# Program Specific Effects of a Semester Abroad on the Likelihood of Pursuing a PhD

Laura Urgelles[1]

[1] Department of Management, University of Paderborn, Paderborn, Germany

Correspondence: Laura Urgelles, Department of Management, University of Paderborn, Warburger Str. 100, 33098 Paderborn, Germany. Tel: 49-151-1799-3138. E-mail: urgelles.laura@gmail.com

## Abstract

The present paper provides an analysis of the impact of a semester abroad during university studies on a students' likelihood of pursuing a PhD. I use a sample of 66 812 German university students and analyze the program-specific subsamples. Propensity score matching reveals that business students who go abroad during their studies have higher intentions to pursue a PhD than their non-mobile peers. The findings are robust across matching estimators. In addition, I find positive and significant effects for cultural and social studies, whereas the effects for medical and law students are insignificant. When splitting the sample at the median grade, a semester abroad has a significant positive impact on below-median grade natural sciences students' PhD decision. In contrast, for engineering students there is a positive and significant effect of a semester abroad only for above-median performers. I build on existing findings concerning the existence of a positive correlation between mobility and the intention to pursue a PhD.

**Keywords:** academic career, academic performance, doctoral studies, PhD, propensity score matching, student mobility, university major

## 1. Introduction

The various benefits of student mobility are mostly taken for granted, while little research exists to empirically prove the actual outcomes of studying abroad. Political parties in Germany have supported internationalization at universities and have explicitly aimed to increase the mobility of students (Coalition Contract, 2013). As a result, both the German government and the European Union subsidize programs for internationalization. In fact, in 2015, 42 680 German mobile students received support through EU grants (e.g., ERASMUS) and 59 310 German students received grants from the German Academic Exchange Service (DAAD, 2017). For the next period (2021-2027), the European Commission has increased its grants to 26.2 billion Euros, which is almost twice as much as during the previous period (European Commission, 2021). However, the outcomes of a semester abroad have not been analyzed sufficiently yet.

In line with existing studies addressing the outcomes of student mobility (e.g., Messer & Wolter, 2007; Petzold, 2020; Van Mol, Caarls, & Souto-Otero, 2020), the present study is based on human capital theory (Becker, 1964). According to human capital theory, utility maximizing individuals invest into their skills with the aim to increase their human capital. An increase in human capital leads to higher productivity and is rewarded with an increase in salary (Becker, 1964). Thus, education constitutes a human capital investment. In addition, the knowledge acquired outside of school as a form of human capital investment has received recognition (e.g., Heckman, 2000). During a semester abroad a student acquires both, the education during the regular course work at the host institution, and several social skills acquired by adapting to an unfamiliar situation. Thus, mobile students may be equipped with a skill set which is desirable in certain positions, giving them an advantage when competing against their non-mobile peers on the labor market. In fact, several empirical studies find a salary advantage (e.g., Favero & Fucci, 2017; Iriondo, 2020; Messer & Wolter, 2007) and a higher likelihood to fill leadership positions (Euler, Rami, & Glaser, 2013) for formerly mobile graduates.

While mobility may be an investment in human capital skills it may also constitute a signal to future employers (Spence, 1973). According to Spence's (1973) job market signaling theory, the acquisition of knowledge does not necessarily lead to a higher level of productivity but is a signal of productivity to future employers. Signals are most effective if they stand for something that is hard to be measured. Therefore, when trying to prove their

flexibility, intercultural interest, or their degree of perseverance to future employers, students may rely on their semester abroad as a credible signal for these hard to be measured social skills (Van Mol et al., 2020; Petzold, 2017). Especially when competing for doctoral research assistant positions, human capital skills acquired during a semester abroad may be relevant and demanded skills. Mobility may therefore be an investment in human capital and a powerful signal particularly to future employers in academia.

Existing research on the relation between mobility and the intention to pursue a PhD is scarce. One of the few studies, addressing this relation was conducted by Messer and Wolter (2007). The authors conduct a probit and instrumental variable (IV) regression to analyze the effect of mobility on the intention to write a dissertation. Their instrument is whether a student resides in the Swiss canton where the university is located (as opposed to having to move to another area). The probit regression shows a positive correlation between mobility and the intention to pursue a PhD. The IV regression, however, does not yield significant effects and thus does not support the hypothesis of causality. The authors conclude that mobile graduates' advantages in the labor market and the academic career are selection effects and 'simply attributable to the better capabilities of these graduates and not to the fact that they have studied in an exchange program' (Messer & Wolter 2007, p.661). The present study builds on Messer and Wolter's (2007) findings by analyzing the causal effect of mobility and the intention to pursue a PhD.

Only few studies consider the university programs when analyzing effects of mobility. In terms of salary, however, positive effects of mobility have been shown for social sciences and engineering students but not for education, arts and humanities, mathematics, or medicine graduates (Rodrigues, 2013; Kratz & Netz, 2016). In addition, writing a dissertation may be more common in certain fields (e.g., natural sciences, medicine) than in others (e.g., business, cultural studies). I therefore add to the scarce knowledge on program-specific mobility effects on individual career decisions.

## 2. Method

### 2.1 The Student Survey by the University of Konstanz

The subsequent analysis is based on the Student Survey ('Studierendensurvey') conducted by the University of Konstanz. The survey has been conducted at 25 German universities and colleges of applied sciences, starting in 1982. Important topics of the survey are, among others, study motives and student expectations; study intensity and duration; requirements and examinations, contacts; social climate and consultancy; as well as study strategies, paths and qualifications. The survey also includes a question on experience abroad (e.g., in the form of studying abroad, internships, or language courses). The present paper focuses specifically on the study abroad experience.

Although the data initially consisted of 100 420 observations, adjustments had to be made due to missing values or nonsensical answers. For example, individuals who were under the age of 18 and older than 45 were excluded. In addition, I excluded individuals who indicated they had studied for more than 30 semesters. The final data set contains 66 812 observations.

### 2.2 Descriptive Statistics

Since the start of the survey during the winter semester 1982/83 the number of mobile students has increased. While only about 4.3 % of students reported being mobile on the 1982/83 survey, the number increased and was as high as 30.1 % in 2006/07. In the most recent wave (2012/13) this number decreased to about 21.8 %. Overall, the data set consists of 18.8% mobile students.
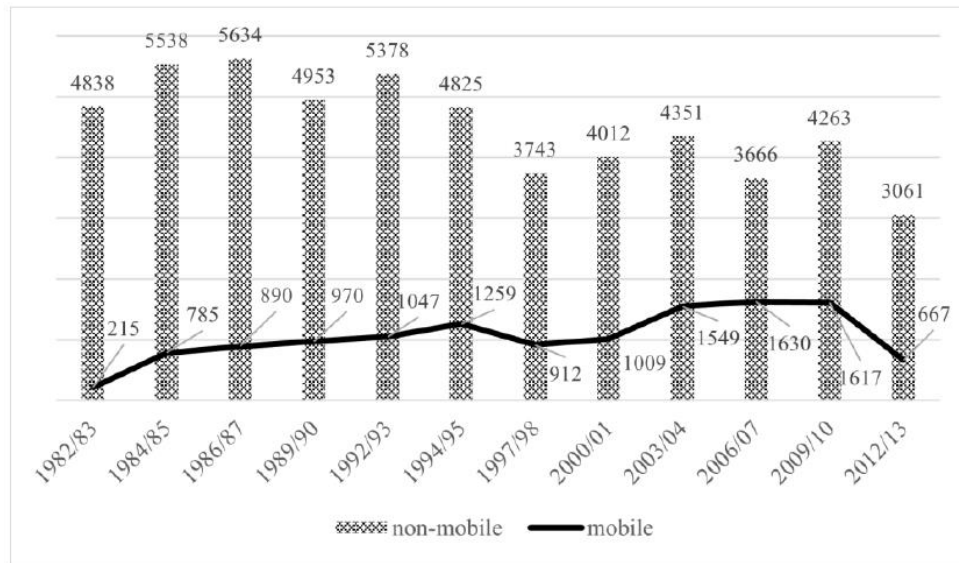
Figure 1. Mobile and non-mobile students across waves in the Student Survey

Most mobile students are enrolled in cultural studies programs. Business students make the second largest group of mobile students. Fewer mobile students study law, natural sciences or social sciences (e.g., pedagogy, sociology, political sciences or journalism). This partly contradicts Chieffo and Griffiths' (2004) findings that natural science students are more likely to go abroad than those studying the humanities. However, it also partly supports their findings because natural sciences make a larger group of mobile students than social sciences students.

The descriptive statistics show that the median grades vary considerably between the different programs. Cultural and social studies students receive the best grades, while law and business students receive the worst grades regardless of whether they went abroad or not. The grades are given on a scale from 1 through 5, where 5 is the best grade and 1 represents a failure. Mobile students are on average almost one year (0.82 years) older than non-mobile students. A t-test reveals that significantly more female than male students are mobile (p<0.000). In both, the control and treatment groups, more students are in a relationship than are single and the minority is married. In terms of their previous education, mobile students on average had better grades in high school than their non-mobile peers. Treatment students indicate to have studied 1.4 semesters longer than control students and university students are over represented compared to students from colleges of applied sciences, in both the control (non-mobile) and the treatment (mobile) groups. The difference is again significant (p<0.000).

Mobile students on average are more active in all leisure activities apart from the participation in religious groups. The majority of both groups finances their studies mainly with their parents' help. However, mobile students rely significantly more often on their parents' financial help than non-mobile students (p<0.000). A total of 8.4 % of the mobile students state that they mainly finance their studies with federal grants. Significantly more non-mobile students (14.5 %) receive federal grants (p<0.000). Though parental background may be a predictor for mobility (Waibel, Petzold, & Rüger, 2018), I do not include it because the question was not part of the survey in the initial wave and would thus reduce the sample by the 1982/83 subsample.

*2.3 Empirical Approach*

To answer my research question, I apply propensity score matching (PSM). The aim is to determine how the outcome for an individual would be with and without a treatment. In my case the outcome is a student's intention to pursue a PhD, the treatment condition is mobility and the control condition is non-mobility. Researchers have argued that treatment and non-treatment individuals usually differ in terms of other covariates even in the absence of a treatment. The covariates for the present treatment, were the previously mentioned predictors (e.g., age, gender, university program, etc.). Program evaluations address this challenge and aim to overcome the selection bias (Caliendo & Kopeinig, 2005). Due to the fact that conditioning on all relevant covariates is hardly possible, Rosenbaum and Rubin (1983) suggest balancing scores. Balancing scores have been defined as 'functions of the relevant observed covariates X such that the conditional distribution of X given b(X) is independent of the assignment into treatment' (Caliendo & Kopeinig, 2005, p. 1). Rosenbaum and Rubin (1983)

introduced propensity score as a balancing score and a means to approximate treatment effects and reduce bias when using observational data sets as it is the case with the Student Survey data.

The most important requirement when conducting PSM is to abide by the Conditional Independence Assumption (CIA). Abiding by CIA ensures the causal interpretations of a matching. The CIA states that differences in the outcomes between the treatment and control group are attributable entirely to the treatment. This means that variables that influence the participation decision as well as the outcome variable should not be included in the calculation because they might confound results. Not complying with CIA may lead to the estimation of non-robust treatment effects (Heckman, Ichimura, & Todd, 1997). Though literature hints on the grade being a predictor for mobility, I do not include it as a predictor in my analysis. Its inclusion would be against CIA as the grade is a variable that could be an outcome itself, making it a 'bad control'. Good controls on the other hand are variables that are constant over time (e.g., gender), exogenous variables (e.g., age), or variables determined before the treatment (e.g., high school grade). In order to control the differences in students' average grades, I conduct a grade-based median split prior to the analysis. However, since grades also vary across programs, I conduct a median split for the different subsamples based on the programs.

For the sake of brevity, I will report the analysis for business students in the following subsections and only briefly report the results of the PSM for the other programs. Business students are a particularly interesting sample because they are the second largest group of mobile students (cultural studies students are the largest group), yet not as many programs require a semester abroad whereas more cultural study programs include a compulsory semester abroad.

*2.4 Estimation of the Propensity Score*

The first step in this analysis is to calculate the propensity scores based on the covariates. By means of a logit or probit regression the binary outcome variable (treatment vs. control) can be calculated. Before using the propensity scores for the matching, I exclude values outside the region of common support. The assumption of common support states that regarding the covariates, units in both the treatment and control groups should be similar. Thus, by excluding those units outside the average maximum and minimum propensity score values I trim the distribution and avoid bad matches (i.e., of different propensity scores). 62 observations with the values outside the common support region have to be excluded.

Probit and logit regressions are equally suitable when calculating the propensity score. I take Euler and colleagues' (2013) paper as an example and report the logit regression results (see Appendix A for the complete regression results). I find that female business students are significantly more likely to go abroad than their male peers. As students get older their probability of going abroad decreases. Having children also decreases mobility and single students are more likely to go abroad than married students. In terms of previous education, high school grades are an indicator for a semester abroad. In other words, the better the high school grade the more likely a student was to go abroad. Further, students in colleges of applied sciences are more likely to be mobile than university students.

In line with the findings for age, the likelihood of spending a semester abroad decreases the longer a student has been enrolled. With respect to leisure activities, both highly and lowly active athletes tend to be less mobile than moderately active students. Politically active students also tend to be more likely to go abroad. Unlike in the athletes' case, however, there was not a quadratic function for the relationship.

Students who receive money from their parents or partly finance their studies through a scholarship tend to be more mobile. In line with existing findings (Salisbury, Paulsen, & Pascarella, 2010), students who receive federal grants are less likely to be mobile. However, very few students in the sample indicate that they receive a scholarship and most students finance their studies partly or mainly through their parents.

*2.5 Matching the Treatment and Control Units*

There are various matching methods which essentially differ based on the matching ratio (e.g., 1:1; 1:5, 1:k) and the tolerance concerning the size of the difference in propensity scores that are matched. In other words, when choosing a matching method, the trade-off lies between variance and bias. Among other suitable methods when analyzing cross-sectional data Todd (2006) names nearest neighbor matching and kernel matching. I conduct several different variations of nearest neighbor matchings as well as kernel matchings to make sure that results are similar across matching estimators.

Nearest neighbor matching is a matching method in which every treatment unit is matched to a non-treatment unit with the closest propensity score. The 1:1 nearest neighbor matching with replacement reveals the lowest average treatment effect (ATT). Matching with replacement means that propensity scores from the control group

can be matched more than once to a treated case. The difference between the means of the treatment and control group is positive and significant in the matched and unmatched case (t-value unmatched: 6.79>|1.96|; t-value matched: 4.91>|1.96|). The positive value for the difference (Δ=0.138) reveals that mobile students are more likely to aim at an academic career than non-mobile students. The difference is larger for the unmatched case which means that I would have overestimated the effect of a semester abroad on the intention to pursue a PhD had I not matched the individuals. I calculate a 1:1 nearest neighbor matching without replacement as well as a 1:5 ratio matching. The average treatment effect remains positive and significant, and changes only slightly in size.

In addition to matching the closest treatment and non-treatment propensity scores, caliper matching creates a tolerance zone. The propensity scores of the non-treatment units have to be within this tolerance zone to be matched to a treatment unit. While in caliper matchings bad matches are less likely to occur, certain treatment units may not have a suitable match (i.e., a propensity score within the tolerance zone) at all and are excluded. I choose a relatively narrow caliper of 0.001 and support my earlier findings of a positive significant average treatment effect (Δ=0.140). This difference is slightly higher than in the 1:1 nearest neighbor matchings with replacement.

Kernel matching uses the weighted means of almost all individuals in a control group in order to calculate the counterfactual outcome. The advantage is the high amount of information being used which decreases variance. A disadvantage is that so-called bad matches (i.e., observations with relatively different propensity scores) are still being matched which may cause an increase in bias. The exclusion of propensity scores outside the common support region is therefore even more important for kernel matching than for nearest neighbor matching. I calculate kernel matchings with two different bandwidths and, again, support my findings of a positive and significant average treatment effect on the treated.

To sum it up, although the ATT sizes vary slightly in the various estimation methods, this difference is relatively small. Across matchings, the difference is positive and significant. Table 1 depicts an overview of the main results for the business students subsample.

Table 1. Overview of PSM results for using different matching estimators (business students subsample)

| Matching estimator | ATT, size of the difference (mobile vs. non-mobile) | Significance of the difference (ATT) |
|---|---|---|
| Nearest neighbor (1:1, with replacement) | 0.138 | 4.91>|1.96| |
| Nearest neighbor (1:1, without replacement) | 0.158 | 6.60>|1.96| |
| Nearest neighbor (1:5, with replacement) | 0.156 | 6.79>|1.96| |
| Caliper matching [caliper (.001)] | 0.140 | 6.44>|1.96| |
| Kernel matching (bw .03) | 0.140 | 6.37>|1.96| |
| Kernel matching (bw .01) | 0.144 | 6.50>|1.96| |

*Note.* The business subsample includes n=9 826 observations. In the kernel matching (bw .01) 4 treatment units outside the common support region are excluded.

### 2.6 Quality of the Matching

The last step in the PSM process consists of the evaluation of the matching quality. There are different ways to ensure that the matching was efficient. For example, percentage reduction in bias (PRB) shows the relative reduction of the differences in covariates among the groups. In addition, a sensitivity analysis is useful when trying to estimate the robustness of the treatment effect by taking into account unobserved confounding factors. The sensitivity analysis is an instrument that enables researchers to evaluate estimated treatment effects in potential error scenarios. In other words, the sensitivity analysis creates scenarios in which unobserved covariates create different amounts of bias. This way researchers can test the robustness of the estimated treatment effects considering the confounding factors which may occur due to selection bias (Caliendo & Kopeinig, 2005).

While the propensity scores for the control group are mostly between 0.1 and 0.3, the propensity scores for the treatment group are higher. In fact, the majority is between 0.2 and 0.45. This shows that the individuals matched differ slightly on the covariates. In order to evaluate the quality of the matching statistically, I calculate the PRB as suggested by Rosenbaum and Rubin (1984). The aim is to achieve a small difference between the predictors, because a small PRB reflects a good matching quality. In my case the mean bias after matching (in the 1:1 nearest neighbor matching with replacement) is 2.3 % on average, which is below the threshold of 5 % and thus

satisfactory. In addition, none of the predictors separately exceeds this threshold.

The sensitivity analysis by means of 'Rosenbaum bounds' serves to analyze the robustness of the treatment effects. In case of my analysis, I find a strong positive treatment effect. It is possible that the positive effect is overestimated (e.g., due to confounding factors). At a value of $\Gamma < 1$ the Rosenbaum bounds calculation assumes there were no selection and confounding processes. In this case the estimated treatment effect would be significant. In addition, I find that until $\Gamma \leq 1.1$ the treatment effects are significant at a level of 5 %. In other words, even if a certain confounding factor would influence the chances of a person being in the treatment group at a 1:1.1 ratio, the effect would still not be attributable to the confounding factors. The null hypothesis stating that the treatment effect is attributable entirely to the confounding factors can be revoked until a level of $\Gamma = 1.1$. When $\Gamma$ increases, the likelihood of finding a positive treatment effect decreases.

To further check robustness I calculate the matching for different wave-specific subsamples. The effect remains positive and significant for the six subsamples participating before 2001 and after 2001. The effect size for the nearest neighbor matching (1:1, with replacement) is slightly larger for the sample of the earlier waves (0.220 vs. 0.153). This variance in size could be interpreted as a slightly different selection of students who went abroad before ERASMUS was introduced (in 1987) and enabled a larger group of students to become mobile. To be more precise, it is possible that with fewer scholarships available, a smaller group of better students was selected. Thus, this group may have been more motivated to pursue a PhD. Nevertheless, though the effect remains significant and positive when calculating the matching for the first three waves (i.e., 1982-1987), it is slightly lower and subject to the limitation that only 257 mobile students can be matched.

## 3. Summary of Propensity Score Matching Results

As previously stated, it is evident that certain programs include compulsory semesters abroad, and grades may vary across programs. In addition, pursuing a PhD is more common in different programs. While program coordinators may consider the grade as a signal for a student to successfully manage mobility, I also know that a semester abroad might in turn affect the grade. As it would have been a bad control, grades were not included in the logit regression. I run different propensity score matchings for subsamples which I divide based on the programs and, in addition, by means of a median split. This way I aim to control any differences the grade might make.

The PSM reveals that business students who go abroad during their studies have higher intentions to get a PhD than their non-mobile peers. There are positive treatment effects for cultural and social studies as well. Distinguishing between above or below median grades does not make a difference regarding the level of significance.

The effects for medical and law students are insignificant. In addition, the mean PRB for the law students sample is relatively high, which means that the matched groups are different to a relatively high degree. Further, a semester abroad has a significant impact on below-median grade natural sciences students' PhD decision. In contrast, for engineering students there is a positive and significant effect of a semester abroad only for above-median performers. Table 2 provides an overview of the results for the PSM calculated separately for the different programs.

Table 2. Overview of PSM results for all programs (1:1 nearest neighbor matching, with replacement)

| program | median split (grade) | results | mean PRB |
|---|---|---|---|
| business studies | >= 4.2 | sig. positive ATT | 2.7 |
| | < 4.2 | sig. positive ATT | 2.9 |
| cultural studies | >= 5 | sig. positive ATT | 3.9 |
| | < 5 | sig. positive ATT | 3.5 |
| social studies | >= 5 | sig. positive ATT | 3.6 |
| | < 5 | sig. positive ATT | 4.1 |
| medicine | >= 4.4 | insig. | 2.7 |
| | < 4.4 | insig. | 4.4 |
| law | >= 4.0 | insig. | 5.9 |
| | < 4.0 | insig. | 8.3 |
| natural sciences | >= 4.7 | insig. | 3.7 |
| | < 4.7 | sig. positive ATT. | 5.2 |
| engineering | >= 4.3 | sig. positive ATT | 2.6 |
| | < 4.3 | insig. | 4.0 |

The visualization in figure 2 indicates that for business, social and cultural studies, the intention to pursue a PhD is stronger for the treatment group (i.e., mobile students). For business students the median grade hardly makes any difference with respect to the intention to pursue a PhD. Compared to social studies and cultural studies, business students are also the group with the lowest intentions of pursue a PhD. The figure also illustrates that for social and cultural studies students the median grade makes a difference. To be more precise, above median students in both programs have stronger intentions to pursue a PhD than below median students, in both the treatment and control groups.
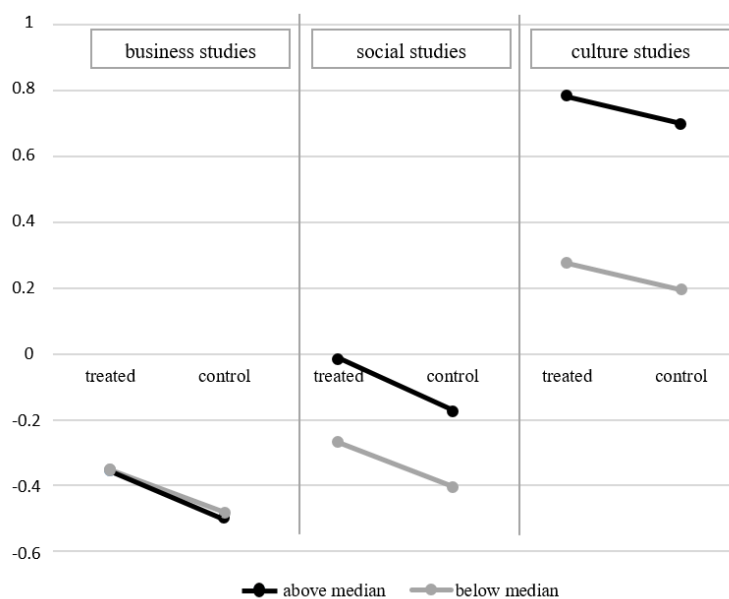


Figure 2. Visual illustration of the differences in the outcome variable (intention to pursue a PhD, standardized)

## 4. Discussion

### 4.1 Discussion and Implication of Results

Across university majors, my analysis reveals higher intentions to pursue a PhD for formerly mobile students. My findings are in line with existing studies showing that formerly mobile students feel an increase in their motivation and passion for their chosen career (Potts, 2015). In addition, existing research shows a correlation between educational attainment and mobility. In an earlier study more than half of the mobile students aimed at another degree after completing their Bachelor's degree (Paige, Fry, Stallman, Josić, & Jon, 2009).

The present study adds to existing knowledge the program-specific effects of mobility. The weaker intentions to pursue a PhD for business students are not surprising. In fact, on the German labor market students who finish their studies in business or economics can expect a higher entry salary in the industry than in academia. Research on the decision to pursue a PhD has also shown, that a lack of knowledge about the type of work a PhD student does influences the decision negatively (Ehrenberg, 2005). It may be hard for students to anticipate how long a PhD may last and thus deduct the opportunity cost of a forgone salary. Therefore, business students may prefer a higher immediate salary than the prospect of a higher salary after finishing a PhD. It is therefore a challenge for universities to attract business and economics graduates for an academic position. Business schools or faculties at German universities should consider informing students earlier on about the possibility of an academic career as an alternative to a labor market entry in the industry. As a semester abroad generates research interest for business students regardless of their average grade, student mobility should continue to be encouraged. A study by Naffzinger, Bott and Mueller (2008) shows that a major obstacle for business students' mobility is the fear of not receiving credit for classes taken abroad. By reducing these obstacles, universities may increase mobility and the number of students intending to pursue a PhD.

The social sciences sample includes students who seek to become school teachers, study politics, psychology, journalism or sociology (among others). For the social sciences student group, going abroad positively influences the intentions to pursue a PhD. Even if the grades are below median, a semester abroad raises the intentions to pursue a PhD above those for the business students. It is surprising that this group of students has

higher intentions to pursue a PhD than the business students, because teachers in the German education system hardly have any perspective of receiving promotions. Mertens and Röbken (2013) show that when comparing salaries between German doctorate holders in different programs, the salary advantage attributed to a PhD in education is smallest. Psychology and journalism students seeking to work in the industry may, however, be more likely to interpret the PhD as a signal to employers. As a conclusion, although the social science student group in the present study might be a heterogeneous group of students the positive effect of mobility on an academic career is evident and should encourage universities to invest further into the internationalization of their social sciences departments.

For engineering and natural sciences students, the effects of mobility differ in the median-grade subsamples. To be more precise, there is a positive effect of formerly mobile above-median performers in engineering. Mobility positively affects the intention to pursue a PhD for natural sciences students with below-median grades. Existing research shows that doctorate holders in engineering earn significantly higher salaries than doctorate holders in other programs, including natural sciences (Mertens & Röbken, 2013). It is possible that above-median performing engineering students are optimizing their human capital by investing not only in their academic career but by adding international experience to their skill set. From the university perspective, this might be the desired outcome and the target group for academic positions. Yet, it is questionable whether the grade is a suitable selection criterion for future PhD candidates. In the case of the natural sciences students, mobility academically motivates students with below-median grades. While it might not seem desirable for universities to attain below-median performers for academic positions, it is possible that these students performed well in the classes relevant for scientific work but poorer in others, thus decreasing their average grade. It is also possible that below-average natural sciences students intent to make up for their grade-deficit by proving other skills attained during their mobility, and thus following signaling theory. As future employers, universities should focus on additional selection criteria during interviews with PhD candidates. For example, the intrinsic motivation for an academic career may compensate for a poorer performance. In summary, especially in those programs where PhDs are not as common (e.g., business or cultural studies) a semester abroad can motivate students to consider an academic career. Though my study cannot answer how many mobile students actually obtain a PhD, universities could nurture the academic interest further by integrating research projects in mobility programs. Most programs require a report of the mobility experience which is often merely descriptive. The study abroad program which is described in Black and Duhon's (2006) study includes an examination and a research paper at the end of the stay. Having follow-up seminars where students can present research-related insights resulting from mobility programs would be another mean to encourage research interest and create reasonable expectations towards an academic career. It would be desirable if students who go abroad use the knowledge acquired in future projects.

### 4.2 Limitations and Future Research

The present study builds on Messer and Wolter's (2007) findings and adds to their understanding of a positive correlation between mobility and the intention to pursue a PhD, the possibility to infer causality. Additionally, the study contributes to existing studies the knowledge about program-specific mobility outcomes. Future studies could analyze the program-specific effects of mobility more precisely. For example, grades for research related classes could be valuable predictors of mobility as well as the intention to pursue a PhD. In addition, a particular focus on below-median performers could help identify and explain different motives for mobility and expectations towards the future academic career. Another aspect is the lack of knowledge about non-mobile students' semesters. It is possible that non-mobile students invest their time in other human capital skills it would therefore be desirable to know more about how the non-mobile students spent their corresponding semesters at home.

According to CIA, omitting important variables is problematic because it may cause an increase in bias (Heckman et al., 1997). In relevant literature, individual and personality-related variables have been shown to influence the decision to go abroad (Black & Duhon, 2006; Chieffo & Griffiths, 2004). Though the data contain questions on the personality traits, they could not be included in the present analysis because they would have been bad controls. Future studies should include measurements for personality traits at different points in time, allowing researchers to match students with similar personalities.

### References

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education.* Chicago, IL: Chicago University Press.

Black, H. T., & Duhon, D. L. (2006). Assessing the Impact of Business Study Abroad Programs on Cultural

Awareness and Personal Development. *Journal of Education for Business, 81*(3), 140-144. https://doi.org/10.3200/JOEB.81.3.140-144

Caliendo, M., & Kopeinig, S. (2005). *Some Practical Guidance for the Implementation of Propensity Score Matching*. DIW Discussion Papers, 485.

Coalition Contract, CDU Deutschlands, CSU-Landesleitung, & SPD (2013). Deutschlands Zukunft neu gestalten: Koalitionsvertrag zwischen CDU, CSU und SPD. 18. Legislaturperiode. Retrieved from https://www.cdu.de/sites/default/files/media/dokumente/koalitionsvertrag.pdf

Chieffo, L., & Griffiths, L. (2004). Large-Scale Assessment of Student Attitudes after a Short-Term Study Abroad Program. *Frontiers: The interdisciplinary Journal of Study Abroad, 10*(1), 165-177. https://doi.org/10.36366/frontiers.v10i1.140

Crossman, J. E., & Clarke, M. (2010). International Experience and Graduate Employability: Stakeholder Perceptions on the Connection. *Higher Education, 59*(5), 599-613. https://doi.org/10.1007/s10734-009-9268-z

DAAD. (2017). *Wissenschaft Weltoffen. Daten und Fakten zur Internationalität von Studium und Forschung in Deutschland*. Retrieved from http://www.wissenschaftweltoffen.de/publikation

Douglas, C., & Jones-Rikkers, C. G. (2001). Study Abroad Programs and American Student Worldmindedness. *Journal of Teaching in International Business, 13*(1), 55-66. https://doi.org/10.1300/J066v13n01_04

Ehrenberg, R. G. (2005). Graduate education, innovation and federal responsibility. *Council of Graduate Schools Communicator, 38*, 1-5.

Euler, H. P., Rami, U., & Glaser, E. (2013). Lohnt sich ein Auslandsaufenthalt während des Studiums? Ergebnisse der Evaluierung eines Förderprogrammes. *DBW, 73*(5), 425-447.

European Commission. (2021). *Verdoppeltes Erasmus+-Budget macht 10 Millionen Europäerinnen und Europäer mobil*. Retrieved from https://ec.europa.eu/germany/news/20210325-erasmus-budget_de

European Commission. (2017). *Erasmus+: EU programme for education, training, youth and sport*. Retrieved from https://ec.europa.eu/programmes/erasmus-plus/about_en

Favero, L., & Fucci, A. (2017). The Erasmus effect on earnings: a panel analysis from Siena. *Quaderni del Dipartimento di Economia Politica e Statistica, 762*, 1-45.

Heckman, J. (2000). Policies to Foster Human Capital. *Research in Economics, 54*, 3-56. https://doi.org/10.1006/reec.1999.0225

Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The Review of Economic Studies, 64*(4), 605-654. https://doi.org/10.2307/2971733

Iriondo, I. (2020). Evaluation of the impact of Erasmus study mobility on salaries and employment of recent graduates in Spain. *Studies in Higher Education, 45*(4), 925-943. https://doi.org/10.1080/03075079.2019.1582011

Kratz, F., & Netz, N. (2018). Which mechanisms explain monetary returns to international student mobility? *Studies in Higher Education, 43*(2), 375-400. https://doi.org/10.1080/03075079.2016.1172307

Marcotte, C., Desroches, J., & Poupart, I. (2007). Preparing Internationally Minded Business Graduates: The Role of International Mobility Programs. *International Journal of Intercultural Relations, 31*(6), 655-668. https://doi.org/10.1016/j.ijintrel.2007.05.002

Mertens, A., & Röbken, H. (2013). Does a Doctoral Degree Pay Off? An Empirical Analysis of Rates of Return of German Doctorate Holders. *Higher Education, 66*, 217-231. https://doi.org/10.1007/s10734-012-9600-x

Messer, D., & Wolter, S. C. (2007). Are student exchange programs worth it? *Higher Education, 54*(5), 647-663. https://doi.org/10.1007/s10734-006-9016-6

Naffziger, D. W., Bott, J. P., & Mueller, C. B. (2008). Factors influencing study abroad decisions among college of business students. *International Business: Research, Teaching and Practice, 2*(1), 40-52.

Paige, R. M., Fry, G. W., Stallman, E. M., Josić, J., & Jon, J. (2009). Study abroad for global engagement: The long-term impact of mobility experiences. *Intercultural Education, 20*(sup1), S29-S44. https://doi.org/10.1080/14675980903370847

Petzold, K. (2020). Heterogeneous Effects of Graduates' International Mobility on Employers' Hiring Intentions - Experimental Evidence from Germany. *Higher Education*. https://doi.org/10.1007/s10734-020-00524-3

Petzold, K. (2017). Studying Abroad as a Sorting Criterion in the Recruitment Process. A Field Experiment Among German Employers. *Journal of Studies in International Education, 21*(5), 412-430. https://doi.org/10.1177/1028315317697543

Rodrigues, M. (2013). *Does student mobility during higher education pay?.* Evidence from 16 European countries. JRC scientific and policy reports. Luxembourg: Publications Office of the European Union.

Rosenbaum, P. R., & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika, 70*(1), 41-55. https://doi.org/10.1093/biomet/70.1.41

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing Bias in Observational Studies Using Subclassification on the Propensity Score. *Journal of the American Statistical Association, 79*(387), 516. https://doi.org/10.2307/2288398

Potts, D. (2015). Understanding the early career benefits of learning abroad programs. *Journal of Studies in International Education, 19*, 441-459. https://doi.org/10.1177/1028315315579241

Salisbury, M. H., Paulsen, M. B., & Pascarella, E. T. (2010). To See the World or Stay at Home: Applying an Integrated Student Choice Model to Explore the Gender Gap in the Intent to Study Abroad. *Research in Higher Education, 51*(7), 615-640. https://doi.org/10.1007/s11162-010-9171-6

Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics, 87*, 355-374. https://doi.org/10.2307/1882010

Van Mol, C., Caarls, K., & Souto-Otero, M (2020). International student mobility and labour market outcomes: an investigation of the role of level of study, type of mobility, and international prestige hierarchies. *Higher Education*. https://doi.org/10.1007/s10734-020-00532-3

Waibel, S., Petzold, K., & Rüger, H. (2018). Occupational status benefits of studying abroad and the role of occupational specificity - a propensity score matching approach. *Social Science Research, 74*, 45-61.

**Notes**

Note 1. I use the standardized values for the intention to pursue a PhD and rescale phd_plan to have a mean of zero and a standard deviation of one. The mean value of plan_phd for the business student sample is 1.62 (std. dev 0.82).

**Appendix**

**Appendix A**

Table A1. Logit regression results

| dependent variable: abroad (0=non-mobile, 1=mobile) | coeff | standard error |
|---|---|---|
| female (dummy; ref: male) | 0.350*** | (0.058) |
| age | 0.766*** | (0.107) |
| age2 | -0.013*** | (0.002) |
| wave1982_83 | 0.000 | (.) |
| wave1984_85 | 1.860*** | (0.264) |
| wave1986_87 | 1.665*** | (0.262) |
| wave1989_90 | 2.079*** | (0.258) |
| wave1992_93 | 2.228*** | (0.257) |
| wave1994_95 | 2.263*** | (0.257) |
| wave1997_98 | 2.067*** | (0.264) |
| wave2000_01 | 2.196*** | (0.260) |
| wave2003_04 | 2.562*** | (0.255) |
| wave2006_07 | 2.822*** | (0.256) |
| wave2009_10 | 2.662*** | (0.257) |
| wave2012_13 | 2.212*** | (0.273) |
| family status (dummy; ref: married) | | |
| relationship | 0.169 | (0.141) |
| single | 0.289** | (0.144) |

| | | |
|---|---|---|
| children | -0.374*** | (0.134) |
| high school diploma type (dummy; ref: regular diploma) | | |
| subject related diploma | -0.489*** | (0.158) |
| job-related diploma | -0.428*** | (0.100) |
| other diplomas | -1.327 | (1.078) |
| high school grade | 0.433*** | (0.047) |
| university type (dummy; ref: university) | | |
| college of applied sciences | 0.357*** | (0.069) |
| number_of_semesters | 0.171*** | (0.024) |
| number_of_semesters2 | -0.007*** | (0.001) |
| sports_involvement | 0.860*** | (0.239) |
| sports_involvement2 | -0.171*** | (0.061) |
| music_involvement | 0.115 | (0.302) |
| music_involvement2 | 0.019 | (0.082) |
| fraternity | 0.217 | (0.618) |
| fraternity2 | -0.022 | (0.161) |
| religious_group_involvement | 0.014 | (0.587) |
| religion_ group_involvement2 | 0.036 | (0.158) |
| student_administration_involvem | -0.632 | (0.678) |
| student_administration_involvem2 | 0.146 | (0.179) |
| political_group | 0.790** | (0.390) |
| political_group2 | -0.136 | (0.106) |
| finance_parents=1 (dummy; ref: no financial aid from parents) | | |
| finance_parents=2 (partly) | 0.208** | (0.081) |
| finance_parents=3 (mainly) | 0.436*** | (0.087) |
| finance_bafoeg=1 (dummy; ref: no financial aid from state) | | |
| finance_bafoeg=2 (partly) | -0.275*** | (0.093) |
| finance_bafoeg=3 (mainly) | -0.451*** | (0.101) |
| finance_scholar=1 (dummy; ref: no financial aid from scholarship) | | |
| finance_scholar=2 (partly) | 1.256*** | (0.158) |
| finance_scholar=3 (mainly) | 0.634*** | (0.233) |
| constant | -19.428*** | (1.679) |

*Notes.* * $p<0.10$, ** $p<0.05$, *** $p<0.01$. 9 826 observations. The high school grade is reported on a scale from 1(worst grade)-6 (best grade), while I deleted students who indicated they failed high school, leaving those with grades ranging from 3-6.

**Copyrights**