

Grading the Graders: Investigating the Effect of Professors' Contract Status on Class Averages

Guillermo Parra

Vancouver School of Economics
University of British Columbia
Gparra@student.ubc.ca

December 16, 2023

Abstract

This study investigates the causal relationship between professors' contract status and their class averages at the University of British Columbia (ubc, 2023b) over the academic years 2020, 2021, and 2022. It utilizes a dataset with 20,639 classes taught by 2,546 different professors; this paper uses a fixed effects regression to control for variables such as individual professors and specific courses. The aim is to determine whether professors in the final year of their contracts have an incentive to inflate grades, potentially influenced by tenure reviews or contract renewals. The findings show a statistically significant increase in class averages during the last year of a professor's contract, indicating a potential moral hazard with grading practices. While the study gives insights into university grading practices, its scope is limited by the available data spanning only three years. Future research could explore different datasets and alternative methodologies.

1 Introduction

The "Genesis" of this research paper began with an intriguing observation in a Microeconomics course (ECON*500) at the University of British Columbia (UBC). There was a massive disparity in midterm grades between consecutive cohorts: a drop from an average of 90% last year to just 56% this year. This prompted me to recall a recent change I heard in passing – the professor who teaches that class had recently achieved tenure. In fact, the year in question coincidentally witnessed the highest-ever average in ECON*500, at 88%, significantly deviating from its historical average of 80%. Similar trends were observed in an Econometrics course (ECON*527), where the highest-ever class average was approximately 86% in 2011, which just so happened to coincide with the tenure review of the course's regular professor.

The final observations were made while analyzing the class averages for ECON502 and ECON526, the other term-one courses in the M.A. Economics program at UBC. Examining the highest-ever recorded class averages for these courses revealed a similar pattern; the professors who taught these classes during those years appeared to no longer be part of UBC in the following academic year. This implies that the years with the highest recorded averages coincided with the last year of those instructors' contracts.¹ These observations led to the hypothesis that professors approaching the end of their contracts or undergoing tenure reviews, might inflate grades to influence student evaluations, potentially impacting their tenure decisions and contract renewals.

The study uses publicly available financial data from UBC to derive professors' contract status, using a 13% wage increase or cessation of faculty listing as proxies for contract renewals, indicating when contracts ended. This gives us our treatment group, professors in the last year of their contract, and potential pre and post-treatment effects. This research is grounded in the economic theory of incentives and the concept of moral hazards, hypothesizing that professors might compromise grading standards for favourable student evaluations. A fixed effects regression model was used to control for various factors and isolate the effect of contract status on grading.

There is a lack of research that explores similar themes as this study, including works by Griffith and Sovero (2021), (Gourley and Madonia, 2019), and (Keng, 2018). These studies provide good insights into factors influencing academic behaviour. However, a gap remains in understanding how contract timing and tenure reviews affect grading practices.

2 Assumptions

This study is based on several key assumptions, each of which plays a critical role in the analysis and interpretation of the data. It is important to note that the threshold for an outlier in the data was 13%. This number was chosen after conducting an interquartile range outlier test, finding that a statistical outlier would be any observation outside the 13% threshold. (Tukey, 1949)

2.1 Assumption of Contract Renewal Indication

The primary assumption is that significant wage increases for professors, as indicated in the financial documents from UBC Finance, signal the start of a new contract. This assumption is necessary due to the lack of publicly accessible detailed contact information. Big salary hikes are used as proxies to identify the final year of a professor's current contract. This rationale comes from standard career progression practices, often associated with substantial pay increases. This approach may introduce uncertainty into

¹Information on class averages and faculty transitions was sourced from UBCGrades.ca, the UBC website, and the LinkedIn profiles of the professors.

the analysis, as not all wage increments necessarily reflect contract renewals. Factors such as performance incentives could contribute to increased wages and are not directly linked to contract renewals. This could lead to underestimating the relationship between contract status and grade inflation.

2.2 Data-Related Assumptions and Limitations

The research required significant data cleaning, reducing the dataset from 68,000 to 25,000 classes. This reduction was due to mismatches between professor names and courses. Despite this, it is assumed that the refined dataset still represents the general grading trends at UBC, albeit with a reduced sample size.

2.3 Theoretical and External Factor Assumptions

The study broadly assumes that professors, like other individuals, are motivated by incentives and will adjust their behaviour accordingly, especially when faced with high-stakes situations like contract renewals. This assumption comes from the economic principle of inter-temporal and dynamic utility maximization (Pauly, 1968). The study, however, does not extensively depend on specific theoretical models to explain professors' behaviour or external factors influencing academic grading practices.

3 Literature Review

This research looks into a relatively new area of study. While existing literature offers insights into the influence of tenure status on course difficulty, grading, and the role of faculty evaluations, there remains a gap in understanding grading behaviour, particularly during the tenure review year. This study aims to contribute fresh perspectives on academic standards during this critical period. Shao-Hsun Keng's 2018 (Keng, 2018) study focuses on the Taiwanese tenure system's impact on grading and teaching effectiveness. It finds that tenure-track professors are more lenient in grading, leading to a 15% reduction in class failure rates. However, this leniency does not necessarily result in higher class averages. The study from Griffith and Sovero (Griffith and Sovero, 2021) investigates grading practices in a U.S. university context. Conducted over 1994-2005, it reveals that female instructors facing job uncertainty grade more leniently than their male counterparts. Patrick Gourley and Greg Madoniay's 2019 (Gourley and Madonia, 2019) research examines University of Colorado Boulder data using an instructor-level fixed effects model; the study highlights a consistent decline in student evaluations after professors get tenured.

This study has similarities with a phenomenon observed in political economy. Some political economy literature looks at the behavioural changes of incumbent governors and mayors nearing the end of the term. For instance, (Besley and Case, 1995) and (Rose et al., 2011) have shown there to be significant shifts in economic policies by incumbent governors, while (Ferraz and Finan, 2011) found evidence of increased corruption

among incumbent mayors. These studies highlight how re-election incentives can lead to significant behavioural shifts, such as the potential grading leniency of professors facing tenure review or contract renewal. This shows the influence future career prospects can have on professional conduct.

The issue of grade inflation is examined in various ways. Tyner and Gershenson’s 2020 (Tyner and Gershenson, 2020) study categorizes grade inflation in U.S. high schools, revealing socio-economic disparities. Mathies, Webber Bauer, and Allen’s 2005 (Mathies et al., 2005) study at the University of Georgia illustrates the relationship between increasing GPAs and SAT scores. Sabot and Wakeman-Linn’s 1991 (Sabot and Wakeman-Linn, 1991) research discusses the division between departments due to grade inflation, coining the terms of high and low-grading departments. Wan-Ju Iris Franz’s 2007 (Franz, 2007) study introduces ‘student nuisance’ as a factor driving grade inflation. Significant grade inflation in various disciplines in Ontario is highlighted in a 2000 study by Anglin and Meng (Anglin and Meng, 2000). Ehlers and Schwager’s 2012 (Ehlers and Schwager, 2012) paper argues that honest grading can enhance a school’s reputation. The anti-grade inflation policy at Wellesley College, studied by Butcher, McEwan, and Weerapana (Butcher et al., 2014), provides insights into combating grade inflation and its effects on racial grade gaps, enrollments, and student evaluations.

This body of literature offers a comprehensive understanding of the factors influencing academic grading practices, focusing on tenure and faculty evaluations. The literature also agrees with our hypothesis that there is a change in behaviour during a critical and uncertain time in professional life.

4 Data

This study is based on two distinct datasets: one containing professors’ information and another with class-related details.

4.1 The Datasets

The first dataset includes professor names, wages, and university expenses for each fiscal year. This data was sourced from the University of British Columbia’s (UBC) public financial documents and covers the fiscal years from 2020 to 2023 (ubc, 2023a). However, due to formatting challenges, the data from the fiscal year 2020 was deemed unusable, limiting the analysis to 2021-2023. The second dataset has class-related information like course titles, sections, subjects, sessions, academic years, campus, instructors, enrollment numbers, grade distributions, and other metrics. This data was obtained from ubcgrades.ca, a comprehensive and publicly accessible source with records from even before 2000. This study’s focus was narrowed to academic years 2020 to 2022.

4.2 Data Processing

Transforming these datasets into a usable form was both highly time-consuming and labour-intensive. Initially in PDF format, the professor dataset required extensive cleaning and reformatting. The original data included approximately 20,000 different professors and nearly 70,000 courses over three years. A significant challenge was matching professor names from financial documents to class data. Due to computational intensity and frequent errors of fuzzy matching, the focus was shifted to exact matches. The final dataset now pairs professors' teaching assignments with their wage information and includes various courses across all university departments. Inevitably, discrepancies and missing data led to unmatched professors and courses, introducing a potential bias in the final dataset. These limitations are acknowledged and should be considered when interpreting the study's findings.

5 Methods

This study uses a fixed effects regression model to analyze the impact of professors' contract status on the average grades in classes at the University of British Columbia. The model is specified below:

$$\text{Average Grade}_{ijt} = \alpha + \beta \text{ContractStatus}_{ijt} + \gamma_j + \delta_i + \omega \text{Enrolled}_{ijt} + \epsilon_{ijt} \quad (1)$$

Where:

- Average Grade_{ijt} is the dependent variable representing the average grade in class *i*, taught by professor *j* at time *t*.
- α is the intercept; it represents the expected average grade, contract status, professor and course effects, and class size when they are at their reference values.
- ContractStatus_{ijt} is the independent variable of interest, a dummy indicating whether they are in the final year of their contract. Giving contract status of professor *j* teaching class *i* at time *t*.
- β is the coefficient measuring the impact of being in the final year of a contract on the class average grade.
- Professor fixed effects (γ_j), controlling for individual differences among professors that are consistent over time but vary across professors.
- Course fixed effects (δ_i), accounting for inherent characteristics of each course that could affect the average grades. As there is a short time frame, we can assume that most or all classes and their structures changed minimally.
- Enrolled_{ijt} controls for the number of students enrolled in each class

- ω is the coefficient measuring the impact of the number of students in a class on the average grade.
- ϵ_{ijt} is the error term.

The fixed effects regression model benefits this study because it can control for unobserved, time-invariant characteristics. These fixed effects are necessary to reduce bias and help identify the causal relationship between contract status and average grades.

The model's focus on individual variations is needed for educational data where differences between individuals can be substantial. This approach allows for examining changes in grading practices within the same professor or course, aiding in causal inference. The fixed effects model also aligns with the study's Directed Acyclic Graph (DAG), addressing potential confounders.

The study's DAG illustrates the potential confounding paths affecting the relationship between professors' contract status and class average grades. By including fixed effects for professors and courses, the regression model blocks these confounding paths, aligning with the structure of the DAG. This ensures the analysis accounts for variables like professors' grading tendencies and the courses' inherent difficulty level. See figure 6

The model was selected before reviewing the literature, and after reviewing similar studies, the choice of the model was confirmed to follow published work. Similar research, such as Griffith and Sovero's study on faculty gender and contract uncertainty (Griffith and Sovero, 2021) and Gourley and Madoniay's (Gourley and Madonia, 2019) research on tenure and faculty course evaluations, also uses fixed effects models in their respective analyses.

6 Results

6.1 Initial Analysis

The initial phase of the analysis focused on the raw average treatment effect of professors' contract-ending status on class averages. The first set of results, including a histogram and regression table, indicated an average treatment effect of -0.1533. However, this effect did not achieve statistical significance, with a standard error of 0.185 and a p-value of 0.407, suggesting no significant impact of contract ending on class averages across the whole dataset. See Figure 1 and Table 1

6.2 Controlling for Individual Professor Characteristics

Introducing professor-fixed effects into the regression model led to a notable change in results. This adjustment yielded a positive coefficient of 0.31, reaching statistical significance at the 5% level. This supports the hypothesis that the average class grade increases by about one-third of a percent when professors are in the final year of their

contract after accounting for individual professor characteristics. See Figure 2, 3 and Table 2

6.3 Incorporating Additional Fixed Effects

Further refining the model to include both professor and class fixed effects, known as the average treatment effect in the treated (ATT), resulted in an increased coefficient of 0.4. With high statistical significance at the 0.1 % level, this coefficient suggests a nearly half-percent increase in class averages for professors in their final contract year. See Figure 4, 5 and Table 3

6.4 Impact of Class Size

The analysis also factored in the influence of class size, revealing an even higher coefficient of 0.41 while maintaining the same level of statistical significance. This suggests that the grading leniency for professors in their final contract year persists even when considering the number of students enrolled. See Table 4

6.5 Comparison with Related Studies

Comparing these findings with related studies, such as the research by (Keng, 2018) and Griffith and Sovero (2021), reveals some agreement, particularly in grading leniency. Although Keng's study did not find an impact on class averages, it reported a reduction in failure rates, suggesting a trend toward leniency in grading. Griffith and Soveros' results closely match the studies, although only from female professors. The main difference between my findings and the literature is that the effects are significantly smaller at about 0.5%, whereas the effects on pre-tenure treatment groups from other papers are between 5-15%

6.6 Last Thoughts of the Results

In summary, the analysis indicates a statistically significant tendency towards leniency in grading for professors in the final year of their contracts. This trend is observed even when controlling for both professor and course characteristics, suggesting that contract status significantly influences grading patterns.

7 Conclusion

This study provides valuable insights into the grading practices at the University of British Columbia, specifically focusing on the influence of professors' contract status. A key finding of this research is the noticeable increase in class averages during the final year of professors' contracts. This trend suggests a tendency towards grading leniency, potentially linked to the pressures surrounding contract renewals or tenure reviews. Such

findings raise important questions about the potential moral hazards in academic grading practices.

The implications of this study extend beyond the context of the University of British Columbia, suggesting several avenues for future research:

- Investigating similar dynamics in more teaching-centric institutions could offer better insights, considering the research-intensive nature of UBC.
- The results are dampened by the inability to separate temporary and tenure-track professors; splitting between them would provide more in-depth results and conclusions.
- Incorporating gender differences into the analysis could shed light on potential disparities in grading practices.
- Examining class failure rates, as explored in studies like (Keng, 2018), could provide a deeper understanding of grading leniency and its relation to overall academic performance.
- Similar models to (Besley and Case, 1995), (Rose et al., 2011), (Ferraz and Finan, 2011) could add a different layer to this research, as it would connect well-known ideas in political economy to the ideas in this paper.
- Extending the analysis over a longer period could enhance the robustness of the findings.
- Conducting an event study using the data could add further validity to the study's results.
- Implementing robustness checks, such as using alternative thresholds for dummy variables or employing different models like difference-in-differences and instrumental variable regressions, could provide additional and varied insights.
- The study could benefit from segmenting into more subgroups, like the age of professors and faculty, to see how results change.

Grading the Graders contributes to the research on grading practices and the inherent stress associated with early academic life. Being a relatively novel area of research, it opens doors for further investigation and aims to better our understanding of the factors influencing academia and its pressures. This study encourages more research on the realities and challenges within academia, mainly focusing on grading standards and faculty evaluation practices.

References

- (2023a). Ubc finance. Accessed 2023-11-27.
- (2023b). Ubc grades. Accessed 2023-11-27.
- Anglin, P. M. and Meng, R. (2000). Evidence on grades and grade inflation at ontario's universities. *Canadian Public Policy / Analyse de Politiques*, 26(3):361–368.
- Besley, T. J. and Case, A. (1995). Incumbent behavior: Vote-seeking, tax-setting, and yardstick competition. *American Economic Review*, 85(1):25–45.
- Butcher, K. F., McEwan, P. J., and Weerapana, A. (2014). The effects of an anti-grade-inflation policy at wellesley college. *Journal of Economic Perspectives*, 28(3):189–204. Summer.
- Ehlers, T. and Schwager, R. (2012). Honest grading, grade inflation and reputation. CEGE Discussion Papers Number 143, Available at SSRN: <https://ssrn.com/abstract=2164322>.
- Ferraz, C. and Finan, F. (2011). Electoral accountability and corruption: Evidence from the audits of local governments. *American Economic Review*, 101(4):1274–1311.
- Franz, W.-J. I. (2007). Grade inflation under the threat of students' nuisance: Theory and evidence. Available at SSRN: <https://ssrn.com/abstract=1028011>.
- Gourley, P. and Madonia, G. (2019). Tenure and faculty course evaluations. *DOI: 10.13140/RG.2.2.11695.18087*.
- Griffith, A. L. and Sovero, V. (2021). Under pressure: How faculty gender and contract uncertainty impact students' grades. *Economics of Education Review*, 83:102126.
- Keng, S.-H. (2018). Tenure system and its impact on grading leniency, teaching effectiveness and student effort. *Empirical Economics*, 55(3):1207–1227.
- Mathies, C., Webber Bauer, K., and Allen, M. (2005). An exploratory examination of grade inflation at the university of georgia. Paper presented at the 2005 AIR Forum, San Diego, CA. Office of Institutional Research, The University of Georgia.
- Pauly, M. V. (1968). The economics of moral hazard: Comment. 58(3):531–537.
- Rose, J. E., Bueno de Mesquita, E., and Shanna (2011). Political economics and public finance. *Handbook of Public Economics*, 5:61–110.
- Sabot, R. and Wakeman-Linn, J. (1991). Grade inflation and course choice. *Journal of Economic Perspectives*, 5(1):159–170.
- Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics*, 5(2):99–114.

8 Appendices

Table 1: Simple OLS Regression Results

R-squared:	0.000					
F-statistic:	0.6872					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	81.3107	0.052	1574.846	0.000	81.210	81.412
ContractEnded	-0.1533	0.185	-0.829	0.407	-0.516	0.209

Table 2: OLS Regression Results with Professor Fixed effects

R-squared:	0.612					
F-statistic:	11.21					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	82.8159	1.579	52.436	0.000	79.720	85.912
ContractEnded	0.3103	0.158	1.966	0.049	0.001	0.620

Table 3: OLS Regression Results with Professor and Course Fixed Effects

R-squared:	0.867					
F-statistic:	13.83					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	97.1817	7.637	12.725	0.000	82.212	112.151
ContractEnded	0.4013	0.114	3.508	0.000	0.177	0.625

Table 4: OLS Regression Results of the Full Model

R-squared:	0.868					
F-statistic:	13.86					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	96.9718	7.630	12.710	0.000	82.017	111.927
ContractEnded	0.4057	0.114	3.551	0.000	0.182	0.630
Enrolled	0.0039	0.001	5.275	0.000	0.002	0.005

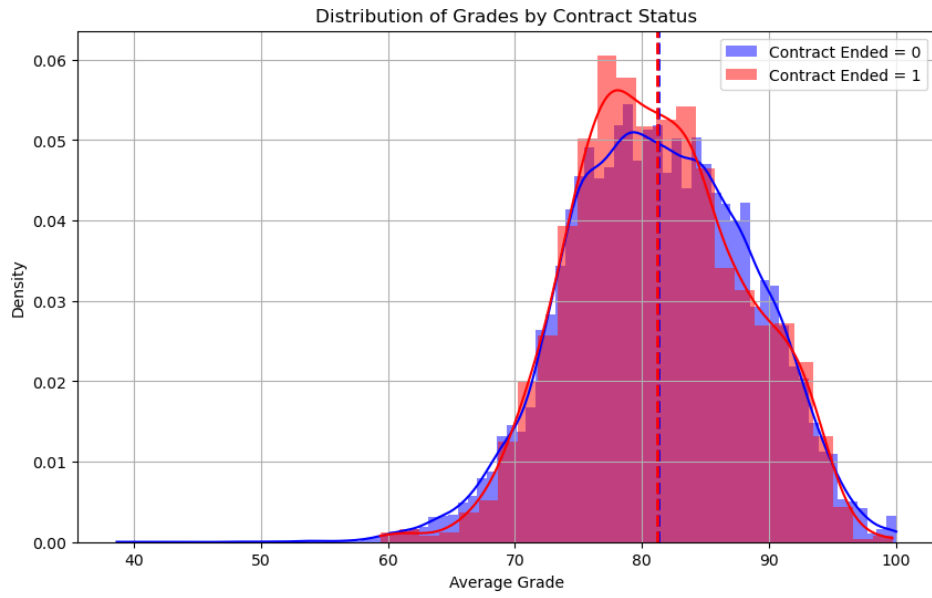


Figure 1: Histogram of the ATE

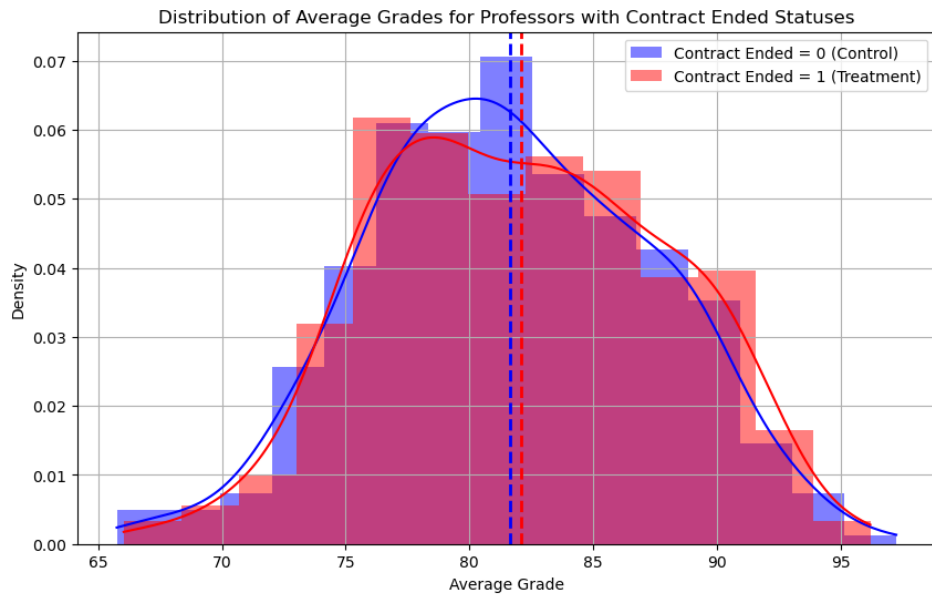


Figure 2: Histogram of the ATT.1

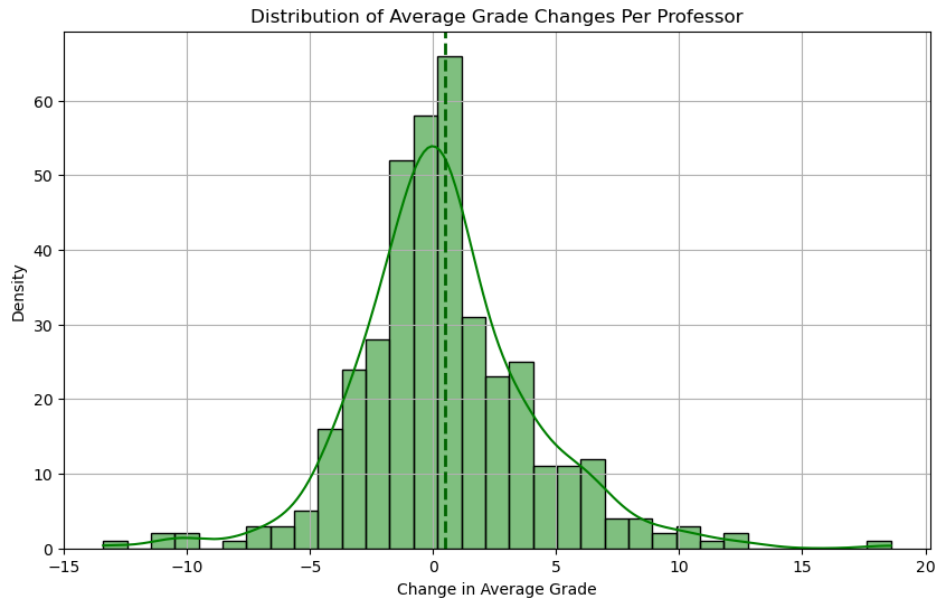


Figure 3: Difference of the ATT.1

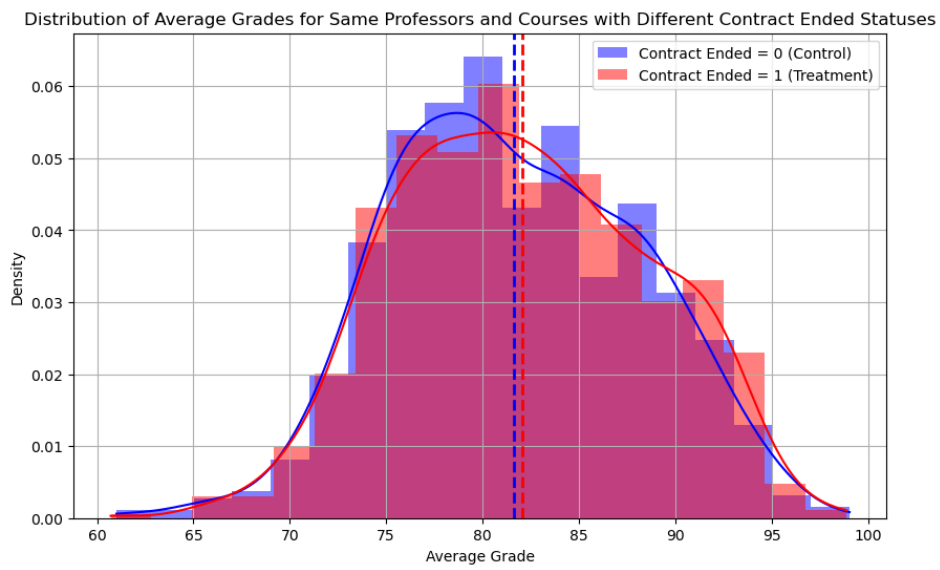


Figure 4: Histogram of the ATT.2

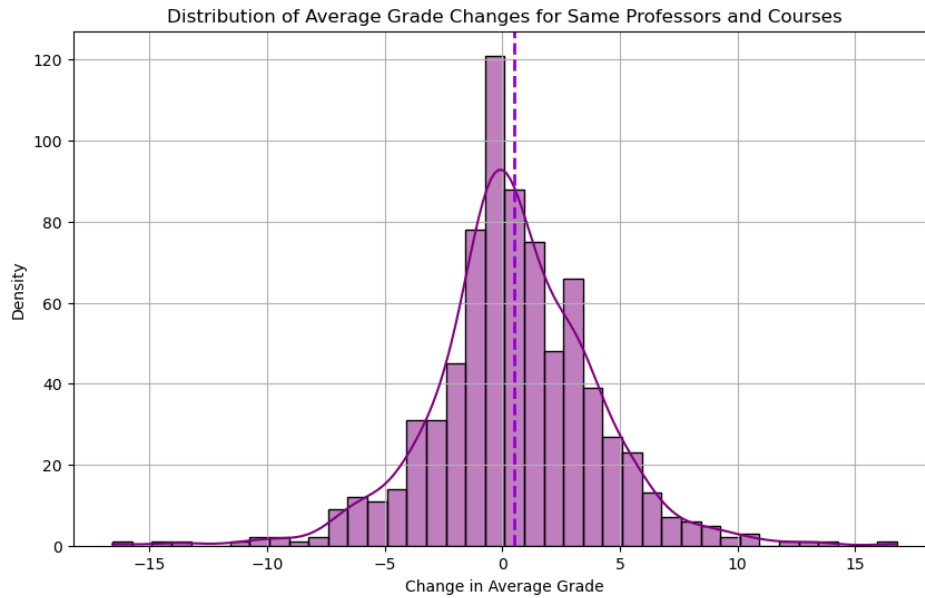


Figure 5: Difference of the ATT.2

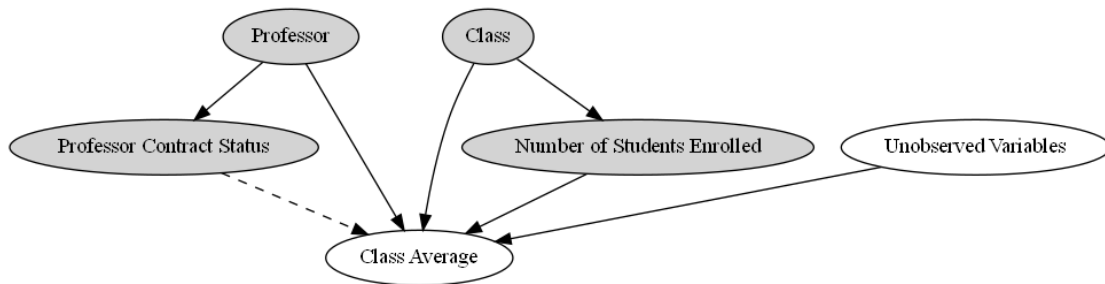


Figure 6: Directed Acyclic Graph

9 Acknowledgments

I want to acknowledge a few people who have helped me in this research project. First and foremost are my ECON*526 professors Jessie Perla, Philip Solimine, Paul Schrimpf and my TA Bruno, who helped me understand many of the techniques and concepts in this study; without them, this paper would have turned out very different. A huge thank you to the creator of UBCgrades.ca, as this project would have been impossible without them. I also want to thank my roommate Etienne, who didn't let me give up on this project even when nothing seemed to go my way and my other roommates, Vicente and Elliot, who helped me better understand many of the ideas written above. And lastly, a thank you to Reid Martel for his insightful comments.