

A Work Project, presented as part of the requirements for the Award of a Master's degree in
Economics from the Nova School of Business and Economics.

**Who benefits from mentoring?
A randomized controlled trial at Tilburg University**

Johanna Stefanie Kary, BSc.

Work project carried out under the supervision of:

Lenka Fiala, PhD

16-12-2021

Abstract

The study evaluates a mentoring program for undergraduate students of the Business Administration Bachelor at the Tilburg University. The program was focused on personal development, especially goal setting. In 2017, the mentoring was introduced to 274 first year students using randomized encouragement and turned mandatory for 296 first year students in 2018. The students were mentored by older peers more advanced in their studies. While the data fails to detect statistically significant effects when looking at the general student body as well as for female versus male students, statistically significant effects are measured for those students that only participated in the mentoring because they were forced to in 2018. In the first three semesters of their bachelor's degree these students failed on average 2 classes fewer compared to their peers who would have signed up anyways on a voluntary basis. These findings suggest that mentoring might be a promising tool for students that are at risk of failing a class and thus help universities to improve retention rates. Considering the costs of mentoring and the limited beneficial effects this approach might not be cost-efficient.

Keywords – Education, Mentoring, Randomized controlled trial, Behavioral Economics

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 1 |
| 2 | Literature | 2 |
| 3 | Experimental Design | 5 |
| 4 | Data | 10 |
| 5 | Methodology | 11 |
| 6 | Results | 14 |
| 7 | Discussion | 19 |
| 8 | Conclusion | 22 |
| | References | 24 |
| | Appendix | 27 |
| | A1 Balance table 2017 cohort control and treatment | 27 |
| | A2 Balance table 2018 IBA and other degrees | 27 |

List of Figures

| | | |
|-----|---|---|
| 3.1 | Outcome of the power test 2017 | 9 |
| 3.2 | Outcome of the power test 2017 and 2018 | 9 |

List of Tables

| | | |
|------|---|----|
| 3.1 | Timeline of mentoring activities in the first and second semester | 7 |
| 4.1 | General demographics 2017 and 2018 cohort only IBA students | 11 |
| 6.1 | Two-sample t-test with equal variances 2017 | 14 |
| 6.2 | Two-sample t-test with equal variances 2018 IBA | 15 |
| 6.3 | The effect of the mentoring on average grades | 15 |
| 6.4 | The effect of the mentoring on graduation and failing | 16 |
| 6.5 | Including gender in the analysis | 17 |
| 6.6 | Two-sample t-test with equal variances 2017 subscribed / not subscribed . | 18 |
| 6.7 | The effect of the change from voluntary to mandatory participation | 19 |
| A1.1 | Balance table 2017 cohort control and treatment group | 27 |
| A2.1 | Balance table 2018 IBA and other degrees | 28 |

1 Introduction

A concern of universities all over the world has been the number of students dropping out of their degrees. The retention rate measures how many students returned to the university at the beginning of the second year of the program versus how many students dropped out of their degree. In the US for example this rate was 81 percent in 2019 (US Department of Education, 2020) and in Germany the retention rate was 68 percent in the same year (Spangenberg et al., 2020).

For universities high dropout rates are a source of various problems. On one hand, it is an economical problem, as recruiting new students is more costly than retaining enrolled ones (Schertzer and Schertzer, 2004). Additionally, students that drop out mean, that university subsidies have been allocated without any return on investment (Letseka and Maile, 2008). On the other hand, a high dropout rate is a problem in the competitive university environment, as it decreases the attractiveness of a university (Hanson et al., 2010). To keep enrollment rates and thus competitiveness high, universities have been focusing on improving the retention rate (Webster and Showers, 2011).

One common approach implemented to improve retention rates is mentoring. With mentoring many universities try to help students face problems that can lead to dropouts, like program related or social difficulties (Fozdar et al., 2006). Many studies have been conducted to measure the effectiveness of this approach, most of them finding small positive results (Eby et al., 2008). However, considering the immense costs associated with the implementation of mentoring at the university level it is also necessary to not only understand **if** the mentoring can have positive benefits, but also to examine **who** are the people that benefit.

The mentoring program studied in this paper is Tilburg University's "International Business Administration Mentoring" program, which was introduced in 2017. In 2018 the program became mandatory for all undergraduate students. I evaluate the programs effect on average grades, the number of failed classes and the graduation time.

2 Literature

Previous literature on the effects of academic mentoring suggests that mentoring can have longterm benefits, for example on satisfaction, career progression and academic achievement (Efstathiou et al., 2018; Mayer et al., 2014).

A vast body of literature is focused on the mechanisms that make mentoring beneficial, for example the role of goal setting in enhancing self-control and preventing procrastination and present bias (Latham and Locke, 1979; Clark et al., 2020; Lent and Souverijn, 2016; Hsiaw, 2013; Duckworth et al., 2010; Moffitt et al., 2011), or increasing self-efficacy and shifting towards an internal locus of control (Cobb-Clark, 2015; Judge and Bono, 2001; Ng et al., 2006). Another mechanism of mentoring programs that has been analyzed is the importance of role models. Some studies suggest that the effectiveness of a role model is dependent on their similarity to the targeted person as well as on the information they communicate (Fiala, 2020; Gladstone and Cimpian, 2020). A successful role model will therefore for example provide information about payoffs in uncertain situations like future career prospects (Breda et al., 2018). While there are many studies examining the reasons why mentoring can be beneficial, fewer studies are being conducted on finding out who are the people that benefit from mentoring. In general, two main aspects are being considered in research about educational outcomes: differences in gender and students' ability.

Gender differences

A large body of literature focuses on gender differences in ability and performance. Gender differences in ability and performance in a school or university context are noted in the literature from an early age. Logan and Johnston (2009) studied the reading ability and attitude towards reading of 232 ten year-old children and find that girls have a small advantage in reading ability, read more frequently and have a significantly more positive attitude towards reading. In university, these differences in performance seem to increase. Except for engineering, manufacturing and construction more women than men graduate in all degree types all over the world (OECD, 2020). Women also earn the majority of doctoral degrees (Zhou and Gao, 2021). In a study at the University of Jordan, Khwaileh and Zaza (2011) show that female undergraduate students consistently outperform male counterparts. Dayioğlu and Türüt-Aşık (2007) support these findings for the Middle East Technical University (METU), a large public university in Turkey. McNabb et al. (2002)'s findings for university students in England and Wales replicate the same pattern; however, the authors also show that female students are significantly less likely to graduate with honours.

While men seem to perform generally worse in university, the reasons for this are not clear. Differences in work ethic, language ability, course taking behavior and cognitive processing have been discussed (Wilberg and Lynn, 1999; Byrnes et al., 1997; Young and Fisler, 2000). In their randomized study on student mentoring Bettinger and Baker (2014) found that male completion rates lagged behind, but mentoring had a larger effect on male students and thus indicated that student mentoring could reduce the gender gap in university performance.

Based on these findings I want to add to the literature and explore if mentoring can help

to close the gender gap in academic achievement and hypothesize:

Hypothesis 1: The mentoring program benefits female and male students to a different extent.

Students at risk

With the aim to improve college graduation rates, differences between high- and low-performing students have been studied. Rodríguez-Planas (2012) provides a meta-analysis of the impact of mentoring, educational services and financial rewards. They find evidence that mentoring can have modest but positive effects for at risk or disadvantaged students. Different reasons for these effects are being discussed: mentoring might help students without the necessary support of their family to increase the mentee's self-worth perception and competency (Rhodes et al., 2000). Students that are considered to be at risk might also be more in need of positive role models than high achieving students (Rodríguez-Planas, 2012). A randomized study by Bettinger and Baker (2014) on the effectiveness of student mentoring supports that theory, as the authors found that mentored students had a lower dropout rate than students without mentoring. Salinitri (2005) evaluated the effect of mentoring on first-year bachelor's students and found that the mentored students had higher GPAs and higher retention rates. Reeves (1996), however, finds no effect of mentoring on GPA and retention rate. The author notes that the mentoring relationship evolved in the study was more crisis-oriented than personal-development-oriented, as the students reached out to their mentors for help mostly when being confronted with social or academic failure.

The existing studies on the impact of mentoring on "at risk" students are thus ambiguous, and further research is needed. With this paper I aim to contribute to the literature and

hypothesize:

Hypothesis 2: The mentoring program benefits high and low performing students to a different extent.

3 Experimental Design

Program design, content and timeline

The study was conducted as a randomized control trial (RCT) at Tilburg University. 274 students from the bachelor's degree of International Business Administration who started their studies in the Fall of 2017 were offered the chance to register for a mentoring program as they received basic information materials about an intensive tutoring program called "IBA's Personal Development Program" with the aim to help the students to improve their personal and professional development. The students were stratified into 16 bins and then randomly assigned to a treatment and a control group. 133 students were assigned to the treatment group, 141 to the control group. The treatment consisted of additional encouragement to join the program, for example in the form of a marketing video. The additional material was targeted and personalised in order to increase the motivation of the students to join the mentoring program.

The main output of the mentoring program is a personal development plan that each student has to write in order to participate in the program. The plan includes the current strengths and weaknesses of the students as well as their goals and the conditions needed to achieve them. The personal development plan is supposed to be updated throughout the program and has three different themes depending on the year of university the student

is currently attending. In the first year the main topic is studying and study skills. The following year focuses on professional skills and the labor market and in the last year of the students bachelor's degree the main topic of the personal development plan is studying abroad, following a minor or doing an internship. The plan is supposed to create self-awareness and a sense of direction and motivation. With this plan the program aims to benefit the students through the mechanism of goal setting. Another goal of the program is to help students become acclimatised with the transition from high school to university with the support of their mentors. The students are expected to benefit from their mentors through signaling and social proof, as they are supposedly seen as successful older peers. The itinerary of the program consisted of an information session in the introduction week for all students, followed by a kick-off session organized by the mentors and training sessions throughout each semester. The training sessions cover the topics of study skills and self reflection. Twice a year the students attend assessment sessions where they assess their personal development plan. For the 2017 cohort, these assessment sessions took place in November 2017 and February 2018. Feedback sessions complete the program setup.

The detailed timeline of the mentoring program for year 1 of the cohort 2017 is described in table 3.1. The separator indicates the division between the first and second semester, respectively the short and medium-run of the program.

For the students starting their university degree in 2018 the university extended the program to more degree types and made the attendance compulsory. These changes in the setting allow me to measure if there was a difference in the impact of the program based on its voluntary or compulsory participation design by comparing the participating students from the 2017 cohort with the 2018 International Business Administration cohort.

Table 3.1: Timeline of mentoring activities in the first and second semester

| Activity | Date |
|----------------------------------|--------------------|
| Student mentor training 1 | 15 June 2017 |
| Mentor kick off | 13 September 2017 |
| Program subscription | 15 September 2017 |
| Student mentor training 2 | 3-4 October 2017 |
| Student kick off (decentralized) | 4-13 October 2017 |
| Student Workshop 1 | 24-27 October 2017 |
| PDP Assignment 1 | 1 November 2017 |
| Exit interview | 13 November 2017 |
| PDP Assignment 1 feedback | 15 November 2017 |
| Consulting hour 1 | December 2017 |
| Preliminary evaluation | January 2018 |
| PDP Assignment 2 | 1 February 2018 |
| PDP Assignment 2 feedback | 15 February 2018 |
| Consulting hour 2 | March 2018 |
| Exit interview | March 2018 |
| Student Workshop 2 | Mid-April 2018 |
| PDP Assignment 3 | 1 June 2018 |
| PDP Assignment 3 feedback | 15 June 2018 |
| Consulting hour 3 | July 2018 |
| End of the year activity | July 2018 |
| Final evaluation | July 2018 |

Recruitment and training of the mentors

Overall, six academic mentors and 16 student mentor pairs worked as mentors in the program. These student mentors were chosen by the Mentorship System of the Tilburg School of Economics and Management (TiSEM), while the academic mentors were chosen by the heads of the departments. The mentors were trained on coaching, cross-cultural differences and feedback as well as meeting conduction. The mentors coached the students in the areas of individual development, motivation, study related activities, student life and study progress problems. The mentoring was also expected to create a positive and cooperative group-feeling. The academic mentors were supposed to supervise the students mentors and take action in case of study-related problems. For their work as a mentor the students received a 400€ reward, the mentoring training and an official certificate.

Estimation

The set up of this randomized control trial in the 2017 cohort allows for the Intention to treat Effect (ITT) to be captured, as I have the sign-up data of the students but no information about their actual participation. Because of noncompliance and withdrawal I can only measure the ITT. By comparing the treatment and the control group I see the impact on those "moderately interested students", meaning those students that only signed up for the program because of the additional encouragement while their equivalents in the control group did not sign up as they did not receive the treatment. As 52 percent of the treatment group signed up for the mentoring in the 2017 cohort and only 29 percent of the non-treated students signed up, I can conclude that the share of "moderately interested students" is 23 percent in 2017. The share of students who would have signed up without any additional encouragement is 29 percent and the students who did not sign up even with the encouragement is 48 percent.

The power of this RCT in the 2017 cohort is 55 percent as shown in figure 3.1, thus falling below the commonly accepted power of 80 percent. The calculation was done by Optimal Design software and is based on the minimum detectable effect size of 0.5 standard deviation, $p < 0.05$ and the sample size of "moderately interested students" of 63. The experiment is underpowered, meaning that if I fail to detect an effect it is impossible to determine whether this stems from an absence of an effect or the absence of evidence.

The power of the comparison between the 2017 and 2018 cohort is 85 percent as shown in figure 3.2. Hereby considered is the sample size of "not at all" interested students of 48 percent, meaning 142 students. These are the students that did not want to attend the program even with the additional encouragement and thus were affected by the change of the program design from voluntary to mandatory.

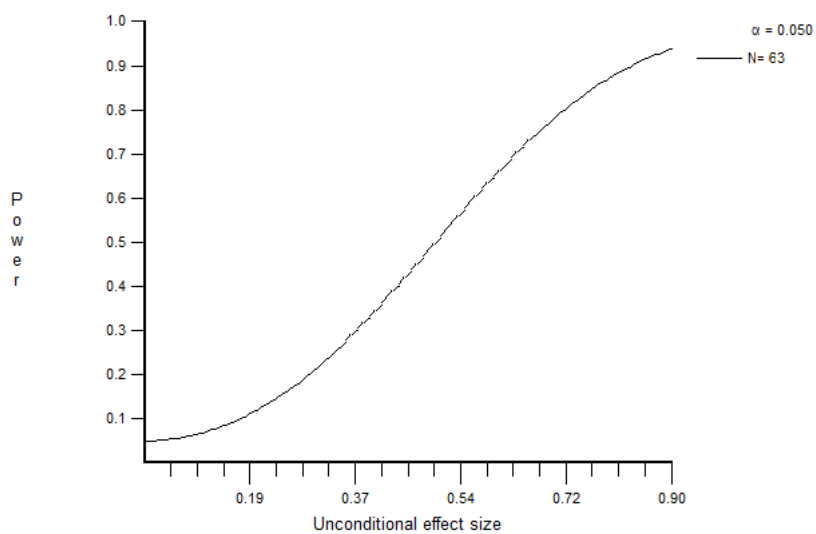


Figure 3.1: Outcome of the power test 2017

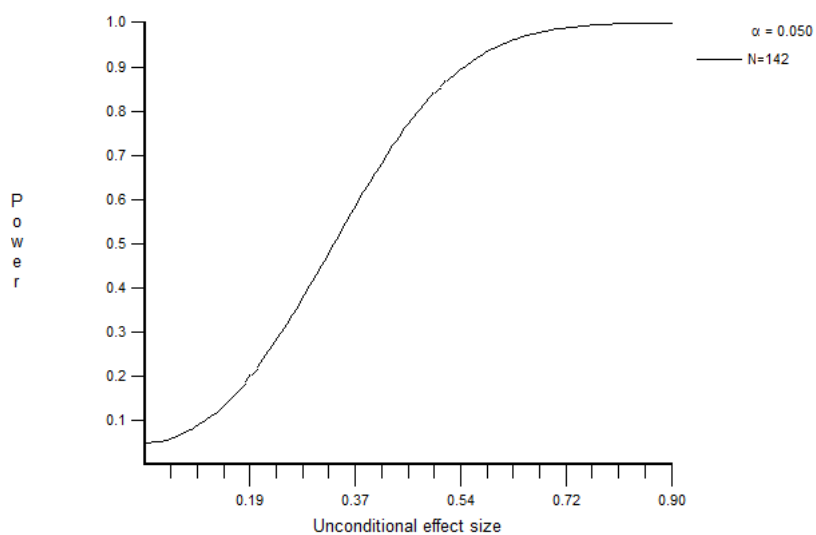


Figure 3.2: Outcome of the power test 2017 and 2018

4 Data

The available data contains information on the 274 students in the 2017 cohort and the 296 International Business Administration students in the 2018 cohort. 133 of the students in the 2017 cohort received the treatment, 141 did not. The data also contains information about the gender, nationality (Dutch/Non-Dutch), treatment status and average grade of the students in each semester, as well as a dummy that indicates if the students graduated in three years and the rate of failed exams throughout the evaluation period. Information about each course the students took and the final grade achieved are also available.

The data was recorded from the beginning of the first semester of the 2017 cohort, September 1st of 2017 up until the end of August 2021. As the program came to a stop because of the Covid-19-pandemic by March 2020, all observations after this moment are dropped. This means that for the 2018 cohort I only consider the first three semesters, and for the 2017 cohort the analysis stops after five semesters.

The treatment and the control group in the cohort 2017 are balanced, thus sufficiently similar in their general characteristics. Considered in the analysis were the nationality, the average evaluation score of the application package, the ranking of the application evaluation score, the application date, the gender and the age. The results of the analysis can be found in table A1.1 in the Appendix.

All observations in the 2017 and 2018 cohort are part of the International Business Administration degree. Table 4.1 shows that the gender ratio of the observations from the 2018 cohort is sufficiently similar to the 2017 cohort. However, there is a difference in nationality ratio, which is why I will control for nationality in my analysis.

Table 4.1: General demographics 2017 and 2018 cohort only IBA students

| Cohort | Share of Female Students | Share of Non-Dutch Students |
|----------------|--------------------------|-----------------------------|
| 2017 | 39.58 | 15.90 |
| 2018 | 37.62 | 26.34 |
| t-test p-value | 0.6229 | 0.0018 |

The numbers displayed are percentages.

5 Methodology

The experiment was conducted using an encouragement design, meaning the individuals in the treatment group received additional encouragement to join the mentoring program. The mentoring program was available for every student however, which is why I estimate the effect of the mentoring by using a 2SLS approach, using the encouragement as an instrument for participating in the mentoring program. In the encouraged group, 52.6 percent of students signed up for the mentoring. In the non-encouraged group, 30.4 percent did. I can compare both groups, as the take-up rate is higher in the encouraged group. Since encouragement was randomly assigned, any difference in outcomes can be attributed to the difference in the take-up rates of mentoring.

The first stage of the 2SLS approach is following:

$$\widehat{participation} = \delta + \beta * encouragement_i + \varepsilon_i \quad (5.1)$$

I then estimate the Intention-to-treat (ITT) using equation 5.2:

$$GPA_{i,t} = \alpha + \beta * \widehat{participation}_i + \mu_i \quad (5.2)$$

where GPA is an outcome for individual i . i equals to the 274 individuals in the 2017 cohort and t equals to the five semesters that are being analysed. Participation measures if the individual participated in the mentoring program, which is being instrumented by encouragement, a variable that indicates whether an individual received the encouragement or not. The standard errors control for heteroskedasticity. The outcome of interest is the student's average of exam grades per semester, their GPA. Respectively, for the outcome of interest "Graduation in three years", a dummy variable, I estimate equation 5.3:

$$Y_i = \gamma + \beta * \widehat{participation}_i + \eta_i \quad (5.3)$$

The variable Y in this equation is the graduation dummy, equal to one if a student graduated in three years and zero otherwise. As previously mentioned the program came to a stop after five semesters due to the Covid-19 pandemic. However, assuming that a majority of the progress towards a degree completion has already happened during the first five semesters, and considering that there is no reason to assume endogeneity between the pandemic and the encouragement treatment, I still decided to measure if mentoring had an effect on the time needed to complete the degree. To capture the effect of the program on the number of classes that were not passed throughout all five semesters, I adjust the outcome of interest Y so that it represents the sum of a students failed classes. To see whether the mentoring had a different impact on male than female students, I will add a gender dummy to the previous equations as well as an interaction term, exemplified in equation 5.4 and 5.5,

$$\widehat{participation} = \delta + \beta * encouragement_i + \kappa * Female_i + \lambda_i \quad (5.4)$$

$$Y_i = \theta + \psi * \widehat{participation}_i + \chi * Female_i + v * \widehat{FemaleParticipation} + \sigma_i \quad (5.5)$$

To compare the impact of the program in 2017 and 2018, when the program switched to being compulsory, I estimate the effect by considering only the students in the 2017 cohort that participated in the program and compare it with all students of the 2018 IBA sample. Overall, 297 students are included in this sample. The comparison of the 2017 and the 2018 cohort will give me the estimates on the "not at all" interested students that were forced to attend the program in 2018. I also consider only the grades of courses that both cohorts attended, thus excluding the observations of the courses "Marketing Management" and "Project International Business Administration" that only some students of the 2017 cohort attended, as well as the courses "International Business Research" and "Philosophy of Science" which only students of the 2018 cohort attended. Additionally, I control for nationality.

$$GPA_i = \delta + \beta * cohort2018_i + \gamma * X_i + \varepsilon_i \quad (5.6)$$

As with the 2017 evaluation, I control for gender and add the interaction term of being part of the 2018 cohort as a female student.

$$GPA_i = \delta + \beta * cohort2018_i + \gamma * X_i + \alpha * Female_i + \lambda * Female2018 + \varepsilon_i \quad (5.7)$$

For the outcome of interest "number of classes that were not passed" I estimate the equations following the previous pattern. Considering that the students in the 2018 cohort

experienced the covid-19 pandemic for one more year of their studies, I assume the outcome of interest "Graduation within three years" cannot be estimated between the 2017 and 2018 cohort as other factors, like the change to online learning, probably affected the students likelihood of graduating within 3 years.

6 Results

Based on the existing literature on the gender performance gap I hypothesized that the mentoring program might have a different impact for male and female students. I therefore start by testing if a gender performance gap exists at Tilburg University.

In the 2017 cohort 41 courses were recorded. The grades are given on a scale from 0-10 with the minimum score to pass being 6. The average grade in the 2017 cohort is 6.61 out of 10. The average grade of female students in the 2017 cohort is 6.87, while the average grade of their male counterparts is 6.43 as displayed in table 6.1. As the p-value is very small, I reject the hypothesis that the two means are equal.

Table 6.1: Two-sample t-test with equal variances 2017

| Gender | Average Grade | Std. Error |
|--------------------|------------------------|---------------------------|
| Male | 6.431 | 0.022 |
| Female | 6.871 | 0.023 |
| combined | 6.611 | 0.016 |
| | | t = -13.2647 |
| | | degrees of freedom = 8906 |
| Ha: diff < 0 | Ha: diff != 0 | Ha: diff > 0 |
| Pr(T < t) = 0.0000 | Pr(T > t) = 0.0000 | Pr(T > t) = 1.0000 |

In 2018 318 students from the "BSc International Business and Administration" degree were registered. Their average grade over 34 recorded courses is 6.45 out of 10. The

average grade of female students in the 2018 IBA cohort is 6.62 and the average grade of male students is 6.34 as displayed in table 6.2. As the p-value is very small, I reject the hypothesis that the two means are equal. I thus find evidence for a gender performance gap at university in both cohorts.

Table 6.2: Two-sample t-test with equal variances 2018 IBA

| Gender | Average Grade | Std. Error |
|--------------------|------------------------|---------------------------|
| Male | 6.336 | 0.021 |
| Female | 6.624 | 0.025 |
| combined | 6.448 | 0.016 |
| | | $t = -8.448$ |
| | | degrees of freedom = 9567 |
| Ha: diff < 0 | Ha: diff != 0 | Ha: diff > 0 |
| Pr(T < t) = 0.0000 | Pr(T > t) = 0.0000 | Pr(T > t) = 1.0000 |

The impact of mentoring on grades

Overall, 118 students subscribed to the mentoring program in 2017. 71 from them received the additional encouragement, 47 did not. To measure the impact of the program in the short and the long-run, I estimate the effect of the mentoring on the average grade of the students in each semester. In neither of the five semesters the impact of the mentoring is statistically significant different from zero. The results are displayed in table 6.3.

Table 6.3: The effect of the mentoring on average grades

| Semester | control group average | treatment group average | p-value |
|----------|-----------------------|-------------------------|---------|
| 1 | 5.457 | 6.762 | 0.116 |
| 2 | 6.488 | 6.449 | 0.998 |
| 3 | 4.807 | 6.017 | 0.256 |
| 4 | 5.703 | 6.619 | 0.248 |
| 5 | 3.299 | 6.452 | 0.728 |
| overall | 5.140 | 6.410 | 0.318 |

This table is based on the results from regression 5.2.

The impact of mentoring on other outcome variables

As table 6.4 shows, the mentoring also did not have a statistically significant effect on the number of exams that were failed throughout the studies, nor on the number of students who finished their bachelor's within 3 years. The results confirm the short term analysis by Van Soest et al. (2018).

Table 6.4: The effect of the mentoring on graduation and failing

| | constant | subscribed | p-value |
|------------------------|----------|------------|---------|
| Graduation in 3 years | 0.456 | 0.053 | 0.839 |
| Amount of Failed Exams | 8.160 | -0.946 | 0.798 |

This table is based on the results from regression 5.3.

Heterogeneity in mentoring effects

The results of the previous analysis seem to provide no evidence for the effectiveness of the mentoring program as displayed in table 6.5. Examining these results additionally for heterogeneous effects on gender, I find no evidence that the mentoring has different effects on male and female students. The results thus do not indicate that the treatment might be effective in closing the gender performance gap at university. However, I observe the gender performance gap with statistical significance, as women perform better in all analysed categories on at least 10 percent significance level.

2017 and 2018 comparison

I analyse my second hypothesis, which states that students at the margin of failing classes benefit most from the mentoring, by comparing the results from the 2017 mentored students with the 2018 mentored students. As the program was compulsory in the latter, I will measure the effect of those students that were "not at all interested" in participating in

Table 6.5: Including gender in the analysis

| Overall Grade | coefficient | St. Error | p-value |
|--------------------|-------------|-----------|---------|
| Subscribed | 0.765 | 1.508 | 0.612 |
| Female | 0.624 | 0.338 | 0.065 |
| Treated*female | 0.318 | 0.519 | 0.540 |
| Constant | 5.036 | 0.573 | 0.000 |
| Failed Classes | coefficient | St. Error | p-value |
| Subscribed | -0.636 | 4.544 | 0.889 |
| Female | -3.004 | 1.040 | 0.004 |
| Treated*female | -0.310 | 1.467 | 0.833 |
| Constant | 9.302 | 1.624 | 0.000 |
| Graduation 3 Years | coefficient | St. Error | p-value |
| Subscribed | -0.011 | 0.301 | 0.970 |
| Female | 0.219 | 0.076 | 0.004 |
| Treated*female | 0.033 | 0.110 | 0.763 |
| Constant | 0.388 | 0.112 | 0.001 |

This table is based on the results of the regressions 5.4 and 5.5.

the program, not even with additional encouragement. Condition for making assumptions about the effect of mentoring on students at risk of failing is thus, that these are the students that are "not at all interested" in participating in the mentoring program.

I tested this for the 2017 cohort by testing the hypothesis that the students who subscribed to the program and the students who did not subscribe to the program entered with the same application ranking. The university evaluates all prerequisites of students that applied for the bachelors and ranks them according to their achievements. Students with lower grades in highschool, lower test scores in the english test etc. have a higher ranking number. The average ranking numbers are displayed in table 6.6. As the p-value is very small, I reject the hypothesis that the two means are equal. I thus find evidence that the students who do not subscribe to the mentoring program have weaker starting

conditions. It is reasonable to assume that the students that showed a lower performance in highschool and thus were at the margin of getting accepted into the undergraduate degree will also struggle more at university and be at risk of failing. In the 2018 cohort, these students were forced to enter the mentoring program. My comparison of the 2017 and 2018 effects of the mentoring program thus allows me to make inferences on the effect of mentoring on the students at risk of failing.

Table 6.6: Two-sample t-test with equal variances 2017 subscribed / not subscribed

| | Average Ranking Number | Std. Error |
|--------------------|------------------------|--------------------------|
| Subscribed | 267.745 | 14.772 |
| Not Subscribed | 334.1 | 12.402 |
| combined | 306.492 | 9.685 |
| | | t = 3.443 |
| | | degrees of freedom = 272 |
| Ha: diff < 0 | Ha: diff != 0 | Ha: diff > 0 |
| Pr(T < t) = 0.9997 | Pr(T > t) = 0.0007 | Pr(T > t) = 0.0003 |

The estimation of the change of the program from being voluntary to compulsory, thus estimating the effect of mentoring on the students who were "not at all interested" in attending the program, is displayed in table 6.7. Making the program mandatory does not have a statistically significant effect on the average grade of the students, but does have a statistically significant effect at 10 percent level for the number of classes failed by the students. On average, the students who only participated in the program because they were forced to, failed around 2 classes fewer than their peers in 2017 who were not mentored.

Table 6.7: The effect of the change from voluntary to mandatory participation

| Overall Grade | coefficient | St. Error | p-value |
|----------------|-------------|-----------|---------|
| 2018cohort | -0.001 | 0.270 | 0.997 |
| Female | 0.164 | 0.328 | 0.617 |
| Female2018 | 0.393 | 0.376 | 0.297 |
| Constant | 5.558 | 0.325 | 0.000 |
| Failed Classes | coefficient | St. Error | p-value |
| 2018cohort | -2.159 | 0.951 | 0.024 |
| Female | -0.637 | 1.252 | 0.611 |
| Female2018 | 0.491 | 1.442 | 0.734 |
| Constant | 9.105 | 1.075 | 0.000 |

The results in this table are based on the results of the regression 5.6 and 5.7. I added nationality as a control.

7 Discussion

This research contributes to the literature about the effects of mentoring. The analysis of the randomized experiment in 2017 fails to provide proof that mentoring has a beneficial effect on the students overall or that it has a different impact on male and female students. As the study is underpowered, I cannot conclude if the lack of statistically significant results stems from an absence of effect or from limited power. It also has to be noted that these results only account for the "moderately interested students" and measure the intention-to-treat. As noncomplying students are also analysed as being treated, the results are likely to be underestimating the effect of the mentoring. Data on the actual participation of the students throughout the mentoring phase would be needed to have greater interpretation power. More research in this field is needed to find if mentoring benefits female and male students to a different extent.

When analysing the change of the mentoring design from voluntary to mandatory, I find

no statistically significant effects on average grades. I do find statistically significant effects however on the number of failed classes at the five percent level. In the first three semesters of their studies the students that only participated in the mentoring in 2018 because they were being forced to failed on average two classes fewer. These "not at all interested" students start out with a worse average application score than their peers that signed up to the mentoring on a voluntary basis. The 2017/2018 comparison includes data that was recorded outside of the randomized control trial and thus issues of self-selection possibly play a role. Unobservable characteristics might have influenced students in the decision to self-select to the mentoring in 2017. In light of the results, the self selection might have negatively affected the outcome: If the return on these unobservable characteristics is lower for the students who decided to be mentored, a bias was introduced in the estimation. I was able to show for example, that the students who did sign up are starting university with a better application ranking than students who decide not to sign up. They might have a higher intrinsic motivation for studying and thus do not need the mentoring as much as the others. For these students there might be less scope for the mentors to make a difference, and thus the analysis would underestimate the effect of mentoring.

The results indicate however that the mandatory mentoring program at Tilburg University has the chance to improve the university's retention rate by reducing the number of classes that are being failed. As the overall grade averages do not improve statistically significant, it seems like the program is working specifically at the margin of failing or passing a class. The students who benefit are those who did not want to take part in the program on a voluntary basis. The reasons of this are unclear. Possible explanations include that these students might be less motivated to achieve a good academic performance in the first place and when being forced to write down their goals, develop new motivation. Another

explanation is that these students are on average less focused on their future and thus less concerned with their university performance. The mentoring program with its main focus on the personal development plan where students had to write down their goals, strengths and weaknesses might have introduced the students to goal setting and created intrinsic motivation to overcome the challenges. The students might have been motivated not to give up and thus fail fewer courses. On the other hand, a possible explanation might be that these students are not aware that they are the ones at risk of failing and thus think they do not need the mentoring as well as that they do not need to increase their studying commitment. The mentoring through their older peers and the information these mentors provide might help them to realise otherwise. The program might also help the students to structure their study activities and thus be more productive. Further research is needed to explain this mechanism and examine if there are underlying characteristics that have influence on this effect. For future research, data on the mentors' perspectives would provide a more complete picture of the mentoring relationship and its outcomes.

There were certain limitations to this study. The randomization of the experiment in the cohort 2017 ensures internal validity. This guarantees that subsequent changes in the outcomes measured are due to the program. But when I compare the 2018 cohort with the 2017 sample that subscribed to the mentoring, this no longer holds true: The students in 2018 were automatically subscribed to the mentoring, while the 2017 sample selected themselves into the sample. The estimation is thus potentially biased through self-selection. Furthermore it would be interesting to have data on the actual commitment of the students in the mentoring program. Another limit of the research is the lack of monitoring of the quality of the mentor-mentee relationship. The quality might not have been consistent and thus have impacted the results.

Finally, the experiment has only limited interpretation power outside of the specific university context as this study was conducted only at a single academic institution. As the sample consists only of bachelor students, this sample is not representative of the population which weakens the study's external validity. Even within the university context external validity is not ensured as the students were all from the same degree. As shown in table A2.1 in the Appendix, the students of the IBA degree differ to the students of other degrees in nationality and gender composition on a statistically significant level of one percent. Additional studies are needed to confirm this study's findings.

8 Conclusion

This study examined if the IBA Mentoring program had a positive effect on undergraduate students at Tilburg University and if there are certain groups of students that benefit more than others. While I could not find statistically significant beneficial effects on average grades and graduation time overall, I could also not confirm that the mentoring had a different impact on female and male students. As the power of this study was low, I cannot draw conclusions from these findings.

The randomized control trial at Tilburg University however provided evidence that mentoring is most beneficial for those students that did not initially want to take part in the mentoring, even with additional encouragement. Even though these students grade average does not improve on a statistically significant level, the number of classes that are failed decreases. These students already started with a worse application ranking than the students that were at least moderately interested in the program.

Mentoring might thus be a promising tool to improve university's retention rates. Further

research is needed to confirm the findings. As mentoring seems to have beneficial effects only on a very small number of students, the cost-effectiveness of mentoring should also be further evaluated. There might be different approaches targeted specifically at the students at risk of failing a class that might be less costly.

Implications of this result could be that universities should focus their mentoring activities towards students at risk, for example students that enter university without a good academic background. Finding the threshold might be very difficult however, as well as the ethical argumentation behind providing or even forcing mentoring based on application quality. A possible explanation for universities could be to introduce probation periods, as it is being practiced in some universities in the US for example. In this probation period students who show a weak academic performance are forced to attend a mentoring program. In light of the discussed findings, this might be a targeted tool to reach the ones that benefit from the approach. This way the mentoring might also be more cost-efficient compared to a full-mandatory mentoring approach.

References

- Bettinger, E. P. and Baker, R. B. (2014). The effects of student coaching: An evaluation of a randomized experiment in student advising. *Educational Evaluation and Policy Analysis*, 36(1):3–19.
- Breda, T., Grenet, J., Monnet, M., and Van Effenterre, C. (2018). *Can Female Role Models Reduce the Gender Gap in Science?: Evidence from Classroom Interventions in French High Schools*. Paris-Jourdan Sciences Economiques.
- Byrnes, J. P., Hong, L., and Xing, S. (1997). Gender differences on the math subtest of the scholastic aptitude test may be culture-specific. *Educational Studies in Mathematics*, 34(1):49–66.
- Clark, D., Gill, D., Prowse, V., and Rush, M. (2020). Using goals to motivate college students: Theory and evidence from field experiments. *Review of Economics and Statistics*, 102(4):648–663.
- Cobb-Clark, D. A. (2015). Locus of control and the labor market. *IZA Journal of Labor Economics*, 4(1):1–19.
- Dayioğlu, M. and Türüt-Aşık, S. (2007). Gender differences in academic performance in a large public university in turkey. *Higher Education*, 53(2):255–277.
- Duckworth, A. L., Tsukayama, E., and May, H. (2010). Establishing causality using longitudinal hierarchical linear modeling: An illustration predicting achievement from self-control. *Social psychological and personality science*, 1(4):311–317.
- Eby, L. T., Allen, T. D., Evans, S. C., Ng, T., and DuBois, D. L. (2008). Does mentoring matter? a multidisciplinary meta-analysis comparing mentored and non-mentored individuals. *Journal of vocational behavior*, 72(2):254–267.
- Efstathiou, J. A., Drumm, M. R., Paly, J. P., Lawton, D. M., O’Neill, R. M., Niemierko, A., Leffert, L. R., Loeffler, J. S., and Shih, H. A. (2018). Long-term impact of a faculty mentoring program in academic medicine. *PloS one*, 13(11).
- Fiala, L. (2020). Statistical role models. working paper.
- Fozdar, B. I., Kumar, L. S., and Kannan, S. (2006). A survey of a study on the reasons responsible for student dropout from the bachelor of science programme at indira gandhi national open university. *International Review of Research in Open and Distributed Learning*, 7(3):1–15.
- Gladstone, J. and Cimpian, A. (2020). Role models can help make the mathematics classroom more inclusive. *OSF Preprints*.
- Hanson, M., Engström, E., Kairamo, A., and Varano, M. (2010). Enhance the attractiveness of studies in science and technology. In *Joint International IGIP-SEFI Annual Conference, IGIP-SEFI, Trnava, Slovakia*.
- Hsiaw, A. (2013). Goal-setting and self-control. *Journal of Economic Theory*, 148(2):601–626.
- Judge, T. A. and Bono, J. E. (2001). Relationship of core self-evaluations traits—self-esteem, generalized self-efficacy, locus of control, and emotional stability—with job

- satisfaction and job performance: A meta-analysis. *Journal of applied Psychology*, 86(1):80.
- Khwaileh, F. M. and Zaza, H. I. (2011). Gender differences in academic performance among undergraduates at the university of Jordan: Are they real or stereotyping. *College Student Journal*, 45(3):633–648.
- Latham, G. P. and Locke, E. A. (1979). Goal setting—a motivational technique that works. *Organizational dynamics*, 8(2):68–80.
- Lent, M. and Souverijn, M. (2016). Goal setting and raising the bar: A field experiment. tinbergen institute discussion paper 2017-001/vii.
- Letseka, M. and Maile, S. (2008). *High university drop-out rates: A threat to South Africa's future*. Human Sciences Research Council Pretoria.
- Logan, S. and Johnston, R. (2009). Gender differences in reading ability and attitudes: Examining where these differences lie. *Journal of research in reading*, 32(2):199–214.
- Mayer, A. P., Blair, J. E., Ko, M. G., Patel, S. I., and Files, J. A. (2014). Long-term follow-up of a facilitated peer mentoring program. *Medical Teacher*, 36(3):260–266.
- McNabb, R., Pal, S., and Sloane, P. (2002). Gender differences in educational attainment: The case of university students in England and Wales. *Economica*, 69(275):481–503.
- Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., Houts, R., Poulton, R., Roberts, B. W., Ross, S., et al. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the national Academy of Sciences*, 108(7):2693–2698.
- Ng, T. W., Sorensen, K. L., and Eby, L. T. (2006). Locus of control at work: a meta-analysis. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 27(8):1057–1087.
- OECD (2020). Education at a glance 2020: Gender differences in university graduates. OECD Publishing. Paris
- Reeves, K. E. (1996). *Formal mentoring for college freshmen: Who participates and what happens*. Northern Illinois University.
- Rhodes, J. E., Grossman, J. B., and Resch, N. L. (2000). Agents of change: Pathways through which mentoring relationships influence adolescents' academic adjustment. *Child development*, 71(6):1662–1671.
- Rodríguez-Planas, N. (2012). Mentoring, educational services, and incentives to learn: What do we know about them? *Evaluation and program planning*, 35(4):481–490.
- Salinitri, G. (2005). The effects of formal mentoring on the retention rates for first-year, low achieving students. *Canadian Journal of Education/Revue canadienne de l'éducation*, pages 853–873.
- Schertzer, C. B. and Schertzer, S. M. (2004). Student satisfaction and retention: A conceptual model. *Journal of Marketing for Higher Education*, 14(1):79–91.
- Spangenberg, H., Ohlendorf, D., & Föste-Eggers, D. (2020). Studienberechtigte 2012 – Qualifizierungswege und Bildungserträge. (DZHW Brief 6|2020). Hannover: DZHW.

- U.S. Department of Education (2020). National Center for Education Statistics. Fall Enrollment component; and Fall 2018, Institutional Characteristics component. Digest of Education Statistics 2020, table 326.30.
- Van Soest, D., Dalton, P., Fiala, L., and Van der Heijden, E. (2018). The impact of iba's personal development program on students' study performance and personal development.
- Webster, A. L. and Showers, V. E. (2011). Measuring predictors of student retention rates. *American Journal of Economics and Business Administration*, 3(2):301–311.
- Wilberg, S. and Lynn, R. (1999). Sex differences in historical knowledge and school grades: A 26 nation study. *Personality and Individual Differences*, 27(6):1221–1229.
- Young, J. W. and Fisler, J. L. (2000). Sex differences on the sat: An analysis of demographic and educational variables. *Research in Higher Education*, 41(3):401–416.
- Zhou, E. and Gao, J. (2021) International Graduate Applications & Enrollment. Council of Graduate Schools.

Appendix

A1 Balance table 2017 cohort control and treatment

Table A1.1: Balance table 2017 cohort control and treatment group

| Characteristic | Control Group | Treatment Group | t-test p-value |
|---|---------------------|---------------------|----------------|
| Dutch | 0.837 (0.030) | 0.850 (0.030) | 0.766 |
| Average evaluation score of IBA application package | 13.542 (0.202) | 13.762 (0.213) | 0.455 |
| Ranking of application evaluation score | 313.439 (13.352) | 299.127 (14.087) | 0.461 |
| Application date | 20846.52 (2.504) | 20841.67 2.690 | 0.1881 |
| Female | 0.405 (0.040) | 0.388 (0.042) | 0.767 |
| Average age (in years) | 18.536 (0.131) | 18.541 (0.104) | 0.1581 |

The table shows that the control and treatment group in the 2017 cohort are similar in their observable characteristics. The "average evaluation score of IBA application package" is the score that the students got based on the information they provided in their application. This includes for example the high school certificate, the language tests and cover letters. The better the score, the better the application was evaluated.

A2 Balance table 2018 IBA and other degrees

The table shows that the students from the 2018 International Business Administration degree and the students from other degrees are different in nationality and gender composition on a statistically significant level of one percent.

Table A2.1: Balance table 2018 IBA and other degrees

| Characteristic | IBA | other degrees | t-test p-value |
|---------------------------|------------------|------------------|----------------|
| Dutch | 0.736 (0.024) | 0.908 (0.009) | 0.000 |
| Female | 0.376 (0.027) | 0.249 (0.014) | 0.000 |
| Average age (in years) | 19.01 (0.069) | 19.11 (0.050) | 0.285 |