AMERICAN UNIVERSITY OF BEIRUT

THE EFFECT OF SPORTS ON ACADEMIC PERFOMANCE: A STUDY OF THE AUB COMMUNITY

by
LEA GEORGE BOU FADEL

A thesis
submitted in partial fulfillment of the requirements
for the degree of Master of Arts
to the Department of Economics
of the Faculty of Arts and Sciences
at the American University of Beirut

Beirut, Lebanon
April 201

AMERICAN UNIVERSITY OF BEIRUT

THE EFFECT OF SPORTS ON ACADEMIC PERFORMANCE:

A STUDY OF THE AUB COMMUNITY

by
LEA GEORGE BOU FADEL

Approved by:

Dr. Leila Dagher, Associate Professor and Chair Economics	Advisor
Dr. Hossein Radmard, Assistant Professor Economics	Member of Committee
Dr. Jad Chaaban, Associate Professor Agricultural Sciences	Member of Committee

Date of thesis defense: April 27, 2017

AMERICAN UNIVERSITY OF BEIRUT

THESIS, DISSERTATION, PROJECT RELEASE FORM

Student Name: Bou Fadel Lea George
Last First Middle

☑ Master's Thesis ◯ Master's Project ◯ Doctoral Dissertation

☑ I authorize the American University of Beirut to: (a) reproduce hard or electronic copies of my thesis, dissertation, or project; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

☐ I authorize the American University of Beirut, to: (a) reproduce hard or electronic copies of it; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes
after : **One ---- year from the date of submission of my thesis, dissertation, or project.**
Two ---- years from the date of submission of my thesis, dissertation, or project.
Three ---- years from the date of submission of my thesis, dissertation, or project.

May 8, 2017

Signature Date

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Dagher and my readers Dr. Chaaban and Dr. Radmard for their valuable advice and comments on this thesis.

Special thanks to my astoundingly supportive teammates and my coach, Dr. Gaby Issa El Khoury, for the inspiration and their wonderful collaboration.

I must express my very profound gratitude to my parents, my brothers, and my friends who have provided me with unfailing support and continuous encouragement. This accomplishment would not have been possible without them.

AN ABSTRACT OF THE THESIS OF

Lea George Bou Fadel for Master of Arts
Major: Economics

Title: The Effect of Sports on Academic Performance: A Study of the AUB Community

The ongoing debate about the importance of sports and its effect on the human mind has left people with questions about the effect on the academic performance of students, which this paper aims to answer. Researchers have found positive relationships between the two variables, seeing the advantages that sports can achieve. Others have found a negative relationship explained by the increased time and commitment required that reduces available time for studying. Some have found no significant relationship between the two. In this study, a linear regression model is used, and after controlling for various variables, the effect of being an athlete on academic performance turns out to be significantly negative. However, sports will enhance the students' health, and increase their self-esteem and well-being levels, which will in turn increase GPA. This makes the overall effect a positive one, where the negative impacts are balanced out by the more optimistic ones.

CONTENTS

Page

ACKNOWLEDGEMENTS v

ABSTRACT vi

LIST OF ILLUSTRATIONS ix

LIST OF TABLES x

Chapter

I. INTRODUCTION 1

II. LITERATURE REVIEW 4

III. BACKGROUND INFORMATION 10

IV. DATA AND METHODOLOGY 12

A. The survey 13

B. Data 14

C. Methodology 15

V. RESULTS 17

V. CONCLUSION 36

Limitations 37

Directions for future research 38

APPENDIX 39

REFERENCES 41

ILLUSTRATIONS

Figure Page

5.1. Pearson Correlation Test 28

5.2. Predicted versus Residuals Model... 35

5.3. Variance Inflation Factor 35

TABLES

Table | Page

5.1. Regressing GPA on Academics 17

5.2. Controlling for gender 18

5.3. Controlling for major and year of study 21

5.4. Controlling for time spent commuting 22

5.5. Controlling for parents' education 23

5.6. Controlling for pre-college differences 25

5.7. Introducing hours spent on varsity and non-varsity sports 26

5.8. Combining sports into one variable 27

5.9. Adding components of the self-esteem index 30

5.10. Adding the self-esteem index 32

5.11. Removing insignificant variables 33

CHAPTER I

INTRODUCTION

According to the World Health Organization, adults aged between 18 and 64 should get at least 150 minutes per week of moderate-intensity physical activity and two muscle-training sessions (WHO, 2011). Unfortunately, many students fail to meet these minimum requirements, either due to laziness or to the misconception they would be “wasting” their time.

In reality, it all depends on how the time is allocated, and what fraction is given to each activity. The concept is more understandable through Becker’s (1965) allocation of time model, which implies that time is allocated in such a way that increasing more of an activity would decrease the time left for other activities. Particularly, it is believed among college students that time spent on sports is time “wasted” for it could have been otherwise spent on studying. This common belief is not necessarily true. A typical student’s time will include attending classes, commuting to and from campus, studying, exercising, and engaging in leisure activities. When students allocate a larger portion of their time to sports, they could be actually using their time more efficiently. By doing so, they decrease their time spent on negative leisure activities, like watching TV, drinking and partying, or just hanging out with friends.

It is essential not to ignore the positive effects of sports. First, physical activity ameliorates people’s health, which would in turn increase their productivity. Second, it trains both cognitive and non-cognitive skills, such as motor skills, communication skills, leadership, teamwork, and socialization. Third, physical activity is associated with enhanced efficiency of learning through increased discipline, motivation, responsibility, confidence and self-esteem (Cornelissen and Pfeifer, 2007).

Through the Design to Move campaign, it was shown that European children are becoming 50 percent less active at the age of 15 than they were at the age of 9, and children in the US are becoming 75 percent less active. This is resulting in major negative effects, as "today's children may be the first generation to have a shorter life expectancy than their parents" (Designed to Move, 2012). The campaign aims at encouraging people, especially children, to invest a larger portion of their time in sports, in light of the crucial benefits it realizes.

Seeing the importance of physical activity and how neglected it has become, I focus in this study on the effect of sports participation on academic achievement, through the notion of the allocation of time model. The direction is unclear, as time spent on sports, depending on the student, may substitute either time spent on studying or time spent on negative leisure activities, which would decrease or increase academic outcomes respectively. The results of previous research, as will be discussed in chapter II, have been greatly ambiguous. Most of the researchers have verified positive relationships between sports and academics, but several have found negative effects, or failed to find any relationship between the two.

While most studies are concentrated around students in the United States of America, some investigate students in Canada or Europe, but none has ever targeted those in Lebanon, let alone the Middle East. Clearly, what works in one country does not necessarily apply to other parts of the world. My curiosity to figure the exact effects in

Lebanon has led me to limit my study around students at the American University of Beirut. The reason why this paper revolves around such a specific sample is mostly due to practicality. I found it only logical to choose a community to which I belong and can understand and relate to because even within the same country, lives on different campuses employ divergent values, norms, and beliefs.

The hypothesis I am testing for is whether or not student athletes on campus exhibit differences in academic achievement when compared to non-athletes. For that, I construct a model, using R Studio, using linear regressions. Possible confounding variables are added gradually to the model, testing the effect of each. The final model removes all the variables that show no significant effect on academic outcome. These factors indirectly determine the extent to which the students are able to allocate their time efficiently and balance their life activities, without affecting their level of academic achievement negatively. More specifically, I focus on those athletes who are spending a large portion of their time training and competing, and how sports is affecting their outcomes, both on the personal and professional levels. The results prove that athletes lose, on average 1.46 GPA points.

As no previous research was available, data had to be collected from scratch. This was done through surveys distributed among a sample of 274 students, of both athletes and non-athletes. The challenge lay in finding participants who were willing to engage in the research and provide truthful and accurate responses.

This paper is divided into 6 chapters. The literature review is discussed in chapter II, and some background information is given in chapter III. Chapters IV and V touch on details relating to the data, the survey, methodology, and results of the study. The final chapter includes the conclusion and discussion of the research.

CHAPTER II

LITERATURE REVIEW

Going through the timeline of previous research, we find that early studies showed no association between sports and higher grades. In 1973, Lueptow and Kayser reported no difference in academic achievement between athletes and non-athletes in the United States of America. In 1978, Hauser and Lueptow found that by the end of the academic year, athletes in Wisconsin had higher GPAs compared to the beginning of the year, but the gains were not as high as those of non-athletes. But these results failed to prove any causal relationship between athletic participation and school grades.

In later years, even after pre-college differences were accounted for, researchers still found no difference in academic achievement between athletes and non-athletes; yet, this did not indicate that the former underperformed or underachieved when compared to the latter. (Stuart, 1985; Pascarella and Smart, 1991; Hood et al., 1992; Marsh, 1993; Jefferson, 1999).

Around the same time, some authors did succeed in proving that athletic participation negatively impacted athletes' GPAs (Maloney and McCornick, 1993). Meanwhile, others argued for the positive impacts brought about by sports. These include growth in interpersonal and non-cognitive skills, peer relationships, teamwork and leadership abilities, the athletes' commitment to their academic institution, as well as their personal and social well-being, all of which rounded up to generate human capital and academic success (Ryan, 1989; Astin, 1993; Cantor and Prentice, 1996; Ewing, 1998).

To support the later positive findings, the National Education Longitudinal Study of 1992 attested the students' commitment and engagement with their institutions. It found that 50.4% of athletes never skipped a class and had a minimal rate of unexcused absenteeism, and 30.6% had a GPA greater than 3.0, compared to 36.2% and 14.2% respectively for non-athletes (O'Brien and Rollesfon, 1992). Moreover, the relationship between the participation in extracurricular activities and academic success was proven small but statistically significant, increasing as the level of activity increased, and suggesting that the lack of participation had negative effects on academic outcomes (Marsh, 1992; Neish, 1993; Klesse, 1994).

In fact, the more students were involved in sports, the more they felt committed to their institutions and obeyed rules, the less discipline problems they faced. Hence, their more meticulous behavior initiated increased efforts and higher grades (Fejgin, 1994). Taking part in extracurricular activities, and especially sports, showed the greatest influence on academic performance (Broh, 2002; Branch, 2003). This was verified by McCarthy (2000) who proved that athletes had higher GPAs and less frequent absences than non-athletes, and female students in general performed better than male students. Also, Zaff, Moore, Papillo, and Williams (2003) used longitudinal data for students in grades 8-12 and found that over time students' participation was linked with higher levels of academic achievement.

Cross-sectional studies showed the beneficial effects of physical activity on the brain, and specifically on the students' intelligence, cognition, and academic achievement. The results proved that students who are physically fit have the ability to "perform tasks more rapidly and display pattern of neurophysiological activity indicative of greater mobilization of brain resources than those who are less fit" (Tomporowski, Davis, Miller, and Naglieri, 2007). Participation in sports also improves cognitive, social, emotional, and intellectual competencies, such as confidence, character, and perseverance. Sports teams provide a gateway for optimal development, while maintaining an environment of enjoyment, challenge, and positive energy while gratifying the need to belong (Felfe, Lechner and Steinmayr, 2011; Heilman, 2011; Hille and Schupp, 2015; Cabane, Hille and Lechner, 2016).

Even though the studies were recurrent over the years, the results remained ambiguous but most agreed to the positive effects of sports. For instance, Stencel (2005) found no statistically significant relationship between participation in athletics and academic achievement. On the other hand, several researchers employed instrumental variables to estimate the influence of athletes' participation in sports on their academic achievement and found small and significant positive effects. As a matter of fact, sports participation improves academic achievement, namely grades, attendance, and completion of degrees, as well as labor market outcomes (Eide & Ronan, 2001; Lipscomb, 2007; Ewing, 2007; Rees and Sabia, 2010; Stevenson, 2010; Pfeifer and Cornelissen, 2010; Lechner and Sari, 2015; Cabane and Clark, 2015). "Interscholastic sports instill a sense of pride in community, teach lifelong lessons of teamwork and self-discipline and facilitate the physical and emotional development of our nation's youth" (Yeung, 2015).

In addition, it was agreed upon the fact that differences between athletes and non-athletes as well as the differences across genders were observable. Athletes consistently outperformed non-athletes and females outperformed their male peers on their GPAs. In fact, 80.5% of athletes reported a GPA of 3.0 and above, while only 69.5% of non-athletes reported this same level of academic performance (Lumpkin and Favor, 2012). This is in line with Cole's results revealing a significant difference in academic performance between athletes and non-athletes, with athletes having almost half a point higher average GPA than non-athletes (Cole, 2014).

Compared to female non-athletes, 12% more female athletes reported having a GPA above 3.0 (78 over 100). The difference was more pronounced for GPAs above 3.5, as 62% of female athletes reported this level, while only 44% of female non-athletes did. As for male students, 74% of athletes achieved a GPA of 3.0 or above compared with 64% of non-athletes. Again, the difference is more striking at the 3.5 GPA level; 43% of athletes reported a GPA of 3.5 or above, while 34% of non-athletes did (Lumpkin and Favor, 2012).

When taking into consideration SAT scores, GPA, and self-rated academic ability, athletes turned out to have statistically lower values than non-athletes, by approximately 0.038 points. This is explained by the fact that non-athletes reported spending 8.07 hours per week studying, while athletes reported only 6.87 hours. The difference was even more pronounced for student-athletes playing football or basketball, as these are viewed high-profile sports, and thus recommend a greater level of training and commitment. Also, a difference was observed across genders, as male athletes reported a deficit of 0.047 GPA points, while female athletes exhibited 0.029 points, when compared to their non-athlete peers (Routon and Walker, 2015).

Yeung (2015) asserted a positive relationship between time spent on studying and academic performance; students who studied more than 15 hours a week scored 4.829 to 6.774 points (GPA calculated over 100) better than those who did not study at all. "There are general advantages for students who come from wealthy and educated families" (Yeung, 2015).

Lechner and Sari (2015) prove that, for the treatment year of 1996, participating in a sport of moderate intensity for 38 additional minutes each day generated about 10% to 20%

greater earnings during the study period. The effects, however, did not show any statistically significant difference on labor market outcomes such as employment status or hours worked.

In a recent study, Schultz (2017) answers the question whether an athlete's academic performance becomes better or worse during the sports season in which they participate. Athletes in varsity teams exhibit small negative in-season effects, whereas junior varsity athletes exhibit small positive in-season effect. This is explained in terms of the allocation of time model, where the time requirement of participation in sports determines the change in academic performance. Varsity sports require more hours of training per week, including "supplementary weight training sessions", as well as more hours spent competing in matches and tournaments.

If we assume participating in sports removes either time from leisure or academics, then athletes will first substitute their leisure time, which is beneficial to them and to their achievement. They will use their time more efficiently and remove any negative activities such as smoking, drinking, and partying. But the more athletes train, the more they will have to substitute time away from their leisure activities until eventually they reach a threshold after which they start substituting time away from academics. Hence, their academic performance worsens as they increase their training hours every week. To that extent, it is clear that the relationship between participation in sports and academic outcome may be nonlinear.

The effect of being in-season is negative and significant for varsity athletes, for their GPA decreases by 0.0281 points. On the other hand, the effect is positive and significant for junior varsity athletes, with an increase of 0.0216 points in GPA. This suggests that participation in sports is associated with costs to both time and academics. The size of the effect brought about by time cost is relatively small and is dependent upon the type and level of participation. The academic costs to participation to varsity athletes are also small, even though their GPA decreases slightly; the academic costs to junior varsity athletes are non-existent. It is noteworthy that effects may differ across genders, year of study, and whether the sport played is a team or an individual sport (Schultz, 2017).

CHAPTER III

BACKGROUND INFORMATION

To start with, “athletes” in this paper simply refers to students who belong to any varsity or junior varsity team. It does not take into consideration the type of sport participated in, nor the amount of required training per week. It should be noted here that some students may be part of off-campus sports teams, but these will be included as non-athletes in this study.

Varsity teams all have a limit to the number of athletes they can accept every year. Therefore, students must take part in several tryout sessions before they make it to the team. Naturally, their physical abilities determine whether or not they are qualified to become “athletes”. This causes an endogeneity problem, which will be further discussed in Chapter IV.

As students come from different schools and backgrounds, many might find it difficult to excel during their first semester due to their lack of acquaintance with the system. Hence, their GPA increases gradually as they become more at ease with their educational program, regardless of the amount of time they put into sports. Naturally, they might seem like they are performing worse academically when compared to non-athletes who are more familiar with the system.

GPA is compared for all students equally, but we should keep in mind that grades differ greatly across majors and across years of study. It is relatively easier to achieve higher grades in majors that deal more with quantitative problems than majors that require a bigger amount of papers and projects that cannot be aced. It also depends on the difficulty of courses each student decides to take and the load in each semester. These factors are not accounted for in the study, as GPA is reported without details about the courses. For instance, while some students choose the easiest courses like drawing and music as their electives, others might choose more technical and beneficial courses related to their major.

Lastly, the questions included in the survey behind this study appear in the appendix. Further details about the survey and the process of data collection are discussed in the next chapter.

CHAPTER IV

DATA AND METHODOLOGY

Data was collected through a survey conducted on a stratified sample of 275 respondents. Students were categorized as members of varsity teams, or athletes, and non-athletes. Based on these two strata, the students were randomly chosen based on their willingness to participate in the study. Clearly, the study is of an observational cross-sectional nature, where student's performance is compared to their peers' at the same moment in time. Athletes are not differentiated on the basis of the type of sport they are engaged in, in terms of the sport's intensity and the time and effort required. The participation in the survey is completely voluntary and is associated with minimal risk.

It should be understood that the participation of athletes in varsity teams is not random; their decision is dependent upon their physical strength and speed, and willingness to engage and commit. They usually have better time management and are relatively more competitive, which translates into better performance on the field and might be also reflected in the classroom. Making the team requires that they be among the best, as every team can only accept a limited number of players. Students who had already been on the team and trained with them previously will obviously have an advantage over the new students, which makes the availability of recruiting new athletes even more limited. Here, students who are "second best" are left out, though their abilities might allow them to join the team in another year when more vacancies are available.

Therefore, when considering the variable “athlete”, we have to keep an eye out for endogeneity. The students’ abilities will grant them a place on the desired team, and with regular trainings, their abilities will be enhanced further. With that, their willingness to achieve both physically and academically will be enriched too. The more they win and prosper, the more they will want to train and develop their capabilities, the more they will succeed in doing so.

Furthermore, the simple cross-sectional estimation of GPA will result in omitted variable bias. The level of academic achievement may be contiguous upon variables that cannot be measured or cannot be observed. While studying, exercising, and self-esteem may be the most common variables, they are certainly not the only ones. It is important to control for pre-college differences, the student’s major and year of study, in addition to shocks that bring about either positive or negative implications to a student’s life.

A. The survey

The surveys were distributed in hard copies to ensure the fair participation between athletes and non-athletes. The questions of the survey appear in the appendix. For athletes, I would make it to the court or field during their training times, and let the coaches know why I am doing this. People were surprisingly very responsive and very helpful. Students were friendly and sociable and often engaged in conversation, not necessarily related to the study. Several times, and especially in teams with male athletes, the coaches would be amazed how I got them all to sit and complete the survey.

Many people gave me positive feedback about my topic, and students always tried to predict the final results. Some would say “GPA definitely goes up”, while others believed it is too time-consuming. A few students even offered to tell me their life stories and experiences with sports and their GPA, if I ever decided to include case studies in my research. However, the stories all showed positive relationships between the two variables, which contradict the results of this study.

The more difficult part was finding non-athletes participants, from the diverse majors and different years of study. For that reason, I targeted the common areas such as the several cafeterias on campus and the libraries. As expected, people were reluctant when I asked if they could fill out the survey, but once I let them know about the purpose, they became more responsive. Also, students were often sitting in groups. Hence, once one student offered to take the survey, the others instantly became more willing to do so as well. To my surprise, several offered to take a bunch and pass them out to any friends they knew around, which was truly helpful.

The last step before conducting the regressions and building the model was manual data entry and manipulation. For me, data collection was the most exciting and entertaining part of my whole study, besides finding out the exact results of the final model. Also, the process of collecting my own data turned out to be an enriching experience.

B. Data

The self-esteem index, in this study, refers to 19 aspects of the students' lives, each rated on a scale of 1 to 5, 1 being "poor" and 5 being "very good". The categories include happiness and time-management skills, as well as writing, analytical, and computer skills. The "drive" category includes grade consciousness, the students' concern about their future, and whether they describe themselves to be ambitious and academically focused.

The “control” category depicts how well students are able to effectively cope with important life changes, are able to handle their personal problems, and are able to control the difficulties in their life. The fourth category “talent” shows the extent to which students perceive themselves as smart, intellectual, creative and artistic. Last but not least, the “extraversion” category is, as its name suggests, a description of the degree of extraversion, meaning the participants’ social skills, outgoingness, confidence, and good leadership.

Of the 274 students surveyed, 49% are female, and 45% are athletes. 55 out of the 124 athletes are female, which is equivalent to 44%. On average, females earn 1.186 points higher than males (with a standard deviation of 4.761) and athletes reported scoring 1.203 points lower than non-athletes. When comparing males alone, athletes score 1.049 points lower than non-athletes. The difference is more pronounced for females, where athletes reported a GPA level lower by 1.457 than their non-athlete peers.

C. Methodology

When the topic of academic achievement and GPA is brought up, an immediate association is made between studying and the student’s results. For that, my initial model starts with the variables of hours spent on studying every week, the percentage of completed assignments and the number of classes skipped.

However, GPA is not determined by these variables only. Many other factors should be controlled for, as will be discussed further in the next chapter. Controls for gender, major, and year of study are added progressively. Then the factor that indicates the hours spent commuting weekly is added, and pre-college differences are controlled for, such as the parents’ highest level of education and the students’ SAT scores. In addition, controls for leisure activities are added, of which sports is amongst. The effect of being an athlete and belonging to a varsity team is tested, and lastly I add a self-esteem index that encompasses how the students perceive themselves, their abilities, and their level of self-confidence.

Last but not least, the model is made stronger by removing statistically insignificant variables, and the robustness is tested. Nevertheless, the omitted variable bias may not be completely eliminated, as some variables that affect GPA may still be unobserved. Details of all the regressions and their results appear in the Chapter V.

CHAPTER V

RESULTS

In this chapter I construct a model to test the effect of being an athlete on the level of academic achievement, while controlling for variables that may affect students' grades. I start by testing the most common belief that more studying will enhance achievement. For that, I include the number of hours spent studying per week, the percentage of completed assignments, and the number of classes skipped every week. The first regression is given by the equation below, and the results are displayed in Table 1.

$$GPA = \beta_0 + \beta_1*Studying + \beta_2*Classes.Skipped + \beta_3*Assignments + \varepsilon \quad (1)$$

β_0, β_1, β_2, and β_3 represent the coefficients of the variables under study. ε, which is the error term, captures all the factors that have not been controlled for.

```
Call:
lm(formula = GPAAv ~ Studying + Classes.Skipped + Assignments,
    data = RESULTS)

Residuals:
     Min       1Q   Median       3Q      Max
-19.6095  -3.2912  -0.1963   3.8994  14.6656

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      79.0222     1.6958  46.598  < 2e-16 ***
Studying          0.5914     0.2902   2.038  0.04251 *
Classes.Skipped  -1.5825     0.5478  -2.889  0.00418 **
Assignments       0.6424     0.3143   2.044  0.04191 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.52 on 270 degrees of freedom
Multiple R-squared:  0.09562,   Adjusted R-squared:  0.08557
F-statistic: 9.516 on 3 and 270 DF,  p-value: 5.381e-06
```

Table 5.1: Regressing GPA on Academics

In effect, the three variables are linked with significant changes in GPA. Hours spent studying and completing an additional 20% of assignments are associated with significant increases of 0.5914 and 0.6424 points in GPA. On the other hand, skipping classes decreases GPA by 1.5825 points. . It should be noted that classes skipped are measured in increments of 4 classes per week. The figures are only logical; attending classes is correlated with a higher GPA that ranges from 0.24 to 0.34 standard deviations (Gottfried, 2010). For practicality reasons, the three variables are grouped into a new variable "Academics".

However, changes in the level of academic performance are not merely determined by studying. Many variables have to be controlled for, the first of which is gender. As shown previously by the means calculated from the survey results, females earn better grades on average than males. For that, I control for gender through the dummy variable which is equal to 1 if the student is female and 0 if he is male. As shown by Table 2 below, the decreases seen earlier turn out to be statistically insignificant.

```
Call:
lm(formula = GPAAV ~ ACADEMICS + Gender, data = RESULTS)

Residuals:
     Min       1Q   Median       3Q      Max
-20.4614  -3.8939  -0.1422   3.9362  16.0386

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  77.0089     1.5061  51.130  < 2e-16 ***
ACADEMICS     0.5777     0.1978   2.921  0.00378 **
GenderM      -1.1689     0.6858  -1.704  0.08945 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.675 on 271 degrees of freedom
Multiple R-squared:  0.04079,   Adjusted R-squared:  0.03372
F-statistic: 5.763 on 2 and 271 DF,  p-value: 0.00354
```

Table 5.2: Controlling for gender

Besides gender, the student's year of study and chosen major play a large part in determining the student's GPA. The years include freshman, sophomore, junior, senior and graduate. For that, four dummy variables are created. The majors, however, are broken down into a large list. I have taken the liberty to group them by their respective faculties. This yields four dummy variables, representing Olayan School of Business (OSB), the Faculty of Engineering and Architecture (FEA), the Faculty of Arts and Sciences (FAS), one dummy for the three faculties of Agriculture and Food Science, Health Sciences, and Medicine (FAFS.FHS.FM). The last dummy is for majorless students, namely those in their freshman year. The students not included in any of these dummies are those working for dual degrees, and may thus belong to more than one faculty at the same time.

The notion that participants should be separated by major and year goes back to the fact that some majors are on average easier than others. The courses might be less demanding, and more straightforward than others. Students in engineering can ace their exams if they are well-prepared, but students in more subjective majors will have research papers and critiques which cannot be aced no matter how good they are. In effect, being in FEA is correlated with an increase of 4.4039 points in GPA, significant on the 95% confidence level, as shown by Table 3. Other factors do not appear significant, due to the small sample size. The initial sample is already not big enough and when divided into separate faculties, the samples of each faculty become too small to show any significance. This brings about a bias reflected by the large standard error for each of the dummy variables.

Also, GPA depends on the number of credits taken per semester, which are not reported in the study. Students who take 18 credits will feel more pressured and stressed than people who take 12 credits per semester. Though this might drive them to focus more on their studies and become more efficient in allocating their time, it is justifiable if it results in a lower grades due to the aforementioned stress and time requirements.

As for the year of study, students who are still in their first year might find it difficult to adapt to the new environment and mode of study. What applies to high school is very different from life in university, hence they need more than one semester to become fully acquainted with the whole system and perform better academically. The regression shows significant decreases in GPA for each of Freshman, Sophomore, and Junior, which become less pronounced respectively. Freshman exhibit a fall of 11.2788 points in GPA, but this number might not be very precise. The standard error associated with it is as high as 4.2120. This large error can be explained by the small number of freshman students in the sample, representing only 3% of the participants. Increasing their number would decrease the error and establish more accurate figures.

Another factor that may affect academic performance is the time spent commuting every week to and from campus. Students who live further away need more time to get to campus and back, which already decreases the time left that they can allocate on either studying or sports or other activities. Moreover, the further they are the more traffic they are prone to have, which eventually increases levels of stress. While travel time typically reduces the hours available to be allocated for other activities and would thus be interpreted as a factor that worsens academic performance, my results have shown otherwise. Adding the variable “Commuting” returns a significant increase of 0.4172 points in GPA for every three extra hours spent traveling, as shown in Table 4. A possible explanation is that the long hours spent commuting would encourage students to make their trip more worthwhile and use their time more efficiently.

```
Call:
lm(formula = GPAAv ~ ACADEMICS + Gender + OSB + FEA + FAS + FAFS.FHS.FM +
    MAJORLESS + Freshman + Sophomore + Junior + Senior, data = RESULTS)

Residuals:
     Min       1Q   Median       3Q      Max
-18.2940  -3.5254  -0.6172   3.3311  18.2060

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  78.0382     2.9360  26.580  < 2e-16 ***
ACADEMICS     0.5411     0.2025   2.672  0.00802 **
GenderM      -1.7400     0.7162  -2.429  0.01579 *
OSB           1.5981     2.0885   0.765  0.44485
FEA           4.4039     2.0876   2.110  0.03585 *
FAS           1.8030     2.0487   0.880  0.37964
FAFS.FHS.FM   0.9544     2.2444   0.425  0.67101
MAJORLESS     8.1343     4.8839   1.666  0.09700 .
Freshman    -11.2788     4.2120  -2.678  0.00788 **
Sophomore    -4.1360     1.6870  -2.452  0.01488 *
Junior       -3.1016     1.6325  -1.900  0.05855 .
Senior       -1.9919     1.5760  -1.264  0.20739
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.5 on 262 degrees of freedom
Multiple R-squared:  0.1288,    Adjusted R-squared:  0.09227
F-statistic: 3.523 on 11 and 262 DF,  p-value: 0.0001268
```

Table 5.3: Controlling for major and year of study

For the next step, I control for pre-college differences resulting from the parent's highest education. The number "1" is given to the parent who has not reached more than high school. The number "2" is given to those who have entered college but never finished their degree, and the numbers "3", "4", and "5" are assigned to parents who have completed the requirements for a Bachelor's degree, a Master's degree, and a Doctor of Philosophy, respectively. The variable "Parents" is an addition of the score of both the mother and the father.

```
Call:
lm(formula = GPAAv ~ ACADEMICS + Gender + OSB + FEA + FAS + FAFS.FHS.FM +
    MAJORLESS + Freshman + Sophomore + Junior + Senior + Commuting,
    data = RESULTS)

Residuals:
     Min       1Q   Median       3Q      Max
-17.5404  -3.3658  -0.5297   3.3247  17.2907

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  76.9496     2.9496  26.088  < 2e-16 ***
ACADEMICS     0.5651     0.2011   2.810  0.00534 **
GenderM      -1.8167     0.7111  -2.555  0.01119 *
OSB           1.5939     2.0714   0.769  0.44230
FEA           4.2463     2.0716   2.050  0.04138 *
FAS           1.7055     2.0323   0.839  0.40214
FAFS.FHS.FM   1.0891     2.2267   0.489  0.62519
MAJORLESS     7.4073     4.8540   1.526  0.12822
Freshman    -10.6778     4.1855  -2.551  0.01131 *
Sophomore    -4.2358     1.6738  -2.531  0.01197 *
Junior       -3.2023     1.6197  -1.977  0.04908 *
Senior       -2.1920     1.5654  -1.400  0.16263
Commuting     0.4172     0.1803   2.314  0.02142 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.455 on 261 degrees of freedom
Multiple R-squared:  0.1464,    Adjusted R-squared:  0.1071
F-statistic: 3.729 on 12 and 261 DF,  p-value: 3.227e-05
```

Table 5.4: Controlling for time spent commuting

The regressions (in Table 5) reveal that having more educated parents is associated with an increase of 0.4994 points in GPA, significant on the 95% confidence level. Surely, "children with more highly educated parents develop higher aspirations for their own education and attain more education by age 19, which in turn relates to higher levels of adult educational attainment" (Dubow, Boxer, and Huesmann, 2009).

```
Call:
lm(formula = GPAAV ~ ACADEMICS + Gender + OSB + FEA + FAS + FAFS.FHS.FM +
    MAJORLESS + Freshman + Sophomore + Junior + Senior + Commuting +
    Parents, data = RESULTS)

Residuals:
     Min       1Q   Median       3Q      Max
-17.7788  -3.6041  -0.2917   3.3731  16.9720

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  73.1774     3.2692  22.384  < 2e-16 ***
ACADEMICS     0.5377     0.1993   2.698  0.00744 **
GenderM      -1.7415     0.7042  -2.473  0.01404 *
OSB           2.2077     2.0636   1.070  0.28570
FEA           4.7532     2.0594   2.308  0.02178 *
FAS           2.1831     2.0196   1.081  0.28073
FAFS.FHS.FM   1.7770     2.2197   0.801  0.42410
MAJORLESS     7.6902     4.8044   1.601  0.11067
Freshman     -9.7082     4.1589  -2.334  0.02034 *
Sophomore    -4.0743     1.6574  -2.458  0.01462 *
Junior       -2.9575     1.6055  -1.842  0.06660 .
Senior       -1.9392     1.5521  -1.249  0.21266
Commuting     0.4373     0.1786   2.449  0.01498 *
Parents       0.4994     0.1950   2.561  0.01099 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.398 on 260 degrees of freedom
Multiple R-squared:  0.1674,    Adjusted R-squared:  0.1257
F-statistic: 4.021 on 13 and 260 DF,  p-value: 5.146e-06
```

Table 5.5: Controlling for parents' education

Other pre-college differences that have to be accounted for are SAT scores, which are major determinants of the school and major the students attend. At the college level, the differences may not be observable, but as the requirements to get into each faculty differ, the student might be pursuing a degree in a major that had not been initially his first choice. This makes him enjoy his courses less and thus perform worse than if he had the opportunity to transfer into the desired major of study. Additionally, students who already had high grades in school and on their SATs will most likely still perform better than others in college.

In effect, SAT Math and Reading scores show increases in GPA by a slight and significant 0.0173 and 0.0135 points respectively, as shown by Table 6. However, not all the participants in the study provided their SAT scores, hence when included in the regressions, the number of observations decreases from 274 to 213. In efforts to keep the sample as big as it is and avoid additional bias, the rest of the regressions will not include SAT scores.

I now take into consideration the quality of leisure activities the students engage in. Particularly, I control for those activities that might be more or less directly related to their academic performance. Hence, I account for the student's participation in clubs that require more than 3 hours every week, and create the dummy variable "Clubs". I also include the frequency of books read per year, which have been measured in the survey in increments of 6 books. The regressions reveal that being part of clubs and reading more books will not have significant effects on a student's level of academic achievement.

This brings me to the variable I am most interested in, one that accounts for the number of hours spent on sports every week. These include activities for all students, both athletes and non-athletes, and does not take into consideration the intensity of the physical activity. I conduct two regressions; in the first I introduce hours spent on varsity sports and hours spent on non-varsity sports, regardless of being an athlete. Evidently, varsity hours target athletes exclusively, while non-varsity hours may target both; athletes who take part in trainings outside their team and non-athletes who exercise. The type of sport ranges from fitness classes and gym sessions, to fun football or basketball games with friends, or simply a quiet morning or afternoon jog. The second regression combines the two components into a single "Sports" variable, encompassing all hours the student allocates to physical activity.

```
Call:
lm(formula = GPAAV ~ ACADEMICS + Gender + OSB + FEA + FAS + FAFS.FHS.FM +
    MAJORLESS + Freshman + Sophomore + Junior + Senior + Commuting +
    Parents + SAT.MATH + SAT.READING, data = RESULTS)

Residuals:
     Min       1Q   Median       3Q      Max
-19.0746  -3.2566   0.2442   3.5723  15.5399

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 54.933866   5.929425   9.265  < 2e-16 ***
ACADEMICS    0.390637   0.235042   1.662  0.09810 .
GenderM     -1.716937   0.811110  -2.117  0.03553 *
OSB          3.391859   2.140873   1.584  0.11471
FEA          4.503451   2.109469   2.135  0.03400 *
FAS          2.957912   2.082822   1.420  0.15714
FAFS.FHS.FM  2.390617   2.526764   0.946  0.34524
MAJORLESS    5.397977   4.932286   1.094  0.27510
Freshman    -4.823581   4.539875  -1.062  0.28931
Sophomore   -3.081782   1.918929  -1.606  0.10987
Junior      -3.123881   1.879739  -1.662  0.09812 .
Senior      -2.594294   1.811473  -1.432  0.15368
Commuting    0.389788   0.203188   1.918  0.05650 .
Parents      0.448682   0.239309   1.875  0.06228 .
SAT.MATH     0.017332   0.006083   2.849  0.00484 **
SAT.READING  0.013461   0.005367   2.508  0.01294 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.445 on 198 degrees of freedom
  (60 observations deleted due to missingness)
Multiple R-squared:  0.2376,    Adjusted R-squared:  0.1799
F-statistic: 4.114 on 15 and 198 DF,  p-value: 1.51e-06
```

Table 5.6: Controlling for pre-college differences

The first regression yields a fall in GPA by 0.1803 points for more hours spent on varsity sports, but the decrease is only significant on the 90% confidence level. Hours spent on non-varsity sports turn out to be insignificant, as shown by Table 7 below. On the other hand, the second regression reveals a slightly significant decrease of 0.1614 points in GPA for every 3 additional hours spent on sports every week (Table 8).

```
Call:
lm(formula = GPAAV ~ ACADEMICS + Gender + OSB + FEA + FAS + FAFS.FHS.FM +
    MAJORLESS + Freshman + Sophomore + Junior + Senior + Commuting +
    Parents + Clubs. + Books + Varsity.hours + Non.Varsity.hours,
    data = RESULTS)

Residuals:
     Min       1Q   Median       3Q      Max
-18.6042  -3.4391  -0.5095   3.2250  17.7987

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)       73.01404    3.46634  21.064  < 2e-16 ***
ACADEMICS          0.52334    0.20122   2.601  0.00984 **
GenderM           -1.40473    0.72423  -1.940  0.05352 .
OSB                2.63093    2.12369   1.239  0.21654
FEA                5.30817    2.12473   2.498  0.01311 *
FAS                2.66167    2.08537   1.276  0.20299
FAFS.FHS.FM        2.40100    2.26942   1.058  0.29106
MAJORLESS          7.94539    4.87381   1.630  0.10428
Freshman          -9.77981    4.13286  -2.366  0.01871 *
Sophomore         -4.00812    1.65930  -2.416  0.01641 *
Junior            -2.85670    1.61770  -1.766  0.07860 .
Senior            -2.02719    1.55727  -1.302  0.19417
Commuting          0.35726    0.19039   1.876  0.06173 .
Parents            0.52764    0.19652   2.685  0.00773 **
Clubs.            -1.15169    0.67758  -1.700  0.09040 .
Books              0.53880    0.43294   1.245  0.21445
Varsity.hours     -0.18030    0.10635  -1.695  0.09124 .
Non.Varsity.hours -0.07432    0.22678  -0.328  0.74339
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.362 on 256 degrees of freedom
Multiple R-squared:  0.1911,    Adjusted R-squared:  0.1373
F-statistic: 3.557 on 17 and 256 DF,  p-value: 5.038e-06
```

Table 5.7: Introducing hours spent on varsity and non-varsity sports

```
Call:
lm(formula = GPAAV ~ ACADEMICS + Gender + OSB + FEA + FAS + FAFS.FHS.FM +
    MAJORLESS + Freshman + Sophomore + Junior + Senior + Commuting +
    Parents + Clubs. + Books + SPORTS, data = RESULTS)

Residuals:
     Min       1Q   Median       3Q      Max
-18.6698  -3.3137  -0.5146   3.2200  17.5964

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 72.98037    3.45989  21.093  < 2e-16 ***
ACADEMICS    0.51631    0.20021   2.579  0.01047 *
GenderM     -1.36103    0.71571  -1.902  0.05834 .
OSB          2.68633    2.11628   1.269  0.20546
FEA          5.36198    2.11755   2.532  0.01193 *
FAS          2.72599    2.07652   1.313  0.19043
FAFS.FHS.FM  2.45296    2.26249   1.084  0.27930
MAJORLESS    8.18615    4.83291   1.694  0.09151 .
Freshman    -9.77346    4.12624  -2.369  0.01860 *
Sophomore   -3.97812    1.65515  -2.403  0.01695 *
Junior      -2.82263    1.61313  -1.750  0.08135 .
Senior      -1.98997    1.55231  -1.282  0.20102
Commuting    0.38183    0.18107   2.109  0.03594 *
Parents      0.53483    0.19548   2.736  0.00665 **
Clubs.      -1.13211    0.67493  -1.677  0.09468 .
Books        0.53053    0.43181   1.229  0.22034
SPORTS      -0.16136    0.09637  -1.674  0.09528 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.353 on 257 degrees of freedom
Multiple R-squared:  0.1905,    Adjusted R-squared:  0.1401
F-statistic:  3.78 on 16 and 257 DF,  p-value: 2.672e-06
```

Table 5.8: Combining sports into one variable

In an attempt to gain insight on more precise effects of being an athlete, I add the dummy variable "Athlete", which equals 1 if the student belongs to a varsity team and 0 if he does not. The regression then returns insignificant results for both hours spent on sports and being an athlete. This goes back to the fact that the two factors may be dependent or correlated. I test for independence between the variables, using the chi squared test. The p-value returned is 0.2086, which means that I do not reject the null hypothesis, stating that the two variables are independent. To test for correlation, the Pearson correlation test is used, which is given by the following formula, where x and y represent the variables "Sports" and "Athlete", and m_x and m_y represent their corresponding means.

$$r = \frac{\sum(x - m_x)(y - m_y)}{\sqrt{\sum(x - m_x)^2 \sum(y - m_y)^2}} \qquad (2)$$

The Pearson test is then conducted in R, using the cor.test () function. The results are given by figure 1. The p-value returned is very small, and less than the significance level alpha 0.05. The correlation coefficient of 0.7097. To interpret the results, I make use of the fact that values closer to 0 indicate no association between the two variable, while values closer to 1 indicate a strong correlation. Therefore, it is safe to conclude that "Sports" and "Athlete" are significantly correlated. In order to minimize the bias for the next regressions, only one of the two variables will be used.

```
        Pearson's product-moment correlation

data:  SPORTS and Athlete
t = 16.612, df = 272, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.6454451 0.7639236
sample estimates:
      cor
0.7096661
```

Figure 5.1: Pearson correlation test

I now test for the individual effects of being an athlete in a team sports versus an athlete in an individual sport, thus GPA is regressed on both variables. Team sports are namely Basketball, Futsal, Handball, Rugby, Soccer, Volleyball, and Waterpolo. Individual sports are Swimming, Table tennis, Tennis, and Track and Field. The decrease in GPA is more prominent for athletes of individual sports with a decrease of 1.6115 points, as opposed to 1.3707 points for athletes belonging to team sports. However, the results are not statistically significant.

Up to this point, almost all observable factors have been accounted for. However, none of the variables describe the student's emotional well-being and self-esteem. More unhappy and less confident students will undoubtedly face more difficulties and hence achieve less, when compared to students with the same abilities but who are more happy and confident in their lives. I first add the components of my self-esteem index as separate variables, to study the effect of each variable independently.

Happiness is linked to a small significant rise in GPA by 0.6405 points, whereas time management, "control", and "mind" show no significant effects. However, "abilities" and "future" incur a significant increase of 0.6345 and 0.8245 points in GPA for every additional rank given to these variables; the results are depicted by Table 9. People who show more concern for their academics and their future are more likely to perform better than those who don't. Also, when they feel like they have more control over their lives and the problems and difficulties that might arise, then their emotions are less likely to undergo fluctuations that they cannot quickly adapt to.

Extraversion, on the other hand, relates to a drop of 0.3423 points in GPA. Indeed, more extraverted and sociable people will want to allocate a considerable portion of their time to engage in social activities and hang around their friends. They find it more difficult to isolate themselves from the people around them in order to spend several hours focusing on their studies. Furthermore, extraverted people enjoy being part of varsity teams. Sports can act as a way to help them become more involved with others and create connections that would be beneficial on both the physical and academic levels.

```
Call:
lm(formula = GPAAV ~ ACADEMICS + Gender + OSB + FEA + FAS + FAFS.FHS.FM +
    MAJORLESS + Freshman + Sophomore + Junior + Senior + Commuting +
    Parents + Clubs. + Books + Athlete + Happiness + Time.mngt +
    Control + Mind + Abilities + Future + Extraversion, data = RESULTS)

Residuals:
     Min       1Q   Median       3Q      Max
-17.4809  -3.0297  -0.0741   2.8926  14.7351

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  60.45887    3.95778  15.276  < 2e-16 ***
ACADEMICS     0.26087    0.19414   1.344  0.18025
GenderM      -1.01464    0.65143  -1.558  0.12060
OSB           2.24211    1.91458   1.171  0.24268
FEA           4.11241    1.91515   2.147  0.03273 *
FAS           2.86373    1.88105   1.522  0.12917
FAFS.FHS.FM   2.18545    2.05557   1.063  0.28872
MAJORLESS     9.26708    4.41272   2.100  0.03672 *
Freshman     -6.43766    3.75960  -1.712  0.08808 .
Sophomore    -2.34618    1.50280  -1.561  0.11974
Junior       -1.89157    1.44918  -1.305  0.19300
Senior       -1.44438    1.38314  -1.044  0.29737
Commuting     0.35724    0.17066   2.093  0.03734 *
Parents       0.36313    0.17746   2.046  0.04177 *
Clubs.       -0.63110    0.62453  -1.011  0.31322
Books         0.50079    0.39622   1.264  0.20743
Athlete      -1.18739    0.67226  -1.766  0.07857 .
Happiness     0.64045    0.35429   1.808  0.07186 .
Time.mngt    -0.30613    0.27490  -1.114  0.26651
Control      -0.03543    0.16117  -0.220  0.82620
Mind         -0.15394    0.18379  -0.838  0.40306
Abilities     0.63454    0.19096   3.323  0.00102 **
Future        0.82447    0.14318   5.758 2.48e-08 ***
Extraversion -0.34227    0.12385  -2.763  0.00614 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.764 on 250 degrees of freedom
Multiple R-squared:  0.3764,    Adjusted R-squared:  0.3191
F-statistic: 6.562 on 23 and 250 DF,  p-value: 1.3e-15
```

Table 5.9: Adding components of the self-esteem index

All the self-rated values are joined together to reflect the self-esteem index. In general, higher self-esteem is correlated with 0.1823 additional points on GPA, as proven by the results in Table 10. Though the difference is not substantial, it is highly significant. As explained, more confidence will help students achieve more; their self-esteem is positively linked to their academic success as they feel more comfortable and at ease with their university. In a study on first-year college students, it was shown that there exists a significant and substantial effect of self-efficacy on academic performance. “Academic self-efficacy is significantly and directly related to academic expectations. Students who enter college with confidence in their ability to perform well academically do perform significantly better than do less confident students. Likewise, students who have higher expectations for academic success show higher performance” (Chemers, Hu, and Garcia, 2001).

From this table, the effect of the variable Athlete is shown to be statistically significant on the 95% confidence level. Being an athlete corresponds to a decrease in GPA by 1.5900 points. The decrease can be explained by the assumption that being committed to a team is time-consuming. Also, athletes are prone to celebrate and party and drink together. This affects GPA in two ways. First, students will allocate a large portion of their time to trainings and competitions as well as negative leisure activities, which will leave them with little time to focus on their studies, as explained by Becker’s allocation of time theory. Second, “drinking causes a statistically significant reduction in performance on the order of approximately one-tenth of a standard deviation” (Carrell et al., 2011).

```
Call:
lm(formula = GPAAV ~ ACADEMICS + Gender + OSB + FEA + FAS + FAFS.FHS.FM +
    MAJORLESS + Freshman + Sophomore + Junior + Senior + Commuting +
    Parents + Clubs. + Books + Athlete + Self.esteem, data = RESULTS)

Residuals:
     Min       1Q   Median       3Q      Max
-18.6930  -3.2596  -0.1632   3.5021  17.5457

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 62.77968    4.18637  14.996  < 2e-16 ***
ACADEMICS    0.41768    0.19652   2.125   0.0345 *
GenderM     -1.72138    0.68562  -2.511   0.0127 *
OSB          1.49398    2.08250   0.717   0.4738
FEA          4.19141    2.07845   2.017   0.0448 *
FAS          1.94632    2.02964   0.959   0.3385
FAFS.FHS.FM  1.26471    2.21898   0.570   0.5692
MAJORLESS    7.00771    4.73799   1.479   0.1404
Freshman    -7.95154    4.03942  -1.968   0.0501 .
Sophomore   -2.81496    1.62505  -1.732   0.0844 .
Junior      -1.87868    1.57912  -1.190   0.2353
Senior      -1.66438    1.50681  -1.105   0.2704
Commuting    0.34812    0.18533   1.878   0.0615 .
Parents      0.43718    0.19143   2.284   0.0232 *
Clubs.      -1.29089    0.66118  -1.952   0.0520 .
Books        0.56274    0.42013   1.339   0.1816
Athlete     -1.59003    0.72225  -2.201   0.0286 *
Self.esteem  0.15976    0.03964   4.030 7.37e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.208 on 256 degrees of freedom
Multiple R-squared:  0.2368,    Adjusted R-squared:  0.1861
F-statistic: 4.672 on 17 and 256 DF,  p-value: 1.372e-08
```

Table 5.10: Adding the self-esteem index

To test the robustness of the model, I create a plot showing the predicted values of the regression versus the residuals. The points appear to be symmetrically distributed, tending to cluster toward the middle of the plot. They do not follow a clear pattern, indicating homoscedasticity, and low variance.

In an attempt to make the model a stronger one, I remove variables that have shown no significant effects on the grade point average, namely the major, year of study, Clubs, and number of books read. Running the regression yields figures significant on the 95% confidence level. The regression of the final model is represented by equation 3 and Table 11 below.

$$GPA = \beta_0 + \beta_1 \text{*Academic} + \beta_2 \text{*Gender} + \beta_3 \text{*Commuting} + \beta_4 \text{*Parents} + \beta_5 \text{*Athlete} + \beta_6 \text{*Self.Esteem} + \varepsilon \quad (3)$$

β_0, β_1, β_2, β_3, β_4, β_5, and β_6 represent the coefficient of the variables being studied. ε captures all the other unobserved variables that have not been controlled for in this final model.

```
Call:
lm(formula = GPAAv ~ ACADEMICS + Gender + Commuting + Parents +
    Athlete + Self.esteem, data = RESULTS)

Residuals:
     Min       1Q   Median       3Q      Max
-19.7636  -3.6289   0.1358   3.4213  15.1268

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 61.02902    3.11875  19.568  < 2e-16 ***
ACADEMICS    0.44364    0.18903   2.347   0.0197 *
GenderM     -1.41016    0.65200  -2.163   0.0314 *
Commuting    0.42816    0.18533   2.310   0.0216 *
Parents      0.43612    0.19037   2.291   0.0227 *
Athlete     -1.46619    0.70485  -2.080   0.0385 *
Self.esteem  0.18226    0.03826   4.764 3.12e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.325 on 267 degrees of freedom
Multiple R-squared:  0.1677,    Adjusted R-squared:  0.149
F-statistic: 8.966 on 6 and 267 DF,  p-value: 6.281e-09
```

Table 5.11: Removing insignificant variables

Again, I test the robustness of the new model using the predicted versus residual values plot and find no clear pattern. The colors depict, as explained by the legend, the value of the residuals. Values close to zero are white, while more positive values become gradually more saturated in red color, and the negative values in blue. Figure 2 below shows the predicted values versus the residuals of the model.

To further validate the results I use the variance inflation factor (VIF), to test how much of the variance is caused by multicollinearity. The latter is a problem that indicates variables are correlated to other predictors in the model. The higher the multicpllોinearity, the more sensitive the coefficient estimates become to any minor changes in the model. This makes the coefficients unstable and difficult to interpret, making the model very weak.

If the explanatory variables return a high VIF, meaning between 5 and 10, then this variable should be removed as it indicates high collinearity. The accepted values that minimize collinearity are understood to be close to 1, but may range between 1 and 5. At this point, the model will have moderate multicollinearity, which is not considered as a major problem.

Applying the VIF to my model, all values returned appear to be close to 1 (Figure 3), which suggests that the variables are slightly correlated, thus multicollinearity is low, and so is variance, indicating the model is strong. No further predictors should be removed or linearly combined together into one new index.

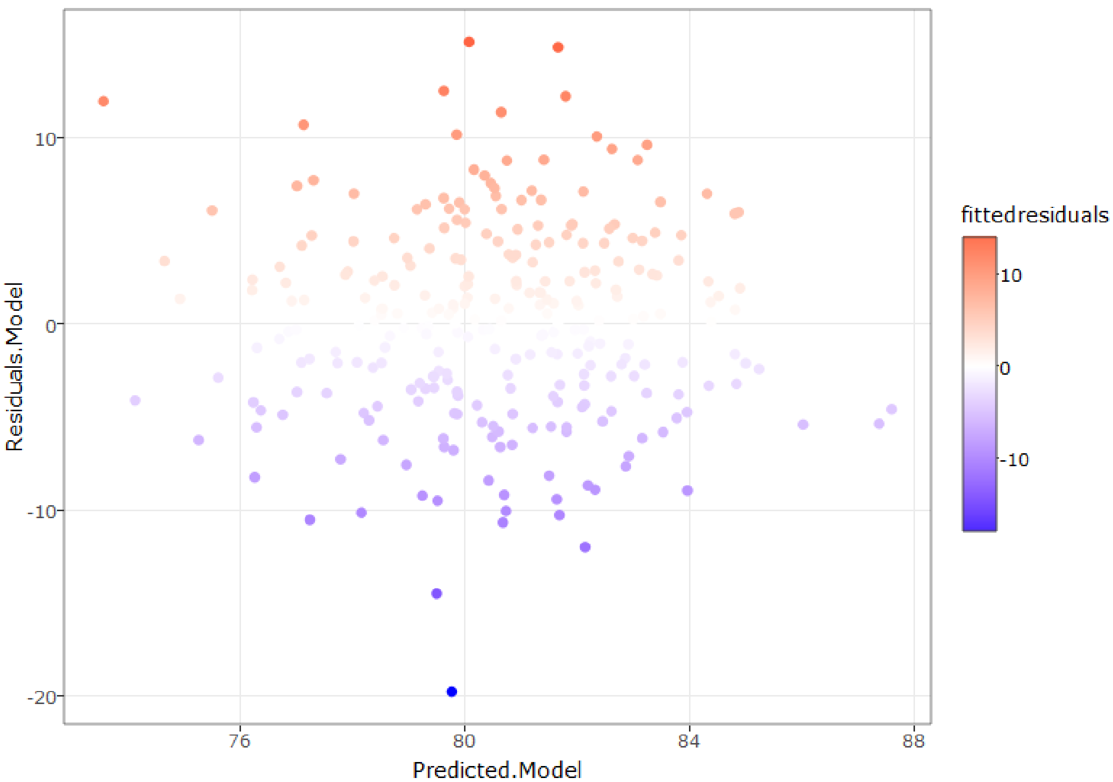

Figure 5.2: Predicted versus Residuals Model

```
 Athlete     Gender  Commuting  ACADEMICS    Parents Self.esteem
1.189139   1.026268   1.151875   1.037600   1.021528    1.089874
```

Figure 5.3: Variance Inflation Factor

CHAPTER VI

CONCLUSION

The results have proved that engaging in physical activity and being part of a varsity team will worsen a student's academic performance. Yet, the decrease in GPA might be too acute when compared to the benefits that sports will bring. Better health, increased self-confidence, less stress, and a sense of belonging are worth the 1 point lost.

GPA is affected by so many uncountable factors, that the effect of sports might be attenuated when other more influential factors take over. Grades are affected by the student's gender, their other activities, the time spent studying, the time spent commuting, the parents' highest education levels, and their perceived self-esteem levels.

It is true that females will, on average, perform better than males, and so will students with more highly educated parents, but this should only encourage the students who don't have these "advantages" to make up by increasing their efforts. Although sports would not, on its own, help them achieve better, it can help them focus more when studying, relieve their stress, and raise their self-esteem, which would all be reflected positively on their grades. The positive effect will offset the minor drop in grades that sports triggers.

At the end of the day, a student's academic achievement does not only depend on his academic performance, but on all the values and qualities he has learned and strengthened during his university experience, which can all be acquired through participation in sports.

Limitations

First of all, the sample size is not big enough to eliminate sampling error, which yields results that are not completely accurate. They are even less so when the sample is categorized into athletes and non-athletes as the number of observations is halved for each, and when sorted into individual versus team sports. This also happens when students are grouped by faculties and years of study, which leaves every category with as little as 50 students.

Second, all answers are self-reported resulting in measurement error thus response bias. The questions in the survey might appear as leading the respondents to the surveyor's more desirable answers. Moreover, as people were completing the survey in groups, they might be affected by each other's responses.

Third, the study does not take into account whether the student is involved in an on-campus work-study program or takes on a part-time job off campus. Wherever the job, it will require significantly more time and commitment than unemployed students. This might also include volunteering activities and camps that take up the student's weekends.

Last but not least, the students are not compared to themselves. The differences between participants are too numerous to be able to control for each, and they may include several unobservable factors that result in the model's omitted variable bias. It also cannot account for any changes in lifestyle that the student might have undergone throughout his university years, which would have impacted his grades.

Directions for future research

A stronger design would involve a longitudinal study as opposed to a cross sectional one. The students could be observed before and after having introduced the treatment, which is in that case belonging to a varsity team. That way, the results will show clear and precise effect that compare the student to himself. A regression discontinuity design would make the stronger and fittest model.

APPENDIX

Survey Questions included:

- Gender
- Age
- Mother’s highest level of education
- Father’s highest level of education
- Major
- Year of study
- Do you belong to a varsity team? If yes, which one?
- Average number of hours spent on varsity practice per week for every semester
- Average number of hours spent per week on non-varsity sports
- Average number of hours spent per week commuting to and from AUB
- Do you take part in other activities and club that recommend more than 3 hours a week?
- Average number of books read per year for non-academic purposes
- Average number of hours per week spent studying
- Percentage of completed assignments
- Average number of classes skipped per week
- SAT Scores - Reading and Math)
- GPA for every semester
- Rate each of the following on a scale of 1 to 5, 1 being poor and 5 being very good
    - Level of happiness

- Writing Skills
- Analytical Skills
- Computer Skills
- Grade consciousness
- Concern about future
- Time management
- Coping effectively with important life changes
- Ability to handle personal problems
- Ability to control difficulties in life
- Academically focused
- Ambitious
- Smart
- Intellectual
- Artistic/ creative
- Socially skilled
- Outgoing
- Confident
- Good leader

REFERENCES

1. Astin, A. W. (1993). What Matters in College? Four Critical Years Revisited. San Francisco: Jossey-Bass.
2. Branch, J. L. (2003). Extracurricular activities and academic achievement (Doctoral dissertation, University of Southern Mississippi). Retrieved from ProQuest Dissertations and Theses.
3. Broh, B. A. (2002). Linking extracurricular programming to academic achievement: Who benefits and why? *Sociology of Education, 75,* 69-91.
4. Cabane, C., Clark, A.E. (2015). Childhood sporting activities and adult labour-market outcome. *Annals of Economics and Statistics, GENES, issue. 119-120* pp. 123–148.
5. Cabane, C., Hille, A., & Lechner, M. (2016). Mozart or Pelé? The effects of adolescents' participation in music and sports. Labour Economics, 41, 90-103. doi:10.1016/j.labeco.2016.05.012
6. Cantor, N. E., and Prentice, D. A. (1996). The Life of the Modern-Day Student-Athlete: Opportunities Won and Lost. Paper presented at the Princeton Conference on Higher Education, Princeton University, Princeton, NJ.
7. Cole, A. (2014). Sports Participation and Academic Achievement: Does Self-Efficacy Play a Role? (Doctoral dissertation, University of Arizona). Retrieved from ProQuest Dissertations and Theses.

8. Cornelißen, T., & Pfeifer, C. (2007). The impact of participation in sports on educational attainment—New evidence from Germany. *Economics of Education Review, 29*(1), 94-103.
9. Eide, E. R., & Ronan, N. (2001). Is participation in high school athletics an investment or a consumption good? Evidence from high school and beyond. *Economics of Education Review, 20,* 431-442.
10. Ewing, B. T. (1998). Athletes and work. *Economics Letters, 59*(1), 113-117. doi:10.1016/s0165-1765(98)00006-8
11. Ewing, B. T. (2007). The Labor Market Effects of High School Athletic Participation: Evidence From Wage and Fringe Benefit Differentials. *Journal of Sports Economics, 8*(3), 255-265. doi:10.1177/1527002505279348
12. Fejgin, N. (1994). Participation in high school competitive sports: A subversion of school mission or contribution to academic goals? *Sociology of Sport Journal, 11,* 211-230.
13. Hauser, W. J. and Lueptow, L. B. (1978). Participation in Athletics and Academic Achievement: A Replication and Extension*. *The Sociological Quarterly*, *19*(2): 304–309. doi:10.1111/j.1533-8525.1978.tb01172.x
14. Heilman, C.A. (2011). A Mixed Methods Approach Examining Alpine Ski Racing as a Context for Positive Youth Development. *The University of Utah, Salt Lake City*, Utah.
15. Hille, A., & Schupp, J. (2015). How learning a musical instrument affects the development of skills. *Economics of Education Review, 44*, 56-82. doi:10.1016/j.econedurev.2014.10.007

16. Hood, A. B. (1992). The impact of athletics, part time employment, and other activities on academic achievement. *Journal of College Student Development*, *33*, 447-453.

17. Jefferson, C. (1999). Educational performance of athletes and nonathletes in two Mississippi rural high schools (Doctoral dissertation, University of Southern Mississippi). Retrieved from ProQuest Dissertations and Theses.

18. Lechner, M., & Sari, N. (2015). Labor market effects of sports and exercise: Evidence from Canadian panel data. *Labour Economics, 35*, 1-15. doi:10.1016/j.labeco.2015.04.001

19. Lipscomb, S. (2007). Secondary school extracurricular involvement and academic achievement: a fixed effects approach. *Economics of Education Review*, *26*, 463-472.

20. Lueptow, L., and Kayser, B. (1973). Athletic Involvement, Academic Achievement, and Aspiration. *Sociological Focus*, *7*(1), 24-36.

21. Lumpkin, A., & Favor, J. (2012). Comparing the academic performance of high school athletes and non-athletes in Kansas in 2008-2009. *Journal of Sport Administration & Supervision* *4*(1), 41-62.

22. Maloney, Michael T. and McCormick, Robert E. (1993). "An Examination of the Role that Intercollegiate Athletic Participation Plays in Academic Achievement: Athletes' Feats in the Classroom," *Journal of Human Resources* 28: 555-570.

23. Marsh, H. W. (1992). Extracurricular activities: Beneficial extension of the traditional curriculum or subversion of academic goals? *Journal of Educational Psychology*, *84*, 553-562.

24. Marsh, H. W. (1993). Stability of Individual Differences in Multiwave Panel Studies: Comparison of Simplex Models and One-Factor Models. *Journal of Educational Measurement, 30*(2), 157-183. doi: 10.1111/j.1745-3984.1993.tb01072.x

25. McCarthy, K. J. (2000). The effects of student activity participation, gender, ethnicity, and socio-economic level on high school student grade point averages and attendance. Retrieved from http://www.eric.ed.gov/PDFS/ED457173.pdf

26. Neish, M. A. (1993). Student activities and academic achievement at Marysville High School. Reston, VA: National Association of Student Activity Advisers.

27. O'Brien, E., & Rollefson, M. (1992). Extracurricular participation and student engagement. Washington, DC: National Center for Education Statistics.

28. Pascarella, E. T., and Smart, J. C. (1991). Impact of intercollegiate athletic participation for African-American and Caucasian men: Some further evidence. *Journal of College Student Development 32*(2): 123–130.

29. Pfeifer, C., & Cornelissen, T. (2010). The impact of participation in sports on educational attainment: New evidence from Germany. Economics of Education Review, 29, 94–103.

30. Rees, D. I., & Sabia, J. J. (2010). Sports participation and academic performance: Evidence from the National Longitudinal Study of Adolescent Health. *Economics of Education Review, 29*, 751–759.

31. Routon, P. W., & Walker, J. K. (2014). Student-Athletes? The Impact of Intercollegiate Sports Participation on Academic Outcomes. *Eastern Economic Journal, 41*(4), 592-611. doi:10.1057/eej.2014.32

32. Ryan, F. J. (1989). Participation in intercollegiate athletics: Affective outcomes. *Journal of College Student Development 30*(2): 122–128.

33. Steinmayr, A., Felfe, C., & Lechner, M. (2011). The closer the sportier? Children's sports activity and their distance to sports facilities. *European Review of Aging and Physical Activity, 8*(2), 67-82. doi:10.1007/s11556-011-0090-0

34. Stencel, J. R. (2005). A study of the relationship between interscholastic athletic participation and academic achievement for a group of Tennessee high school students (Doctoral dissertation, Union University). Retrieved from ProQuest Dissertations and Theses.

35. Stevenson, B. (2010). Beyond the classroom: Using Title IX to measure the return to high school sports. *The Review of Economics and Statistics, 92*(2), 284-301.

36. Stuart, D. L. (1985). Academic preparation and subsequent performance of intercollegiate football players. *Journal of College Student Personnel*, *26*, 124-129.

37. Tomporowski, P., Davis, C., Miller, P., & Naglieri, J. (2008). Exercise and children's cognition and academic achievement. *Educational Psychology Review, 20*, 111-131. doi: 10.1007/s10648-007-9057-0

38. Yeung, R. (2015). Athletics, Athletic Leadership, and Academic Achievement. *Education and Urban Society, 47*(3), 361-387. doi:10.1177/0013124513495277

39. Zaff, J. F., Moore, K. A., Papillo, A. R., & Williams, S. (2003). Implications of extracurricular activity participation during adolescence on positive outcomes. *Journal of Adolescent Research, 18,* 599-630.