NSCG), Campus Safety and Security (CSS), and the Integrated Postsecondary Education Data System (IPEDS). Using the DEA, I calculate efficiency scores for universities based on the defined input/output set. We then test the effects of several institutional- and state-level factors on training minority students.

The dissertation proceeds as follows. In Chapters Two, Three, and Four, we present each of the three essays. Each essay is a stand-alone paper. In Chapter Five we discuss the main findings and conclusions and point to future avenues for research.

References

1. Bitters BA, Team EM. Useful definitions for exploring educational equity. Wisconsin Department of Public Instruction. 1994.
2. Fiske EB, Ladd HF. Elusive equity: Education reform in post-apartheid South Africa: Brookings Institution Press; 2004.
3. Garces LM. Understanding the impact of affirmative action bans in different graduate fields of study. *American Educational Research Journal*. 2013;0002831212470483.
4. Garces LM, Mickey-Pabello D. Racial diversity in the medical profession: The impact of affirmative action bans on underrepresented student of color matriculation in medical schools. *The Journal of higher education*. 2015;86(2):264.
5. Holding a Four-Year College Degree Brings Blacks Close to Economic Parity with Whites. *The Journal of Blacks in Higher Education*. 2005(47):6-8.
6. Ceci SJ, Ginther DK, Kahn S, Williams WM. Women in academic science A changing landscape. *Psychological Science in the Public Interest*. 2014;15(3):75-141.

Chapter 2. Simulation of the Leaky Pipeline: Gender Diversity in US K- Graduate Education

Abstract

What is the future trend of gender (dis)parity in US higher education and workforce? We take a systems approach and develop a dynamic simulation model of gender diversity in US higher education, from kindergarten through elementary, high school, undergraduate and graduate studies. The model consists of the education pipeline for male and female and includes several mechanisms that reinforce interest in higher education, or affect university admission or dropout. We simulate the model and examine its fidelity in replicating the historical data of US higher education (1970-2014). Then, we simulate the model to forecast the future trends of gender composition at different stages of the education pipeline for the next two decades. The model forecasts that women will be the majority of university degree holders in US in 2035 (female to male ratio of 1.7 of advanced degrees) with consistently lower rates of dropout in high school, undergraduate, and graduate level of education. We discuss the findings and implications.

Keywords: Gender diversity; US higher education; education pipeline; system dynamics; simulation.

Introduction

In an equitable education system, we expect people from different racial, ethnic, and gender groups with any sexual orientation and disability level to have equal chances of academic success. Academic success is translated into receiving college admission, graduating university, achieving higher standardized test scores, and landing technical and scientific jobs [1]. In such a society, as opportunities are solely distributed based on skills, and skills are independent from demographic characteristics, we expect to see the same distribution of demographics in different fields that we see in the population. However, data suggest that we are far from this ideal situation, and our initial demographics may predict the level of education that we will receive and our future economic situation.

One particular dimension of diversity, which is the focus of this paper, is gender diversity. An education system is gender segregated at the higher education level if female students are less represented at undergraduate or graduate levels than their male counterparts [2]. Women's lower representation in higher education was historically due to the issues related to access to higher education and their lower rates of degree completion. Dealing with childcare, family and labor market commitments with lack of supporting mechanisms makes receiving a higher degree less attractive for women. Cultural norms and gender stereotypes are two widely cited reasons for the second aspect of gender segregation [2]. It is also argued that current gender composition of university students and faculty members further influences the interest of women in higher education and their retention rate by reinforcing gender stereotypes [3].

Recent evidence points to changes in gender composition of higher education in favor of more parities and less "leakage" in education pipeline. We call the education pipeline leaky, when it fails to retain female students at different levels of education. It seems that the trends are changing

and women's demand for higher education is growing [3]. The pipeline leakage however still exists at the later stages of higher education. The low rate of transition from doctorate degrees to faculty positions in women is an example of this phenomenon [4]. An important question is how the future trend of gender (dis)parity in US higher education will look like. This is our focus in this research.

A major shortcoming of previous studies of gender diversity is the common focus on a subset of the education pipeline rather than analyzing the entire pipeline of education. Dropout in high school [5], college [6], and faculty stages [7] are important factors that affect gender diversity. However, these factors are interdependent, and for example, low quality of education at high school can only later manifest itself in the form drop out from college. From a systems perspective, one should look at the entire system of education [8] and examine long-term trend of gender diversity.

In this paper, we develop a system dynamics model of the US education pipeline to examine the trends of gender diversity in the US higher education. We consider a broad boundary for our model, one that includes pre-high school and high school education as well as university level education. The model simulates the flow of male and female students through the education pipeline. The simulation model will help examine the future trends of gender composition in workforce and higher education.

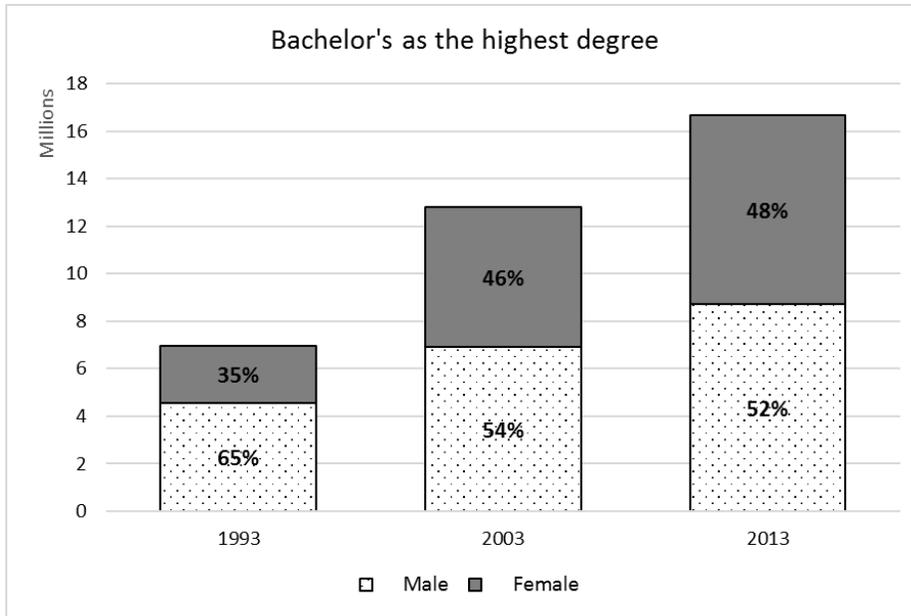
Background

Improving diversity in education is valued by itself but also is expected to increase the overall achievement of the entire education system [9]. More diversity results in higher representation of different perspectives, cultures, ideas, and strengths, which all improve our understanding of the society as one entity [10, 11]. As a societal value, we also expect that every child regardless of his/her family's socio-economic status and demographic characteristics should be given similar

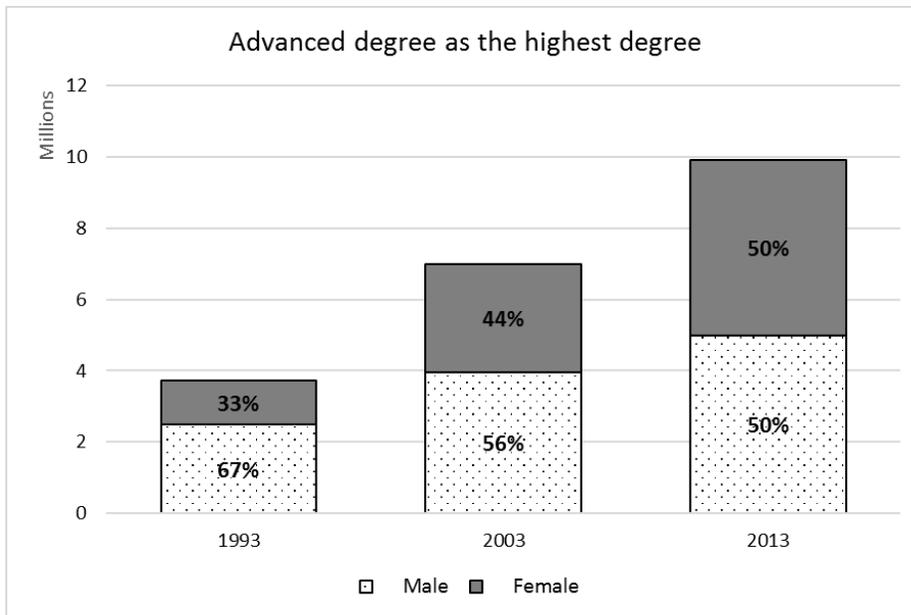
opportunities for flourishing. An equitable education is a corner stone in increasing opportunities for under-represented groups. Such education systems ensure that historically underrepresented populations obtain academic achievements [9].

Recent trends of gender diversity in US higher education have been promising. Overall, and across all fields, women's representation in educational ladders have been gradually improving during the past two decades. Figure 1 depicts changes in gender composition of science and engineering workforce by the level of highest degree. During the past two decades, the number of science and engineering (S&E) workforce has grown for both genders, but in a higher rate for women than men. As the figure shows, in 1993, only 35% of the population of US workforce with bachelors as the highest degree were women, but the ratio increased to 46% and 48% in 2003 and 2013. Number of women with Bachelor's degrees changed from about 2 million in 1993 to about 6 and 8 million people in 2003 and 2013.

The gender composition of the US advanced degree holders is also moving from the state of women's underrepresentation toward gender parity. According to Figure 1.b, only 33% of the advanced degree holders were women in 1993. Women's representation in the population of advanced degree holders increased over the last two decades, as they composed 44% and 50% of this population in 2003 and 2013 respectively. The number of women with advanced degrees also grew from about 1 million in 1993, to 3 and 5 million in 2003 and 2013.



(a) Bachelor's as the highest degree



(b) Advanced degree as the highest degree.

Figure 2.1: Gender diversity in the US science and engineering workforce by the highest degree. From National Science Foundation, Scientists and Engineers Statistical Data System (SESTAT).

Significant improvements are also observed in women's academic success measures [12]. In the past, one mostly cited reason of the relatively low number of women entering higher education was their under preparation throughout high school [4]. Now, female students compared to their male peers perform better in reading tests [12], and have higher college degree attainment rates [13]. Over years females also improved their performance in mathematics [14], and their representation in male dominated fields [15]. Women reached to parity in non-math-intensive fields such as life science, psychology, and social sciences [16].

Despite all of the described improvements in different aspects of gender equity in higher education, gender gaps are observable in some areas, like attrition of female students in math-intensive fields, fewer employment of female faculty members, especially as associate or full professors [17].

The identified causes of gender gaps in higher education are multiple and interconnected. Many studies have attempted to identify these causes, which change over decades. Historically, gender discrimination against women was the main reason of their low representation in academia, however, currently it is not among the most contributing factors [18]. Some of the reasons that adversely affect women in academia are role modeling [19], general beliefs about the appropriate behaviors of men and women (i.e. gender-role attitudes), faculty composition, campus climate, (i.e. self-selection of students) [20], lower expectations [4], and negative stereotypes that lead students underperform in standardized tests [21]. Lack of female role models is factor that plays an important role in both retention and recruitment. Even the enrolled female students might perceive that they are not following a correct educational path because there is not many women faculty or graduates in their fields [4]. Career choices of men and women are influenced by the

gender stereotypes that pictures men with higher mathematics abilities, and ultimately yields females to non-math intensive fields [16].

The gender gap persists at the latest stages in academia; faculty positions. Literature reports that number of required working hours and quality of female applicants does not contribute to the existing gender disparity in academia [22]. Gender bias against female PhD degree holders in academia is the main cause for undervaluing women and hiring them with lower salaries [23]. The life choices, and career preferences of female PhDs, as well as the general belief that faculty positions are not compatible with married life and raising children are other factors that lead them to opt-out of tenure-track positions [16].

In this research, we specifically aim at understanding the future trends of gender disparities at different stages of the US education pipeline in the next 20 years. We model the dynamics of men and women's educational attainment in the US using previous empirical findings from the literature including social and personal factors that affect recruitment and retention of students in higher education institutes.

Method

Procedure

We develop a system dynamics [24] model that represents the US education pipeline. The causal relationships in this model are mostly informed by empirical research in the area of education policy and studies of gender diversity in education and scientific workforce. The model captures the dynamics of population at different educational levels from kindergarten to graduate school, from 1970 to 2014, and for different genders. Most model parameters are estimated from the past empirical research. For the unknown parameters, we conducted partial model calibration [25]. Details of parameter values and calibration procedures are presented in the Appendix.

We conduct various tests for model validation including structural and behavioral validity tests as described in Business Dynamics [24] including a comparison of the simulation trends with the historical trends of US men and women enrollment in K-12 and higher education. The model boundary is around the education system, and US birth rate is an exogenous variable in this model. We also run the model for twenty years in the future (2016-2035) to forecast the enrollment and graduation trend of each gender group. We also conduct sensitivity analysis for a wide range of changes in estimated or assumed parameters, presented in the Appendix.

Data

US Census Bureau publishes school enrollment data by educational attainment level and by different demographic characteristics. They collect the data through current population survey (CPS) which is the main source of labor force and population statistics in the US. CPS provides school enrollment from pre-kindergarten to higher levels of education; graduate school. We use the longitudinal enrollment data for men and women from 1970 to 2014 as a reference for replicating the historical data.

Current Population Survey (CPS) postsecondary education enrollment data includes both US-born and non-US-born students in its estimations. We need to include the international students that temporarily reside in the US to study in undergraduate or graduate programs. Institution of International Education (IIE) Open Doors provides a longitudinal data of the number of enrolled international students in the US by education level. We use Open Doors data, which is available online from 1960 to 2015.

Male and female birth rates are the other two exogenous variables. We use the data reported by Census to provide the model with these rates from 1970 to 2014. Furthermore, Census publishes

a projection of the US population as well as birth, and death rates. We use Census population projections to estimate US birth rate for the next twenty years.

Modeling

The pipeline(s)

The model relies on two major aging chains, which represent two parallel education pipelines, one for women and one for men [24]. Figure 2 is a simple representation of the US education pipeline for the female population. Individuals enter the pipeline upon their birth and go through different stages. The students can attain high school diploma, bachelors, or advanced degrees (e.g. masters and doctorate) or dropout of school at any stage and work with their highest attained degree until retirement or death. Individuals enter the pipeline through birth rate; then they go to school after a period of “Under School Age” (in Figure 2, first stock from the left). For simplicity, education prior to university is aggregated by two variables, “Enrolled in PreK-9th Grade” and “Enrolled in High School.” Some may dropout from school; others will ultimately graduate from high school and join the stage of “with High School Diploma as Highest Degree.” Some of this population stay in this stage until retirement/death, and some may decide to be “Enrolled in College” (in the Figure, the most left stock variables among the second row of stock variables). This population may also dropout, complete education and start a job, or decide to continue education at the graduate level.

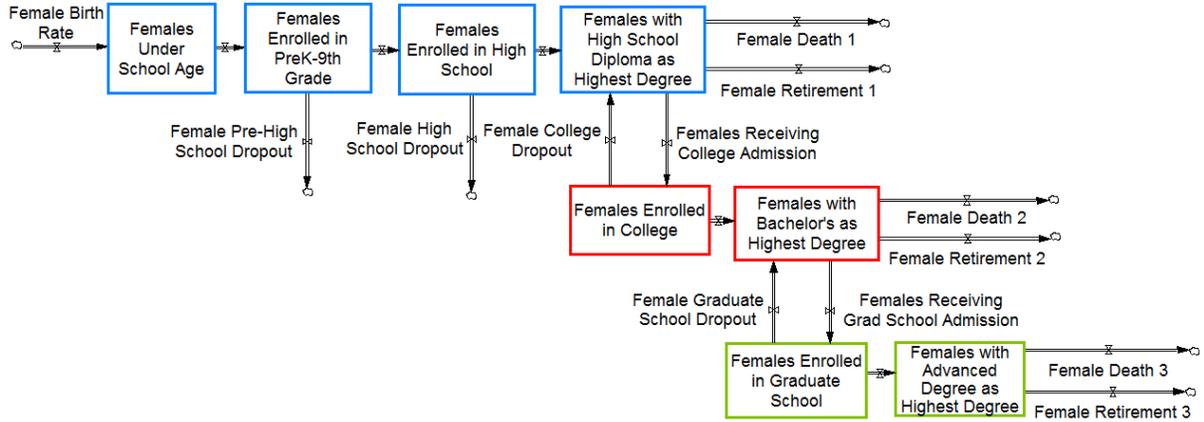


Figure 2.2: Levels of the US Education Pipeline for female population.

A model parallel to the one depicted for females in Figure 2, represents the pipeline of education for males, which of course may have different dropout rate or different tendency to continue education to higher levels.

These two parallel pipelines of course are interdependent. In the next sections, we describe the main feedback structures that connects male and female education pipelines.

Recruiting Students

We first estimate total college admission rate, and then split it between female and male based on the past trends and mechanisms described here. Figure 3 shows the structure and here we discuss major equations.

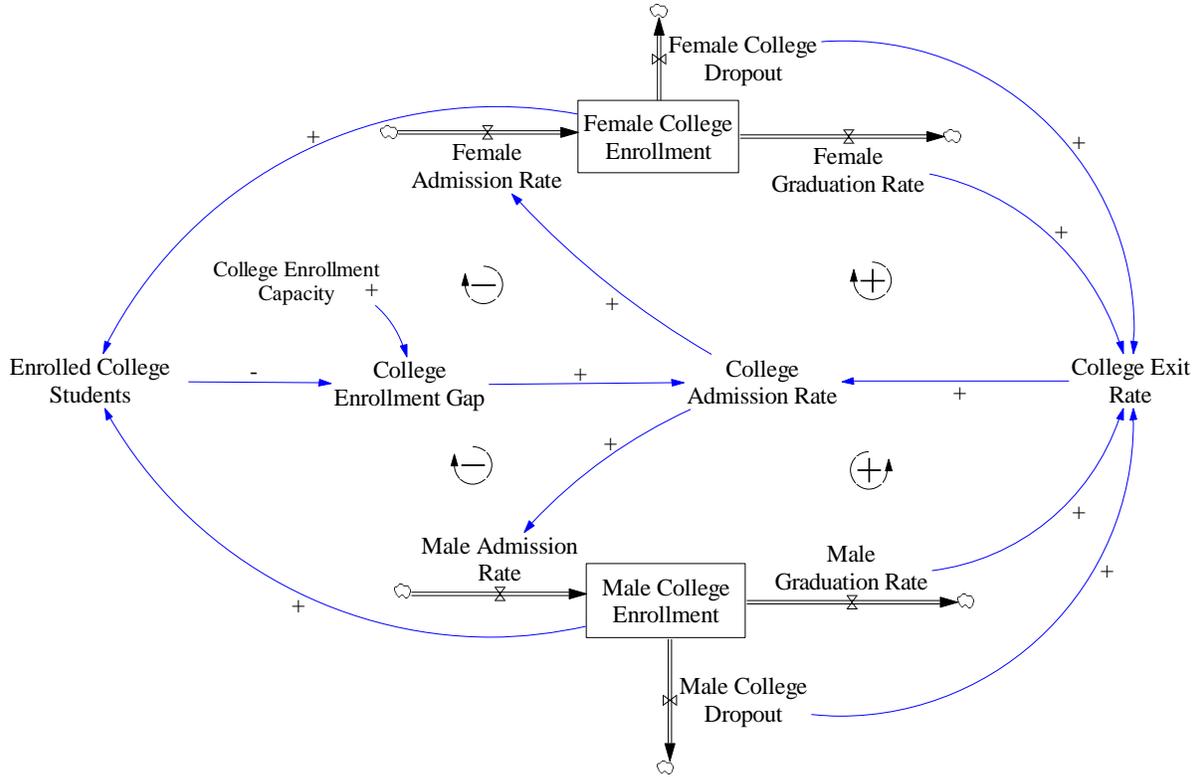


Figure 2.3: Recruiting feedback structure.

In a steady state, universities' admission rate is equal to college exit rate that is new students are replacing students who leave college. College exit rate is equal to graduation rates and dropout rates of both foreign-born students, and US-born.

$$\begin{aligned}
 \text{College Exit Rate} = & (\text{Female Graduation Rate} + \text{Male Graduation Rate}) + \\
 & (\text{Female College Dropout} + \text{Male College Dropout})
 \end{aligned}$$

in steady state where College Enrollment is constant, we have

$$\text{College Admission Rate} = \text{College Exit Rate}$$

In general, when College Enrollment Capacity changes, we will have:

$$\text{College Admission Rate} = \text{College Exit Rate} + \text{Change in College Enrollment}$$

where

$$\begin{aligned} \text{Change in College Enrollment} &= \text{College Enrollment Capacity} - \text{College Enrollment} = \\ &= \text{College Enrollment Capacity} - (\text{Female College Enrollment} + \text{Male} \\ &= \text{College Enrollment}) \end{aligned}$$

After estimating admission rate, then, the total admission rate splits between male and female applicants at different fractional rates (R_{Female} , and R_{Male}). We estimate the fractions from model calibration.

$$\text{Female Admission Rate} = R_{\text{Female}} * \text{College Admission Rate}$$

$$\text{Male Admission Rate} = R_{\text{Male}} * \text{College Admission Rate}$$

where $R_{\text{Female}} + R_{\text{Male}} = 1$, Female Admission Rate < Females with High School Diploma as their highest degree, and Male Admission Rate < Males with High School Diploma as their highest degree.

Similar mechanisms exist for graduate admission rates. For high school and lower grades, we assume infinite capacity, so any one can enroll as long as they do not drop out.

Reinforcing Students

We identified several feedback mechanisms that reinforce students' enrollment in universities. These mechanisms improve students' enrollment through either increasing admission rate or decreasing dropout rate. The following is a list of the reinforcing mechanisms.

- *Parental financial support*: Individuals with higher university degrees receive higher incomes on average. When women's education level increases, families would be better prepared to support their children financially because mothers can contribute more [26].
- *Mentoring*: Families would be able to provide more support to their children through their education path when parents have any or more college experience [27].
- *Role modeling*: Existence of female role models encourages students of the same gender to pursue higher degrees, improve their perception of chance of successfully finishing their studies, and reduce their dropout rates [28].

- *Established image of female highly educated professionals*: Existence of more women with university degrees in the society makes the image of a female highly educated employee more established [28].

We model the positive effect of having more women with university degrees (i.e. increased education level of women) on the representation of female students at universities through two aggregate loops. These two reinforcing loops represent all previously described mechanisms, which reinforce highly educated women in the society. First, increased education level of women positively affects younger women’s interest in higher education by increasing their inflow to the universities. Second, increased education level of women improves female student’s persistence through getting their degrees.

Figure 4 shows the effect. It represents the change of college dropouts based on the population of individuals holding bachelor’s degree and higher.

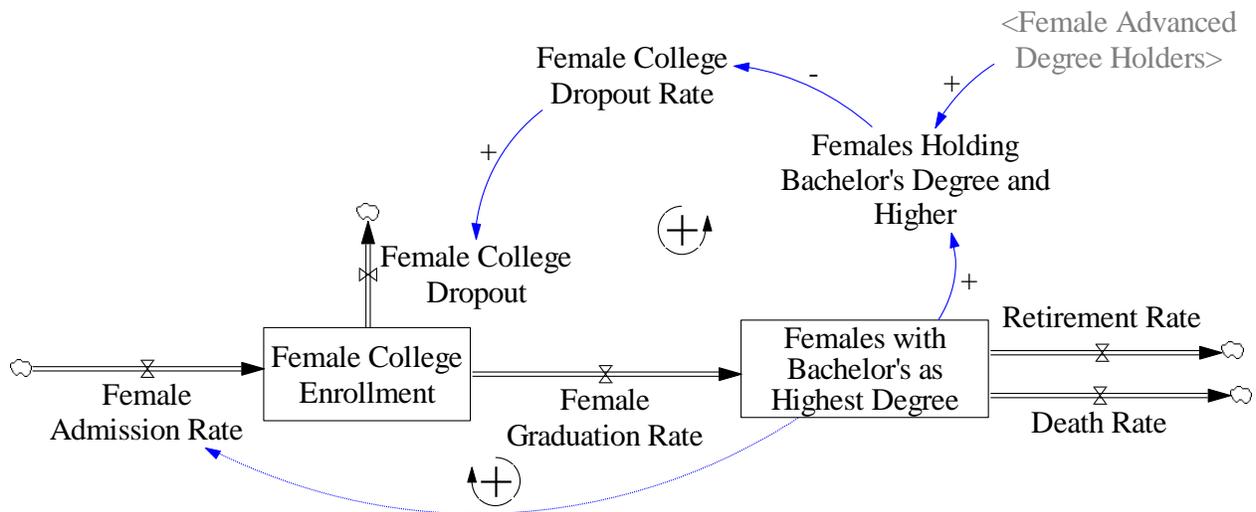


Figure 2.4: Reinforcing feedback structure.

Here, we only described the recruiting and retaining feedback structures for college level. Similar structure also exists for other levels of the model (e.g. high school and graduate school).

International Students

The previously described aging chain structure is designed to model the educational attainment level of US-born students using the birth rate and university capacity at both undergraduate and graduate levels. However, the goal of this model is to replicate the US enrollment data, which includes both foreign-born and US-born students. Therefore, we need to incorporate the annual incoming international students into the model at university level and adjust the enrollment rates to account for foreign-born in primary and secondary education levels. The trends of incoming international students at both educational levels are derived by multiple complex mechanisms. Modeling the dynamics of international students is out of the boundaries of the current system dynamics model. Therefore, we import the number of international students enrollment by education level and decompose it with a constant rate of 56% male and 44% female [29]. The population of international students affects the dynamics of the current model in the following ways:

1. *Enrolled Students by Education Level:* The stocks of the number of enrollments include only US-born students. We add the number of international students by gender and education level to the number of domestic students to count the total enrollments and replicate the historical enrollment data.
2. *Exit Rate by Education Level:* We use the college exit rate to calculate the available seats for the next cohort of students in both undergraduate and graduate schools. We add the number of graduating international students to the graduating domestic students to reach to a more accurate estimation of next year's admission rate.
3. *Degree Holders by Educational Attainment Level:* Past studies state that 45% of international students want to stay in the US. Therefore, we assume that 45% of the graduating international students enter the stocks of degree holders and affect the retention mechanisms.

Based on our estimations from the Scientists and Engineers Statistical Data System (SESTAT) from NSF, the ratio of international graduate students who did their undergraduate studies in the US was 12% on average. We use this ratio to avoid double counting the number of international students staying in the US with bachelor's as their highest degree in the stock of international students who stay in the US with advanced degrees.

The Entire Model

Figure 5 shows a simple representation of the whole model. Both gender groups share the same stock-and-flow structure. Here, we use M and F acronyms for differentiating between male and female variables. We import two types of variables into the model: birth rate and university capacity. We divide the first exogenous variable, birth rate, into two rates by gender, which are the inflow to the first stock of the education pipeline, pre-high school enrollment. Male and female students enter the education pipeline upon their birth and enter pre-high school. Then they proceed through the aging chain and can receive higher degrees if they do not dropout or decide to work with a lower degree. Number of new admissions to college and graduate school is a function of other exogenous variables; college and graduate school capacities. They determine the extra available seats in universities each year and are exogenous variables. These variables drive the aggregate admission rate (for both male and female students) and the portion that goes to each gender group is calculated by recruiting students feedback loops (described in previous sections). Finally, individuals exit the pipeline when they retire or die. The rates have different values for men and women. We documented detailed formulation and parameter values of the model in the appendix.

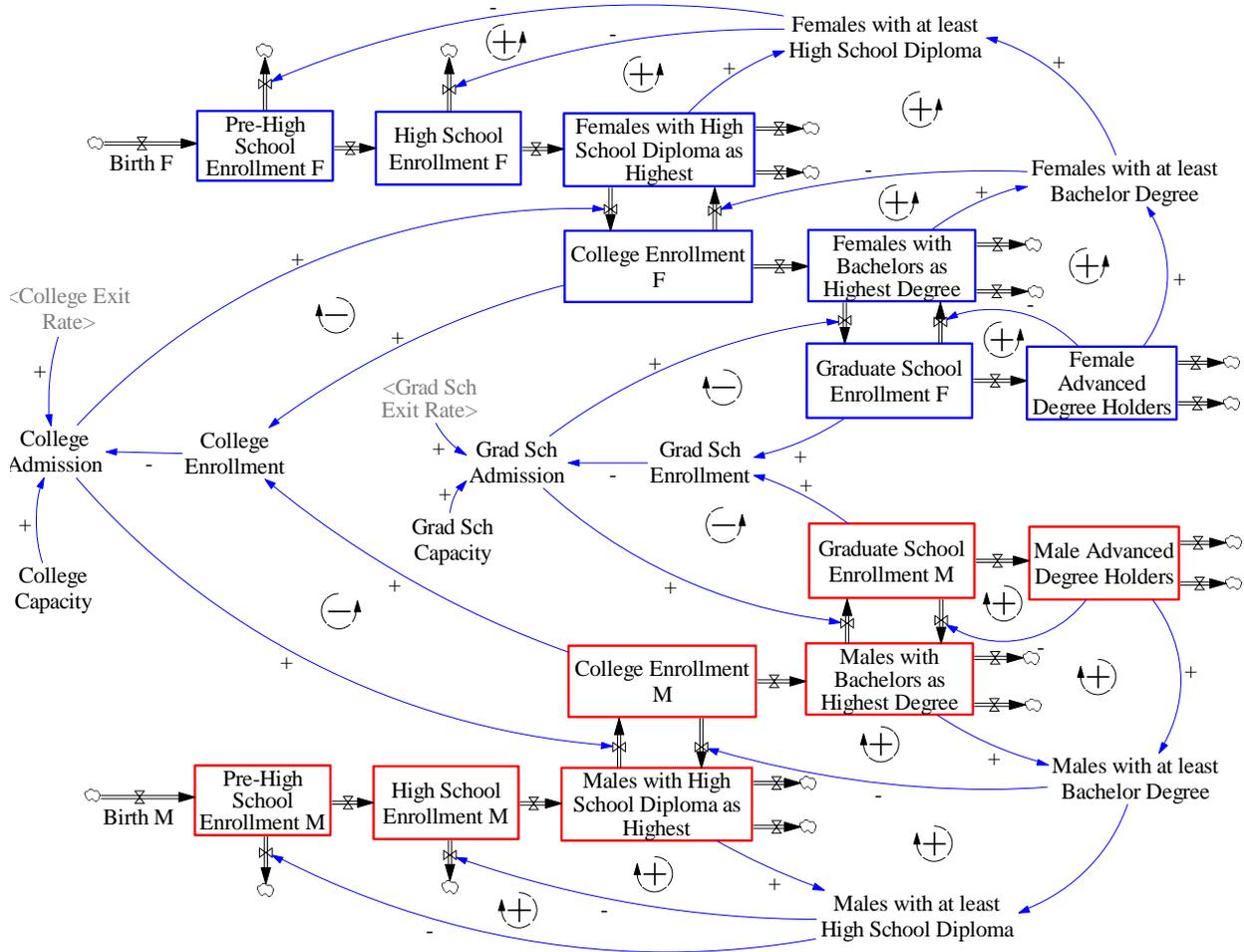


Figure 2.5: Feedback structure of education pipeline model.

Base Run Simulation

Figures 6 through 9 depict simulation results for the number of pre-high school and high school enrollments in the US from 1970 to 2035 for male and female students. They also compare the results to the pre-actual enrollment data for the period of 1970--2014 for males and females, which are both quite successful replications. As Figures 6-9 show, the model fairly replicates the data and predicts a relatively steady trend for male and female enrollment in pre-high school and high school for the next twenty years.

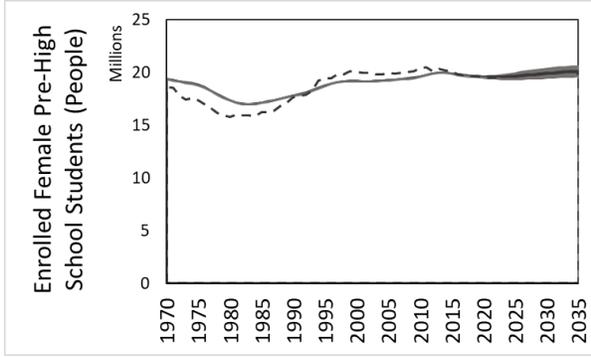


Figure 2.6: Female pre-high school enrollment.

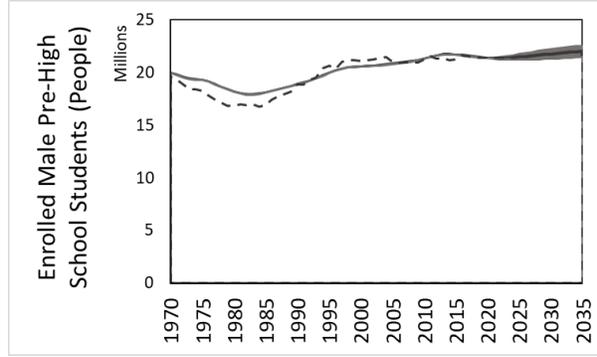


Figure 2.7: Male pre-high school enrollment.

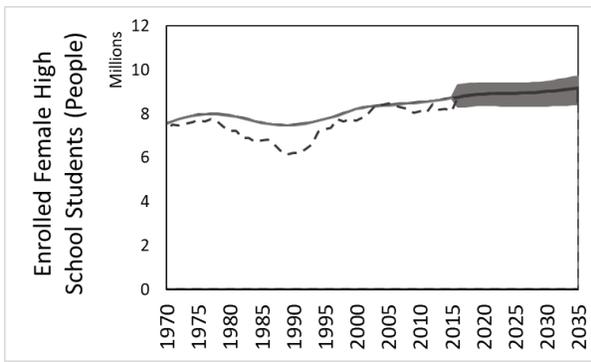


Figure 2.8: Female high school enrollment.

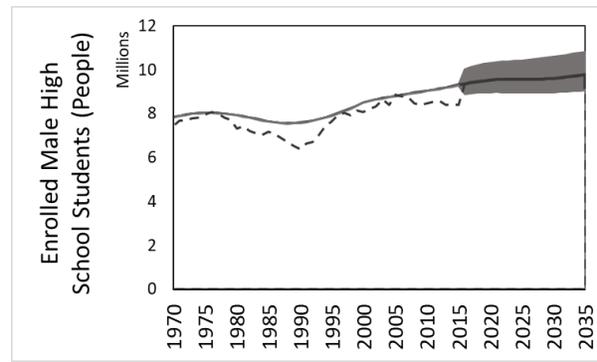


Figure 2.9: Male high school enrollment.

Figures 10 and 11 depict the simulated and the actual number of college enrollments in the US from 1970 to 2035 for female and male students respectively. The figures compare the results of simulation to the actual enrollment data for both males and females, which is a quite successful replication. The model also forecasts the enrollment trends for the next twenty years (until 2035). The grey band around the enrollment graph shows the sensitivity of the model to the calibrated parameters. In other words, there is a confidence interval for the forecasts of the population model.

The simulation model replicated the past forty five years growth in college enrollment: It more than doubles from 6.2 million to 15.4 million students. In 1970, the population of female college students were 1 million smaller than the male population. Since the number of female

undergraduate students historically grew faster than males, the population of female college students exceeded males in 1974. In 2014, the number of female college students is approximately 3 million more. The model forecasts that the gender gap in college enrollment continues to widen with a slower pace compared to 1994-2014 decade. However, the model predicts that the gap is reversed in the next two decades with 1.6 million more male college enrollments. The main reason is that a large proportion of the population has already received college degrees. Enrollment for male, in fact, might stay steady.

Figures 12 and 13 also depict graduate school enrollment for males and females. Population of graduate students in the US grew from 1.1 to 3.8 million students from 1970 to 2014, which translates to five percent annual growth. The number of female graduate students was half of male enrollment (0.4 million female and 0.8 million male graduate students). The population of female graduate students grew faster than males. Therefore, the intersection point of male and female enrollment graphs happens in 1977 with about 2 million total graduate school enrollments. Female enrollment increased from 366 thousands to 2.1 million in 1970-2014, which is higher than male enrollment (1.6 million male students). The gender gap in graduate school grew annually by 8 percent in 1994-2014, but the model predicts that this pace increases from 2014 to 2035 (13 percent annual growth).

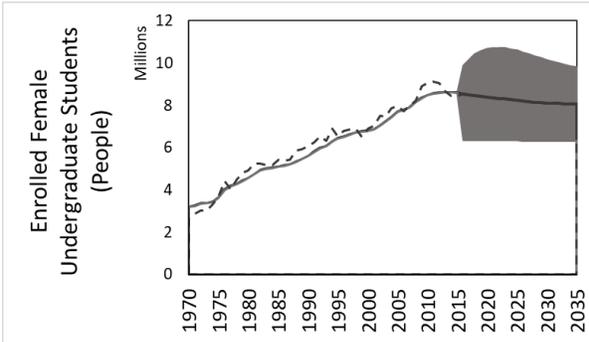


Figure 2.10: Female college enrollment over time.

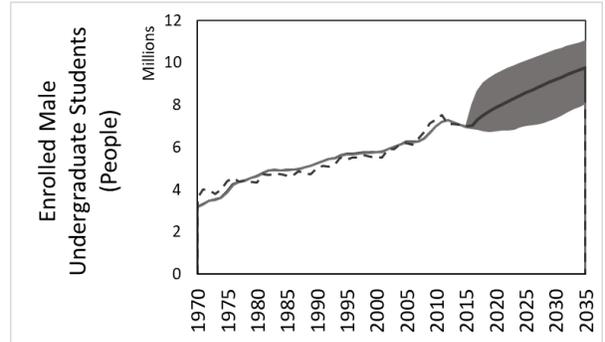


Figure 2.11: Male college enrollment over time.

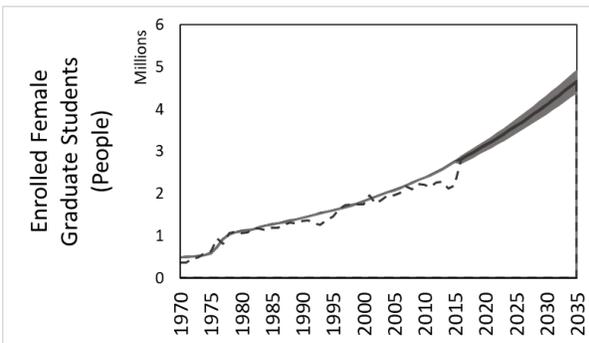


Figure 2.12: Female graduate school enrollment.

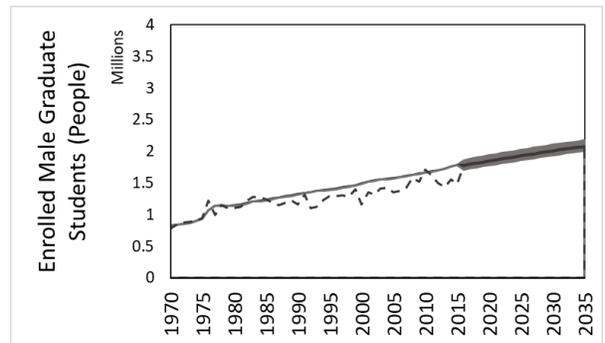


Figure 2.13: Male graduate school enrollment.

Gender Parity Trends Prediction

We investigate model's prediction of gender parity. Figure 14 depicts the number of men and women by their highest level of educational attainment from 1970 to 2014. It also predicts the population of similar groups from 2015 to 2035. We present three levels of highest educational attainment (e.g. high school diploma, bachelor's degree, and advanced degree) according to model properties.

According to the figure, both men and women pursue their studies at some level of higher education. The population of men and women with high school diploma as their highest level of education slightly increases until 1980-1990, and then it slightly decreases over time. However, the decline rate is steeper in women's population with high school diploma as highest degree. Men and women continue their education up to different levels and their tendency to stay in school changes over time. For example, we have almost equal number of men and women working with their bachelor's degrees until 2020. In 2035, the gap is around 10 million people in favor of men with bachelor's as their highest degree increases. The model also forecasts that more women will choose to stay in school and get advanced degrees. In 2035, this will lead to around 7.7 million more women holding advanced degrees and consequently larger population of men holding undergraduate degree and high school diploma.

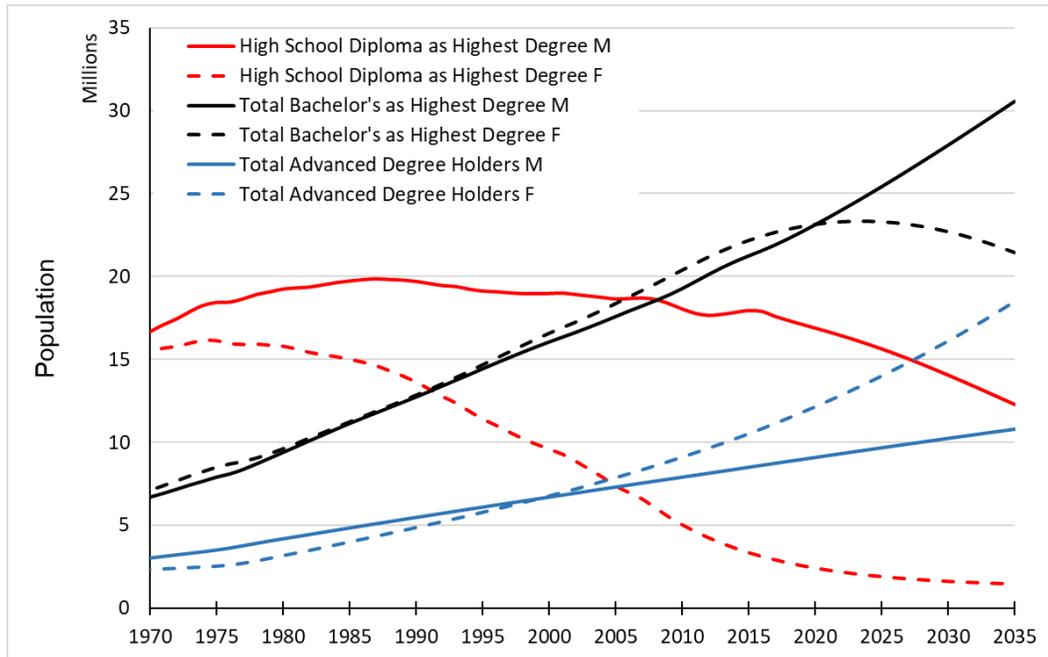


Figure 2.14: Male and female population by highest level of educational attainment from 1970 to 2035.

Figure 15 depicts the ratio of female to male population by their highest degree. Female to male ratio of one shows gender parity in each group. Any point below the dashed line (female to male ratio equals 1) shows male dominated situation, and points above the dashed line are female dominated ones. Female to male ratio of individuals with bachelor's degree as their highest were slightly more than one in 1970 (male dominated situation), then it stays at a steady state until 2020. After 2020, the female to male ratio of individuals with bachelor's degree as their highest degree slightly becomes a male dominated situation (female to male ratio around 0.7 in 2035). Looking at the population with advanced degrees shows that this group is moving toward being female dominated by passing the gender parity level in 1998. Female to male ratio of people with advanced degrees were historically less than one (around 0.75 in 1975). Then the ratio grows over time and the model forecasts it to be around 1.7 in 2035. Population of high school diploma holders was close to parity in 1970 (female to male ratio around 0.9). The ratio decreases over time and becomes male dominated with female to male ratio around 0.15 in 2035.

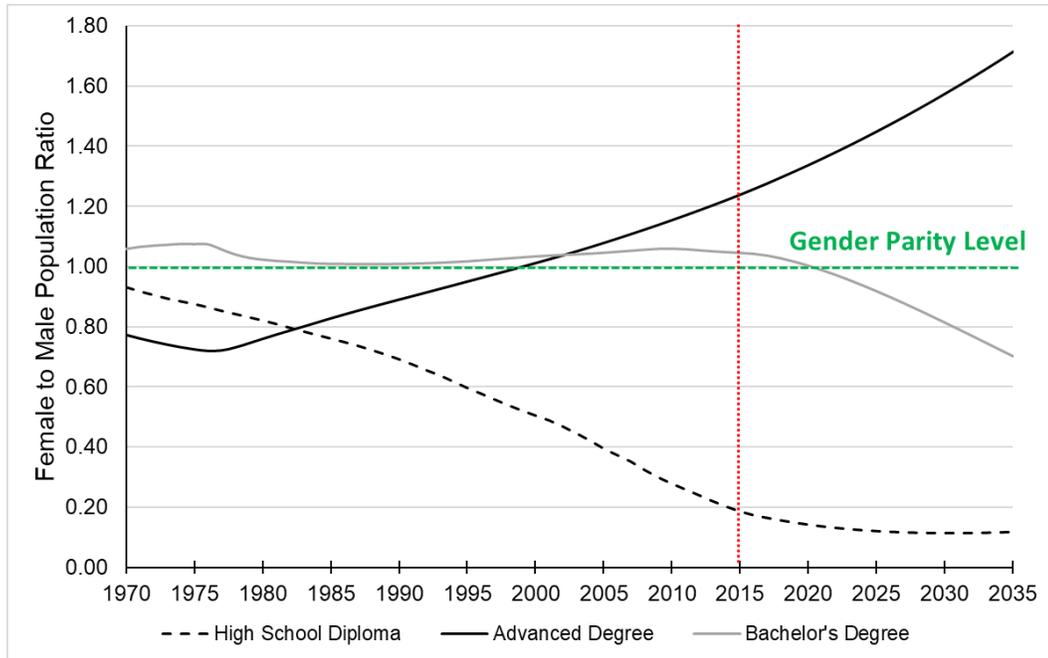


Figure 2.15: Female to male population ratio by highest level of educational attainment from 1970 to 2035.

Discussion

We developed a system dynamics model of the US education system from its earliest stage (pre-kindergarten) to the latest stage (graduate school). We use recruiting, retaining and role modeling as the main feedback structures to develop the model with a broader scope compared to other studies that modeled a smaller section of the education pipeline. The developed model captures the population dynamics of the students and education workforce, focused on the gender composition from 1970 to 2014. We use the same model to predict the gender composition of the educated workforce by degree attainment level (e.g. high school diploma, bachelor's degree, and advanced degree) from 2015 to 2035.

Results of the model show that in the next two decades, women hold 50% of bachelor's and 64% of advanced degrees. Meaning that the gender gap is in favor of women among bachelor's and graduate degree holders. Number of bachelor's and advanced degree holders of both gender

groups will increase, although at different rates. Consequently, the number of females with high school diploma as highest degree decreases over this period since many will pursue higher educations.

The population forecasts of this study are close to the estimations of National Center for Educational Statistics (NCES). NCES performs periodic projections of education statistics using statistical methods, such as regression analysis. The forecasted numbers of bachelor's and advanced degrees awarded is on average 8 percent different between these two studies. For example, they estimated that around 1.91 million bachelor and associate degrees would be awarded to females, while our model forecasts 2.03 million degrees, which are very close estimation.

One policy implication of forecasting the population by educational attainment level and gender is to estimate the future demand for jobs [30]. If population projection model predicts a downward trend in any subgroup of the educated workforce, we can test preventive policies to prevent the creation of further imbalance in the education system. We can also apply the model to estimate the future racial composition of the educated workforce by field of study. Furthermore, forecasting the gender composition or demographics of educated workforce or generally education statistics helps the policy makers to take the required actions to support each group's special needs.

This study contributed to the literature of education policy and gender studies at different levels. First, we offered a systems view of the pipeline of education as affected by several feedback loops. The model has a broad boundary of analysis and does not focus on one or two stages, but includes different education stages from elementary to graduate studies. Then, we used the model to predict the future trend of gender parity in the US. Our findings corroborates with other resources [31, 32] and show that the gender gap in higher education is reversed, which were historically in favor of men.

This study has several limitations that suggest future studies. First, this was a first attempt to model the entire education pipeline, so we did not attempt to model bachelor and advanced degree holders in different fields, and rather aggregated them. We understand that the future trend for different fields will be different, and we suggest future studies to take our model and disaggregate the undergraduate and graduate studies based on fields. Looking at similar gender composition variables at field level might reveal that the gaps between male dominated and female dominated fields are getting wider. We also focused on gender parity; one can further develop the model to study racial disparities and the intersection of racial and gender groups. A valid argument is one that workforce dynamics affect individuals' interest in pursuing higher degrees in long-term. We assumed that workforce dynamics does not change students' interest in higher education in the next two decades and modeling that phenomenon is out of the boundaries of the model, which is limited to the education system. Modeling this effect can be a future addition to the model. Moreover, many questions are remained for future studies. For example, will the change in gender composition translate to more women representation in academic leadership, and full professor positions? Will more representation of women lead to mitigating salary gaps? We hope this study triggers future dynamic modeling works that utilize methods such as system dynamics to study gender disparities in science workforce.

Acknowledgement

The National Institute of General Medical Sciences and the Office of Behavioral and Social Sciences Research of the National Institutes of Health (NIH) supported this work (Grant 2U01GM094141-05). We thank our research team members Ran Xu and Arash Baghaei Lakeh for useful comments on the earlier drafts.

References

1. Skrla L, Scheurich JJ, Johnson Jr JF, Koschoreck JW. Accountability for equity: can state policy leverage social justice? *International Journal of Leadership in Education*. 2001;4(3):237-60.
2. Charles M, Bradley K. Equal but separate? A cross-national study of sex segregation in higher education. *American Sociological Review*. 2002;573-99.
3. Altbach PG, Reisberg L, Rumbley LE. Trends in global higher education: Tracking an academic revolution. UNESCO Pub.; Sense; 2009.
4. Council NR. To recruit and advance: Women students and faculty in science and engineering: National Academies Press; 2006.
5. Vallerand RJ, Fortier MS, Guay F. Self-determination and persistence in a real-life setting: toward a motivational model of high school dropout. *Journal of Personality and Social psychology*. 1997;72(5):1161.
6. Bean JP. Dropouts and turnover: The synthesis and test of a causal model of student attrition. *Research in higher education*. 1980;12(2):155-87.
7. Bettinger EP, Long BT. Do faculty serve as role models? The impact of instructor gender on female students. *The American Economic Review*. 2005;95(2):152-7.
8. Ghaffarzagdegan N, Larson R, Hawley J. Education as a Complex System. *Systems Research and Behavioral Science*. 2016;34:211–5.
9. Bitters BA, Team EM. Useful definitions for exploring educational equity. Wisconsin Department of Public Instruction. 1994.
10. Harvey CP, Allard M. Understanding and managing diversity: Pearson; 2015.
11. Page Scott E. How the power of diversity creates better groups, firms, schools and societies. Princeton; 2007.
12. Corbett C, Hill C, St Rose A. Where the Girls Are: The Facts about Gender Equity in Education: ERIC; 2008.
13. DiPrete TA, Buchmann C. Gender-specific trends in the value of education and the emerging gender gap in college completion. *Demography*. 2006;43(1):1-24.
14. Baker DP, Jones DP. Creating gender equality: Cross-national gender stratification and mathematical performance. *Sociology of Education*. 1993:91-103.
15. Bradley K. The incorporation of women into higher education: Paradoxical outcomes? *Sociology of education*. 2000:1-18.
16. Ceci SJ, Ginther DK, Kahn S, Williams WM. Women in academic science A changing landscape. *Psychological Science in the Public Interest*. 2014;15(3):75-141.
17. Hur H, Andalib MA, Maurer JA, Hawley JD, Ghaffarzagdegan N. Recent trends in the US Behavioral and Social Sciences Research (BSSR) workforce. *PLoS one*. 2017;12(2):e0170887.
18. Ceci SJ, Williams WM. Sex differences in math-intensive fields. *Current Directions in Psychological Science*. 2010:0963721410383241.
19. Buchmann C, DiPrete TA, McDaniel A. Gender inequalities in education. *Annu Rev Sociol*. 2008;34:319-37.
20. Tinto V. Dropout from higher education: A theoretical synthesis of recent research. *Review of educational research*. 1975;45(1):89-125.
21. Garces LM. Aligning diversity, quality, and equity: The implications of legal and public policy developments for promoting racial diversity in graduate studies. *American Journal of Education*. 2014;120(4):457-80.
22. Leslie S-J, Cimpian A, Meyer M, Freeland E. Expectations of brilliance underlie gender distributions across academic disciplines. *Science*. 2015;347(6219):262-5.
23. Moss-Racusin CA, Dovidio JF, Brescoll VL, Graham MJ, Handelsman J. Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*. 2012;109(41):16474-9.
24. Sterman JDJD. Business dynamics: systems thinking and modeling for a complex world 2000.
25. Homer JB. Partial-model testing as a validation tool for system dynamics (1983). *System Dynamics Review*. 2012;28(3):281-94.

26. Fortin NM. The gender wage gap among young adults in the United States: the importance of money versus people. *Journal of Human Resources*. 2008;43(4):884-918.
27. Hahs-Vaughn D. The impact of parents' education level on college students: An analysis using the beginning postsecondary students longitudinal study 1990-92/94. *Journal of College Student Development*. 2004;45(5):483-500.
28. Berger JB, Milem JF. The role of student involvement and perceptions of integration in a causal model of student persistence. *Research in Higher Education*. 1999;40(6):641-64.
29. Statista. Share of international students in the United States in 2012/13, by academic level & gender. The Statistics Portal [cited 2018 4/15/2018]. Available from: <https://www.statista.com/statistics/297146/international-students-us-by-academic-level-and-gender/>.
30. Day JC, Bauman K. Have We Reached the Top?: Educational Attainment Projections of the US Population: Population Division, US Census Bureau; 2000.
31. Berman J. Millennial Women Are More Educated Than Men, But Are Still Paid Less: Status of the Women in the States. Available from: <https://statusofwomendata.org/coverage/millennial-women-are-more-educated-than-men-but-are-still-paid-less/>.
32. Francis DR. Why do women outnumber men in college? : The national bureau of economic research.

Chapter 3. Happy with a baby and a declined tenure:

Effects of tenure-clock extension policies on career outcomes of new parent faculty members

Abstract

Tenure-track faculty members often cope with high levels of stress before promotion to tenured positions. Childbearing and childrearing, which happens in the tenure probationary periods of around 45% of the tenure-track faculty, adds to the stress. Meeting tenure and promotion criteria and raising a newborn child turns these tenure-track professors to servants of two masters, who eventually have to spread their time and energy between two demanding responsibilities. Higher education institutions adopt policies such as parental leaves, reduced teaching and service loads, and tenure-clock extensions to support their employees in one of their most stressful life events. In this paper, we look at the tenure-clock extension policies of the top 250 US universities based on the US news ranking and investigate their characteristics, adoption period, and their effects. We specifically focus on automatic approval or grant of one-year extension to tenure-clock for birth or adoption of a child. We investigate effects of these policies on career outcomes of tenure-track faculty including promotion to tenured position, tenure-track duration, and job satisfaction. The results show that tenure-track faculties stay an average of 6 years on their probationary period. This duration increases by 0.9 years in new parent faculty who are employed in universities with automatic TCE policies compared to their colleagues without a newborn. Staying more on tenure-track does not necessarily lead to improved chance of getting tenure compared to leaving high status academic positions. We observe that faculty members in universities with automatic TCE

policies are generally happier about their jobs than universities without the same policy. We do not find any evidence for a positive effect of automatic TCE policies on the chance of getting tenure.

Keywords: Work-life balance programs, tenure-clock extension, higher education, tenure-track faculty, newborn.

Introduction

Since mid-18th century, when women started attending colleges and earning university degrees, the nation have made many advancements in providing equal educational and employment opportunities for women. As today, women earn an almost equal share of university degrees and even the gender gap is in favor of women in colleges. While there are still gender imbalances in science and engineering fields against women, the overall trend is promising. Women receive half of the doctorate degrees awarded annually by the US higher education institutions (i.e., 46.2% of doctorate recipients in 2015 were female). Field variation, though, still exists: among doctorate degree recipients of math-intensive fields, women are significantly less represented than men. Figure 1 depicts the ratio of female doctorate recipients from 2006 to 2016 by field of study. Field of psychology graduates the highest ratio of female PhDs in the US; health slightly falls below psychology; Biological Sciences relatively awards a lower ratio of the PhD degrees to women (i.e., around 52%), however it is still above the national average (dashed line). The ratio of female doctorate recipients in other fields is less than the average (i.e., around 30%), which includes engineering and sciences.

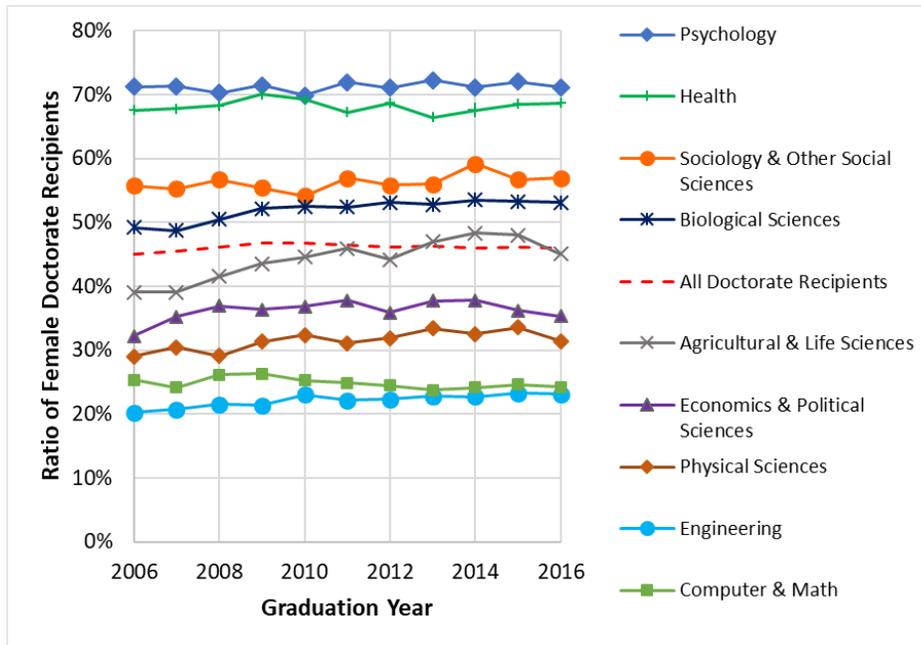


Figure 3.1: Ratio of female doctorate recipients from 2006 to 2016 by field of study.

Despite the narrowed down gender gaps in educational outcomes, women are still among minorities in higher-level managerial and high-status faculty careers [1-4]. They get promoted to managerial jobs less frequently than their male counterparts [1]. Even in academic positions, fewer female tenure-track faculty members are employed. However, the socio-economic processes that lead to educational segregation and those that foster job segregation are different. The gap against female doctorate recipients holds true for all fields; however, the magnitude of this gap varies [5]. Getting a tenure-track position is very competitive, considering the relatively long waiting times in temporary positions such as postdoc (3 years on average) and low chance of getting a tenure-track position (around 17%) [6]. As depicted in Figure 2, there is a declining trend in women's representation as we go from doctorate recipients to tenured faculty. As the figure shows there is also field differences in female PhDs representation. The gap is wider in biological sciences compared to psychology and social sciences. In engineering, there is less than 10% female faculty

among tenured faculty members which is alarming. The disproportionate distribution of tenure-track positions might be due to men’s higher intention in applying for faculty positions than women’s interests [7]. Moreover, women who succeed to land a tenure-track job in academia tend to opt-out of their positions with higher rates, before getting promoted to a tenured faculty position [2]. The pipeline leakage in the transition of female doctorate recipients to professoriate careers [8, 9] can be also a reason for the observed trend. However, we should be careful that the depicted figure is a snap shot picture of the representation, and a dynamic trend may exist since it takes time to see the impact of change in new graduates on more senior faculty positions. In simple words, the composition of more senior faculty members may change as time passes as more PhD students graduate and join tenure-track positions, and more senior faculties (who are disproportionately male) are retired. This is called the “cohort effect” in the literature.

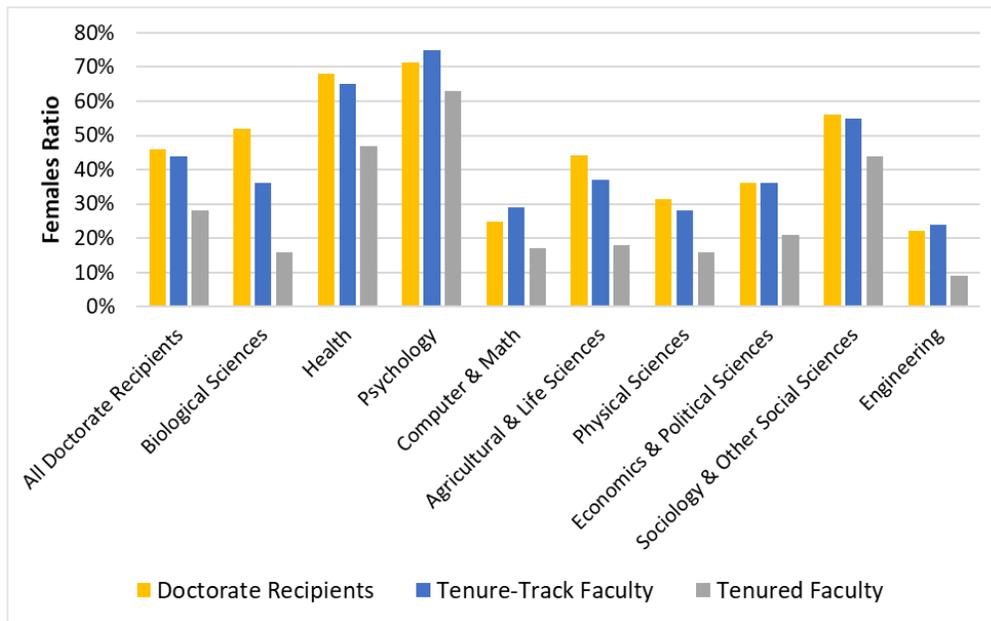


Figure 3.2: Ratio of women among doctorate recipients, new-hire, and all tenure or tenure-track faculty- averaged over 2006-2016.

The historical context of gender representation and discrimination in the society and academia can provide insights into the underlying causes of today's gender gaps. Effects of stereotypical gender roles show a continued presence in women and men's education and employment choices. In the past, women were expected to stay at home, and raise children, while men supposed to be socially and economically active and spend long hours out of home [10]. The social structure of US has changed due to the changes in the economy, family structure, and education [11]. Traditional gender roles gave their place to more egalitarian attitudes [11]. Single-earner families turned into dual-earner families. Women's education level and employment rate improve as the society adopts more egalitarianism gender attitudes [11]. Women's perception of happiness also evolved through the last century. Currently, women who have multiple roles (e.g. childcare, family, and employment) report higher levels of well-being, while traditionally having domestic roles was sufficient to be considered as successful women [10].

The literature attributes the gender gap in high-status academic positions (e.g., tenure-track and tenured faculty) to several factors; stereotypes, gender bias, and family settings. *Stereotypes* about careers in sciences requires individuals who want to become a faculty, to have certain characteristics like innate talent [12]. Gender stereotypes create a mismatch between women and high-status faculty and scientific careers and perceive men to be more compatible with such careers [12, 13]. Furthermore, small portion of female faculty strengthen the stereotypes of women's lack-of-fit to careers in sciences [14]. Women's marginalized presence in science and technology fields is another example of such reinforcing stereotypes and the misinformation [9, 15]. *Gender bias* in interviewing, hiring, publishing, and funding puts women among minorities in high-status academic positions [16]. Male faculty tend to hire less female students and postdocs, especially in elite research groups [17]. Female PhD students tend to receive less financial rewards and

encouragement from faculty during their studies. Therefore, all of the mentioned negative experiences during doctorate studies lead women to *self-select* non-academic careers [16]. However, there are variations in the interest of female PhDs in tenure-track positions and their chance of receiving the offer among different fields.

Family settings including marital status, parental status, taking care of elder family members are widely cited as the reason behind underrepresentation of female faculty in academia. The chance of having a spouse/ partner in women faculty is lower than men faculty. Spouses/partners of those women faculty spend more time on household tasks [18]. There are contradictory evidence regarding the effect of children on women's career choices and advancements. A group of studies report that children do not negatively affect women's tenure and promotion status as well as their research productivity [19], while having a child benefits male faculty promotion [20]. On the other hand, a second group of studies report that having children and the career choice that it imposes put a downward pressure on mothers' rather than fathers [9, 20, 21] and parenthood eventually leads to self-exclusion from academic careers [7].

Higher education institutions have interest in improving women's presence among their faculty members. They design and adopt policies to attract and improve retention rates of faculty and assist them in balancing their work-life responsibilities. The NSF ADVANCE programs are designed to promote women's presence in academic science and engineering careers. Between 2001 and 2016, more than 160 US higher education institutions received the ADVANCE grants. Work-life support programs such as dual career support, parental leave, and tenure-clock extension (TCE) policies are among these policies. The goal of dual career support is to retain the newly hired tenure-track faculty and moreover attract another faculty to the department. A study on dual career academic couples states that for these academic couples, job satisfaction is higher, working

hours is higher, stress level is lower and balancing work-life responsibilities is easier [22]. However, the support mechanisms do not always function in the desired way. There is stigma or bias against the spouse who is hired after his/her spouse. The second partner is viewed as being inferior to the spouse as well as to the rest of the colleagues, called the “trailing spouse.” Tenure-clock extension policies (TCE) are being implemented by higher education institutions for around 40 years [23]. Delaying tenure review might cause a backlash among faculty against the colleague who is using these policies [23, 24]. This might lead to a similar stigma and fear against using TCE policies due to child birth, child care or adoption [25]. Recently universities started to adopt a new level of TCE policies due to childbearing and childrearing. They automatically extend the tenure probationary period for one year as the tenure-track faculty notifies them of birth or adoption of a child. Their goal is to fight the previously explained stigma among faculty members [25]. There are some variations in the details of the automatic TCE policies and the way that universities operationalize them. In the next section, we will explain the differences by bringing examples and clarify what we mean by auto TCE policies.

Past studies investigated the effects of extending tenure-clock on different career outcomes of tenure-track faculty. In a large research university, delaying tenure review had a salary penalty for the faculty. However, the tenure-clock extension did not have a negative effect on their promotion rate to tenured positions [23]. Another institution-level research focused on tenure-track faculty of university of Wisconsin-Madison in 2003. The results showed that job satisfaction even declines in faculty who use tenure-clock extensions and the negative effect exists for both gender groups [25].

In this paper, we seek the answer to two questions; (1) which universities adopt the automatic tenure-clock extension (TCE) policy? (2) How does automatically extending the tenure

probationary period of new parents affect their career outcomes? To answer these questions, we conduct the study in two steps. First, we take a statistical analysis approach to investigate the characteristics of universities which adopt automatic TCE policies. For this purpose, we construct a unique dataset of the levels of TCE policies among the top 250 US universities according to the US news ranking. In the second part, we also take a statistical approach to test the effect of automatic TCE policies on the career outcomes of tenure-track faculty who birth or adopt a child. For this purpose, we use the Survey of Doctorate Recipients (SDR), conducted by National Science Foundation (NSF), of 2008, 2010, 2013, and 2015 and connect the dataset to our institution-level dataset. We discuss policy implications.

Automatic Tenure-Clock Extension Policy

Before moving to methods, we clarify what we mean from automatic TCE policy in higher education institutions. Figure 3 depicts the variations on TCE policies of US universities. In most of universities that have a tenure system, faculty who encounter any difficult family circumstances including birth or adoption of a child can request an extension to their tenure-clock (shown by box A1 in Figure 3). In other words, they can request a one-year delay in their tenure review date. Some universities have a case-by-case review process for these requests and it has to be approved by the department head, the appropriate dean and the provost (box B2 in Figure 3). For instance, new-parent faculty in Georgia State University should request the stoppage and it must be approved by the Provost. On the other hand, others automatically extend the tenure probationary period for new parent faculty (box B1 in Figure 3). We acknowledge that the policies of universities that we considered as automatic TCE adopters are not exactly identical. The first group of auto TCE adopters require the faculty to request the extension and then the university approves the extension automatically (box C1 in Figure 3). For example, Virginia Tech modified its stop-

the-clock policy and implemented automatic TCE policy in 2005 for both genders but requires a request form to be submitted to the office of the provost [24]. The second group of universities even do not require any formal request and documentation from the faculty (C2 in Figure 3). They extend the probationary term of the tenure-track faculty member as soon as the faculty takes a family leave; however, the faculty may later deny the granted extension. For example, Princeton University automatically extends assistant professors' tenure-clock by one year for each childbirth or adoption of a child, upon notification. The policy is the same for men, women and same-sex partners [26]. The University of Missouri System is another example of the second group of policies. The tenure-clock of new-parent faculty members will be automatically extended by one-year when they take family and medical leaves [27]. In this paper, we take both of these groups as automatic TCE adopters or the treatment group of the analysis. We take universities as automatic TCE policy adopters if they automatically approve the request regardless of their requirement for submitting formal requests.

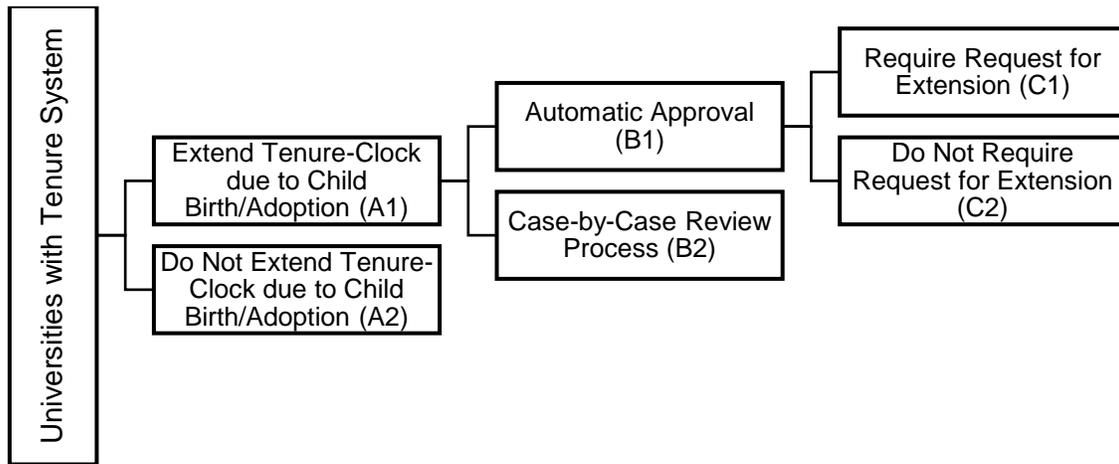


Figure 3.3: Variations in tenure-clock extension policies in the US universities.

Method

Institution-level Dataset

We construct a distinct dataset of TCE policy adoption in the US higher education institutions. We specifically focus on the top 250 universities based on the US news ranking. We used a mix of online resources, email communications and short interviews to collect the data about TCE policies in universities. The data are collected according to faculty handbooks or other online resources describing the related policies in each university. In cases of ambiguity in online documents, we contacted either the office of human resources or the office of the provost and requested more information. In several cases, we had a chance to have short interviews with the provosts of some of the universities which resulted in more insights on the policy and its operationalization on campuses. We accomplished to collect data on 204 universities policies. In short, the constructed dataset includes the following variables at institution-level: ranking, institution type (private vs. public), geographic region (e.g., West, South, Northeast, and Midwest), political orientation of the state in 2016 presidential elections (democratic vs. republican), institution size, and Carnegie classification (framework of classifying colleges and universities in the U.S. based on factors such as degree level and research intensity). Another variable, called “AutoExt” shows whether each university adopts automatic TCE policy. We consider “AutoExt” (automatic TCE) to be equal to 1 if a university automatically extends the tenure probationary period of tenure-track faculty who birth or care for a newborn or adopted child regardless of the requirement for formal request. Furthermore, we include whether each university requires request for extension and the approval is automatic or they will automatically stop the tenure-clock when the faculty member takes a family leave. We also collected the year when the policy first was in

place in each university. Descriptive statistics of the institution-level dataset is shown in Table B.1 of Appendix B.

Individual-level Dataset

We build the second part of the dataset, individual-level dataset, by connecting the institution-level dataset to Survey of Doctorate Recipients (SDR) conducted by National Science Foundation (NSF). The data include individual-level information about tenure-track faculty who participated in the Survey of Doctorate Recipients (SDR) by National Science Foundation (NSF). We include individual data if they meet two criteria. First, the individuals should work in one of the institutions of the institution-level dataset during 2008-2015 period. Therefore, we know whether they had the option of using automatic TCE before 2008 or not. Second, they should have responded to the SDR in years 2008, 2010, 2013 and 2015. The latter allows us to generate measures of progress in each individual's career outcomes. These measures include their promotion to tenured position, salary raise, likelihood of receiving government funding, and change in their job satisfaction level. Furthermore, we control for other independent variables including age, gender, race, having a newborn baby, time since graduation, and hours worked during each week.

Table 1 shows the descriptive statistics of the individual-level dataset. Out of total 487 data points of this dataset, 205 are female and 282 are male. We compare the mean values of different variables between male and female subjects and present the p-values. Hours worked and probability of being married is significantly higher in the male sub-sample compared to females.

Table 3.1: Descriptive statistics of individual-level dataset and t-test between gender groups.

	Female (N=205)		Male (N=282)		p-value*
	Mean	SD	Mean	SD	
Salary	90,539.70	69,724.75	96,485.08	44,389.70	0.252
Job Satisfaction	3.39	0.73	3.36	0.65	0.663
Any Government Funding	0.42		0.47		0.211
Promotion					
Tenured	0.74		0.77		0.452
Left TT	0.11		0.10		0.781
Still on TT	0.15		0.13		0.491
Tenure-Track Duration	6.14	1.63	6.02	1.35	0.427
Race					
White	0.60		0.60		0.864
Asian	0.11		0.17		0.074
Hispanic	0.13		0.13		0.922
Black	0.09		0.07		0.402
Other Races	0.07		0.02		0.011
Age (2008)	37.40	6.39	37.11	6.24	0.621
Newborn	0.41		0.47		0.242
Married	0.82		0.88		0.065
Cohort	2002.68	3.51	2002.45	3.85	0.495
Start Year	2005.69	2.31	2006.00	2.13	0.125
Hours Worked	203.40	35.05	212.77	38.63	0.006

* p-value for the null hypothesis of no gender difference.

Figure 4 illustrates the rate at which individuals in the sample get promoted to a tenure position or leaving high status jobs for other employment opportunities over the study period. We have four data points per individuals from years 2008, 2010, 2013, and 2015. In the first data point all individuals are tenure-track faculty and over time. With different starting points, except a small percentage, they either get promoted or leave the position. According to the figure, in 2015, 9% of tenure-track faculty members are still in the same position (not yet promoted to tenured), 75% are promoted to tenured position and 16% leave the position for other types of employment in academia, industry or government.

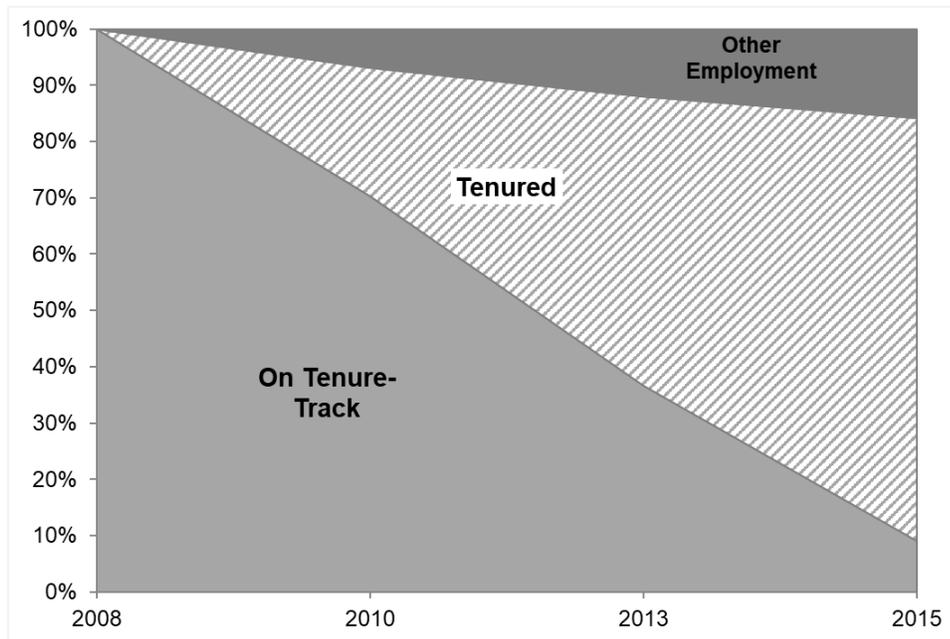


Figure 3.4: Change in tenure status and employment from 2008 to 2015 according to individual-level dataset.

Model

We conduct two rounds of analysis to answer the research questions. In the first round, we investigate the likelihood of adopting automatic TCE vs. non-automatic TCE using a logistic regression model. We exclude universities that do not have any TCE policy. The dependent variable is binary and takes a value of zero if the university reviews the tenure-clock extension files on a case-by-case basis regardless of its requirement for formal request of tenure-clock extension by the faculty member. We are interested to know whether different characteristics of higher education institutions correlate with the probability of adopting auto TCE policy. Therefore, we take the following variables as the independent variables of the analysis: being located in a state with democratic political orientation, ranking, campus size (e.g. number of students enrolled), being private, type of campus (e.g., urban, rural, etc.), and Carnegie classification affects their policy adoption level.

In the second round of the analysis, we use the individual-level dataset to investigate the effect of automatic TCE policy on career outcomes using different regression models. Our outcome variables are duration of tenure-track, promotion to tenured position, salary raise, likelihood of receiving government funding, and their job satisfaction level are the dependent. We run logistic regression for promotion to tenured position which is a binary variable. It takes a value of 1 if the faculty gets promoted to a tenured position and zero if he/she leaves the tenure-track position for temporary academic or non-academic positions. For continuous variables (e.g., tenure-track duration, salary raise and likelihood of receiving government funding), we use multivariable linear regression. Finally, the job satisfaction variable is in a 1-4 scale (4 highest). We run ordinary least-square regressions for job satisfaction. Moreover, we run each regression for female and male subgroups separately. It is worthwhile mentioning that we do not use the weights of the survey of doctorate recipients' based on the suggestion of U.S. Bureau of Labor Statistics regarding connecting multiple waves of a longitudinal survey [28].

Results

Which institutions adopt automatic tenure-clock extension policy?

Let us first start with answering the following question: “Which institutions adopt automatic tenure-clock extension policy?” As mentioned before, we collected data related to TCE policy in the top 250 universities. We accomplished to collect 204 data points through online documents and university employees. Out of these 204 universities, 112 institutions (55%) have the tenure-clock extension option for new parent faculties but it has to be requested by faculty and approved by the dean and the provost. Eighty two universities (40%) have adopted some form of automatic TCE policy and 10 institutions (5%) do not have any form of TCE. A larger group of the institutions require the new-parent faculty to request the tenure-clock extension and the approval

is automatic (N=55). While, a smaller group in our dataset automatically extend the tenure-clock when the faculty member takes a family or medical leave (N=27). We also collected the date that the automatic TCE policy was first in place in each university. We received answers from 74 universities. Therefore, we miss 8 adoption dates from total 82 universities that adopt automatic TCE policy. It is interesting to investigate the timeline of adopting automatic TCE policy. Figure 3.5 shows the number of universities that adopted the auto TCE policies over time. According to the policy adoption graph, in late 1980's there were fewer universities with automatic TCE policy, which are operated under the University of California System. A relatively larger number of the higher education institutions joined the adopters around 2007- 2008 that reported to be a result of American Association of University Professors (AAUP) collective bargaining.

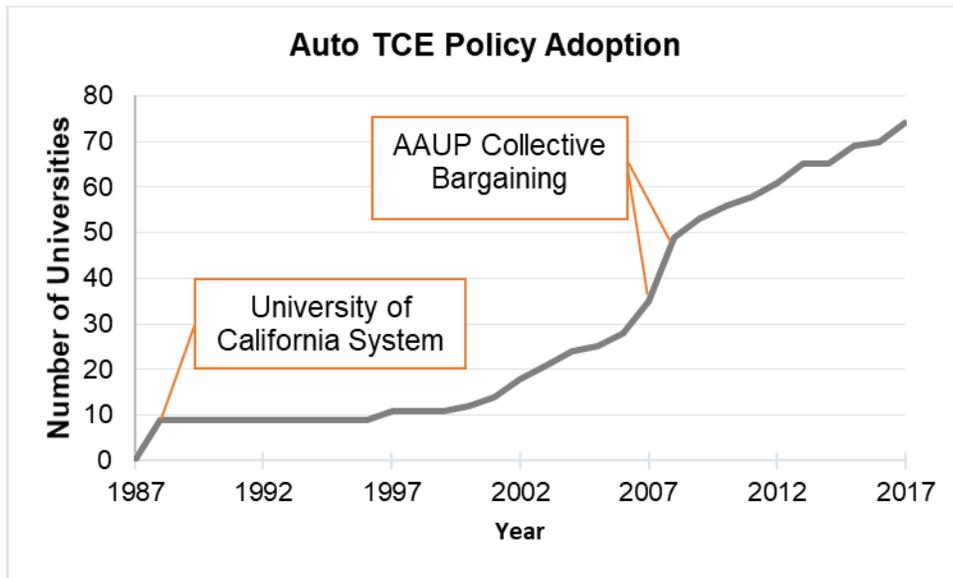
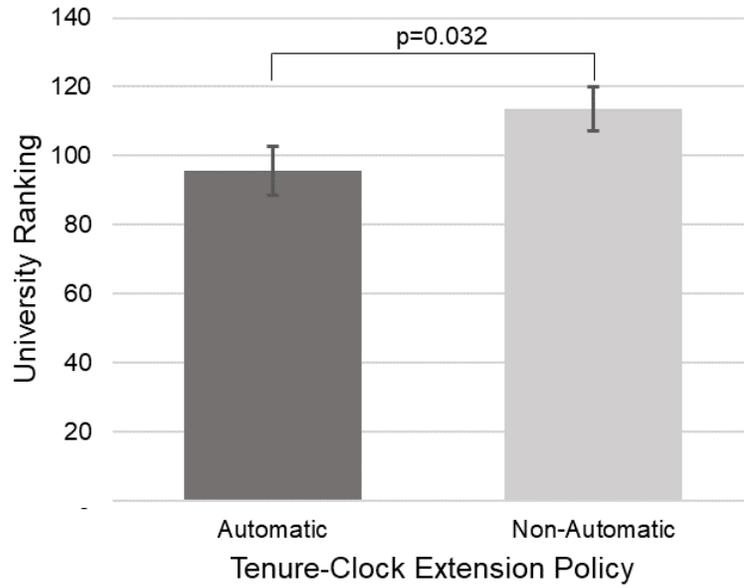


Figure 3.5: Number of automatic TCE policy adopter universities over time- policy adoption.

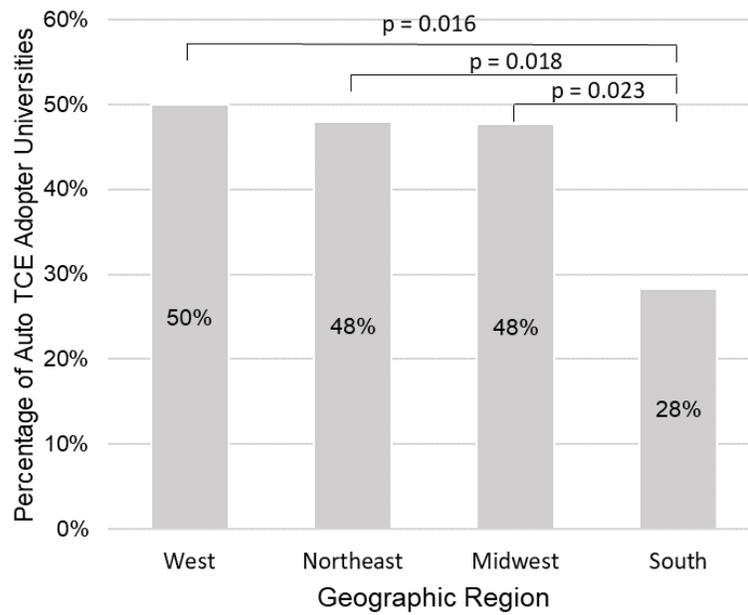
Note: AAUP: American Association of University Professors.

As mentioned in the method section, we run a logistic regression for different combinations of independent variables (ranking, geographic region, Carnegie classification, institution size, private indicator, and location). We choose the best logistic model based on the results of a

stepwise regression analysis. The final model has ranking of universities and geographic region as its only two independent variables. Pseudo R^2 is 0.041. Figure 6 depicts the selected characteristics of universities that adopt automatic TCE. Ranking of the university is significantly correlated with the likelihood of adopting automatic tenure-track extension policy. The odds ratio is less than one (i.e., equal to 0.996) which means that universities with a higher number for ranking (low ranked universities) are less likely to adopt automatic TCE policy (Figure 6.a). In other words, the probability of adopting automatic TCE policy drops by 5% if university's ranking is 10 ranks worse. The average ranking for adopters and non-adopters are 96 and 113 respectively ($p=0.032$). Moreover, if we add the geographic region of the state to the regression, ranking stays significant (p -value is 0.076). The odds ratio of adopting automatic TCE policies in universities located in south geographic region is 0.41 compared with northeastern institutions ($p=0.036$). As depicted in Figure 6.b, fewer southern universities adopted automatic TCE policies. The figure depicts the p -values from the t -test between each pair of geographic regions.



(a) Average university ranking by level of TCE policy adoption



(b) Percentage of universities with auto TCE policy by geographic region

Figure 3.6: Selected characteristics of universities by level of tenure-clock extension (TCE) policies.

Effect of automatic TCE on career outcomes

Next interesting question is whether automatically extending the probationary period of new parent tenure-track faculty, helps faculty members in achieving better career outcomes or at least prevents their productivity from dropping. We investigate the effect of automatic TCE on tenure-track duration, likelihood of promotion (i.e., getting tenure), job satisfaction change, likelihood of receiving government funding, and salary raise in a seven years period (i.e., 2008 to 2015). For this purpose, we use the constructed individual-level dataset.

One of the goals of adopting automatic TCE policies is to remove the stigma attached to extending tenure probationary period due to birth or adoption of a child. We hypothesize that if the automatic TCE policies are successful in removing the bias or stigma, the average time in tenure-track position should be longer for individuals with a newborn baby. We run a multivariable linear regression to test the effect of automatic TCE policies on the duration that a tenure-track faculty stays in the position until promotion to a tenured position or leaving tenure-track position. Table 2 shows the results of the regression analysis. The sample is limited to individuals who are either promoted to tenured position or left tenure-track position. Regressions M1 to M3 include different sets of independent variables other than the core independent variables which are AutoExt (automatic TCE policy indicator), Newborn, the interaction effect of the former variables, and university ranking. We include ranking of the university as an indicator of the difficulty of meeting tenure criteria for pre-tenure faculty or tenure-track pressure. In all three models, the main effect of AutoExt and the interaction of AutoExt and Newborn are significant ($p < 0.05$).

We take M3 with highest R^2 of 0.053 as the final model and discuss the results accordingly. The intercept of the model is 6.2 years. Average tenure-track duration of four groups are depicted in figure 7. The mean time in tenure-track position for (a) faculty with *no newborn* and employed

in a university with *auto TCE policy*, (b) faculty with *newborn* and employed in a university with *auto TCE policy*, (c) faculty with *no newborn* and employed in a university with *non-auto TCE policy*, and (d) faculty with *newborn* and employed in a university with *non-auto TCE policy* are relatively 5.7, 6.3, 6.4 and 6.1 years. The interaction effect of AutoExt and Newborn which is shown by the coefficient of AutoExt X Newborn is 0.92. This means that a tenure-track faculty with newborn baby who works in an institution with automatic TCE policy stays 0.92 years longer than another faculty who works in the same institution and does not have a newborn.

The mean chance of getting tenure for (a) faculty with *no newborn* and employed in a university with *auto TCE policy*, (b) faculty with *newborn* and employed in a university with *auto TCE policy*, (c) faculty with *no newborn* and employed in a university with *non-auto TCE policy*, and (d) faculty with *newborn* and employed in a university with *non-auto TCE policy* are relatively 0.94, 0.82, 0.89 and 0.92.

Table 3.2: Results of multivariable linear regression for tenure-track duration.

	M1	M2	M3
	Coef. (Std. Err.)	Coef. (Std. Err.)	Coef. (Std. Err.)
AutoExt	-0.68 ** (0.30)	-0.76 ** (0.32)	-0.80 ** (0.33)
Newborn	-0.26 (0.25)	-0.28 (0.27)	-0.27 (0.28)
AutoExt X Newborn	0.82 * (0.43)	0.84 * (0.45)	0.92 ** (0.26)
Female		0.10 (0.22)	0.05 (0.23)
University Ranking		0.0006 (0.0017)	-0.0005 (0.0017)
Hours Worked			0.001 (0.003)
Race (ref: White)			
Asian			-0.21 (0.31)
Black			-0.02 (0.49)
Hispanic			0.20 (0.32)
Native Hawaiian			-0.30 (0.99)
Multiple Race			0.48 (0.52)
Intercept	6.38 *** (0.18)	6.48 *** (0.30)	6.17 *** (0.74)
AIC	595.8	565.7	575.4
BIC	608.4	587.3	612.5
R-squared	0.032	0.040	0.053
Observations	173.0	162.0	162.0

* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

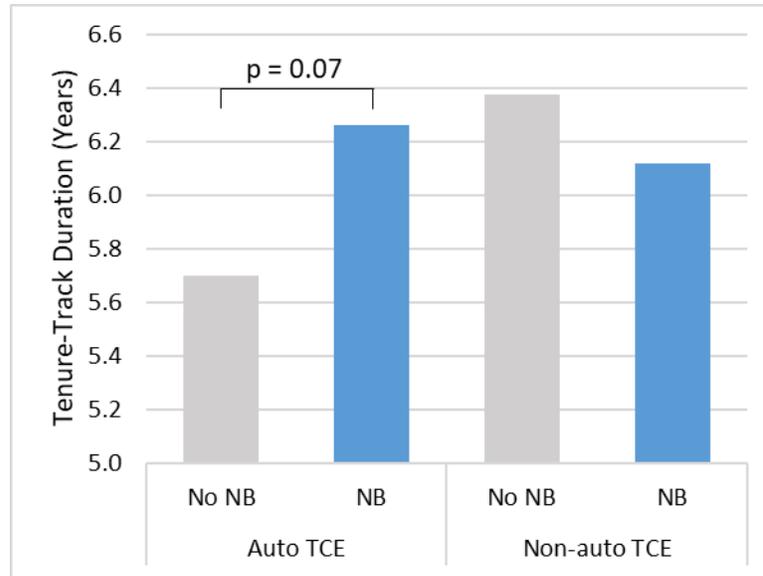


Figure 3.7: Average tenure-track duration by newborn and auto TCE policy.

Note: NB: faculty who have a newborn, No NB: faculty who do not have a newborn.

We are also interested to know if adopting automatic TCE policy in an institution improves the likelihood of getting tenure for tenure-track faculty. We also test whether staying longer on a tenure-track position improves the chance of getting tenure compared to those faculty who leave these positions. We run binomial logistic regression for a binary dependent variable called “Tenured” which takes 1 if the faculty is promoted to a tenured position and takes 0 if the faculty leaves tenure-track position to other jobs in academia, industry or government. Table 3 shows the results of the regression analysis for three different combinations of independent variables. M1 only tests for the main effects of AutoExt, having newborn as well as the interaction effects. We also control for gender, hours worked, and university ranking in M1. Model M2 is similar to M1 except it controls for tenure-track duration too. In M3, we also add the effect of racial group to model M2. Automatic TCE policy and its interaction effects with newborn and older child variables do not have a significant effect on the likelihood of getting tenure. Also, the duration of

tenure-track does not show a significant correlation with the chance of getting tenure. Figure 8 depicts the chance of getting tenure by different combinations of AutoExt and Newborn. No significant difference is observed between the means.

Table 3.3: Binomial logistic regression results for getting tenure.

	M1	M2	M3
	Odds Ratio (Std. Err.)	Odds Ratio (Std. Err.)	Odds Ratio (Std. Err.)
AutoExt	2.20 (1.93)	2.16 (1.92)	2.13 (1.90)
Newborn	2.20 (1.52)	2.26 (1.59)	2.34 (1.65)
AutoExt X Newborn	0.14 * (0.16)	0.17 (0.20)	0.17 (0.20)
Female	0.75 (0.41)	0.71 (0.39)	0.70 (0.40)
Start Year	1.12 (0.14)	1.19 (0.18)	1.18 (0.18)
Hours Worked	1.00 (0.01)	1.00 (0.01)	1.00 (0.01)
University Ranking	1.00 (0.004)	1.00 (0.004)	1.00 (0.004)
Race (ref: White)			
Asian			0.66 (0.49)
Black			0.76 (0.87)
Hispanic			0.67 (0.50)
Multiple Race			0.74 (0.87)
Tenure-Track Duration		1.11 (0.25)	1.11 (0.25)
AIC	120.4	117.9	125.0
BIC	145.2	145.7	165.0
Pseudo R-squared	0.044	0.044	0.048
Observations	164.0	164.0	160.0

* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

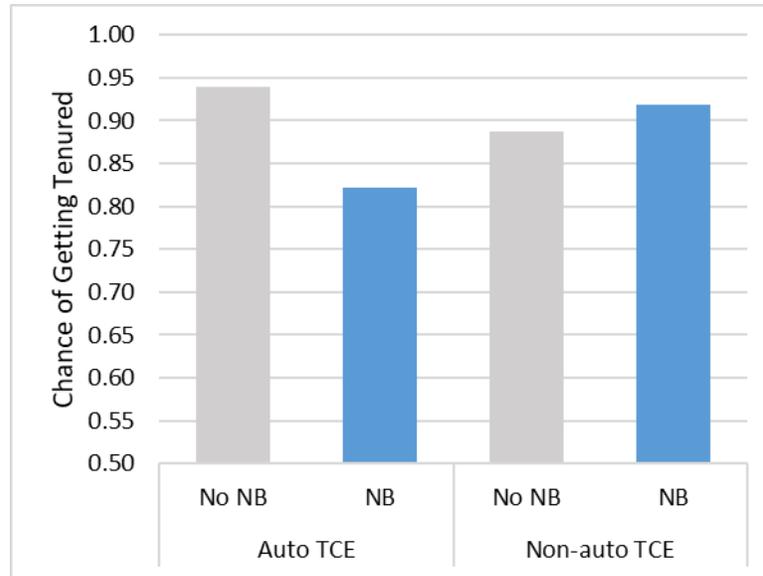


Figure 3.8: Chance of getting tenure (compared to leaving tenure-track position) by having newborn and auto TCE policy.

NB: faculty who have a newborn, No NB: faculty who do not have a newborn.

The results of the previously described analyses show that automatic TCE policy significantly improves the time that a faculty stays on tenure-track but it does not significantly affect the chance of getting tenure compared to leaving high status academic positions. Job satisfaction of tenure-track faculty is also another interesting outcome in a sense that work-life balance programs seek the well-being of the university employees and the community. SDR’s questionnaire assesses the job satisfaction of each individual on a 1 to 4 scale. Therefore, we use an ordered logistic regression to test the effect of automatic TCE policies and other variables of interest (Appendix B, Table B.2). Table 4 shows the results of the regression for job satisfaction in 2015. All three models test the effect of core independent variables; AutoExt, newborn, the interaction effect and gender. We also control for initial job satisfaction reported in 2008 and geographic region of university. Models M1 to M3 are different in the set of control variables.

We take M2 with R^2 of 0.22 and AIC of 373.6 as the final model and discuss the results accordingly. Overall, faculty who work in universities that adopt auto TCE policy report 0.28 points higher job satisfaction on average ($p < 0.01$). Moreover, having newborn children improves the job satisfaction in tenure-track faculty by 0.25 points on a 1 to 4 scale ($p < 0.05$). The interaction term of AutoExt and Newborn is -0.33 ($p < 1.0$). This means that faculty who have a newborn and work in an institution with automatic TCE policy report 0.33 points lower job satisfaction than their colleagues who do not have a newborn during their pre-tenure period. Figure 9 shows the average job satisfaction by AutoExt and Newborn combinations. New parent Tenure-track faculty who work in non-adopter universities significantly report higher levels of job satisfaction compared to their colleagues with no newborn children ($p = 0.01$). The mean job satisfaction for (a) faculty with *no newborn* and employed in a university with *auto TCE policy*, (b) faculty with *newborn* and employed in a university with *auto TCE policy*, (c) faculty with *no newborn* and employed in a university with *non-auto TCE policy*, and (d) faculty with *newborn* and employed in a university with *non-auto TCE policy* are relatively 3.6, 3.5, 3.2 and 3.5. No gender heterogeneity is observed in the models. It is worthwhile mentioning that we test the effects of the same sets of independent variables on two other career outcome variables (e.g., individual's salary in 2015 and an indicator for having government funding support). No significant effect of AutoExt, children and gender is observed on salary and funding. Tables B.3 and B.4 of Appendix B show the results of the regression analyses.

Table 3.4: Multivariable linear regression for job satisfaction.

	M1		M2		M3	
	Coef. (Std. Err.)		Coef. (Std. Err.)		Coef. (Std. Err.)	
AutoExt	0.33 (0.11)	***	0.28 (0.12)	***	0.26 (0.12)	**
Newborn	0.27 (0.10)	***	0.25 (0.10)	**	0.24 (0.10)	**
AutoExt X Newborn	-0.38 (0.17)	**	-0.33 (0.17)	*	-0.29 (0.17)	*
Female	0.13 (0.08)		0.11 (0.09)		0.10 (0.09)	
Race (ref:white)						
Asian			-0.31 (0.11)	***	-0.26 (0.11)	**
Black			-0.07 (0.18)		-0.04 (0.18)	
Hispanic			-0.13 (0.13)		-0.14 (0.13)	
Native Hawaiian			0.03 (0.42)		0.11 (0.42)	
Multiple Race			-0.08 (0.22)		-0.02 (0.22)	
Hours Worked			-0.002 (0.001)	*	0.00 (0.001)	*
Ranking			-0.001 (0.001)	**	0.00 (0.0006)	**
Initial Job Satisfaction	-0.30 (0.07)	***	-0.27 (0.07)	***	-0.27 (0.07)	***
Major Field (ref:Computer & mathematical sciences)						
Biological, agricultural and environmental life sciences					0.33 (0.17)	**
Physical sciences					0.13 (0.16)	*
Social sciences					0.09 (0.18)	
Engineering					0.37 (0.22)	
South Region	-0.22 (0.09)	**	-0.20 (0.09)	**	-0.18 (0.09)	**
Intercept	3.68 (0.13)	***	4.36 (0.29)	***	4.15 (0.34)	***
AIC	409.2		373.6		374.1	
BIC	432.9		420.1		437.3	
R2	0.155		0.221		0.256	
Observations	220		206		206	

* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

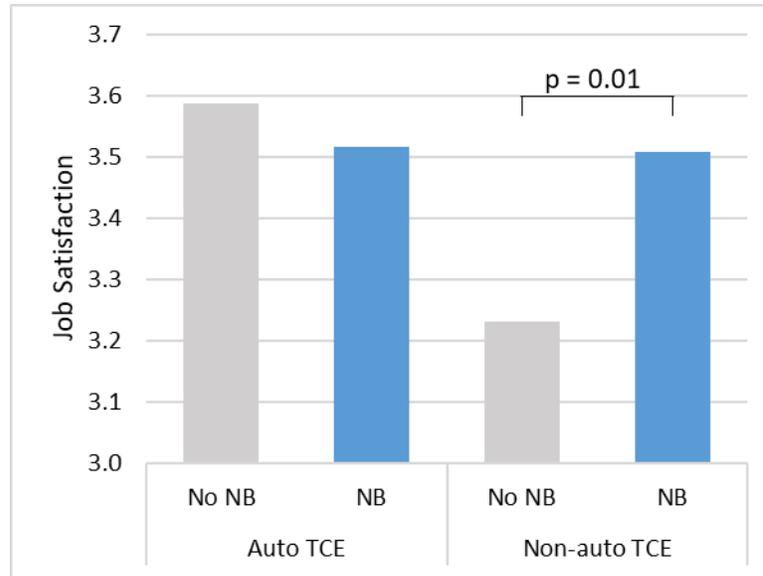


Figure 3.9: Average job satisfaction by having newborn and auto TCE policy.

Note: NB: faculty who have a newborn, No NB: faculty who do not have a newborn.

Discussion

Work-life balance can be challenging for new-parent faculty who are in their pre-tenure period. Tenure-clock extension policy gives them the option of postponing the date of their tenure review for each child birth or adoption. However, new-parent tenure-track faculty might be reluctant to apply for tenure-clock extension because of the possible backlash from their colleagues. Universities automatically grant a one-year extension to the tenure probationary period of their pre-tenure faculty upon birth or adoption of a child to remove the stigma attached to requesting extensions. In some universities, the automatic extension of tenure-clock is defined as an automatic approval of the extension requests, and the faculty members who have a new child should request the extension in order to be automatically approved. While in the other group of universities, tenure-clock is extended automatically when the faculty member takes a family leave. Therefore, the faculty member does not need to fill an extension request form or take a parental

leave to use the tenure-clock stop option. According to the results of our short interviews with provosts of several universities, the difference is mainly due to the various ways that universities operationalize the flexibility in TCE policies. Human resources department is often responsible for processing the leaves (including parental and pregnancy leaves) and tenure-clock extension/stopping is related to the office of the provost. Sometimes the reason that the university requires a request is the difficulty in coordination of these two offices. In other cases, universities might not have any leave specific to new parents or family leaves. They have sick leave and in case of adoption, the new parent cannot use sick leave. So, they require the faculty members to request an extension, unless there would be no other way that the university determines the occurrence of life changes. Considering how universities state such flexibility in their policies and operationalize it, there is variation in the procedure of extending tenure probationary period for new parent faculties. A large group of universities (around 67 percent of all adopter institutions) require a request for tenure-clock extensions. In this paper, we consider both of the previously described groups as universities with automatic TCE policy.

Twelve higher education institutions started adopting automatic TCE policy in late 1980's including the University of California System, Michigan State University and University of Wisconsin-Madison. The adoption rate picks around 2007-2008 and according to our data; 21 universities adopted the policy in these years. Finally, 74 institutions among the 204 collected data points adopted the policy since 2017. Our analysis of the characteristics of higher education institutions that adopt automatic TCE policy show that higher ranked universities are more likely to follow these policies. In accordance with the results of this study, universities with best work-life balance programs are more high-ranked universities such as Brown University and Yale University [29], the Universities of Notre Dame, University of Texas at Austin, University of

Illinois at Urbana Champagne and Virginia Tech [30]. Moreover, looking at the geographic region of the adopter institutions shows that those located in southern states of the United States are less likely to use these policies, even controlling for their ranking.

Other than investigating the characteristics of these universities, it is important to understand the impact of the policy on the career outcomes of new parent faculty. According to our current knowledge, this study was the first attempt to investigate the effect of automatic TCE policy on career outcomes of faculty at national level. We studied the duration of tenure-track until getting tenure or leaving high status academic positions as the initial outcome of such policies. Furthermore, these policies are designed to improve the retention rate of tenure-track faculty. Thus, we investigated the likelihood of getting promoted to a tenure position in our sample. Finally, well-being of faculty is another important objective of work-life balance programs including automatic TCE policies. So, we investigated the reported job satisfaction based on the Survey of Doctorate Recipients.

The results show that the average tenure-track duration in universities that adopt the automatic TCE policies is smaller than the ones that require approval. Tenure-track faculties spend on average 6 years in their position until getting tenure or leaving high status faculty positions to temporary academic, industry or government employment opportunities. Those who birth or adopt a newborn child during their probationary period and work in institutions that adopt automatic TCE policy stay around 0.92 years more than their colleagues without a newborn. We did not observe any gender heterogeneity in these effects. However, one of the widely mentioned objectives of making automatic TCE policies gender neutral is to encourage faculty of both genders to contribute to childcare responsibilities in their families. Higher education institutions

that automatically extend the tenure-clock of faculty who birth or adopt a child are successful in removing the stigma attached to requesting the extension.

New parent tenure-track faculty employed in universities with automatic TCE policies stay longer on tenure-track, however, the results indicate that increasing the duration of pre-tenure position does not affect the probability of ultimately getting tenure and meeting the tenure criteria. We expected that universities with automatic TCE policies have higher faculty retention rates the same as the results of a study of work-life policies in top-ten medical schools [31]. However, we do not find any significant of automatic TCE policies in our sample. A study on tenure-track faculty in a doctorate-granting university found no significant effect of TCE policy on faculty retention rate [23]. Although, we cannot determine that the reason of leaving high status academic positions and starting a non-academic job is necessarily the person's lower ability in meeting tenure criteria or him/ her being unable to keep a desired work-life balance. Faculty who leave academia might have different concerns like change in spouse's job location, and desire for higher salary. We acknowledge that one might say that for observing the effect of automatic TCE policies on the likelihood of getting tenure the timeframe of applying automatic TCE policy should be more. Future research can further investigate the effects of automatic TCE policies or other work-life balance programs on the chance of getting tenure which highly depends on data availability. Performing the same analysis on a sample of faculty who left academia because of family concerns can be a better indicator of these effects.

As mentioned before, one of the main objectives of work-life balance is to improve the well-being of employees and the community. According to the results, generally having a newborn child improves job satisfaction of tenure-track faculty regardless of their gender. Furthermore, faculty members that work in universities with automatic TCE policy, report higher levels of job

satisfaction. These universities might have a set of work-life balance programs that improve the overall job satisfaction among all faculty members. New parent faculty who work in higher education institutions that do not adopt automatic TCE policies report higher job satisfaction, while having a newborn among faculty who are employed in adopter universities slightly decreases job satisfaction. Another study found that job satisfaction in faculty members who use TCE policies declines [25], which is in accordance with our finding in universities with automatic TCE policy. Working more hours also reduce the reported job satisfaction in tenure-track faculty.

In conclusion, automatic TCE policy, which is a part of work-life balance programs, assists new parent faculty to overcome the existing stigma attached to requesting tenure-clock extension. It helps the junior faculty to allocate time to both their job and family responsibilities. Especially, it is important because junior faculties are at the age of starting families and having children. In other words, their tenure-clock and biological clock for having children overlaps and such policies can assist them in handling the conflicting responsibilities. These policies do not necessarily improve the retention rate of tenure-track faculty. It might be due to different reasons behind tenure-track faculties' opt out other than inability to meet tenure criteria such as relocation and other family and career related concerns. Overall, faculties who can benefit such policies spend more time in the tenure-track position and are happier with their career.

We recognize that our study provides a descriptive view of the correlations between various investigated factors and we do not report any causations. We also acknowledge that ignoring the weights in our regression analyses affect the generalizability of our study. As we previously discussed in the method section, we could not use the provided survey weights by NSF because of connecting several waves of the SDR longitudinal data. Another limitation of this study is that we cannot separate the effects of auto TCE policies from other work-life balance programs (e.g.

modified teaching and service duties) on tenure-track duration and job satisfaction in new parent faculty. As, universities might adopt auto TCE policy under a set of work-life balance programs. Further investigation is required to test whether the observed correlations are a result of auto TCE programs solely or the combination of several work-life balance programs.

Acknowledgements

The National Institute of General Medical Sciences and the Office of Behavioral and Social Sciences Research of the National Institutes of Health (NIH) supported this work (Grant 2U01GM094141-05). We thank Niyousha Hosseinichimeh (VT), Joshua Hawley (OSU), Konstantinos Triantis (VT), Eileen Van Aken (VT) and Peggy Layne (VT) for helpful comments.

References

1. Bain O, Cummings W. Academe's Glass Ceiling: Societal, Professional-Organizational, and Institutional Barriers to the Career Advancement of Academic Women. *Comparative education review*. 2000;44(4):493-514.
2. Ceci SJ, Williams WM. Sex differences in math-intensive fields. *Current Directions in Psychological Science*. 2010;0963721410383241.
3. Ginther DK, Kahn S. Women in economics: moving up or falling off the academic career ladder? *The Journal of Economic Perspectives*. 2004;18(3):193-214.
4. Long MC. College applications and the effect of affirmative action. *Journal of Econometrics*. 2004;121(1):319-42.
5. Hur H, Andalib MA, Maurer JA, Hawley JD, Ghaffarzagdegan N. Recent trends in the US Behavioral and Social Sciences Research (BSSR) workforce. *PloS one*. 2017;12(2):e0170887.
6. Andalib MA, Ghaffarzagdegan N, Larson RC. The Postdoc Queue: A Labour Force in Waiting. *Systems Research and Behavioral Science*. 2016.
7. Van Anders SM. Why the academic pipeline leaks: Fewer men than women perceive barriers to becoming professors. *Sex roles*. 2004;51(9-10):511-21.
8. Ceci SJ, Ginther DK, Kahn S, Williams WM. Women in academic science: A changing landscape. *Psychological Science in the Public Interest*. 2014;15(3):75-141.
9. Ceci SJ, Williams WM. Understanding current causes of women's underrepresentation in science. *Proceedings of the National Academy of Sciences*. 2011;108(8):3157-62.
10. Barnett RC, Hyde JS. Women, men, work, and family. *American psychologist*. 2001;56(10):781.
11. Cotter D, Hermsen JM, Vanneman R. The End of the Gender Revolution? Gender Role Attitudes from 1977 to 20081. *American Journal of Sociology*. 2011;117(1):259-89.
12. Leslie S-J, Cimpian A, Meyer M, Freeland E. Expectations of brilliance underlie gender distributions across academic disciplines. *Science*. 2015;347(6219):262-5.
13. Carli LL, Alawa L, Lee Y, Zhao B, Kim E. Stereotypes about gender and science: Women ≠ scientists. *Psychology of Women Quarterly*. 2016;40(2):244-60.
14. Walton GM, Cohen GL. A question of belonging: race, social fit, and achievement. *Journal of personality and social psychology*. 2007;92(1):82.
15. Nosek BA, Smyth FL. Implicit social cognitions predict sex differences in math engagement and achievement. *American Educational Research Journal*. 2011;48(5):1125-56.

16. Moss-Racusin CA, Dovidio JF, Brescoll VL, Graham MJ, Handelsman J. Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences of the United States of America*. 2012;109(41):16474-9.
17. Sheltzer JM, Smith JC. Elite male faculty in the life sciences employ fewer women. *Proceedings of the National Academy of Sciences*. 2014;111(28):10107-12.
18. Shollen SL, Bland CJ, Finstad DA, Taylor AL. Organizational climate and family life: How these factors affect the status of women faculty at one medical school. *Academic Medicine*. 2009;84(1):87-94.
19. Sax LJ, Hagedorn LS, Arredondo M, DiCrisi FA. Faculty research productivity: Exploring the role of gender and family-related factors. *Research in higher education*. 2002;43(4):423-46.
20. Perna LW. Sex differences in faculty tenure and promotion: The contribution of family ties. *Research in Higher Education*. 2005;46(3):277-307.
21. Fox MF, Fonseca C, Bao J. Work and family conflict in academic science: Patterns and predictors among women and men in research universities. *Social Studies of Science*. 2011;41(5):715-35.
22. Wilson R. Academic Couples Said to Be Happier Working at the Same University. *The Chronicle of Higher Education*. 2002;9(2):02.
23. Manchester CF, Leslie LM, Kramer A. Is the clock still ticking? An evaluation of the consequences of stopping the tenure clock. *ILR Review*. 2013;66(1):3-31.
24. Layne P, Johnson A, Hyer P. Policy Change Is Not Enough: Measuring the Impact of Work/Life Policies at a Research University. *Women in Engineering ProActive Network*. 2008.
25. Pribbenow CM, Sheridan J, Winchell J, Benting D, Handelsman J, Carnes M. The tenure process and extending the tenure clock: The experience of faculty at one university. *Higher Education Policy*. 2010;23(1):17-38.
26. Tenure Clock Extension for New Parents Princeton University Website: Office of the Dean of the Faculty. [cited 2018 2/15/2018]. Available from: <https://dof.princeton.edu/policies-procedure/policies/tenure-clock-extension-new-parents>.
27. Faculty Leave, 070 (2008).
28. Statistics NLSBoL. Sample Weights & Design Effects National Longitudinal Surveys [May 08, 2018]. Available from: <https://www.nlsinfo.org/content/cohorts/nlsy97/using-and-understanding-the-data/sample-weights-design-effects/page/0/0/#practical>.
29. Association TTR. WorldatWork Alliance for Work-Life Progress Announces 2016 Seal of Distinction Award Recipients: Word at Work. 2016 [cited 2018 4/15/2018]. Available from: <https://www.worldatwork.org/docs/worldatworkpressreleases/2015/worldatwork-announces-the-recipients-of-the-work-life-2016-seal-of-distinction-award.html>.
30. Blog TG. 11 Universities With Great Work-Life Balance – Hiring Now! : The Glassdoor Blog. 2017 [4/15/2018]. Available from: <https://www.glassdoor.com/blog/11-universities-with-great-work-life-balance-hiring-now/>.
31. Bristol MN, Abbuhl S, Cappola AR, Sonnad SS. Work-life policies for faculty at the top ten medical schools. *Journal of Women's Health*. 2008;17(8):1311-20.

Chapter 4. What Factors Do Make a University More Successful in Diversity Programs?

Abstract

Despite all the attention and progress to diversify university campuses, many educational institutions are still far from gender and racial parity in their faculty and student populations. An important question is what makes an institution perform better regarding diversity and outcomes of under-represented minorities (URM). Our focus is on the problem at the institutional level. We construct an institution-level dataset from multiple sources and apply the Data Envelopment Analysis (DEA) to systematically benchmark, compare, and contrast organizational performance with respect to the educational and work outcomes of URM. We find that universities located in states with affirmative action policies in place have higher graduation rates of URM, but show no significant difference in average GPA and first job salaries of URM. More competitive universities with lower acceptance rates show better organizational performance regarding diversity. Our analysis shows that many diversity measures are highly influenced by macro-level factors beyond organizational boundaries. Moreover, we explore institutional outliers (very successful or unsuccessful universities), and analyze the differences. This study also provides a methodological contribution by providing an example of how a well-established operations research technique can be used to study a multi-dimensional, social concept, such as diversity, in organizations.

Keywords: Data Envelopment Analysis, Higher Education, Diversity, Under-represented minorities, Gender, Race, Ethnicity.

Introduction

There has been an increasing attention to improve diversity in different public and private organizations including educational institutions. Institutions of higher education have thrived because of the quality of their education programs. The definition of quality in higher education differs depending on the stakeholders' perspectives that includes the products (students), users of outputs (employers of graduating students), the employees and universities [1]. Different stakeholders might have different priorities, however, graduating students' educational and career outcomes are the two mutual outcomes, which are important to all of the stakeholders [2]. Earlier versions of universities' strategic plans put more emphasis on increasing the number of enrollments and graduation rates as well as improving the quality of teaching and research activities. Besides creating knowledge, educating citizens and training the future generation of the workforce are the responsibilities of higher education [3]. Especially in a diverse society like the United States, providing equal educational opportunities is essential. In the last decades, the role of diversity and inclusion have received more attention in the design of higher education institutions' strategic plans. Now, many universities strive to provide equal opportunity for all of their students regardless of their social background [4]. They care about the social composition of their admitted and graduated cohorts. They also seek to improve student minority educational outcomes in addition to improving overall performance. The number of universities designing and implementing diversity strategic plans is increasing due to the issuance of presidential executive orders around the topic of diversity and inclusion.

In 2011, the Obama administration issued an Executive Order to establish an initiative to improve workforce diversity and inclusion across all federal government agencies [5]. Likewise, the department of education sets goals and objectives to promote faculty and staff diversity and

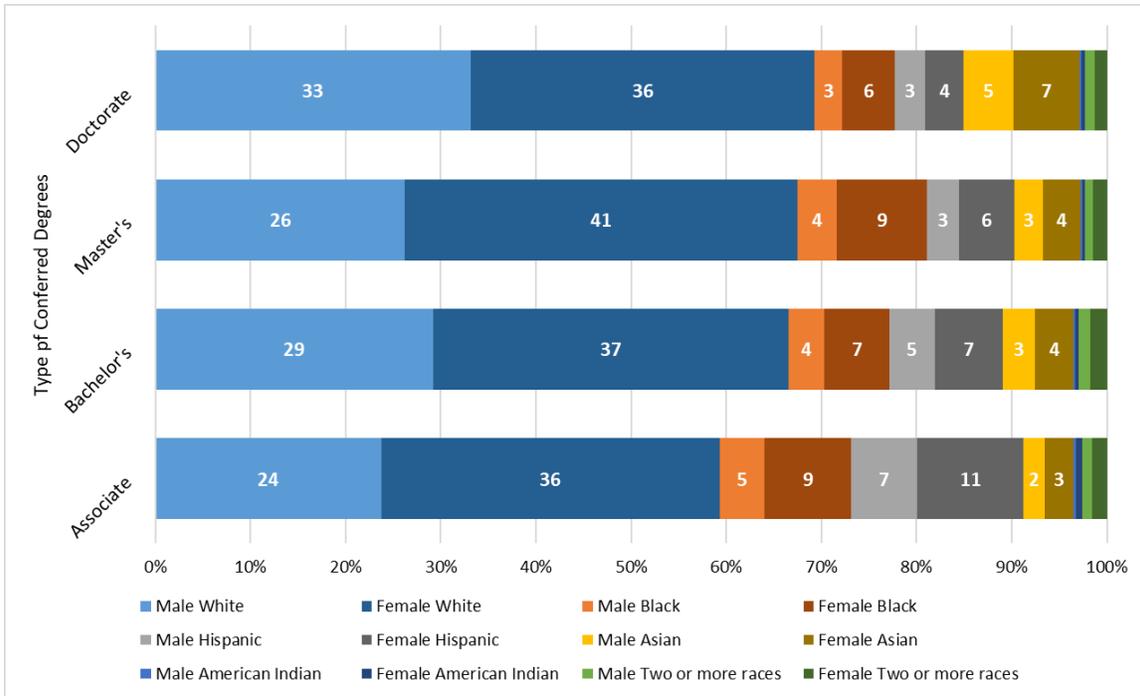
campus inclusion [6]. Diversity and inclusion initiatives now appear in most strategic plans of major organizations, and many managerial-level plans are developed to improve racial and gender parities. Universities, as examples of such organizations, strive to provide equal opportunities for all of their students and professors regardless of their social and demographic background [4]. They pay special attention to the demographic composition of their admitted and graduated students, and many actively try to recruit students and professors from under-represented minority groups. In addition, universities seek to improve educational outcomes of minorities such as graduation rate and GPA. Higher educational outcomes for URM can lead to a better representation of URM in workforce, thus is of critical importance.

Diversity in educational systems is often examined at three main levels of access to education, college experience, and post-graduation outcomes [7]. Historically, racial minorities had a very limited access to higher education. It has been less than two hundred years that the first African-American entered college [8] and less than hundred years that they were granted access to predominantly white institutions (PWI) [9]. Furthermore, their college experience and employment outcomes were not as favorable as their peers. They were subject to negative perceptions regarding their education capabilities and even explicit or subtle insults [10]. Average salary of racial minorities also lags behind their white and Asian peers, which is caused in part by their lower education attainment level [11].

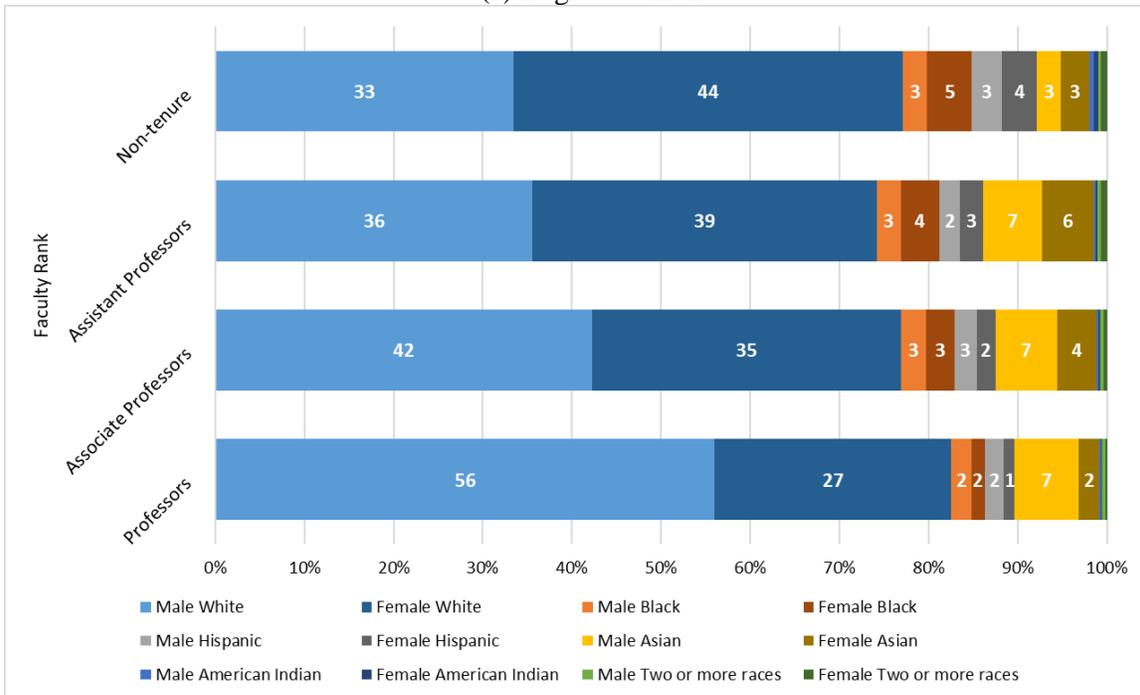
During the last decades, universities have made progress in the aforementioned areas of education equality; they achieved to narrow down the gaps between minority and majority students in access, college experience and post-graduation outcomes. Despite all improvements in educational outcomes of higher education institutions, they have not achieved an equal population representation of students and faculty members from all segments of the society. Most recent

statistics on racial and gender compositions (NCES 2015) show that there are major gaps that remain. Figure 4.1 compares the status of men and women from different racial groups in the US education system in 2015. Figure 4.1.a depicts the gender and racial composition of conferred degrees by universities, and Figure 4.1.b shows gender and racial composition of faculty members by their academic rank. In an ideal situation, the population representation of each social segment should be close to its population representation in the US (i.e., according to the Census population estimates of 2016, the US population were composed of 61% white, 13% Black, 18% Hispanic, 6% Asian, and 2% other races). However, looking at the number of conferred degrees shows that there is an imbalance in the fraction of conferred degrees to different racial groups. The proportion of black and Hispanic students receiving degrees are smaller than their population representation in the US population (13% Black and 18% Hispanic). Disparities further increase as we move to the higher degree levels [12].

The trend is similar for the postsecondary faculty. As depicted in Figure 4.1.b, 83% of professors are white and the remainder distributed unevenly between black, Hispanic and Asians (4, 3 and 9 percent respectively). The racial imbalance in faculty positions worsens at higher faculty ranks [12]. Unlike the number of degrees conferred, women are among the minority in postsecondary faculty positions, especially for associate and full professors. The condition is reversed in lower rank positions such as assistant professor and non-tenure-track faculty.



(a) Degrees conferred



(b) Postsecondary faculty

Figure 4.1: Racial and gender compositions of United States universities, 2015.

Underrepresentation of ethnic and racial minorities is a result of complex individual, institutional and environmental factors [13]. Past studies widely investigated the effect of socio-economic factors on minority students' persistence in higher education and their educational outcomes. College affordability, lack of parental support, and campus integration are examples of these socio-economic factors. In the next section, we will discuss each of these factors in more details to present a comprehensive view of the current literature. However, what is worthwhile mentioning is the small body of literature on the effect of institutional and state-level factors on the success of ethnic and racial minorities in US higher education. Obviously, not all universities are equally successful in designing and implementing campus-wide diversity plans. This can be due to their institutional characteristics, environmental factors or state-level policies. US universities have decentralized decision-making structures, they design their diversity and inclusion improvement plans based on the specific characteristics of their campuses. Heterogeneity exists among universities in their diversity and inclusion efforts, as they are different in their available resources. Some universities might be able to better allocate their resources to support minority students, while others might not be that effective. Thus, identifying the best practices of US universities and understanding why they can better implement their diversity plans is essential for improving the overall diversity of US universities. A major research question is: *what makes a university more successful in improving diversity outcomes?*

In this paper, we compare the effectiveness of the US universities' campus diversity and inclusion initiatives to identify the best practices. We use the Data Envelopment Analysis (DEA) approach to quantify the effectiveness of each institution's efforts in training their minority students, with a focus on African-Americans and Hispanics, based on multiple inputs and outputs. Furthermore, we test several hypotheses to see whether various environmental factors that are

attributed to universities (e.g., campus size, university ranking, location, applying vs. banning Affirmative Action programs, etc.) contribute to diversity inefficiencies. Finally, we introduce environmental factors that hinder (or facilitate) diversity initiatives in higher education institutions.

Theoretical background

Factors contributing to underrepresentation of ethnic and racial minorities in higher education are widely studied. At an individual-level, educational background of students and their socio-economic status are argued to influence education outcomes in universities [13]. Students' academic preparation is in fact one of the most important factors contributing to their underrepresentation. Their high school GPA and standardized test scores are relatively lower than their peers on average [14]. Furthermore, they lack the required financial resources to afford their education expenses. URM students are more likely to be the first generation college students in their families. They lack receiving potential mentoring from their parents who themselves may not have college experience [15].

As mentioned earlier, a group of studies investigate the effect of students' individual characteristics (e.g., students' background and families' socio-economic status) on their graduation rate and educational outcomes [13]. These specific characteristics of ethnic and racial minorities make it difficult to meet white-created educational standards. Also, subtle bias exists in course designs and curriculum, which makes minority students to be less successful in their educational outcomes [9]. Minority students' academic preparation is one of the most important factors contributing to their underrepresentation. Their high school GPA and standardized test scores are relatively lower than their peers on average [14]. Furthermore, they lack the required financial resources to afford their education expenses, which is due to their families' lower socio-economic status. Parents' education level is an important factor that hinders academic

achievements of minority students. Parents, who do not have college experience, are less able to provide financial support and mentoring to their children [15]. However, even if we account for family income, the gaps between students of color and their Asian and White peers are not completely explained [16].

The aforementioned individual factors are not the only drivers of ethnic and racial minority students' underrepresentation in higher education. At macro-level, ethnic and racial minorities' educational outcomes suffer from the negative stereotypes, racial biases and lack of role models. Negative racial stereotypes exist about their inferior merits and intellectual abilities to successfully graduate [10]. These stereotypes have more negative influence on the educational outcomes of minority students as they are more exposed to such environments [17]. Racial bias in evaluating ethnic and racial minorities' academic background also explain a part of URM's historical underrepresentation in universities.

Between these extremes of individual and social factors, one can look at policy initiative at government or organizational levels that can support URM representation and their educational outcomes. Higher education institutions are the environments that can be very influential in boosting minority students to higher economic status.

At intuitional levels it appears to be more feasible to implement policies that help URM. However, the literature is also less developed. Many studies that looked into the effects of institutional factors on ethnic and racial minorities' success, investigated a small sample of institutions, mostly public universities [9, 13, 18-20]. They mostly focused on the details of campus climate and university culture and their effects on minority students' success. In these studies, minority students' success is defined as having the same retention and graduation rates as their white peers. Underrepresented minorities often do not find the campus climate welcoming

[21] and as time goes by, they become isolated and alienated from the majority of students [13]. This issue is especially reported in elite schools, where minority students experience subtle bias [10]. As a result, URM students become less involved in university activities and receive less peer support during their studies [20]. All of these factors ultimately lead to higher attrition and lower graduation rates.

Overall, it seems that at the organizational level, effects of campus climate on minority students' success, especially interaction of students with faculty and staff, are more explored than effects of organizational characteristics and policies [20]. Educational outcomes of minority students are also better in universities where minorities have more interaction with faculty and receive more targeted support from the university [19]. Past studies postulate that ethnic and racial minority students' performance is also influenced by cultural differences with the majority of students which impedes blending with the community of students (e.g., cultural difference theory) [22], and by lower outcome expectations leading to lower outcome [23-25].

In this study we look at the effect of institutional characteristics and its environment on the performance of higher education institutions in training their minority students. Figure 4.2 offers a big picture of competing or supplemental theories of underrepresentation of URM students' underrepresentation in the higher education system. In this paper, our first group of hypotheses (Figure 4.2- top left corner) deal with characteristics of the environments in which universities are operating, such as ratio of minorities living in the city and whether the state banned affirmative action programs. It is expected that in more diverse environments, the cultural differences of URM students are more accepted. As a result, their integration with the academic system including faculty and staff improves and this leads to more commitment to academic goals which is essential for better educational outcomes. Also, higher population of minority students enhances their social

interaction on campus and probability of finding peers. The second group of hypotheses (Figure 4.2- bottom left corner) relates to the effects of institutional characteristics such as universities financial resources, university (public vs. private), and selectivity in the admission process. This group of hypotheses state that the educational outcomes are improved through either better academic integration or social integration. For example, more selective institutions admit students with better academic backgrounds. In the next section, we elaborate more on the theories and offer our specific hypotheses.

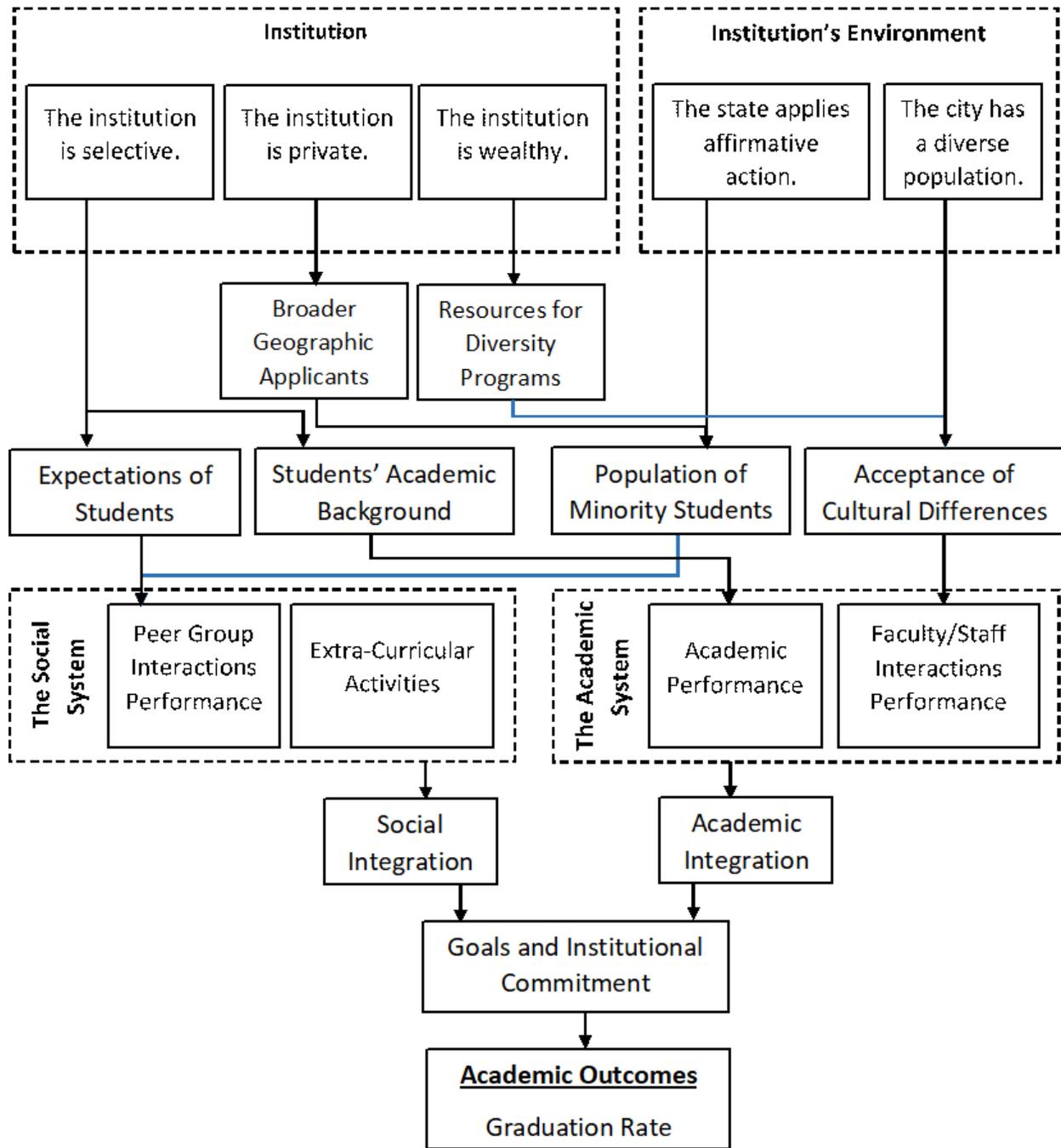


Figure 4.2: Theory of minority students' underrepresentation in higher education.

Institution's Environmental Factors

According to Tinto's theory of departure, students show more commitment to academic goals when they are better integrated into their academic and social communities [26]. One major challenge for entering cohorts is that they are separated from their old social network and have to build up new connections and integrate with their new social group including their peers and faculty [27]. Connecting with the new social network even becomes more challenging in less diverse environments. If URM students fail to connect with the new network they become more vulnerable to academic difficulties that might happen during their studies. Therefore, diversity of the new community is important for improved social and academic integration of minority students and ultimately their educational outcomes [28]. A more diverse community can improve the sense of belonging to campus which is an important factor in decreasing the chance of leaving college for college students [12, 28]. Moreover, it affects URM students' interaction with faculty and staff. URM students' learning styles are often different than the White majority of the class and their cultural differences are better embraced by faculty, staff and students in a diverse environment. Students can also build their social network out of the classroom and their social network does not have to be limited to their classroom [12]. We hypothesize that if a campus is located in a city or town with a diverse population (i.e., higher ratio of ethnic and racial minority residents), faculty, staff and majority students are more exposed to different cultures. Thus, acceptance of cultural differences improves in that university and minority students have higher educational outcomes.

Hypothesis 1: Universities located in cities/towns with a higher ratio of ethnic and racial minority residents are more successful in graduating minority students with high educational and employment outcomes.

State-level decisions to adopt or ban diversity programs such as Affirmative Action also matter. Affirmative Action is a set of programs and policies initiated in the 1960s with the goal of providing equal employment and education opportunities for everyone in the United States, regardless of their race, sex, color, creed, or national origin [29]. Initially, Affirmative Action intended to overcome the past discriminations against African-Americans and then it was followed by 1964's Civil Rights Act. Later, women and other groups of underrepresented minorities were added to Affirmative Action [30]. Affirmative Action includes a broad range of policies and programs from early stages of education to faculty positions and employment [31]. During the last half century of applying Affirmative Action policies, most of the debates focused on racial selection in college admissions [31]. California, Texas, Washington, Florida, Michigan and Nebraska banned Affirmative Action between 1997 and 2008. Between 2009 and 2012, Arizona, New Hampshire and Oklahoma passed legislatures to prohibit preferences based on race, color, sex, national origin or religion in employment and education [32]. These states consider the socio-economic status of applicants to support underrepresented groups. They also have percent plans, which consider the top students of each high school and the rationale is that they admit more students from under-resourced schools [33]. Here, we categorize universities in the six states, which banned Affirmative Action before 2008 (e.g. California, Texas, Washington, Florida, Michigan and Nebraska) in group A and the rest of the universities as group B. We propose the following hypothesis to test the difference between the progresses made in training minority students from 2008 to 2015 by these two groups of universities.

Hypothesis 2: Universities in states which banned Affirmative Action are less successful in graduating minority students with high educational and employment outcomes.

Institutional Factors

Universities plan various initiatives to improve different aspects of diversity. Outreach to high school students is important to improve minority students' access to higher education. Universities also provide high school students and their parents with information on admission and financial aid. They also do campus tours and visits to high schools in their area. Mentoring, coaching, academic advising and on-campus support programs are only a few examples of programs that aim to improve minority students' success. Successful implementation of the previously mentioned initiatives require financial resources. Furthermore, availability of financial aid, reinforces low-income and minority students ability to continue their education [12]. Therefore, we propose the following hypothesis to examine the effect of universities' financial resources on the effectiveness of diversity programs.

Hypothesis 3: Wealthier universities are more successful in graduating minority students with high educational and employment outcomes.

The second institutional factor that we propose is the type of the institution (i.e., public vs. private). We test the following hypothesis to investigate the difference between private and public schools when training underrepresented minorities. Private schools are known to attract students from various other states while the students attending public schools usually come from the state that the public school is located. Larger populations of minority students on campus help reducing their isolation and alienation from the majority of students and improves their educational outcomes [13].

Hypothesis 4: Private universities are more successful in graduating minority students with high educational and employment outcomes.

Top universities are famous for being more selective in their admission processes; they have lower admission rates compared to other universities and stricter admission criteria to choose the best incoming cohort of students. Furthermore, the high ranked universities do not increase their capacity to keep up with the increasing demand for higher education. Therefore, getting admission from one of these universities is getting harder for everyone over time. On the other hand, statistics show that minority students are disadvantaged in getting into elite universities because of their lower average academic preparation [12]. The aforementioned factors contribute to admitting a smaller portion of minorities to the top universities. The question is that whether minority students in elite schools have the same educational and career outcomes as their peers. If top ranked universities admit a diverse cohort of students, they should be able to provide them equal education and training opportunities. Training quality and reputation of high ranked universities contribute to better educational outcomes of their students including minority students [34]. Otherwise, they will fail to either graduate given the usual heavy workload of these schools or graduate with below average education and career outcomes. We propose the following hypothesis to investigate the effect of acceptance rate (i.e., selectiveness) on the successful implementation of diversity practices. In this paper, we define success in graduating minority students with high educational and employment outcomes as higher effectiveness score in the DEA analysis.

Hypothesis 5: Universities that are more selective (i.e., lower acceptance rate) are more successful in graduating minority students with high educational and employment outcomes.

Method

Data Envelopment Analysis in education

Data Envelopment Analysis (DEA) is a non-parametric method of evaluating the performance of business units and its most basic model, CCR model, were first introduced in 1978. DEA

identifies a production frontier including the best practices among the decision-making units (DMUs) by considering multiple inputs and outputs [35]. One characteristic of DEA is that there is no need to assume any relationship between the input and the outputs.

DEA has two advantages over regression analysis. First, it accounts for multiple outcomes for the set of DMUs, while regression analysis considers only one output [36]. The goal of diversity in higher education is not limited to only improving the number of minority students on campus (e.g., superficial diversity), but inclusion and outcomes are also important [37]. Therefore, to compare the success of diversity practices among US universities we need to capture several outcomes to develop a comprehensive model of institution diversity. Second, DEA is able to identify a frontier for the set of DMUs, which shows the maximum possible outcomes, while regression analysis provides the average outcomes [38]. The method is widely used in non-for-profit entities and the public sector with no market price is available for the products [39, 40]. Further, it is also widely applied to benchmark the performance of hospitals and local government functions such as police departments [41]. More specifically, DEA has a broad application in education, as a typical example of a system with multiple inputs and outputs, where market prices are absent [40, 42].

Previous research evaluated the efficiency of both secondary and post-secondary schools through DEA approaches. Based on the application context, the selection of decision-making units (DMUs) and input/output variables are varied among these papers. DEA was applied in education, as a production process of students, to benchmark the education units at both the aggregate-level and individual-level. Most of the studies in education context estimated the relative performance of higher education institutions [43-47], departments within a university [36] and secondary schools [38, 48, 49] using aggregated data. A smaller group of studies benchmark the performance

of individuals by considering them as the DMUs [38, 50]. Individual-level DEA analysis separates the effect of students' characteristics from schools'. On the other hand, it increases the computational load of the DEA analysis. Furthermore, in individual-level DEA analysis, identifying outliers becomes important because of the sensitivity of DEA's frontier to observations [40]. DEA's application is also geographically dispersed in education; efficiency of Australian [44], English [38, 47, 50], Italian [43], Greek [46] higher education institutions were benchmarked in several studies. Selection of input and output measures depends on the goals of the DMUs as well as the availability of data. For example, the number of instructional staff and socio-economic background of students [48, 49] are selected as inputs when educational performance and teaching quality of DMUs are important. Operational costs of universities are also considered as inputs when economic efficiency of DMUs is of interest. In case of outputs, the number of graduates, the quality of research [45], test scores [48], retention rate and employment rate [44] are among the selected variables.

Many authors propose hypotheses which require examining the effect of environmental variables on the non-parametric efficiency scores. For example, authors of a study on Australian universities were interested to know the effects of different environmental variables such as location of university and the proportion of senior professors on the performance of the DMUs [51]. In healthcare, effects of using advanced medical technologies and type of hospital on Greek hospitals' performance were investigated using combined DEA and regression methods [52].

Despite all of the described applications, there are methodological difficulties in performing a hypothesis test using the DEA efficiency scores. Probability distribution of non-parametric estimators are required to perform statistical inference on DEA efficiency scores. Simar and Wilson assume a data-generating process (DGP) and suggest applying a bootstrapping technique

to reduce the bias in the DEA efficiency scores and estimate their confidence intervals [53]. The bootstrapping method regenerates the input/output data and mimics the sample distribution through repeatedly selecting different samples with replacement. Thus, bootstrapped DEA scores provide multiple estimates and make hypothesis testing possible [54]. There is also some debate over the type of regression model to use with bootstrapped DEA scores. Truncated regression analysis is the most widely cited method of hypothesis testing on bootstrapped DEA scores [53].

Model

In this study, we use DEA to rank the US universities in terms of educating and training minority students. Furthermore, we use a bootstrapping approach to estimate the bias-corrected DEA scores and test the effect of environmental variables on the training effectiveness of minority students. We start this section with some basics of DEA and then we continue describing the applied variations of the DEA models.

The CCR model, which was introduced by Charnes, Cooper and Rhodes in 1978, is the most basic DEA model and is useful in describing the logic behind the DEA analysis [39]. Linear programming formulations represented by Eq. 1 and Eq. 2 show the primal and dual problems of the CCR model. Vectors X and Y are the inputs and outputs respectively. Vectors v and u are input and output multipliers. The LP is calculated once for each decision-making unit (DMU) to find a set of weights (v, u) in a way that the efficiency score of DMU_i is maximized. If there are n DMUs, the optimization is done n times. DEA analysis returns a set of efficiency scores (i.e., for each DMU), optimal weights, slacks that show the waste in inputs and the surplus in outputs, peers and performance targets.

(LP_0) $\max uy_0$ Eq.1 <i>Subject to</i> vx_0 $= 1$ $-vX + uY \leq 0$ $v \geq 0, u \geq 0$	(DLP_0) $\min \theta$ Eq. 2 <i>Subject to</i> $\theta x_0 - X\lambda$ ≥ 0 $Y\lambda \geq y_0$ $\lambda \geq 0$
---	---

We choose output-oriented DEA model to maximize the educational and employment outcomes of minority students. In addition, we consider variable-returns-to-scale (VRS) to assume that outputs do not increase proportional to increases in inputs (i.e., $\sum \lambda = 1$). We use “rDEA” R-package to analyze the educational outcomes of minority students at each institution. We refer to the existing well-developed literature for more details on DEA methods. The analytical procedure of this study is as follows:

1. Define the inputs, outputs and transformation process of training minority students.
2. Estimate the DEA efficiency scores of the DMUs using an output-oriented model with variable-returns-to-scale (VRS).
3. Estimate bias-corrected DEA efficiency scores using a double-bootstrap DEA with environmental variables.
4. Perform truncated regression analysis to test the effect of each state-level and institution-level variable on training effectiveness of minority students (i.e., each hypothesis).

Input, output and environmental variables

Input variables: As we mentioned earlier, the higher education system is not equitable in minorities’ access to higher education, as well as their college experience and post-graduation

outcomes. Universities design and implement specific initiatives or diversity improvement plans to tackle each of these issues [55]. We can form four major groups from current diversity improvement practices of universities: (a) improving diversity in applications and admissions, (b) ensuring college is affordable, and (c) providing strong support to help students succeed, and (d) ensuring safe and inclusive campuses [55]. Based on the first three categories we identify inputs to our DEA model. Table 4.1 shows the list of all input, output and environmental variables as well as their units and data sources. We also depict the transformation processes in Figure 4.3 using causal links (positive and negative causations). The box in the figure shows the boundary of the DMUs.

- a. Improving diversity in applications and admissions:** The first step in improving campus diversity is improving minority students' access to higher education. Universities try to improve underrepresented students' access to college through several programs such as high school outreach, enrichment and recruitment programs. Outreach programs provide minority students and their parents with information application process, financial aid availability and other information required to apply to colleges [56]. In this study, we consider the number of minority students enrolled in each university instead of number of application and admissions because of database limitations. Furthermore, a higher number of undergraduate and graduate students enrolled in a DMU leads to increased sense of inclusion in minority students. Graduate students from minority groups can also act as mentors, who understand the specific concerns of their mentees and improve learning (Figure 4.3, solid red links).
- b. Ensuring college is affordable:** One of the reasons for college dropouts in students from low-income families is the lack of financial resources to cover tuition and living expenses. Students from underrepresented racial groups are more affected by these financial issues. According to statistics, parents of minority students have a lower education level compared to well-represented racial groups and as a result, they would have lower-levels of income. Thus, providing financial support to students is one of the means of assisting

minority students in affording their college expenses; decrease their dropout rates and ultimately improving their educational outcomes (Figure 4.3, dashed red links).

- c. **Providing strong support to help students succeed:** Academic and social support that students receive from their family and peers improve their academic success as well as their retention rate. Better integration of students into their social and academic communities also improves their persistence [12]. Minority students are disadvantaged in terms of receiving academic advice from their parents because they have a higher chance of being the first college generation and parents, who do not have college experience, cannot guide their children through their studies as well as the highly educated parents. Universities try to offer mentoring and coaching programs to minority students and assist them in dealing with academic challenges. Existence of minority faculty and their involvement in teaching and research activities improves the academic outcomes of minority students in several ways. First, minority faculty act as the “change agents” in promoting campus diversity culture. The university community can better recognize the value of minority faculty when they are included in academic activities [57]. Second, minority faculty have a better understanding of the challenges that minority students face through their academic path and can provide them with better academic advice. They may also become minority students’ role models and help them to better integrate in academic and social communities [12, 56] (Figure 4.3 solid blue links).

Output variables: Table 4.1 also shows the selected output variables. If universities were successful in training minority students, they would have better educational achievements, which can be translated into higher graduation rates (y_1) and higher GPAs (y_2). The former represent both quality and quantity of graduating cohorts and the latter shows the quality of their education. Another goal of training students in higher education institutions is to help them find well-paid employment opportunities. Therefore, we add another output variable to compare the average salary of minority graduates with the well-represented group (y_3).

The output variables that we discussed so far were desirable outputs; however, we may have undesirable outputs for the DMUs because of increasing chance of conflicts between students from different social segments when the campus diversity improves. As discussed before, ensuring safe and inclusive campuses is one of the major groups of current diversity improvement practices. Students need to be integrated in their academic and social community in order to have desirable academic outcomes. Especially underrepresented minorities should feel welcomed and well treated in the new community to reach the expected educational outcomes. Hate incidents are the events that happen against individuals because of their different characteristics including sex, ethnicity, color, sexual orientation, religion and national origin [58]. These incidents threaten the integrity of campus as a whole [58]. Furthermore, they affect students from underrepresented groups, especially women and African-Americans, who are more sensitive to hate crimes against the members of their minority group [59]. Thus, the hate incidents are considered as an apparent threat to the educational outcomes of minority students. We acknowledge that exceptions may happen on campuses even with a diverse and inclusive environment. Also, we admit that the collected data on campus hate incidents might miss a portion of incidents because minority students may have a fear of being harassed if they report such events [58].

Environmental Variables: We also take the size of the campus (i.e., total number of enrolled students) as another environmental variable (z_1).

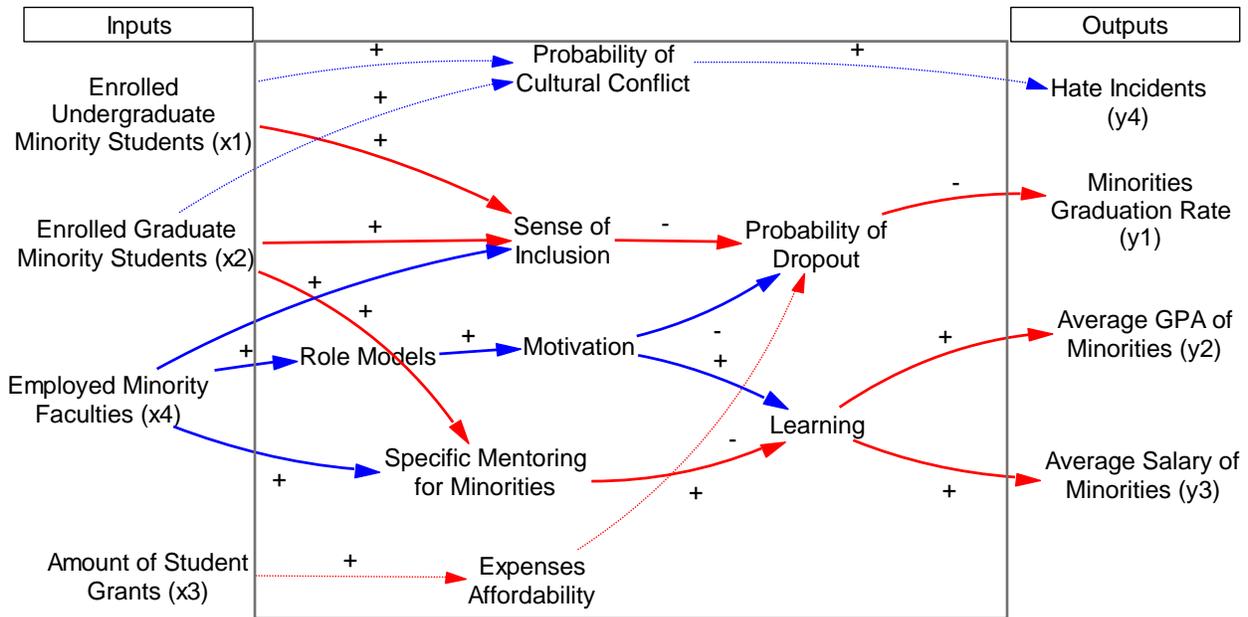


Figure 4.3: Transformation Process.

Table 4.1: Input, output and environmental variables of data envelopment analysis.

	Variables	Units	Source
Input	Number of enrolled minority undergraduate students (x_1)	People	IPEDS
	Number of enrolled minority graduate students (x_2)	People	IPEDS
	Total dollar amount of students grants (x_3)	US Dollars	IPEDS
	Number of minority faculty (x_4)	People	IPEDS
Output	Graduation rate 150% of minority students (y_1)	People	IPEDS
	Average undergraduate GPA of minority students (y_2)	Dimensionless	NSCG
	Average salary of the first job of minority students (y_3)	US Dollars	NSCG
	Number of hate incidents (y_4) - undesirable	Dimensionless	CSS
Environmental	Number of students enrolled (z_1)	People	IPEDS

Note: IPEDS stands for Integrated Postsecondary Education Data System; NSCG stands for National Survey of College Graduates; CSS stands for Campus Safety and Security.

Data

We construct a dataset at the institutional level and from three main sources of the Integrated Education Data System (IPEDS), National Survey of College Graduates (NSCG) and Campus Safety and Security (CSS). IPEDS is a data collection program conducted annually by the Department of Education's National Center of Education Statistics (NCES). The surveys collect

information on different characteristics of higher education institutions such as admissions, enrollment, graduation rate, financial aid and human resources from more than 7,500 colleges, universities and vocational schools in the US. The data is publically available from school year 1980 to 2016. The National Science Foundation's NSCG survey collects employment, education and demographic information of individuals who have a bachelor's degree at least biennially. Office of Postsecondary Education (OPE) of the Department of Education collects CSS data on crime and fire incidents in higher education institutions annually from 2008.

We use IPEDS to extract data related to the university's financial aid, faculty composition of higher education institutions as well as graduation rates and university enrollments by race/ethnicity. Then, we use NSCG to extract average GPA and salary of degree holders by institution and race/ethnicity using the survey weights. We also use CSS to acquire the cumulative number of race-based hate incidents by institution-level. These three databases can be connected using the unique identification numbers of universities (OPE ID). Regarding the temporal setting of the surveys, we use the latest available data (e.g., 2015) of NSCG, IPEDs and CSS. Furthermore, we limit the institutions list to only doctorate/research universities with IPEDS CARNEGIE code value of 15 and 16 (i.e., universities classified as doctorate universities with extensive and intense research activities). The remainder universities are around 210.

The input and output variables are calculated for minorities (i.e., Black and Hispanic racial groups in this study). As mentioned before, the diversity database is constructed using three databases including IPEDS, NSCG, and CSS. After connecting the databases, we came up with 156 universities with complete data sets. Table 4.2 shows the descriptive statistics of the research universities in the US by input, output, and environmental variables.

Table 4.2: Descriptive statistics table, 2008. (Number of units=156)

Variables	Mean	Median	Standard Deviation	Minimum	Maximum	Sum
Enrolled undergraduate minorities (x_1)	3,630	2,420	3,973	135	32,446	566,300
Enrolled graduate minorities (x_2)	888	723	732	29	4,742	138,457
Total student grants (x_3) Million dollars	147.3	116.6	99.4	7.8	523.2	22,973.2
Minority faculty (x_4)	677	539	608	27	4,280	105,572
Graduation rate of minorities (y_1)	292	207	264	9	1,367	45,590
Average GPA of minorities (y_2)	3.29	3.26	0.31	2.59	3.74	513
Average salary of minorities (y_3) dollars	38,990	38,828	17,858	3,500	122,367	6,082,510
Number of hate incidents (y_4)	1.81	0.50	2.85	0	16	283
Total number of enrolled students (z_1)	24,248	24,285	13,436	2,127	61,642	3,782,711
Number of minorities living in the city (z_2)	245,628	42,151	591,842	249	4,240,000	38,317,981

Results

Outlier Analysis

We perform outlier analysis using the principal component analysis (PCA). According to the results, we do not eliminate any data point (i.e., DMU) based on measurement errors because the data sources are public federal data and have reliable measurement systems. However, among the identified outliers, we decide to eliminate two of the outlier DMUs. One of the deleted DMUs is a

historically black university (Howard University) and the other one is located in Puerto Rico (University of Puerto Rico-Rio Piedras Campus) whose population is 98% Hispanic.

The remaining bad leverage-points in our outlier analysis, which we keep them, are located in the states of New York and Florida (e.g., New York University, Rutgers University-New Brunswick, and Florida International University). Furthermore, the six good leverage-points are located in Florida, Texas and Arizona (e.g., University of Florida, University of Central Florida, University of Texas at Austin, Texas A&M University, and Arizona State University). Next, we continue the analysis with the remaining 154 DMUs. Complete details of the PCA analysis is available in Appendix C.

DEA Analysis

In this section, we describe the properties of the DEA model and the results of the analysis to evaluate the effectiveness of training minorities in US universities. Variable returns-to-scale is considered as one of the properties of the current DEA model for several reasons. First, it is intuitive that the larger a campus is the chance of that institution in attracting more minority students is higher than smaller universities. Second, previous studies, which applied DEA approach in education context consider variable returns-to-scale to evaluate the efficiency in schools, universities and other educational programs. Furthermore, we use a test for returns-to-scale developed by Simar and Wilson [60, 61] to test the null hypothesis (i.e., constant returns-to-scale) versus the alternative hypothesis (i.e., variable returns-to-scale). We use an r-package (i.e., rDEA) to run the returns-to-scale test. Our results support variable returns-to-scale ($p < 0.05$). Furthermore, the nature of the problem is output-oriented because the availability of inputs (e.g., number of enrolled minority students, faculty and grants) are limited for each institution.

Moreover, universities are interested to know how much they can improve their outcomes within their limited resources.

Next step in the analysis procedure is to run an output-oriented DEA model with variable-returns-to-scale on 154 DMUs with four input and four output variables. Number of hate incidents is one of the output variables (i.e., y_4) that is considered as an undesirable variable. Past studies suggest several methods to deal with undesirable variables including ignoring the undesirable output, using non-linear DEA model [62] and treating the bad output as an input. Each mentioned approach, imposes some limitations on the results of DEA analysis [63]. For example, considering the undesirable output as an input might not show the true production process [64]. We use the approach suggested by Seiford and Zhu [64], which uses a translation vector to the bad outputs (i.e., linear monotone decreasing transformation). The aforementioned model preserves the linearity and convexity in DEA. The undesirable output (y_4) is translated using the following equation $\bar{y}_4 = -y_4 + w > 0$. Moreover, we iterate several values for w and calculate the efficiency scores. The efficient and inefficient units do not change by changing the value of w . Only the range of efficiency scores slightly change. Thus, we select $w = 17$ because of its widest range of efficiency scores.

Before moving to results of DEA analysis, we try the stochastic frontier analysis (SFA) on the dataset to estimate the ratio of the total variations that are due to inefficiency versus the ratio, which is due to noise (i.e., random variation). A small portion of the inefficiencies is related to the noise. In other words, our DEA model explains most of the variation in the estimated efficiency scores. Therefore, we continue building our analysis on the results of the DEA models. More details of the SFA analysis is available in Appendix D.

We perform the DEA analysis using an r-package called “rDEA”. As depicted by Table 4.3, we consider three different combinations of input and output variables to define models 1, 2 and 3. Model 1 is the full model with four input and four output variables. In model 2, we only consider the output variables related to the quantity of graduating minority students. Moreover, in model 3, we consider only those output variables, which represent the quality of minority students that graduate (e.g., average GPA and salary). In all three models, the undesirable variable (e.g., number of race/ethnicity based hate incidents) is included.

Table 4.3: Input and output variables of DEA models.

	Variables	M1	M2	M3
Inputs	Enrolled undergraduate minorities (x_1)	✓	✓	✓
	Enrolled graduate minorities (x_2)	✓	✓	✓
	Total student grants (x_3) Million dollars	✓	✓	✓
	Minority faculty (x_4)	✓	✓	✓
Outputs	Graduation rate of minorities (y_1)	✓	✓	
	Average GPA of minorities (y_2)	✓		✓
	Average salary of minorities (y_3) dollars	✓		✓
	Number of hate incidents (y_4)	✓	✓	✓

Table 4.4 presents the descriptive statistics of each of the previously described models under CCR and BCC DEA models. CCR models generally yield fewer efficient DMUs and the range of the efficiency scores are broader than the BCC model. In the BCC full model (M1), efficiency scores of 102 of the total 154 DMUs are estimated as 1, however only 45 of them have zero slacks; BCC-efficient. Therefore, 29% and 37% of the DMUs are respectively BCC-efficient and Radial-efficient (i.e., efficiency score is 1 but the slacks are non-zero) in the full-model. Model 2, which only considers the quantity of graduating minorities, improves the discrimination in DEA efficiency scores. Lower portions of the DMUs are BCC-efficient (i.e., 17% BCC-efficient and 41% Radial-efficient). The range of the efficiency scores in the second model is also wider; the minimum score and the standard deviation are 0.58 and 0.10 respectively (Figure 4.4). Model 3,

which only accounts for the quality measures of graduating minorities (e.g., average GPA and salary), has less efficient DMUs than the other two models but the range of the estimated efficiency scores is not as wide as Model 2. In this model, 8% and 51% of the DMUs are BCC-efficient and Radial-efficient respectively. The efficiency scores range from 0.74 to 1.00 with a standard deviation of 0.07. We present a detailed table of the results of the DEA analysis in Appendix E.

Table 4.4: Descriptive statistics of DEA efficiency scores.

Type	DEA Models	Mean	Standard Deviation	Minimum	Maximum	Radial-efficient DMUs	BCC/CCR -efficient DMUs	Inefficient DMUs
BCC	M1	0.97	0.05	0.76	1.00	57	45	52
	M2	0.93	0.10	0.59	1.00	63	26	65
	M3	0.96	0.07	0.74	1.00	78	13	63
CCR	M1	0.69	0.21	0.22	1.00	-	23	131
	M2	0.68	0.21	0.19	1.00	-	20	134
	M3	0.05	0.23	0.26	1.00	-	8	146

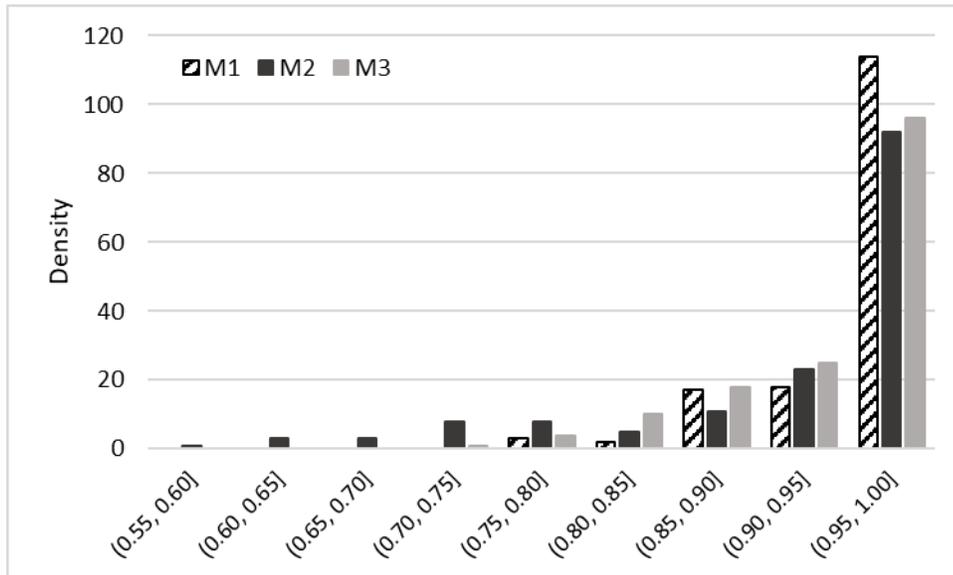


Figure 4.4: Distribution of efficiency scores by DEA model.

Some universities are efficient in all three models including Yale University, Princeton University and Emory University. The second group are efficient in Models 1 and 3 but inefficient

in Model 2, which means that the universities are good in quality measures and URM graduates perform better in terms of GPA and first job salary. Some universities of the second group are Ohio University, University of Wisconsin-Madison, University of Montana, and University of Missouri. The third group are efficient in models 1 and 2 but inefficient in model 3, which means that the universities are good in quantity of graduating URMs. Some universities of this group include University of Virginia, University of Michigan Ann Arbor and University of Texas at Austin.

It is also worthwhile mentioning that the number of times that an efficient DMU is identified as a peer to other inefficient DMUs can be an indicator of its importance as a best practice. We present this frequency for each DMU, which is identified as a peer, in Table E.4 of Appendix E. In this table, the University of Florida is the first ranked with 82 times being identified as a peer to other universities. University of Tulsa, Brandeis University, Tennessee State University and Andrews University are ranked two to five in this list respectively.

We also locate each university on the US map to better communicate the efficiency scores. Figure 4.5 shows the geographical distribution of calculated efficiency scores under Model 1. The darkest green dots show the most efficient units with close to 1 efficiency scores. The red dots are the least efficient DMUs and the yellow dots are average. We will further investigate the characteristics of the efficient and inefficient DMUs in the next sections.

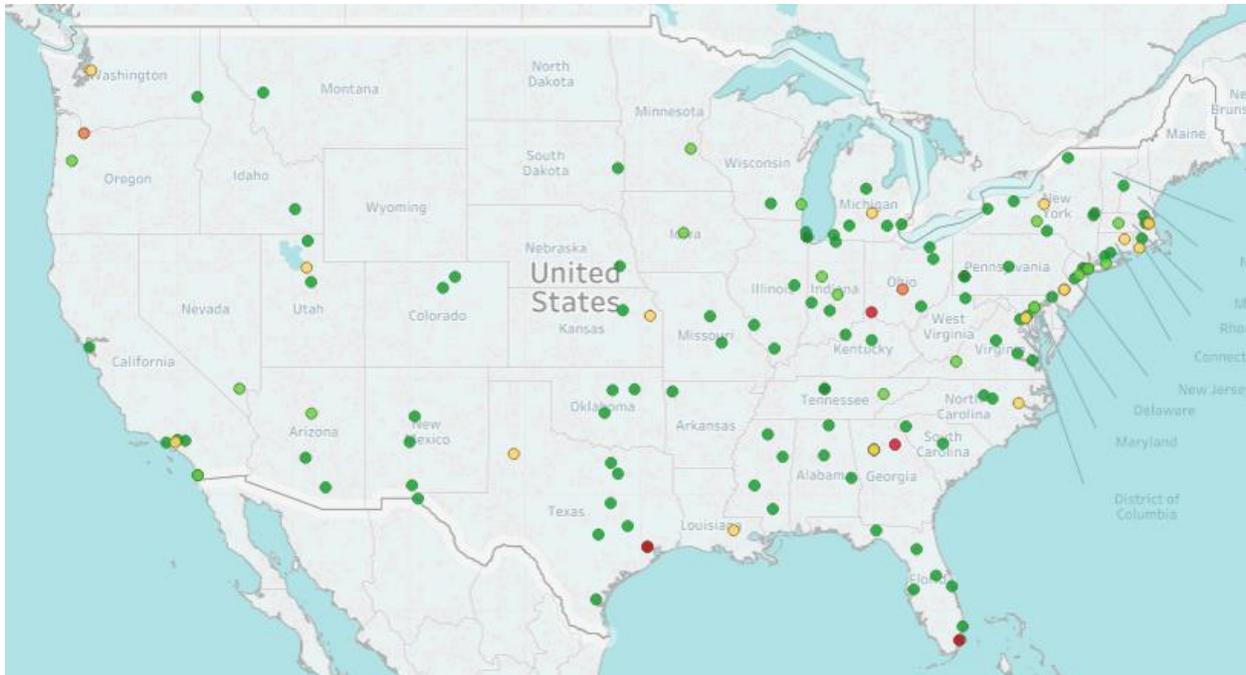


Figure 4.5: DEA efficiency scores distribution on the US map.

It is also interesting to look at the estimated efficiency scores of Ivy League universities. We have the data of seven out of eight Ivy League universities. As depicted by Table 4.5, three universities are identified as BCC-efficient according to Model 1; Yale University, Dartmouth College and Princeton University. The other four also show relatively high efficiency scores.

Table 4.5: Ivy League universities.

Name	Efficiency Score	Freq.	Peers (λ)
Yale University	1.00	13	-
Princeton University	1.00	11	-
Dartmouth College	1.00	7	-
Harvard University	0.961	-	University of Florida (0.04), University of Tulsa (0.33), University of Wisconsin-Madison (0.53), Yale University (0.10)
Brown University	0.957	-	Dartmouth College (0.66), Princeton University (0.07), University of Florida (0.05), University of Illinois at Urbana-Champaign (0.04), University of Tulsa (0.18)
Cornell University	0.947	-	Dartmouth College (0.26), Rice University (0.41), University of Texas at Austin (0.10), University of Florida (0.10), University of Tulsa (0.13)
Columbia University	0.909	-	Princeton University (0.37), University of Chicago (0.21), University of Michigan- Ann Arbor (0.10), University of North Carolina at Chapel Hill (0.32)

Hypotheses testing

As discussed in method section, we use double-bootstrapped DEA analysis to test the effect of several environmental variables on the effectiveness of training minorities. The variables are ratio of minority residents in the geographic location, affirmative action programs indicator, wealth, private indicator and acceptance rate. In the regressions, we also control for campus size. Ratio of minority residents in the geographic location, wealth and acceptance rate are continuous variables, while affirmative action indicator and private indicator are binary. Affirmative action

indicator is zero if the state of the university has banned the affirmative action policies and otherwise it is one. Private/public indicator is one if the university is private and zero if it is public.

We use the same r-package as we used to conduct the DEA analysis; rDEA. It provides the results of truncated regression analysis for the environmental variables. The dependent variable is the bias-corrected efficiency score and the independent variables are the environmental variables. The algorithm is based on the second algorithm of Simar and Wilson's paper [53] , which is a double bootstrap procedure and can be used to provide inference about the regression coefficients. The first step of the algorithm is calculating bias-corrected distance functions with L_1 replications. The second stage of the algorithm obtains a set of bootstrapped estimates to the regression. For calculating the confidence intervals, we assumed the significance level or alpha to take several values and the algorithm does 100 replications for the first loop and 2000 replications for the second loop. We run the analysis for different values of the significance level; 0.01, 0.05 and 0.10 and use the results to determine the level of significance as displayed by the asterisks in Table 4.6. We recognize that according to the assumptions of Simar and Wilson [53] algorithm, the environmental variables should be separable from the designed set of inputs and outputs. Thus, we checked for the correlations between each pair of the variables and the results confirm that environmental variables of our study are separable from the inputs and outputs.

The result of the truncated regression analysis for each combination of DEA models and environmental variables is shown in Table 4.6. Campus size shows a significant effect on the efficiency scores of all three models with a positive coefficient ($p < 0.01$). Ratio of minority residents in the city is significantly and positively correlated with the efficiency scores in all three models ($p < 0.05$ in M1 and M2, $p < 0.10$ in M3). Acceptance rate of the university is significantly and negatively correlated with the efficiency scores in models 2 and 3 ($p < 0.05$ in M2 and $p < 0.01$

in M3). The previously mentioned correlations mean that universities with larger campuses, those located in more diverse locations, and more selective universities have higher minority graduation rates, GPA, and salary. Private indicator is only positively correlated with the efficiency scores of model 2; only quantitative output measures ($p < 0.10$). In other words, private university indicator was associated with higher graduation rate of minorities, while financial resources of universities was positively associated with lower average GPA and salary of minorities. Wealth or financial resources of the institutions is negatively correlated with the efficiency scores of model 3; only the qualitative output measures ($p < 0.05$). The last variable is affirmative action indicator, which is only positively correlated with the efficiency scores of model 2 ($p < 0.01$). This means that Affirmative Action was associated with higher rates of graduation of URM but not necessarily with higher GPA and salary.

Table 4.6- The results of the bootstrapped second-stage regression on the environmental factors.

Environmental Variables	Model 1			Model 2			Model 3		
	Coef.	Lower bound (2.5%)	Upper bound (97.5%)	Coef.	Lower bound (2.5%)	Upper bound (97.5%)	Coef.	Lower bound (2.5%)	Upper bound (97.5%)
Intercept	0.58 *	-0.10	0.88	-14.10 *	-33.72	-8.38	-1.18	-9.29	0.53
Campus size	0.17 ***	0.09	0.35	3.42 ***	2.155	7.81	0.46 ***	0.13	1.81
Minority Residents Ratio	0.31 **	0.05	0.76	3.66 **	0.68	10.81	0.91 *	-0.06	4.33
Acceptance Rate	-0.04	-0.15	0.04	-1.17 **	-3.48	-0.25	-0.41 ***	-1.79	-0.08
Private Indicator	0.08	-0.11	0.31	2.55 *	0.17	7.80	0.32	-0.43	2.35
Wealth	-0.04	-0.19	0.03	-0.57	-3.22	0.10	-0.36 **	-1.80	-0.01
Affirmative Action Indicator	0.03	-0.10	0.20	3.30 ***	1.78	9.15	0.26	-0.20	1.54

* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Discussion

In this study, we applied DEA framework to estimate the effectiveness of higher education institutions in training ethnic and racial minorities. We assumed African-Americans and Hispanics as the two major groups of ethnic and racial minorities. For this purpose, we construct a diversity dataset from several federal data sources (e.g., IPEDS, NSCG and CSS) including four input and four output variables. Number of enrolled minority undergraduate students, number of enrolled graduate students, dollar amount of student grants and number of minority faculties are the inputs. The outputs can be grouped into two categories; desirable and undesirable. Desirable output variables are the number of graduating undergraduate students, average GPA and average salary of minority students. We also consider number of race-based hate incidents as an undesirable output variable. We run an output-oriented DEA with variable returns-to-scale on the constructed dataset including 154 universities. Furthermore, we use Simar and Wilson's double-bootstrap approach [53] to test the effect of several institutional and institution's environmental factors on the estimated efficiency scores. As described in background section, the selected factors will add to the theory of underrepresentation of minority students in the US higher education mainly focused on the effect of environmental and institutional characteristics.

This study employs three DEA models to estimate an overall efficiency score for the success of minority students in universities (M1) besides an efficiency score for the quantity of graduating minorities (M2) and another one for the quality of their educational outcomes (M3). Comparing the number of efficient DMUs and the range of efficiency scores brings out the following conclusions. One major finding of this study is that performance of most universities in terms of improving quantity vs quality of URM graduates differ, and some factors shown to influence only one of these. We find that less than one third of the universities of this study are efficient in the

full model. However, if we only consider quantity and quality output measures, less universities are found to be completely efficient. M3 yields the least efficient DMUs. This means that more universities need to focus on the ability of their minority students to receive good grades and find high paid jobs. However, still a good number of universities need to work on the number of their graduating minorities. Generally, the elite universities received high efficiency scores. Among the identified best practices, we see Ivy League universities that both excel in education quality and diversity practices including Yale University, Princeton University and Dartmouth College. The efficiency scores of other elite schools are relatively high.

The results of the study indicate that macro *environmental factors* can significantly explain the variations in higher education institutions' performance regarding training URM students. Characteristics of the geographic location of universities affects the effectiveness of their diversity practices on higher educational outcomes of minority students. The results of the regression show that ratio of the minority residents significantly explains the pattern of the efficiency scores' estimations. It has a positive correlation with the efficiency scores of all three models. We propose several hypothesis to explain this positive relationship. First, according to “cultural difference theory”, lack of comprehension of the cultural differences between dominant and minority groups causes lower educational outcomes of minority students [22]. In a more diverse city or town, faculty, staff and students are exposed to these cultural differences on a daily basis and it might affect their understanding of these differences. Thus, minority students are better integrated on campuses, which are located in places with higher ratio of minority residents.

The second environmental factor, which significantly explains the pattern of efficiency scores is *affirmative action policies*. Those universities which are located in states that apply affirmative action policies have higher efficiency scores only in model 2, which accounts for quantitative

outcomes of minority students. Applying affirmative action policies helps the number of admitted minorities as well as the number of graduating minorities. This finding is in accordance with the results of another study focused on elite universities [65]. However, it does not have a positive or negative effect on the efficiency scores of model 3, which considers quality outcomes. Maybe applying affirmative action policies in the admission process, reinforces the stereotypes against minority students on one side and improves their social integration through increasing their student population. Therefore, these two phenomena even out and applying affirmative action policies does not have a significant effect on the estimated efficiency scores of model 3.

Furthermore, institution-level factors explain the variation in higher education institutions' performance regarding training URM students. The way institutions admit students and where those applicants are coming from are among the influential institution-level factors. According to the results, acceptance rate of universities has a negative correlation with the estimated efficiency scores. In other words; the more institution selectiveness, the more efficiency score. First way of explaining this relationship is that selective or high ranked universities have more applicants compared to their admissions (i.e., lower admission rates). Thus, their chance of finding and admitting minority students with higher academic preparation from a larger pool of applicants. As a result, the minority students who are enrolled in elite schools are better academically and socially integrated and have higher educational outcomes. This finding contradicts the literature that states the campus climate in elite schools is worse for minority students and they are subject to more negative stereotypes [10]. The second way of explaining this finding is that selective schools have more academic infrastructure to support the minority students through their studies. Furthermore, in elite schools the expectations from the students is higher. This might lead to higher educational

aspirations in the minority students and higher educational outcomes. Moreover, their reputation is also an important factor in helping the graduating students land higher paid jobs.

Wealth and university type (i.e., private versus public) also explain the pattern of the estimated efficiency scores of universities studied in the current research. As mentioned earlier, the tuition in *private* universities is not differently set for in-state and out-of-state students. Therefore, the applicants for the private institutions are more geographically diverse compared to public schools. We think that this geographic diversity of applicants might attribute to higher ethnic and racial diversity of the enrolled and graduating students. Despite how we hypothesized *wealth* and financial resources of universities positively affect their ability to promote diversity on the campuses, the results show that wealth of the institution is negatively correlated the efficiency scores only in model 3- only quality measures of the minority students. Wealthier universities attract students from high economic status families [66]. We hypothesize that when the majority of students are from top of the income scale, the campus climate is less welcoming for ethnic and racial minorities who are on average from lower economic status families [67]. Therefore, it would be more difficult for them to socially and academically integrate on these campuses and get good grades. We can also make another hypothesis about the priority of diversity programs in wealthy universities. Decision-makers at institution-level might find other ways of establishing universities' reputation more attractive and achievable, when they have more financial resources.

Implications

According to our current knowledge, this study is the first that applies DEA to diversity in higher education context. It is also one of the first studies that does a quantitative performance measurement for diversity and inclusion in higher education at institution-level. One of the distinctions of using DEA method is that it enables us to take into account multiple goals and

estimate the efficiency scores of universities based on both diversity and inclusion measures. In comparing the effectiveness of higher education institutions in training ethnic and racial minorities, it is important to avoid only evaluating the superficial diversity or the number of minority students. Other factors should also be considered to achieve a comprehensive evaluation of diversity and inclusion in universities. In this research, other than the number of minority students graduating from the universities, we consider their qualitative outcome measures (e.g., GPA and salary) as desirable outputs. The number of race-based hate incidents that happened on each campus is also considered as an undesirable factor. We constructed our unique database from the public data which were available at institution level. Obviously, the study can be improved as more educational data are available at institution-level for minority students specifically. This current study is a beginning to use performance measurement approaches in higher education diversity context. The line of research can be continued in several ways such as dynamic efficiency measurement of diversity in higher education institutions over time using system dynamics approach.

The results of this study can be informative for decision-makers, especially those at higher education institutions. The outcomes of this study can also be useful for individuals such as students and their parents to identify the best higher education institutions that have a higher potential to help them achieve their educational aspiration. Considering the fact that none of the current university rankings considers a diversity and inclusion ranking and neither there exists a separate diversity and inclusion ranking. This research has the capacity of being further extended to create a diversity and inclusion ranking at a larger scope. Hopefully we can have substantial effect on universities' behavior and help them identify best practices (i.e., DEA analysis peers) and better understand the environmental barriers in front of their diversity initiatives.

References

1. Harvey L, Green D. Defining quality. *Assessment & evaluation in higher education*. 1993;18(1):9-34.
2. Lagrosen S, Seyyed-Hashemi R, Leitner M. Examination of the dimensions of quality in higher education. *Quality assurance in education*. 2004;12(2):61-9.
3. Gutmann A. *Democratic education*: Princeton University Press; 1999.
4. Hurtado S. Linking diversity with the educational and civic missions of higher education. *The Review of Higher Education*. 2007;30(2):185-96.
5. Obama B. Executive Order 13583-Establishing a coordinated government-wide Initiative to promote diversity and inclusion in the federal workforce. *Daily Compilation of Presidential Documents*. 2011:1-3.
6. Office of Management (OM), Office of Human Resources (OHR), (OEEOS) EEOS. U.S. Department of Education Fiscal Year 2016-2019
Diversity and Inclusion Strategic Plan. In: Education USDo, editor. 2015.
7. Jacobs JA. Gender inequality and higher education. *Annual review of sociology*. 1996;22(1):153-85.
8. Harper SR, Patton LD, Wooden OS. Access and equity for African American students in higher education: A critical race historical analysis of policy efforts. *The Journal of Higher Education*. 2009;80(4):389-414.
9. Benton MA. Challenges African American students face at predominantly White institutions. *Colorado State University Journal of Student Affairs*. 2001;10:21-8.
10. Solorzano D, Ceja M, Yosso T. Critical race theory, racial microaggressions, and campus racial climate: The experiences of African American college students. *Journal of Negro Education*. 2000:60-73.
11. Gradín C. Poverty among minorities in the United States: Explaining the racial poverty gap for Blacks and Latinos. *Applied Economics*. 2012;44(29):3793-804.
12. Swail WS. *Retaining Minority Students in Higher Education: A Framework for Success*. ASHE-ERIC Higher Education Report. Jossey-Bass Higher and Adult Education Series: ERIC; 2003.
13. Hall C. *African American College Students at a Predominantly White Institution: Patterns of Success*. 1999.
14. Davies S, Guppy N. Fields of study, college selectivity, and student inequalities in higher education. *Social forces*. 1997;75(4):1417-38.
15. Dennis JM, Phinney JS, Chuateco LI. The role of motivation, parental support, and peer support in the academic success of ethnic minority first-generation college students. *Journal of college student development*. 2005;46(3):223-36.
16. Reardon SF, Baker R, Klasik D. Race, income, and enrollment patterns in highly selective colleges, 1982-2004. Center for Education Policy Analysis, Stanford University Retrieved from <http://cepa.stanford.edu/content/race-income-and-enrollmentpatterns-highly-selective-colleges-1982-2004>. 2012.
17. Fischer MJ, Massey DS. The effects of affirmative action in higher education. *Social Science Research*. 2007;36(2):531-49.
18. Hurtado S, Clayton-Pedersen AR, Allen WR, Milem JF. Enhancing campus climates for racial/ethnic diversity: Educational policy and practice. *The Review of Higher Education*. 1998;21(3):279-302.
19. Museus SD. Generating Ethnic Minority Student Success (GEMS): A qualitative analysis of high-performing institutions. *Journal of Diversity in Higher Education*. 2011;4(3):147.
20. Palmer RT, Maramba DC, Holmes SL. A contemporary examination of factors promoting the academic success of minority students at a predominantly White university. *Journal of College Student Retention: Research, Theory & Practice*. 2011;13(3):329-49.
21. Kobrak P. Black student retention in predominantly white regional universities: The politics of faculty involvement. *The Journal of Negro Education*. 1992;61(4):509-30.
22. Banks JA. Ethnicity, class, cognitive, and motivational styles: Research and teaching implications. *The Journal of Negro Education*. 1988;57(4):452-66.
23. Allen W. The color of success: African-American college student outcomes at predominantly White and historically Black public colleges and universities. *Harvard Educational Review*. 1992;62(1):26-45.
24. Guiffrida D. To break away or strengthen ties to home: A complex issue for African American college students attending a predominantly White institution. *Equity & Excellence in Education*. 2005;38(1):49-60.
25. Owlia MS, Aspinwall EM. A framework for measuring quality in engineering education. *Total Quality Management*. 1998;9(6):501-18.

26. Tinto V. Dropout from higher education: A theoretical synthesis of recent research. *Review of educational research*. 1975;45(1):89-125.
27. Kuh GD, Kinzie JL, Buckley JA, Bridges BK, Hayek JC. What matters to student success: A review of the literature: National Postsecondary Education Cooperative Washington, DC; 2006.
28. Tinto V. Building community. *Liberal Education*. 1993;79(4):16-21.
29. Eisaguirre L. *Affirmative action: A reference handbook*: Abc-Clio Incorporated; 1999.
30. Sander RH, Taylor S. *Mismatch: How affirmative action hurts students it's intended to help, and why universities won't admit it*: Basic Books; 2012.
31. Crosby FJ, Iyer A, Clayton S, Downing RA. Affirmative action: Psychological data and the policy debates. *American Psychologist*. 2003;58(2):93.
32. Hinrichs P. The effects of affirmative action bans on college enrollment, educational attainment, and the demographic composition of universities. *Review of Economics and Statistics*. 2012;94(3):712-22.
33. Backes B. Do affirmative action bans lower minority college enrollment and attainment? Evidence from statewide bans. *Journal of Human Resources*. 2012;47(2):435-55.
34. Arcidiacono P, Aucejo EM, Hotz VJ. University differences in the graduation of minorities in STEM fields: Evidence from California. *The American Economic Review*. 2016;106(3):525-62.
35. Cooper W, Seiford L, Tone K. *Data envelopment analysis: a comprehensive text with models, references and DEA-Solver Software applications*. Boston: Kluwer Academic Publishers; 2000.
36. Sarrico CS, Dyson RG. Using DEA for Planning in UK Universities-An Institutional Perspective. *The Journal of the Operational Research Society*. 2000;51(7):789-800.
37. Hiraldo P. The role of critical race theory in higher education. *The Vermont Connection*. 2010;31(1):7.
38. Thanassoulis E, Da Conceição M, Silva Portela A. School Outcomes: Sharing the Responsibility Between Pupil and School. *Education Economics*. 2002;10(2):183-207.
39. Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. *European journal of operational research*. 1978;2(6):429-44.
40. Thanassoulis E, De Witte K, Johnes J, Johnes G, Karagiannis G, Portela CS. Applications of data envelopment analysis in education. *Data Envelopment Analysis*: Springer; 2016. p. 367-438.
41. Thanassoulis E, Dyson R, Foster M. Relative efficiency assessments using data envelopment analysis: an application to data on rates departments. *Journal of the Operational Research Society*. 1987:397-411.
42. De Witte K, López-Torres L. Efficiency in education: a review of literature and a way forward. *Journal of the Operational Research Society*. 2017;68(4):339-63.
43. Agasisti T, Johnes G. Beyond frontiers: comparing the efficiency of higher education decision-making units across more than one country. *Education Economics*. 2009;17(1):59-79.
44. Avkiran NK. Investigating technical and scale efficiencies of Australian universities through data envelopment analysis. *Socio-Economic Planning Sciences*. 2001;35(1):57-80.
45. Cohn E, Rhine SL, Santos MC. Institutions of higher education as multi-product firms: Economies of scale and scope. *The review of economics and statistics*. 1989:284-90.
46. Katharaki M, Katharakis G. A comparative assessment of Greek universities' efficiency using quantitative analysis. *International journal of educational research*. 2010;49(4):115-28.
47. Thanassoulis E, Kortelainen M, Johnes G, Johnes J. Costs and efficiency of higher education institutions in England: a DEA analysis. *Journal of the Operational Research Society*. 2011;62(7):1282-97.
48. Charnes A, Cooper WW, Rhodes E. Evaluating program and managerial efficiency: an application of data envelopment analysis to program follow through. *Management science*. 1981;27(6):668-97.
49. Portela MCS, Camanho AS, Borges D. Performance assessment of secondary schools: the snapshot of a country taken by DEA. *The Journal of the Operational Research Society*. 2012;63(8):1098-115.
50. Johnes J. Measuring teaching efficiency in higher education: An application of data envelopment analysis to economics graduates from UK Universities 1993. *European Journal of Operational Research*. 2006;174(1):443-56.
51. Lee BL. Efficiency of research performance of Australian Universities: a reappraisal using a bootstrap truncated regression approach. *Economic Analysis and Policy*. 2011;41(3):195-203.
52. Kounetas K, Papathanassopoulos F. How efficient are Greek hospitals? A case study using a double bootstrap DEA approach. *The European Journal of Health Economics*. 2013;14(6):979-94.
53. Simar L, Wilson PW. Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of econometrics*. 2007;136(1):31-64.

54. Tziogkidis P. Bootstrap DEA and hypothesis testing. Cardiff Economics Working Papers, 2012.
55. U.S. Department of Education OoP, Evaluation and Policy Development and Office of the Under Secretary. Advancing diversity and inclusion in higher education. In: Education USDo, editor. Washington D.C.2016.
56. Milem JF, Chang MJ, Antonio AL. Making diversity work on campus: A research-based perspective: Association American Colleges and Universities Washington, DC; 2005.
57. Aguirre Jr A, Martinez R. Leadership practices and diversity in higher education: Transitional and transformational frameworks. *Journal of Leadership Studies*. 2002;8(3):53-62.
58. Downey JP, Stage FK. Hate crimes and violence on college and university campuses. *Journal of College Student Development*. 1999;40(1):3.
59. Lyons CJ. Individual perceptions and the social construction of hate crimes: A factorial survey. *The Social Science Journal*. 2008;45(1):107-31.
60. Simar L, Wilson PW. Non-parametric tests of returns to scale. *European Journal of Operational Research*. 2002;139(1):115-32.
61. Simar L, Wilson PW. Inference by the m out of n bootstrap in nonparametric frontier models. *Journal of Productivity Analysis*. 2011;36(1):33-53.
62. Färe R, Grosskopf S, Lovell CK, Pasurka C. Multilateral productivity comparisons when some outputs are undesirable: a nonparametric approach. *The review of economics and statistics*. 1989:90-8.
63. Dyson RG, Allen R, Camanho AS, Podinovski VV, Sarrico CS, Shale EA. Pitfalls and protocols in DEA. *European Journal of operational research*. 2001;132(2):245-59.
64. Seiford LM, Zhu J. Modeling undesirable factors in efficiency evaluation. *European journal of operational research*. 2002;142(1):16-20.
65. Espenshade TJ, Chung CY. The opportunity cost of admission preferences at elite universities. *Social Science Quarterly*. 2005;86(2):293-305.
66. Aisch G, Buchanan L, Cox A, Quealy K. Some Colleges Have More Students From the Top 1 Percent Than the Bottom 60. *Find Yours*. The New York Times. 2017.
67. Aries E, Seider M. The interactive relationship between class identity and the college experience: The case of lower income students. *Qualitative Sociology*. 2005;28(4):419-43.

Chapter 5. Conclusions

Overall diversity of higher education institutions in the U.S. is a key driver of their excellence in the world. Equal inclusion of people from different backgrounds is essential for any democratic society. A diverse higher education system enhances well-being at the national, institutional, and individual levels. At the national level, it leads to equal inclusion of individuals of different backgrounds in the economic and political life of the country and improves economic competitiveness. To remove negative stereotypes of minority groups and strengthen communities and workplaces, it is necessary to ensure diverse bodies of students and faculties. Studying in a diverse environment benefits all individuals who comprise the future generation of the country's educated workforce. The experience of daily communication with students of different cultural backgrounds enhances one's educational experience.

The gender and racial composition of higher education varies in fields of study and education levels. In many fields, especially engineering, women and racial minorities are less represented. Studies show that women are also underrepresented in math-intensive fields, such as engineering and economics. Moreover, women and racial minorities are less represented in higher levels of employment, such as managerial positions and tenure-track faculty positions. The lack of diversity in all of these areas is a result of historical discriminations against women and racial minorities, as well as the complex interactions of national-, institutional-, and individual-level factors. Universities are seeking ways of improving diversity in their student, staff, and faculty bodies.

In this dissertation, we employed several industrial and systems engineering and management science methods including system dynamics, data envelopment analysis, and statistical analysis to address some of the challenges of gender and racial diversity in higher education. The dissertation included three essays. In essay 1, we presented a systems perspective of the U.S. higher education system and decomposed the educated workforce by gender groups. We developed a system dynamics model to replicate the historical data and forecasted future trends for men and women in higher education.

In essay 2, we focused on work-life balance programs in universities and, more specifically, the effect of tenure-clock extension policies on new-parent tenure-track faculty. We constructed an institution-level data set of tenure-clock extension policies. To the best of our knowledge, this is the first collected data set on tenure-clock extension policies that include major U.S. academic institutions, and such data could be further expanded to include other universities. This paper contributes to studies of the effectiveness of work-life balance programs on career outcomes of tenure-track faculty. Stigma is often attached to using tenure-clock extension for family and childcare responsibilities. Faculty might be reluctant to apply for tenure-clock stop because of possible negative perceptions from their colleagues. Automatic approval of tenure-clock requests help to remove the attached stigma to requesting such extensions. Higher-ranked universities have adopted an automatic tenure-clock extension policy more often than lower-ranked universities, and Southern universities have adopted this policy even less often. Universities first adopted this policy in the late 1980s. A high adoption rate is observed in years 2007 and 2008. Our results indicate that the average length of tenure-track for new parents is greater in universities that adopt automatic tenure-clock extension policies, which is an indicator of

a removed or reduced stigma. Moreover, in universities that adopt the policy, tenure-track faculty generally report higher levels of job satisfaction. Also, although faculty with newborns report higher job satisfaction, in policy-adopter universities, the average job satisfaction is no different between faculty with newborns and faculty without newborns. Our findings also show that the policy does not significantly affect chances of obtaining tenure, as pre-tenure faculty might have various personal or professional reasons to leave tenure-track academic positions.

In essay 3, we focused on the racial diversity of higher education institutions as another important aspect of diversity in higher education. We investigated the relative effectiveness of training racial minority students in U.S. universities. Given that a university's environment affects their operations, we tested the effect of different factors that could facilitate or hinder the effectiveness of universities' diversity programs. This study is one of the first that applies a DEA approach in diversity in a higher education context and compares the training effectiveness of U.S. universities through quantitative measures. We constructed a unique data set of diversity and inclusion measures at the institutional level with the potential of expansion to more data points and variables. Our findings show that universities located in towns or cities with a diverse population, as well as selective universities, are more successful in effectively training minority students. Furthermore, future research could apply the same method to the educational outcomes of a subset of racial groups or a specific field of study. System dynamics modeling could be adopted to continue this line of research with the goal of capturing the dynamics of efficiency scores over time for different universities.

This dissertation as a whole provides a systems view of the diversity trend in U.S. higher education. Inclusion of women in higher education, both at undergraduate and graduate levels, has greatly improved. Over the next decades more women will attend higher education institutions and the number of female bachelor and graduate degree holders will reach parity with men. Despite these improvements, gender disparities exist in different areas of higher education. Women are underrepresented in higher levels of education, managerial jobs, and tenure-track positions. Moreover, they are also underrepresented in some fields of study, such as engineering, while other fields, such as the biological sciences, are female-dominated. In addition to identifying the causes of women and racial minorities underrepresentation in the abovementioned areas and providing a systems view of the contributing phenomena, designing and implementing support mechanisms for minorities is of great importance. The issues of childcare and marriage and a general lack of work-life balance of junior faculty are among the most widely cited hindrances to their representation in faculty positions. Therefore, the support mechanisms that universities develop and implement play a critical role in the success and well being of tenure-track faculty. However, no gender heterogeneity is observed in career outcomes of faculty between those universities that adopt an automatic tenure-clock extension policy and those that do not, the policy generally helps to remove faculty stigma in applying for tenure-clock extensions and improves overall job satisfaction. Racial minorities share some reasons with women for their underrepresentation in higher education and the science workforce, such as role modeling and a welcoming campus climate. However, other reasons are specific to this group, such as lack of parental support and college affordability. Universities have adopted and implemented several policies to

improve the representation of racial minorities. Although universities have adopted similar diversity improvement programs to each other, their environmental characteristics correlate with the effectiveness of these programs. Larger campus size, a higher ratio of minority residents in the city or town (i.e., location of the campus), and lower acceptance rate of universities positively affect minorities' GPAs, first job salaries, and graduation rates. Other environmental characteristics include private (vs. public) university indicator and adopting Affirmative Action programs positively affect the graduation rate of minorities, not their average GPA and first job salary. Finally, minority students who studied at universities with more financial resources (i.e., wealthier universities) show lower GPAs and first post-graduate job salaries.

Contributions

The overall goal of this dissertation is to assist higher education institutions in making better decisions about the design and implementation of initiatives that address the issue of diversity. We also introduced innovative industrial and systems engineering and management sciences methods to evaluate the adopted policies, assess the relative effectiveness of diversity programs, and forecast future diversity trends in higher education. This dissertation makes several major contributions to the literature of gender and racial/ethnic diversity in higher education. Scholars in other fields have studied similar questions, but our modeling approach to the problem of diversity is novel and brings new policy and organizational insights.

Essay 1 contributes to forecasting and other analytical methods that attempt to predict gender composition in the U.S. science workforce. This essay provides more accurate gender-composition estimations by considering the details of the education pipeline from

kindergarten to high school to undergraduate and graduate levels. It also offers a holistic view of the education system and its gender balance and a systems view of several existing mechanisms that drive major dynamic trends in enrollment and graduation rates of male and female students. The model is flexible enough to be simulated for a wide range of scenarios regarding future industry needs.

Essay 2 provides an in-depth view of work-life balance programs in universities, especially tenure-clock extension policies due to child birth/adoption, the policies' adoption year, and level of policy adoption. To the best of our knowledge, this is the first attempt to create an institution-level data set of the characteristics of tenure-track extension policy. The study details the effect of such policies on the career outcomes of new-parent junior faculty and looks at the effects at the national level over time, while other studies have mainly focused on a single campus. The study also has policy implications for designing work-life balance programs for junior faculty members.

Essay 3 contributes to the literature of racial diversity in higher education by comparing two contrasting theoretical explanations for institutional success by expanding diversity and identifies best practices in training minority students among U.S. universities. While many studies have focused on the effects of environmental factors such as policy changes on diversity in higher education, our study goes deeper into the characteristics of each institute. It applies a Data Envelopment Analysis approach to a unique domain—diversity in higher education. Moreover, a unique database is constructed to evaluate the diversity outcomes in universities. Our DEA approach also identifies best practices, which is especially helpful in determining the components of a successful diversity program.

The dissertation as a whole also has methodological contributions. Applying three different methods to interrelated questions help us to identify potential synergies of utilizing different modeling methods and techniques.

Broader impact

I am hoping that the dissertation will help educate policy makers at state or institutional levels on ways of transforming universities to become better places for students and faculty. DEA and system dynamics models could provide policy makers with the opportunity to perform policy tests. System dynamics models have been used as “flight simulators” in the past, providing a venue for learning by simulation. These types of models could also be used in educational workshops. The results of this dissertation could inform policy makers about initiatives for better training of their minority students, including women and racial/ethnic minorities. Furthermore, both individuals and policy makers can gain additional insights into the barriers doctorate holders face in following a successful academic career path. In addition, I see the potential of designing case studies on diversity modeling for system dynamics, DEA, or in management systems engineering classes.

Future avenues of research

In the future, the line of research that the current dissertation has followed could be continued through two main avenues. First, an in-depth study of work-life balance of doctorate students, faculty members, and research workforce is a possible path. Special characteristics of doctorate students’ and doctorate holders’ jobs such as high stress levels and the time-consuming and mentally challenging nature of research projects also make their work-life balance an important and critical issue. Therefore, special attention should

be paid to balancing time allocated to work and other aspects of their lives, which can ultimately affect physical and mental health. For example, family decisions on timing for such decisions as marriage and having children is a challenge for doctorate students and fresh PhDs who seek both a successful career and a family. The effect of family decisions' timing on career outcomes of doctorate holders could be tested in future research. It is also interesting to look at the effect of family settings (e.g., marriage, career characteristics of the spouse, and timing of having a baby) on career choices and changes in their desired career paths. This area of research informs policy makers at institutions and on national levels to better assist the future generation of university professors and researchers in maintaining a healthy life and a successful career. Furthermore, exploring the effect of other, less-investigated diversity improvement programs on the retention rate and career outcomes of minorities could also be an extension of the second essay. One example of these programs is grants that are designed to support the reentry of minority doctorate holders to research careers.

The second avenue of research is an extension of the third essay, which looked at a snapshot of training effectiveness of minorities in U.S. universities. Training-effectiveness trends of minorities over a period of time could also be investigated, which could provide an opportunity to explore why some universities are more successful than others in attracting and training racial minorities. One possibility is to look at this problem from a systems perspective using system dynamics modeling and DEA approaches to measure the change in training effectiveness of minorities over a period of time.