

Differences in Student Debt among Demographic Groups in Those Recently Graduating with a Bachelor's Degree 2011-2012

Bailey Ingraham Lopresto,

¹Christiana Care | Value Institute, 4755 Ogletown-Stanton Road, Newark, DE

Abstract

A bachelor's is considered one of the basic ways to improve an individual's increase of long-term earnings. For many fields, a bachelor's is required to be taken seriously as a candidate for desirable employment, even entry level. The amount of debt to obtain a degree also has major impacts on future earnings, savings, and investment in retirement. Analyses have looked at initial income in recent graduates and described the harsh impacts of lower income for women and minorities on repaying what was assumed equal debt upon graduation. This paper analyzes the assumption of equal debt burden among demographic groups. A discrepancy in debt burden between demographic groups indicates disparities and disadvantages in achieving similar financial success. This study focuses on two major outcomes: 1) Average student debt per borrower and 2) the proportion of each demographic group receiving loans for those bachelor graduates from a 4-year college in Virginia. Outcomes are compared between men, women, Caucasian and Minority subgroups using Bayesian methods and cluster modeling at the college level. We will evaluate our hypothesis by calculating credible intervals around the outcomes.

KeyWords: Bayesian, student debt, cluster analysis, demographic disparities

1. Introduction

A bachelor's degree is considered one of the most basic ways to improve an individual's chances to increase long-term earnings (Fry, 2014), and for many fields, a bachelor's degree is required to be taken seriously as a candidate for desirable employment, even entry level (Carnevale, 2010). In 2014, those with a completed bachelor's had median weekly earnings of \$1,101, higher than the overall median of all workers of \$839 and was also higher than the median earnings of those with an associate's, those with some college and those with no secondary education. Those with bachelor's also had an unemployment of 3.5%, lower than the overall rate of 5%, and lower than those with associates (4.5%), those with some college (6.0%) and those with none (6.0-9.0%) (U.S. Bureau of Labor Statistics, 2014).

The amount of debt accrued while pursuing a degree also has major impacts on future earnings (Minicozzi, 2005), savings or investment in capital, retirement, and family. Previous analyses have looked at initial income in recent graduates and described the harsher impacts of that lower income of women and minorities on current or future financial success (Corbett/AAUW, 2014). These studies assume all demographic groups are paying back equal debt burden upon graduation. This paper addresses that assumption by analyzing differences in two common measures of student debt between demographic groups, specifically between Caucasians and Minorities and differences between genders.

If there significant differences, this would suggest that student loans and therefore access to education is another point in career development were certain demographic are at a disadvantage than their peers. This question is important in ascertaining systematic differences in access to opportunities leading to financial success.

This study focuses on two major outcomes 1) average student loan debt per borrower and 2) the proportion of each demographic group receiving loans for those graduating with a bachelor's in the state of Virginia for the 2011-2012 academic year. These outcomes will be compared in between men and women and Caucasians and minorities. An anticipating confounder that will potentially influence borrowing (how much or not at all) is how much is needed to complete a bachelor's is the Institution providing the degree. Controlling for academic institution should simultaneously control for a myriad of other factors including varying cost of living expenses, what each place is charging for classes or other scholastic expenses, average time to complete a degree, social factors that are associated with that school.

Access to this data is fairly limited and reports that address this specific question are limited with the exception of a Pew Research Center Report in 2014 (Fry, 2014) and the American Association of University Women in 2012 (Corbett, 2012) but even these do not directly answer the simple of question of who borrows more often and requires greater funding support. The state of Virginia does report these descriptive summaries to the public by year, institution, and demographic group. To meaningfully analyze this limited information while controlling for institution as a random variable, Bayesian methods were employed while clustering outcome by institution. Calculated credible intervals around the outcomes were used to evaluate differences between groups and the sensitivity of these intervals were compared with 3 different priors for the first outcome (average debt per borrower) and 2 different priors for the second (proportion of group borrowing).

2. Methods

2.1 Data Sources

Information from College Board for average debt per graduating borrower, percentage of graduates borrowing and total number of graduates was used to create less informative priors (College Board, 2011). These national summary statistics were based off the 2009 year estimates and are the result of summarizing self-reporting school survey information from the Common Data Set (CDS). This dataset compiles survey data from 1,075 colleges (a combination of public four-year and private non-profit four-year institutions) that reported, along with other variables, the proportion of graduates that took out student loans and their average debt. This knowledge is purely from the participating schools, any student loans not received through schools' financial office are unknown.

Virginia data was collected by Virginia's public and private-not-for profit institutions and compiled by the State Council of Higher Education for Virginia (SCHEV) into financial aid data files (FADF)(SCHEV, 2010-2012). Similar to the CDS dataset, loans not processed through each college are not reported. An example of the data format available and used to create the R dataset for each group refer to SCHEV 2011-2012. This dataset included information from the 2007/08-2011/12 academic school years. The focus of the analysis was the graduating class of 2011-2012 and the previous year contributed to informative prior information.

Data tables were collected for each institution selecting for a “Four-Year Bachelors Degree”. An institution’s data was only collected and used if it had available data for both 2010-2011 and 2011-2012 academic years and all four groups of interest: “Majority Students/Men”, “Majority Students/Women”, “Students of Color/Men”, “Students of Color/Women”. Data was then compiled and manipulated in excel to form four datasets, Male, Female, Caucasians, Minorities. The number of graduates was calculated from the percentage of borrows and number of borrowers. Variance was estimated assuming a normal distribution of average debt using the 25% and 75% percentiles (assume the distance from the mean to the 75th percentile estimates $\sim 2/3$ of the standard deviation) with the following formula.

$$\hat{\sigma}^2 = \frac{(\text{Percentile}_{75} - \text{Percentile}_{25})^2 * 3^2}{4^2}$$

The mean of each final group is a weighted average of the mean of each initial group. For example, the final Male mean is a weighted average of the Majority Students/Men and Students of Color/Men original data. A similar process was used to calculate the pooled variance.

2.2 Analysis of Outcome 1: Average Debt per Borrower

A hierarchal model was created using institution attended as a cluster. To keep things relatively simple, variance was assumed to be the same within a group for each institution. A Gibbs sampler was used to generate 5000 iteration estimates (first 500 used as the burn-in period) of the Virginia State mean. Starting values for the sampler were the frequentist estimates for each group. Institutions attended were assumed to be independent and therefore exchangeable. Semi-conjugate priors were used to for prior information of the 3 fixed but unknown parameters of interest and each institution-specific mean. The model priors and sampling models needed for this kind of analysis are (Hoff, 2009):

Model Priors

$$\begin{aligned} \frac{1}{\sigma^2} &\sim \text{gamma}\left(\frac{\nu_0}{2}, \frac{\nu_0 \sigma_0^2}{2}\right) \text{Within - Insitution Variance} \\ \frac{1}{\tau^2} &\sim \text{gamma}\left(\frac{\eta_0}{2}, \frac{\eta_0 \tau_0^2}{2}\right) \text{Between Instituion Variance} \\ \mu &\sim \text{normal}(\mu_0, \gamma_0^2) \end{aligned}$$

Sampling Models

$$\begin{aligned} \phi_j &= \{\theta_j, \sigma^2\}, p(y|\phi_j) = \text{normal}(\theta_j, \sigma^2) \\ \psi &= \{\mu, \tau^2\}, p(\theta_j|\psi) = \text{normal}(\mu, \tau^2) \end{aligned}$$

Three sets of priori information were used to test the sensitivity of the results (Table 1). The first and second priors used the national summary statistics from College Board with a large weight (country sample size) and much smaller weight (Table 1). This same summary was used for each group. The third more informative prior was a summary of Virginia data from 2010-2011. Each subgroup at each institution had their own, specific, third prior.

Table 1: Outcome 1 Prior Information

(Because the average debt cannot be below 0, the average was treated as 3 times the standard deviation to estimate a wide uninformative prior)

Variable	Prior 1		Prior 2		Prior 3	
	National Heavyweight	National Lightweight	Virginia Males	Virginia Females	Virginia Caucasians	Virginia Minorities
v_0	500	1	391	574	679	286
σ^2	49,467,778	49,467,778	9,066	8,931	8,777	9,482
η_0	1,140,000	1	9,388	13,780	16,297	6,871
τ^2	49,467,778	49,467,778	9,121	9,344	8,799	10,357
μ^0	21,100	21,100	24,771	25,289	24,398	26,696
γ^2	49,467,778	49,467,778	9,121	9,344	8,799	10,357

2.3 Analysis of Outcome 2: Proportion of Graduates that Borrowed

This outcome used the same cluster structure. Each individual sample size (n_j) for an institution was used and new one were randomly regenerated from a Poisson distribution with a mean n_j averaged from the n_j 's of the current and previous year. A Gibbs sampler was used to generate 5,000 iteration estimates (first 500 used as the burn-in period) of the Virginia State average proportion of borrowing graduates. Institutions attended were assumed to be independent and therefore exchangeable. To simplify calculations, semi-conjugate priors were used to for prior information of the 2 fixed but unknown parameters of interest and each institution specific average portion and n_j . The model priors needed for this kind of analysis are (Hoff, 2009):

Model Priors

$$\mu \sim \text{beta}(\mu_0 * \zeta_0, (1 - \mu_0) * \zeta)$$

$$\zeta \sim \text{Poisson}(\zeta_0)$$

Sampling Models

$$\phi_j = \{\theta_j, n_j\}, p(y|\phi_j, n_j) = \text{binomial}(\theta_j, n_j)$$

$$\psi = \{\mu, \zeta\}, p(\theta_j|\psi) = \text{beta}(\mu\zeta, (1 - \mu)\zeta)$$

Where ζ is the total population size. Two sets of priori information were used to test the sensitivity of the results (Table 2). The first prior used the national summary statistics from College Board with a small weight. This same summary was used for each group but the sample sizes used were specific for each institution and group from the Virginia 2010-2011 data.

Table 2: Outcome 2 Prior Information

Variable	Prior 1		Prior 2		
	National Uninformative	Virginia Males	Virginia Females	Virginia Caucasians	Virginia Minorities

μ_0	.55	.55	.60	.55	.66
ζ_0	1,140,000	17,206	23,036	29,793	10,456

3. Results

3.1 Outcome 1: Average Debt per Borrower

3.1.1 Men vs. Women

1st and 2nd priors that used national data estimated that men borrowed about \$443 more than women. The previous year Virginia prior estimated smaller difference of men borrowing \$136 more (Figures 1-2). All the credible intervals suggested males had more debt per borrower (Table 5).

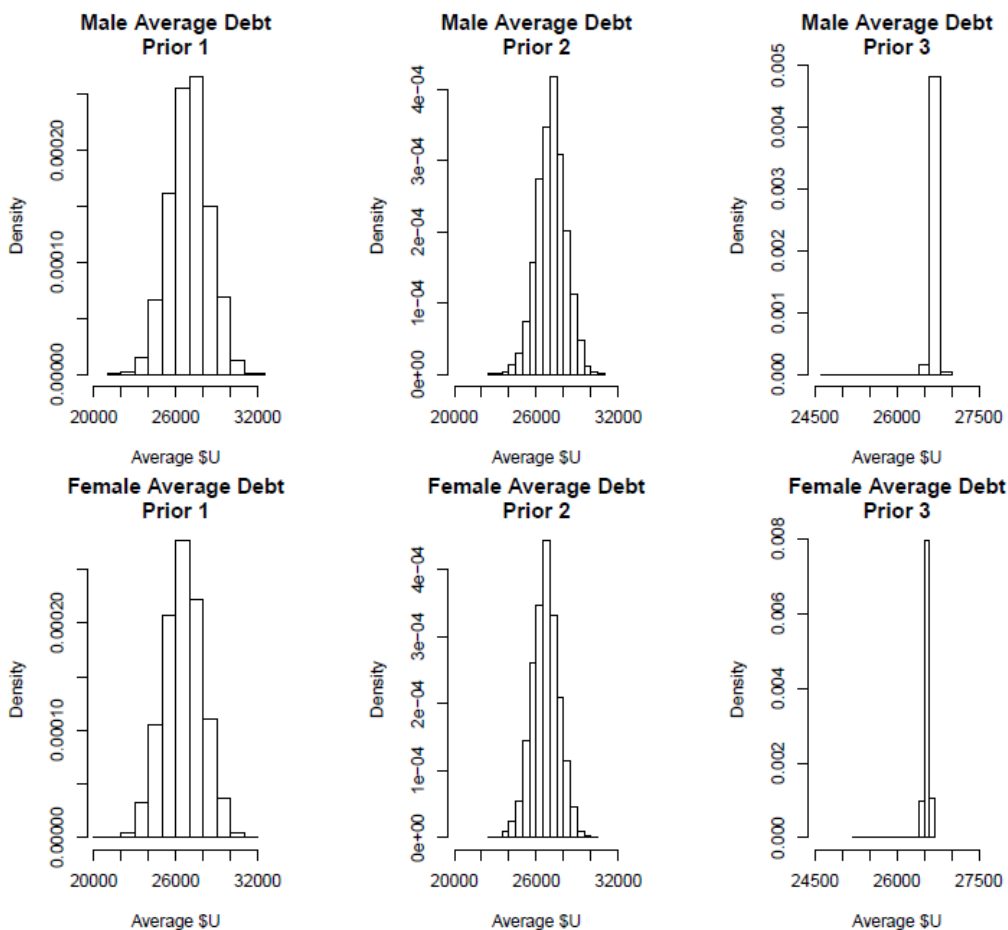


Figure 1: Posterior Distributions of the Average Debt of Males and Females Using Different Priors

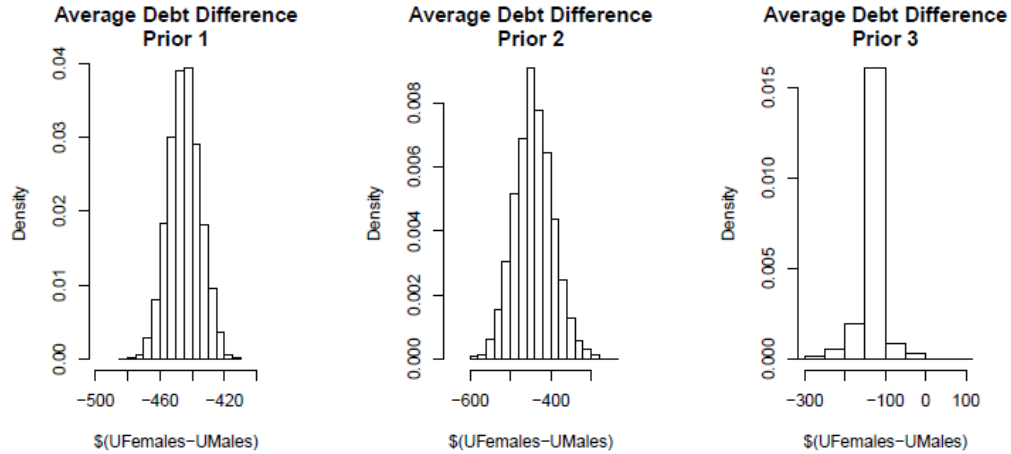


Figure 2:Posterior Distributions of the Mean Difference in Average Debt between Females and Males

Table 3: Mean Difference in Average Debt per Borrower between Males and Females

	<i>Prior 1</i>	<i>Prior 2</i>	<i>Prior 3</i>
<i>Mean Difference</i>	-\$444.87	-\$441.98	-\$136.31
<i>95% Credible Interval</i>	[-\$464.12, -\$425.50]	[-\$532.90, -\$344.03]	[-\$212.37, -\$59.35]
<i>Pr(μ_{Females} > μ_{Males})</i>	0	0	0.002

3.1.2 Caucasians vs. Minorities

1st and 2nd priors estimated that Minorities borrowed about \$2,550 more (mean difference) than Caucasians. The 3rd estimated that Minorities borrowed \$2200 more (Table 4). Credible intervals suggested that minorities graduated with greater debt and the difference is visually more obvious in the posterior distributions (Figures 3-4).

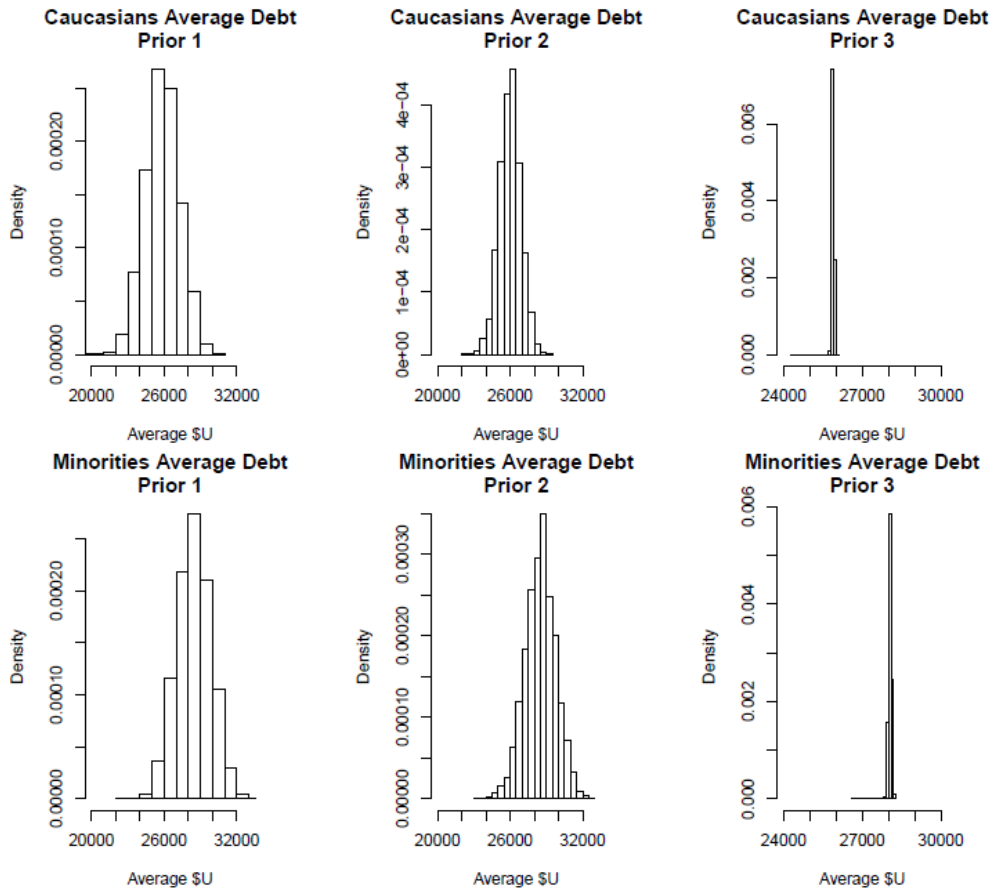


Figure 3:Posterior Distributions of the Average Debt of Caucasians and Using Different Priors

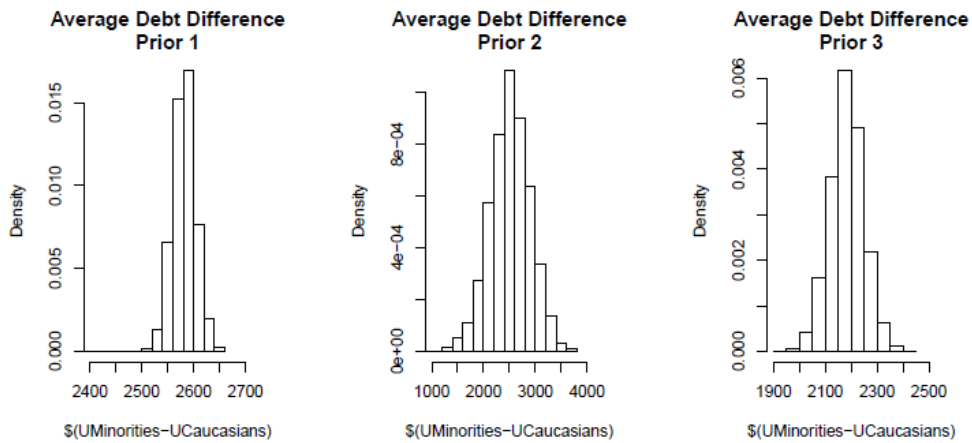


Figure 4:Posterior Distributions of the Mean Difference in Average Debt between Caucasians and Minorities

Table 4: Mean Difference in Average Debt per Borrower between Caucasians and Minorities

	<i>Prior 1</i>	<i>Prior 2</i>	<i>Prior 3</i>
<i>Mean Difference</i>	\$2,581.67	\$2,511.05	\$2,182.99

95% Credible Interval	[\$2,538.44, \$2,625.20]	[\$1,717.28, \$3,249.73]	[\$2,049.49, \$2,314.24]
$Pr(\mu_{Minorities} > \mu_{Caucasians})$	0.99	0.99	0.99

3.2 Outcome 2: Proportion of Graduates that Borrowed

3.2.1 Men vs. Women

The absolute difference in the proportion of female vs male graduates that borrowed was essentially zero for the first prior and .0527 (5.2%) for the second (Table 5). In this case there is an obvious difference in posterior distributions and credible intervals of the second prior but not the first (Figures 5-6).

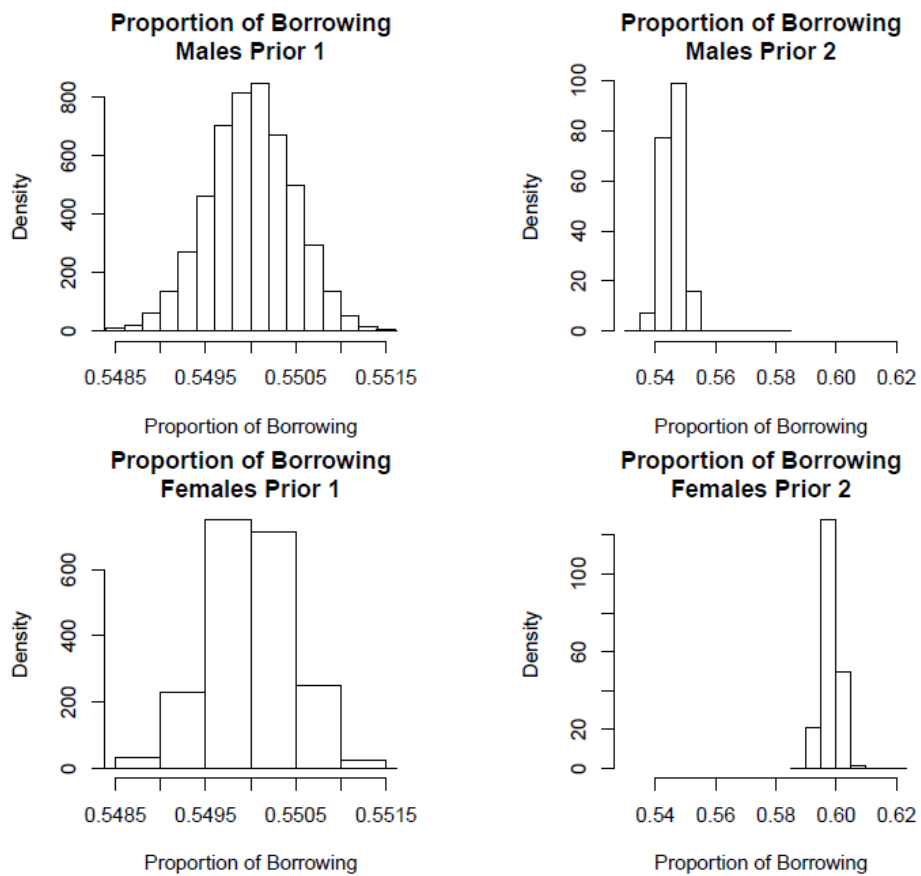


Figure 5: Posterior Distributions of the Proportion of Graduates that Borrowed for Males and Females Using Different Priors

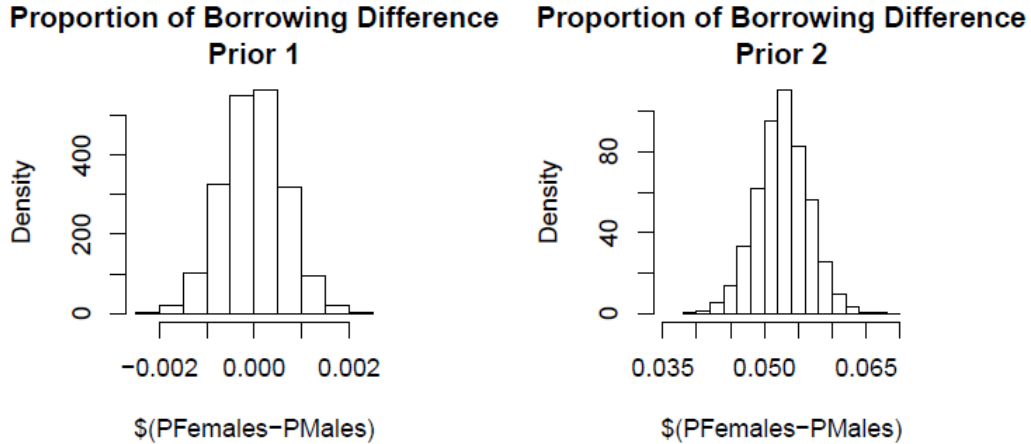


Figure 6: Posterior Distributions of the Mean Difference in Proportion of Graduates that Borrowed Using Different Priors

Table 5: Mean Difference in Proportion of Graduates that Borrowed between Males and Females

	<i>Prior 1</i>	<i>Prior 2</i>
<i>Mean Difference</i>	0.000	0.053
<i>95% Credible Interval</i>	[-0.001, 0.001]	[0.045, 0.060]
<i>Pr($\mu_{Females} > \mu_{Males}$)</i>	0.50	0.99

3.2.2 Caucasians vs. Minorities

The comparison between Caucasians and Minorities was similar, with the absolute difference being very small for the first prior and larger (11% more of Minorities borrowed) for the second (Table 6). Again, the posteriors for the second prior are obviously different than the first and the credible interval of the informative prior does not include zero (Figures 7-8).

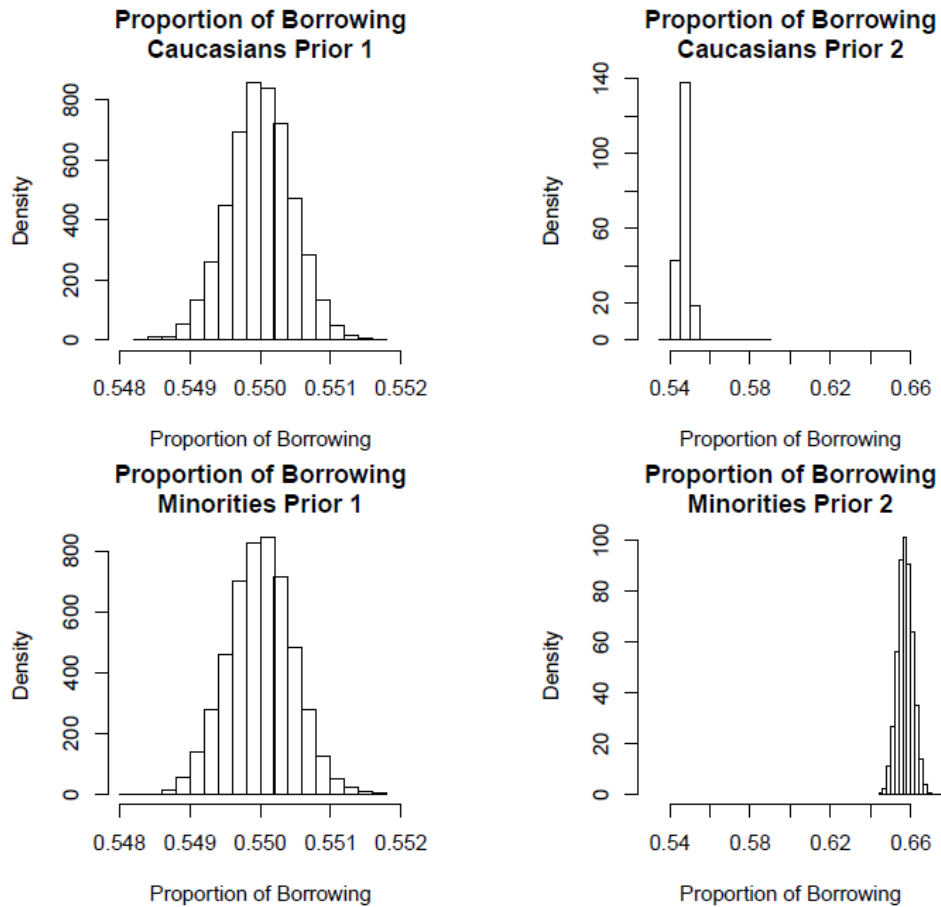


Figure 7:Posterior Distributions of the Proportion of Graduates that Borrowed for Caucasians and Minorities Using Different Priors

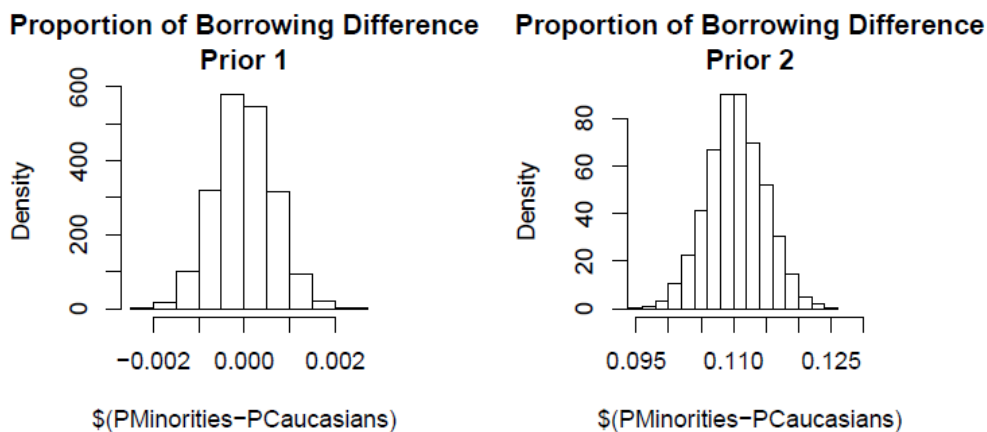


Figure 8:Posterior Distributions of the Mean Difference in Proportion of Graduates that Borrowed Using Different Priors

Table 6: Mean Difference in Proportion of Graduates that Borrowed between Caucasians and Minorities

	<i>Prior 1</i>	<i>Prior 2</i>
<i>Mean Difference</i>	0.000	0.110

<i>95% Credible Interval</i>	[-0.001, 0.001]	[0.102, 0.119]
$Pr(\mu_{Minorities} > \mu_{Caucasians})$	0.49	0.99

4. Discussion

Based on these results, there seems to be a small significant difference in the average debt between males and females; however, from fiscally relevant standpoint a roughly 400 dollar difference in debt probably will not have major impacts on future financial success. The proportion of graduates that borrowed was less sensitive to differences in priori information. This likely due to the first prior's larger weight from the national level. There is potential to try multiple weights, possibly closer to Virginia's total student population and to improve the Gibbs sampler for a better n_j sampling model. Estimates for average debt per borrower are likely more stable due to the outcome's continuous nature and more model assumptions and structure.

This analysis is limited by data availability. While the analysis only includes debt due to a bachelor's degree it does not include debt from any previous schools (College Board, 2011). This means it is likely underestimating average debt and proportion of borrowers for transfer students. In addition, only federal loan data and private loans received through a school were reported. This means that the results will be underestimates of average debt. The amount of missing loans will likely be linked to outcome because the first and most available loans are through the institution. Individuals in need of more funds will be looking towards outside sources that are less likely to be reported. There is also concern about the schools accuracy in reporting. Considering that these reports are easily available and readily used by prospective students to determine preferable schools, there is incentive to under-report debt burden of graduates. This will likely not be related to a proportion of graduates that borrow because, again, the mostly easily received loans are also the most reported.

Other aspects of this work that might be of future interest would be to see how sensitive results are changing the assumptions of the methods or using different distributions. Other things to consider would be to estimate the underestimation and amount of missed student loans, the total burden of debt for a demographic or building a mixed model with gender and race/ethnicity as fixed effects, possibly with an interaction, and institutions as a random effect. Other predictors of interest could also be major, region of school and type of institution.

Overall, the other results suggest that student loan debt is possibly another step of at which Minorities and women experience a systemic disadvantage on the road to wealth accumulation and financial success. These results are consistent to national level reported in Fall 2014 for the academic 2011-2012 year (Fry, 2014). This exemplifies one of the many student loans issues that need to be addressed along potential wage gaps to achieve equality in potential financial success.

Acknowledgements

Thanks to Dr. Tabb at Drexel University for inspiring this project and academic guidance.

Thanks to Seema Sonnad for her encouragements and guidance.

Thanks to Dr. Howard Stauffer (Humboldt State) for his infectious love of Bayesian methods.

Thanks to Value Institute of Christiana Care for giving me time to pursue this project.

References

- Bureau of Labor Statistics. 2014. Earnings and unemployment rates by educational attainment. *Employment Projections*. Washington, DC. United States Department of Labor. accessed 07/15/2015. <http://www.bls.gov/emp/ep_chart_001.htm>.
- Carnevale, A., Smith, N., Strohl, J., 2010. *Help wanted: projections of job and education requirements through 2018*. Washington, DC. Georgetown University. pp. 78-90.
- College Board. 2011. Trends in Student Aid 2011:9a and 10a. *www.collegeboard.org*. New York, NY. The College Board.
<<http://trends.collegeboard.org/student-aid/figures-tables/loans>>.
- Corbett, C., Hill, C., 2012. Graduating to a pay gap: the earnings of women and men one year after college graduation. *AAUW.org*. Washington, DC. American Association of University Women. accessed 03/08/1/2014
<<http://www.aauw.org/files/2013/02/graduating-to-a-pay-gap-the-earnings-of-women-and-men-one-year-after-college-graduation.pdf>>.
- Fry, R. 2014. Young adults, student debt and economic well-being. *Social Demographics and Trends*. Washington, DC. Pew Research Center. accessed 07/15/2015.
<<http://www.pewsocialtrends.org/2014/05/14/young-adults-student-debt-and-economic-well-being/>>
- Fry, R. 2014. The changing profile of student borrowers. *Social Demographics and Trends*. Washington, DC. Pew Research Center. accessed 07/15/2015.
<<http://www.pewsocialtrends.org/2014/10/07/the-changing-profile-of-student-borrowers/>>
- Hoff, P., 2009. A First Course in Bayesian Statistical Methods. Seattle, WA. Chapter 8.
- Minicozzi, A. 2005. The short term effect of educational debt on job decisions. *Economics of Education Review*. Vol 24. (4), pp.417- 430.
- SCHEV. 2010-2012. EOM debt profile: student loans by gender and race 5 years. *Virginia.gov*. Richmond, VA. State Council of Higher Education for Virginia. accessed 04/05/1/2014
<http://research.schev.edu/studentdebt/DebtProfile_SL014.asp>.