Academic Major and Political Attitudes

Movses Musaelian and Matt Vela

Section 1: Introduction

Our survey project regards the relationship between academic major and political attitudes. In gauging this relationship, we would like to see if certain stereotypes regarding people taking certain academic majors really do hold among Carnegie Mellon students. For example, does the notion hold that CFA students are significantly more left wing than Tepper students? Given this, we would like to see if it is possible to make certain inferences regarding quantitative versus more humanities oriented students in terms of political beliefs. In attempting this, we will observe the differences between the various disciplines at CMU, which include CIT,SCS, MCS, Tepper+Econ, CFA, SDS+Stats, and humanities, and see what differences are significant. We recognise that such differences may be minute, for example the difference between CIT and MCS, and thus not give any significance, which could still give us an interesting assessment of the Carnegie Mellon undergraduate population. We hope that a significant result will provide the community a helpful perspective on our diverse population - and in addition provide an interesting comparison with other similar studies.

Our final results showed significant differences in political attitude between MCS and Humanities, CFA and MCS, and finally between Humanities and MCS, which confirmed our expectations. However, we did not see a consistent significant difference between the quantitative and humanities groups, which leads us to give less credence to the stereotype.

There have been a few studies done on this subject matter. One of the most relevant ones was a study done at University of Minnesota which pursued an identical research question to this one, but used different methodology, for example a much shorter and concise questionnaire (Gage et al. 2010). The study implemented ANOVA and subsequently Tukey test to find that a significant difference existed between the business fields and the science/humanities fields (less economically liberal). For social liberalness no significant difference was found. Our results, which had more groups, showed also several significant interactions but similar to the Minnesota study did not see a consistent pattern in differentiation in political beliefs in the areas of study. Another source found was a university article which investigates the liberalisation of students' socio-political orientations (Hastie, Brainne 2007). The article however does not employ a survey, but provides an extensive history and analysis regarding this trend which is crucial to this topic. For example, Hastie discusses two hypotheses given, regarding the "liberalisation in students' socio-political orientations". The first one being self-selection where people will choose disciplines that reflects their beliefs and the other being socialisation where people's beliefs change based on their surrounding peers. The article reports the trend that social science majors are the most liberal while engineering and business students being the least. We did not see such a trend in our data. This third article examines the argument that political orientation is significant in affecting academic success in university (Kemmelmeier et al 2005). The sample units utilised in the paper are undergraduate students at an American university and their respective grades. The statistical analysis includes multilevel regression modelling. While the article doesn't directly address our question, it provides us with a good background in studying this relationship between political orientation and academic success. In all, the study confirms this relationship and the significant role that political orientation plays in predicting academic success.

The fourth paper is regarding voting behaviour of parents and their children at university (Lupkin et al.). The paper includes some useful data regarding political attitudes of students and since this research was done via survey, the questionnaire is shown in the paper. Similarly to the third article, this provides more background to our general topic of study. In short, the article concludes an expected trend that students' political views align with parents more closely than with peers.

Section 2: Methods

The sampling frame for this topic is the CMU student directory includes the undergraduate students at Carnegie Mellon. However, this sampling frame has to be narrowed down to students who have a declared primary major; hence, many first year students may become illegible units. Other illegible units include interdisciplinary majors and part time students. Given this, the units were randomly sampled from CMU's student directory using excel as a numerator and also a random number generator. We also chose not to exclude all the first year students, but we had our first question "filter" students who do not have declared majors. We divided our frame into 7 strata (CIT, MCS, SCS, SDS+Statistics, CFA, Tepper+Econ, Humanities) which represent comparable fields of study. We firstly calculated a SRS sample size with a margin of error of .05 (calculation shown below). We then divided that sample size evenly by 7 among our seven stratums. We did not have any specific prior probabilities of the stratums and did not weight them; though, it may have been wise to apply weights to the sample sizes based on population proportions. We then proceeded to sample from each of the seven stratums, which in all makes our sample a stratified sample.

$$\hat{p} = \frac{1}{2}, z = 1.96\ (assuming\ 95\%\ confidence)$$

$$n \geq \frac{Nn_0}{N + n_0} \rightarrow \frac{(6178)(384)}{6178 + 384} = 362$$

$$\frac{362}{7} \approx 52 \rightarrow sample\ size\ for\ each\ stratum$$

Rather than have the target population be general US undergraduate students, it is confined to just Carnegie Mellon undergraduates. A target population of US undergraduate students would yield substantial coverage errors and given our resources - it is best to make our inference regarding the sphere of Carnegie Mellon. Obviously, our sample selection is a small part of this population and coverage error could come from over representation in a certain school or major. Another source of error could come from the fact that most freshmen will be excluded and hence the input of the younger students will be minimal. Our demographic data will tell us of the extent of such a bias.

The mode of collecting the data was a short questionnaire which will be hosted on one of the online surveying hosts (Surveymonkey). The questionnaires will be sent via email to the sample population. For this topic, this method seems the easiest as one can choose the people of his sample size carefully, representing a variety of majors, and doesn't require much physical involvement. Of course one downside is the lower response rate that will come with this method, thus we will probably have to utilise an incentive scheme.

Response Data

1st Batch

Total Sent	364
Total Responded	76
Total Completed	61
Response Rate	21%

2nd Batch

Total Sent	364
Total Responded	61
Total Completed	49
Response Rate	17%

Total

Total Sent	728
Total Responded	137
Total Completed	110
Response Rate	19%

Strata Response (total completed, response rate)

CIT	(16, 15%)
SCS	(13, 13%)
MCS	(20, 19%)
Humanities	(18, 17%)
CFA	(8, 8%)
Tepper+Econ	(15, 14%)
SDS+Stats	(20, 19%)

Race

White	77 - 70%
Asian	15 - 14%
Black	4 - 4%
Hispanic	7 - 6%
Other	7 - 6%

Geography

Northeast	58 - 53%
South	20 - 18%
Midwest	16 - 15%
West	7 - 6%
International	9 - 8%

Grade:

Freshman	11 - 10%
Sophomore	27 - 25%
Junior	36 - 33%
Senior	36 - 33%

It should be firstly noted that we only counted the individuals that fully completed the survey. Almost all of the incompleted surveys had only the demographic section filled out. We also threw out individual responses that were clearly not genuine. There really wasn't clear non-response information from our sample. It can be seen that CFA was underrepresented in comparison with the other stratums. Demographically, we see a big representation of whites from the Northeast which is expected given the demographics of CMU. In terms of a student's year level, we see a slight under-representation of freshman, which is what we expected, since a good portion of freshmen have not finalised their primary major. We proceed to apply post-stratification weights for major, further work is shown in results section. With that, we can quantify our data via a coding scheme then calculate the weighed means, find variances, and lastly perform statistical inference and gauge significant differences between the stratums.

The questionnaire is relatively short, but still longer and more detailed than the Univ. of Minn. study's. The first things we ask are demographic. For example, we want to know if they have a declared major or not and a question regarding their background (race and homeland). The questions of the content core gauge the individual's attitude towards big/debated questions of political and economic philosophy such as the size of government and government welfare. The other variables gauge the person's views towards the most controversial social issues such as abortion. A combination of these variables gives us a pretty informative picture of the individual as a whole in our sample.

The three sample questions each come from the three sections of the questionnaire. All questions have the same answering scheme:
(I don't know)/(Yes, absolutely)/(Yes probably)/(indifferent)/(probably no)/(no, absolutely not)

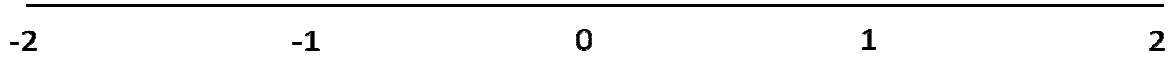
Economic
Is government intervention in the business sector via regulations and monetary stimuli justified outside a recession?

Social
Should traditional gender roles, such as domestic housewives, be broken?

Political
Is it justified to sacrifice liberty for security in a national security crisis?

Section 3: Results

After finishing data collection, we constructing a coding scheme with which we quantified the data. The coding scheme is as follows:



-2: very left | -1: left | 0: Centre | 1: right | 2: very right

Each of our questions was written such that there was a clear political attitude for every possible response. Hence, we assigned an appropriate value to each question answered and in the end averaged all the values to receive a comprehensive value for each respondent. After all the averages were calculated, it was necessary to clean the data. This meant scrutinising each response and deciding whether that response should be kept or should be cleaned out due to non-response or inappropriate answering. After the data was cleaned, we conceived a table of the unweighted averages:

What does "inappropriate answering" mean? Since the questions were multiple choice what made you think an answer was inappropriate?

	Econ	Social	Pol	Total
CIT	-.6860119	-.8303571	.05803571	-.5052057
SCS	-.9025641	-.6190476	-.0970696	-.5364276
MCS	-.6082143	-.7154762	.11642857	-.4159226
Hum	-.6653439	-.944709	.02896825	-.5401037
CFA	-.485119	-1.014881	-.25	-.5896884
Tepp	-.6595238	-.5825397	.13904762	-.3811066
SDS	-.8928571	-.672619	.11261905	-.483145
Total	-.6999477	-.7685185	.0154328	-.4930857

This table displays the raw, unweighted averages for each stratum divided by the three categories and followed by the comprehensive average of each stratum and of each category. While statistical analysis will be conducted on the weighted averages, it's worth looking at the preliminary numbers for trends. Firstly, we see that averages of the stratums are rather close to each other, but relatively Tepper is the most right wing and CFA is the most left wing, which makes intuitive sense. It should be noted though that there are some trends that don't fit our expectations; for example, Tepper being more left wing economically than CFA.

Next came the natural set of constructing our post-stratified weights for our 7 groups. The formula used for constructing weights is $weight = \frac{population\ proportion}{sample\ proportion}$ and hence we proceeded to calculate the respective proportions and subsequence differences, obtaining the following weights:

Stratum	Weight
CIT	1.82
SCS	.752
MCS	.615
Hum	1.04
CFA	2.01
Tepper	.54
SDS	.171

After obtaining the weights, we could proceed to calculate the respective weighted averages. It should be noted that we used the excel program to make the necessary calculations. The equation for a weighted mean with post-stratification weights is:

$$\bar{y}_w = \frac{\sum_i w_i y_i}{\sum_i w_i}$$

For each stratum, we multiplied each respondent's average (y_i) by the respective weight of that stratum. Then we summed all those products into a number and also summed the weights of that stratum into a number. The difference of these two numbers got us the weighted average of the stratum. This process was repeated 7 times. The resulting weighted averages were as follows:

Weighting the strata makes sense if you are computing a mean over the whole population. For looking at the means of individual populations, weighting doesn't make sense since you are just considering a SRS of people within that group.

Stratum	Avg
CIT	-.46955
SCS	-.48884
MCS	-.42049
Hum	-.51721
CFA	-.61526
Tepper	-.37778
SDS	-.49541

After this, we proceeded to calculate the variances. Given the use of the weighed means, we needed to adjust our variances for that and hence used the Taylor Series Method. The formula for the Taylor Series is as follows:

$$Var_{TS}(\bar{y}_w) \approx \frac{1}{(\sum_i w_i)^2}\left[Var(\sum_i w_i y_i) - 2\bar{y}_w Cov(\sum_i w_i y_i\ , \sum_i w_i) + (\bar{y}_w)^2 Var\left(\sum_i w_i\right)\right]$$

We made three separate calculations (the two variances and covariance) for each stratum in order to be able to use the main equation. $Var(\sum_i w_i)$ was calculated by first determining the weight average (which included the weights of all the stratums). Then in each stratum the respective weights were subtracted from the average, squared, and then all the values were summed and finally a population correction was applied. For calculating $ar(\sum_i w_i y_i)$, we first found the $\overline{wy}$ for each stratum, which was basically the average of our $w_i y_i$ values and then we subtracted each $w_i y_i$ from the average and then summed all the values and finally applied the population correction. Finally, for calculating $ov(\sum_i w_i y_i \ , \sum_i w_i)$, we summed the product $(w_i y_i - \overline{wy})(w_i - \overline{w})$ using the different values we had from the previous calculations and then applied the population correction. After the three components were calculated, we used the main equation to calculate the variance for that stratum, and repeated this process 7 times. The subsequent variances obtained are as follows:

Stratum	**Variance**
CIT	.024121
SCS	.026874
MCS	.029429
Hum	.018098
CFA	.057832
Tepper	.172724
SDS	.301804

The following two photos show the set-up that was used for each stratum's calculations:

10	2.01	1.035161		1.753165		1.347147987			
11	2.01	1.035161		1.421031		-1.212846087			
12	2.01	1.035161		0.12572		-0.360749652			
13	2.01	1.035161		1.529377		1.258233576			
14	2.01	1.035161		1.848431		-1.383265375			
15	2.01	1.035161		1.753165		1.347147987			
16	2.01	1.035161		1.464769		-1.231369923			
17	2.01	1.035161	Varw	0.053676	Varyw	0.235717853	Cov	VarTS	
18		*8.281287*	**9.464328**	*9.949333*	**11.37067**	*1.63673E-05*	**1.87055E-05**		**0.057832**

These are not necessary in the body of a paper. If you want to show details of calculations you should put them in an appendix.

0.043478	0.087391	2.01	
-1.20833	-2.42875	2.01	
-0.79167	-1.59125	2.01	
0	0	2.01	
-1.29167	-2.59625	2.01	
0.043478	0.087391	2.01	
-1.21739	-2.44696	2.01	
-0.5	-1.005	2.01	
yi	wiyi	wi	wavg
	-9.89342	*16.08*	**-0.61526**
wyavg	-1.23668		

It should be noted this was the set-up for CFA calculations (hence only 8 rows) and the italicized numbers represent the sum of their column.

After these values were found, the next step entailed the actual statistical inference. In proceeding with this inference, our research topic should be clearly repeated in the context of such inference. What we are trying to gauge is if significant differences, in the average political score, exist between our stratums. However, since we want to know which specific stratums significantly differ between each other, we need to calculate the difference for each possible combination using the small sample test for comparing two means. We hope to see a consistent pattern in the combinations that have a significant difference. For example, CFA having a significant difference with the quantitative schools (e.g. MCS, CIT, etc.). The equation for this comparison and the pooled variance is as follows:

$$t^* = \frac{((\bar{x}_1 - \bar{x}_2) - \mu_d)}{\sqrt{\frac{sp^2}{n_1} + \frac{sp^2}{n_2}}} \qquad sp^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 + n_2 - 2)}$$

There are 21 pairs of combinations of the seven stratums. Given the repetitive amount of calculation needed (21 t*) we used R to quicken our calculations. In R we devised the necessary equations and repeated the program 21 times with each time changing the necessary parameters (means and/or variances). We came up with 21 t* and found the p-value for all 21 t* using R. Our resulting p-values can be seen in this table:

Often a p-value is considered significant only if below 0.05. Why did you choose a higher threshold?

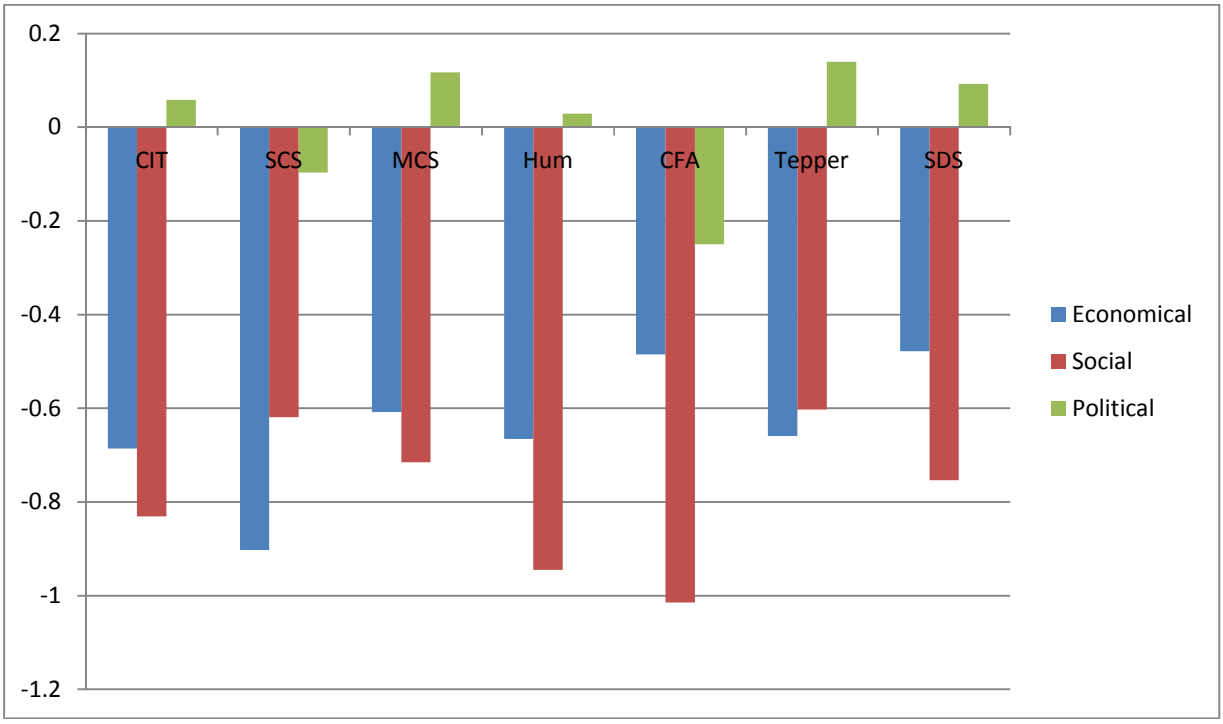
	CIT	SCS	MCS	HUM	CFA	Tepper	SDS
CIT		.748	.380	.344	.0851	.415	.857
SCS			.264	.601	.167	.375	.967
MCS				.0629	.0229	.680	.564
HUM					.193	.1884	.8708
CFA						.154	.560
Tepper							.493
SDS							

The graph shows all the possible combinations and their respective probabilities. There are three p-values at a significant level for the following combinations (MCS & Hum) with humanities being significantly left of MCS, (MCS & CFA) with CFA being significantly left of MCS, and (CIT & CFA) with CFA being significantly left of CIT. While the significant differences in these combinations fit our expectations (quantitative vs. Humanities effect) we do not see a consistent significant difference in the other humanities/quantitative combinations; for example, in (CFA & SCS) or (Hum & SCS) or (CFA & SDS) which gives less credence to the stereotype of humanities students being significantly left of students studying quantitative fields.

The economic, social, and political weighted averages were also calculated separately with the same method. We obtained these averages:

	Econ	Soc	Pol
CIT	-0.68601	-0.83036	0.058036
SCS	-0.90256	-0.61905	-0.09707
MCS	-0.60821	-0.71548	0.116429
Hum	-0.66534	-0.94471	0.028968
CFA	-0.48512	-1.01488	-0.25
Tepper	-0.65952	-0.60238	0.139048
SDS	-0.47818	-0.75313	0.092231

Resulting in this graph:



The graph shows some expected trends, for example, humanities and CFA are the most liberal in social issues. For economic issues, we don't have such a pattern; it is more mixed. It is though interesting that when it comes to political issues such as the power of government, we see that CMU is rather centre in such issues, a contrast to its centre left position on the other issues. While these values don't really help in answering our specific question, they add a level of depth to our study and a comprehensive visualisation of the different issues that we measured from our groups.

It is helpful in visualising our data to construct 95% confidence intervals around the comprehensive political mean of each stratum:

$$\bar{y} \pm z_{1-\frac{\alpha}{2}} \cdot se(\bar{y})$$

CIT

$$-.46955 \pm (1.96)(.1553) = (-.7739, -.165162)$$

SCS

$$-.48884 \pm (1.96)(.1639) = (-.81008, -.16760)$$

MCS

$$-.42049 \pm (1.96)(.1715) = (-.75663, -.08436)$$

HUM

$$-.51721 \pm (1.96)(.1345) = (-.78083, -.25359)$$

These confidence intervals don't seem to support your argument of having significant differences between any of the groups. All the confidence intervals overlap significantly. Your lowest p-value is for the contrast between MCS and HUM, but those confidence intervals are almost entirely overlapping below. Perhaps there is an error in one set of calculations or the other?

CFA

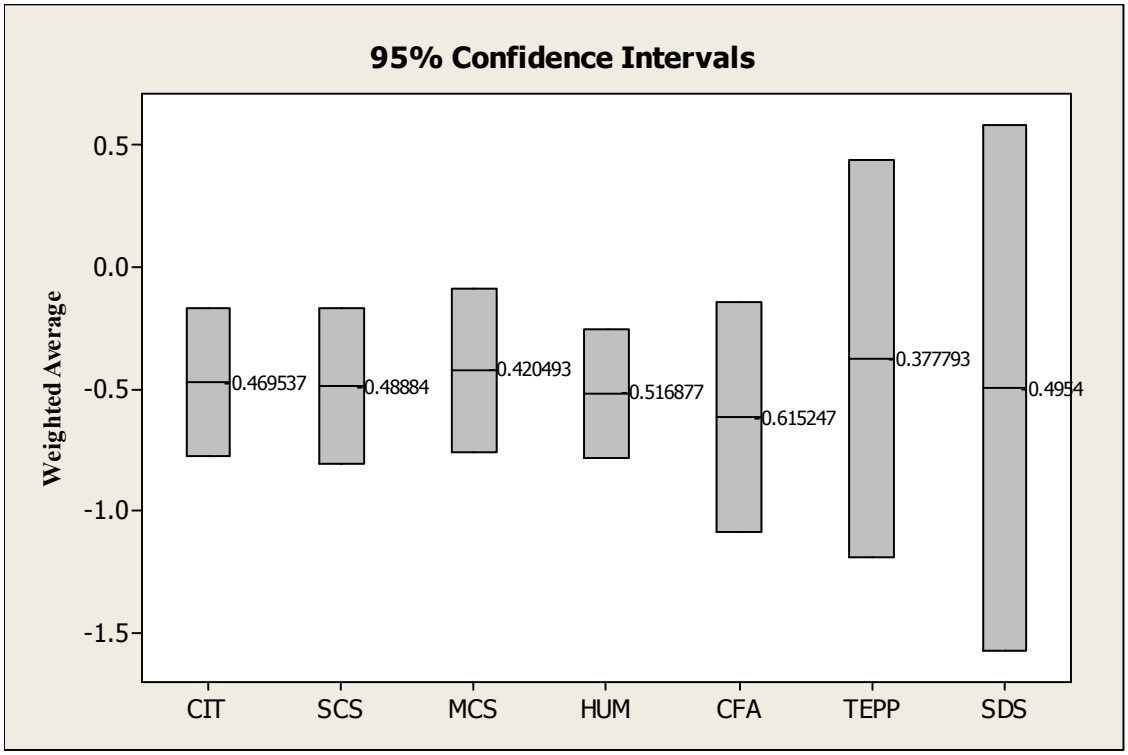
$$-.61526 \pm (1.96)(.2405) = (-1.0866, -.14388)$$

Tepper

$$-.37778 \pm (1.96)(.4156) = (-1.1924, .43680)$$

SDS

$$-.49541 \pm (1.96)(.5494) = (-1.5722, .58141)$$



In our survey, our last three questions posed to the respondent, asked for them to rank themselves politically. While the data from these three questions will not help answer our survey question they provide a helpful and interesting diagnostic for our survey, in seeing how well people rank up with their own perceived political rank.

Stratum	Self-Reported Avg	Weighted Avg
CIT	-0.26667	-0.46955
SCS	-0.15385	-0.48884
MCS	-0.45	-0.42049
Hum	-0.38889	-0.51721
CFA	-0.79167	-0.61526
Tepper	-0.37778	-0.37778
SDS	-0.19444	-0.49541

In all there seems to be a trend of reporting closer to the centre. However, one notable exception is CFA, which reports itself more to the left. One possible explanation may be the effect that this stereotype has on their self perception. It was also an astounding surprise to see the self reported average of Tepper students to be the same as their actual weighted average.

Section 4: Discussion

In all, our study yielded interesting and informative results regarding the political landscape of CMU students, but more importantly shed light on the political relationships between several of CMU's colleges. As stated before, our results do not fortify the stereotype that quantitative majors differ significantly from more humanities majors, with humanities majors said to lean more left. The major reason for this is that our data and inferences do not show a consistent significant difference between the quantitative and humanities groups. Rather, we have only a few combinations that display such a significant difference. The fact that such combinations do exist at the significant level may add to the stereotype between the specific schools (e.g. CIT & CFA), but not to the stereotype in general that quantitative majors are that different from humanities majors. In all, our results show that the CMU student population is as a whole politically centre-left, and in fact our seven groups have rather similar political averages.

Some unexpected results came from the comparison of the several weighted averages. For example, CFA was relative to the others rather close to the right in terms of economic issues and Tepper being relatively left when it came to economic issues. Other such unexpected results came from our p-values in our statistical inference; for example, the p-value of the Humanities and SDS combination being .8708. On the other hand, we did receive many expected results and associations, with some being significant as mentioned.

There were several strengths in our study the first being that we were able to get a response rate of around 20% (normal rate given our method) and successfully infer from those respondents. Our "political diagnostic" of the community was also sensible given our knowledge of the CMU population. We believe also that our questions were interesting and concise, which led to an overall positive experience for our respondents. We were able to avoid large non-response bias and no demographic group was severely underrepresented given the nature of the student population at CMU. At the same

time our study did have some weaknesses. For example, the logistics of sending the email to the students at times did not go very well; for example, our failure to prevent people from replying to the email invitation of the survey. While some of these logistical problems were fixed, they still played a role in harming the image of our study. Another weakness of our survey stems more from the tools that we used to analyse our results. Given the heavy amount of calculation needed for seven stratum, our reliance on programs such as excel made our work greatly inefficient at time, and proficiency in R could have very well helped us at certain points in the data analysis. Another weakness in our study stemmed from the fact that we did not use stratified weights in the sampling. Such weights may have helped with our response rate across the seven stratums.

In conclusion, stereotypes are often exaggerated and used to make rather generalised statements and inferences. Our results show the stereotype of such political beliefs among certain majors to be rather exaggerated, but they do not completely eradicate the stereotype or a significant relationship rather revealing certain significant differences between certain academic groups. We are hence stuck in this "grey area" though often times such generalised stereotypes will fall in such an area.

References

Gage, Rachel, Jennifer Zick, Kaylie Tully, and Elaina Simon. "Choice of Major and Political Attitudes: A Study of University of Minnesota Students." *University of Minnesota Undergraduate Journal of Psychology (Sentience)* 3 (2010): 7-11. Print.

Hastie, Brianne. "Higher Education and Sociopolitical Orientation: The Role of Social Influence in the Liberalisation of Students." *European Journal of Psychology of Education*XXII.3 (2007): 259-74. Print.

Kemmelmeier, Markus, Cherry Danielson, and Jay Basten. "What's in a Grade? Academic Success and Political Orientation." *Personality and Social Psychology Bulletin* 31 (2005): 1386-399. Print.

Lupkin, Alyca, Brianna Knight, and Katy McFadden. "Political Views of College Students in Relation to Their Parents and Friends." *Manchester College Research*. Web.

Appendix

Full Questionnaire:

Demographic

1. What is your grade level?
   Freshman/Sophomore/Junior/Senior
2. What is your primary major if any?
   Fill in
3. In which country have you spent most of your life in? If in the US, please specify your state.
   Fill in
4. What is your race?
   White/Black/Hispanic/American Indian/Asian/Pacific Islander /Other

Answering scheme for all questions:
(I don't know)/(Yes, absolutely)/(Yes probably)/(indifferent)/(probably no)/(no, absolutely not)

Economic Attitude

5. Is government intervention in the business sector via regulations and monetary stimuli justified outside a recession?
6. Is a progressive tax based on income, where the wealthy members of society pay the most, justified?
7. Should trade with foreign nations be restricted in order to protect businesses and workers at home?
8. Do the poorest members of society deserve economic support?
9. Should more goods be made in America even if it means paying higher prices for them?
10. Is government spending on public goods justified (e.g. roads, parks, bridges)?
11. Should the US K-12 public education system be run as a business?

Social Attitude

12. Should marriage be between a man and a woman only?
13. Are abortions for non-rape victims morally wrong?
14. Should traditional gender roles, such as domestic housewives, be broken?
15. Are religious values an integral part of society?
16. Should universities give free contraception to students?
17. Is the recreational use of marijuana acceptable?
18. Is racial profiling for the sake of security sometimes justified by law enforcement?

Political Attitude

19. Is it justified to sacrifice liberty for security in a national security crisis?
20. Is larger government more effective?
21. Should governmental regulation of firearms be stricter?
22. Should the size of the military be reduced?
23. Are people adequately represented by their government?
24. Should taxpayer money be sent for the economic development of 3rd world countries?
25. Should burning your national flag be considered a crime?

26. Where do you stand on social issues?
    Left/Centre-Left/Centre/Centre-Right/Right
27. Where do you stand on economic issues?
    Left/Centre-Left/Centre/Centre-Right/Right
28. Where do you stand on political issues?
    Left/Centre-Left/Centre/Centre-Right/Right

Contact Email:

Hello!

You have been randomly selected to take part in this survey from the official Student Directory. We are undergraduate statisticians conducting a research project for 36-303 (Sampling, Surveys, and Society). Our project concerns the relationship between academic major and political attitudes. We kindly ask for your participation. The survey should only take you about 10 minutes. You can find the link to the survey below:

https://www.surveymonkey.com/s/KHHCP7L

We appreciate your time!

We ask that you do not respond to this email. Thanks.

Informed Consent Form:

This survey aims at gauging the relationship between a student's political attitudes and academic major. The survey is 28 questions long regarding different political, economic, and social issues. The survey takes approximately 10 minutes. The questions in this survey are mild; however, some people may find some questions to be a bit sensitive. Your name or any personal identifiers will not be collected. If you have any questions or concerns please contact mtv@andrew.cmu.edu or mmusaeli@andrew.cmu.edu. Participation in this survey is completely voluntary and discontinuation of the survey can be done at any time.