https://doi.org/10.20378/irbo-54036

Felix Bittmann

Universität Bamberg
Feldkirchenstraße 21
96045 Bamberg
Germany

0049/(0)951-863-1442
felix.bittmann@uni-bamberg.de

Explaining the Mechanisms linking Field of Study and labour Market Outcomes: Focus on STEM

Working Paper Version 1.0 / January 2019

Abstract

The paper analyses labour market outcomes for university graduates of the STEM-fields and aims to explain causally why differences between the fields arise. Based on theoretical considerations, constructs and mechanisms are selected that could explain the relationship in detail. Regression models are run to estimate the effect of different variables on hourly income five years after graduation, job mismatch and risks of ever being unemployed. After controlling for a wide range of factors, two mechanisms (skills of graduates and occupation in the labour market) are tested. The main results are that graduates of the four fields do differ with respect to sociodemographic variables before study and self-selection effects are present. Although both mechanisms can account for some variation, a large part of unexplained variance remains even after adding all control variables. Especially graduates of the Sciences have significantly lower incomes and higher risks of unemployment than any other STEM-field. As these differences cannot be attributed to any variable included in the analyses, field of study does have an independent influence on outcomes in the labour market. The overall source of variation remains unclear and requires further investigation of mechanisms.

Keywords: STEM, Labour market, Unemployment, Wages, Comparative analysis

Word count: 7,133.

Conflict of Interest: The author declares that he has no conflict of interest.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgements: The author wants to thank Hans-Peter Blossfeld for his comments and Viktoria Sophie Zorn for checks on the final draft.

Background

One of the most severe concerns of German policy-makers and economists is the declining interest in STEM-fields among university graduates. Although Germany is highly dependent on a well educated STEM-workforce for its economic prosperity, which is based, for example, on its chemical, automotive or machine-building industry, it apparently faces a stark decline in STEM-graduates and a shortage of expert workers (Anger et al. 2018). The problem becomes even more perplexing when taking into account that the general interest in higher education is unbroken in Germany and statistics show a steady rise in the number of university-students (Statistisches Bundesamt 2018a). Therefore the question arises whether students lack the interest in pursuing *certain* STEM-fields, which could result in a specific shortage. A selective pattern could easily arise when some fields show significantly worse outcomes for graduates, for example with respect to wages, working conditions, risk of unemployment, overeducation or work-life-balance. There are some reasons to believe that, although grouped together in the STEM-scheme, subjects within STEM are not a homogeneous group as suggested by the grouping but differ quite drastically, for example with respect to intelligence and ability needed to complete a study successfully or in usefulness and demand on the labour market. There is enough evidence to believe that the phenomenon of STEM-heterogeneity was hardly research thoroughly in the past, which poses a great obstacle for assessment of the usefulness and profitability of careers in STEM. On a positive note, Germany has one of the highest shares of university graduates in high-skill jobs in Europe, which should foster the general interest in higher education (Green and Henseke 2017, 27). Finally, the results seem also highly relevant in the global context as in 2017 about 360,000 university students were exchange students (Statistisches Bundesamt 2018b) and from these, more than 50% studied a STEM field (Apolinarski and Brandt 2018, 25).

Specifically, the present research attempts to answer three research questions:

1. How similar are labour market outcomes for graduates of the four different STEM-field in comparison?
2. In case of finding any differences, are these actually caused by the field of study or just the result of self-selection?
3. Which mechanisms are accountable for the arising differences?

These research questions seem central not only for prospective students of STEM but especially for economists and policy-makers who are interested in tailoring the workforce of the future for the demands of the economy. The current study attempts to reduce this research gap and adds to the existing research of STEM-education evaluation.

Review of literature

Taking a look at previous studies concerning the outcomes of STEM-graduates it becomes clear that, especially for the German context, a research gap exists. To fill this gap a large number of theories, which aim at the explanation of how education and occupational outcomes are linked, could be applied. One of the most prominent ones is human-capital-theory (Becker 1975), which assumes that every worker has a certain productivity, which can be influenced through education, training or other means, which could also target the personal health or general fitness. The higher the productivity the better the outcomes as more productive workers can produce more goods or services per time, which is rewarded by the employer. Following the human-capital-theory one can conclude that different fields of study could have a different influence on the personal productivity. Differences in productivity and outcomes are therefore (partially) caused by the different fields of study. The drawbacks of this model are that it is very general and for that reason hard to define and operationalize what is meant exactly by productivity, how it should be measured and how it can be influenced in a causal fashion.

The second main theoretical framework comprises signaling- and screening-theories (Spence 1973) and has quite different assumptions. According to signaling, the productivity of a worker is not variable and cannot be changed easily but is relatively constant and different educational paths can hardly influence pre-existing levels of productivity. Different studies and certificates are rather a way to recognize and sort the abler from the less able workers as the real productivity is hidden and not measurable directly. Students self-select into different fields according to their personal abilities and when they complete a field successfully they are awarded a certificate which signals their abilities. Therefore, employers use field of study and the respective certificates as a mean to choose the ablest persons. According to this notion, there is no or only a weak influence of education on the personal abilities, which is very different from human-capital-theory (Arrow 1973). These assumptions seem to be quite unsatisfactory, as they generally deny the possibility that people do learn and acquire new skills in the educational system and higher education.

Finally, the training-costs-model (Glebbeek et al. 1989) is a synthesis of these two main frameworks and furthermore incorporates the job-competition model (Thurow 1975). The main assumptions are that workers compete for positions and not for wages (1), productivity is closely linked to the occupation and not the person (2), skills are acquired predominantly on the job (3) and education is a positional rather than an absolute good (4). This theoretical approach seems most appropriate to describe the current situation of the labour market and helps in explaining differences between fields of study. According to the model, there exists a large number of labour market queues, one for each occupation. In each queue, workers are ranked according to their training costs for the respective occupation and the lower the costs, the more prominent the position of a worker in a queue. Employers hire workers according to their ranking in the queue until all positions are filled. Training costs are determined by the set of skills of a worker, by the match between requirements of the occupation and available skills and by the uncertainty an employer faces when hiring a graduate of a certain field, as skills and abilities are not known perfectly, but only estimable through easily measurable attributes. These factors can be linked directly to education (e.g. certificates or grades), sociodemographic ones (e.g. gender, age or socioeconomic status) or psychological ones (e.g. intelligence and attitudes). Some of these are apparent while others are often labouriously or costly to assess. In detail, there are two prominent themes of the training-costs-model: occupational specificity and selectivity. Graduates with a high specificity will have better outcomes in general, as they have low training costs for the occupation they are specialized in, than graduates with a general set of skills, which have higher training costs for most positions. As employers hire graduates under uncertainty, they choose graduates from fields with a high selectivity, meaning, on average, that graduates of this field show a low variation in skills and abilities, which reduces the probability of hiring an underperforming worker. Thus, previous results, which show that female STEM-graduates as well as members of minority groups earn significantly less than their counterparts in some STEM-fields (Falk 2010, 61; Melguizo and Wolniak 2012; Kirkebøen et al. 2014), suggest that employers attribute higher training costs to women and minority group members. Based on this framework hypothesis one can be formulated, which is quite general, yet should be tested explicitly: Completing a field of study has an effect on occupational outcomes and differences between the fields cannot be explained with self-selection alone. Confirming this hypothesis would rule out that differences in labour market outcomes are solely attributed to personal background characteristics like age, gender or GPA, which can be seen as a rough proxy for intelligence (Kaufman and Lichtenberger 2006, 12).

As one study could already verify most of the implications of the training-costs-model by concluding this from large differences between “soft” fields and the sciences (Klein 2010), it seems promising to adapt its expectations to the own research question. First, it is based on the assumption of a large variation in occupational specificity. As this might be true for some fields, especially the liberal arts and humanities, which seem to incorporate more general than specific skills, it is difficult to argue that within the STEM-fields there is large variation as specificity is high for all. Because of their specialization STEM-graduates are usually not employed outside their subject area but follow suited occupations (Schramm and Kerst 2009, 25). Yet, another study from Canada reports higher mismatch rates in Science and Technology which seems to come along with lower incomes in these fields (Hango 2013). Still, as there are no “generalists” fields in STEM, it is unclear why one field should be better or worse suited for a job than another when only considering the degree of specialization. A solution to this problem is to be more specific and not only difference the degree of specialization but rather assess the actual skills held by graduates. The exact number or degree of skills for each field is quite unclear and would require extensive research because different skills had to be extracted on a very detailed level to difference all STEM-fields thoroughly. Although this cannot be done in the own analysis it seems logical that a better fit between skill requirements and actual skills of a graduate will lead to better outcomes. Considering the imaginary job-queue one can assume that rational graduates with similar skills will choose similar queues for optimal outcomes. For example, when looking for a programmer it should not matter to an employer whether a candidate graduated in biology or mathematics as long as the set of skills of the candidates is identical. People who get hired out of such a queue therefore will have the largest overlap with skill requirements of the occupation. Based on this one can formulate hypothesis two: Skills held by graduates are a mechanism that explains how field of study and occupational outcomes are related.

Although it seems reasonable to assume that a rational employer should focus on the skills of graduates to find the best match for a position, the problem persists that skills are often hidden and costly to assess. Therefore, it is possible that not the actual skills of a person but the perception of these skills by the employer has the largest influence on outcomes. For example, an employer might prefer the mathematician over the biologists for the programming job as she believes that a graduate of math will have better fitting skills (even if the skills are identical in reality). This is in line with the theories of signaling and screening which predict that employers will use strong and obvious signals. To account for this the model needs an extension. A solution is to incorporate labour market occupations into the framework. After choosing one or several queues, a graduate will finally get picked to enter the labour market in a certain occupation. As already explained, this final occupation depends on several factors, like actual skills and the perception of skills by an employer with respect to field of study but also personal preferences of the graduate. For example, some highly skilled graduates might prefer to enter the public sector even if average incomes are lower in exchange for a higher job-security which especially seems to be true for some female STEM-graduates (Falk 2010, 52). This underlines that outcomes also depend on factors that might be completely uncorrelated to field of study or skills. Based on this, we can formulate hypothesis three:

The occupation in the labour market after graduation is a mechanism that explains the relation between field of study and outcomes. Or to formulate it differently: Graduates from different fields who end up in the same occupation should have similar outcomes on average. This assumption is also supported by Thurow who states that graduates compete for jobs and not for wages (Thurow 1979). A desirable property of this approach is that it controls implicitly for supply and demand factors in the labour market which are obviously related to outcomes. As only graduates in the same fields are compared, supply and demand should also be comparable as these are usually linked to certain positions and occupations.

To summarize, the framework predicts three major explanatory factors: First, field of study does influence graduates somehow and self-selection alone is not enough to explain differences in outcomes. Second, the actual skills held by graduates are an important mechanism that explains why outcomes are different between the fields. Lastly, the occupation in the labour market is another mechanism that explains differences in outcomes for graduates with the same set of skills. Furthermore, it can be expected that there is a chain of mechanisms: Field of study does affect skills of graduates which in turn do affect the occupation of a graduate. A causal diagram (directed acyclic graph) of the model presented here can be found attached (**Figure 1**) which follows the general ideas of causal graphs introduced by Pearl (2000).

Although the relationship between fields of study and labour market outcomes was studied in the past, especially with respect to sociodemographic variables like gender or general level of education, there exists a research gap concerning tertiary education STEM-fields and mechanisms of effect transmission.

Method

Data and sample

The analyses make use of the DZHW graduate panel 2005/06 by the German Centre for Higher Education Research and Science Studies. Graduates of all fields all over Germany were surveyed twice, 1.5 (wave 1) and 5 years after graduation (wave 2, around 2011) to assess long-term development. Although this data is already some years old, it is, at the time of analysis, the only German data available which includes a large proportion of STEM graduates. The population comprises all people who received a first degree of graduation from a German university or university of applied science (*Fachhochschulen*) in 2004 or 2005 (not including universities of the

German military or distance-learning universities). Universities were sampled and contacted. Due to privacy concerns, questionnaires were then distributed via the respective universities for wave one, in the second wave directly from the DZHW if people agreed to stay in the sample. There are 11,788 respondents in wave one and 6,459 in wave two (Baillet et al. 2017, 19-25).

The sample was then further restricted for the analyses. Only graduates who participated in both waves were included. All teachers were excluded as they are usually employed by the state and job-market mechanisms are quite different for this group. All graduates from universities of applied sciences were removed as well. This can be justified theoretically as these forms of higher education are focused on practical knowledge that can be directly applied in a job which leads to large differences with universities, which are based on a more scientific approach. Therefore it seems problematic to compare graduates of both universities to each other. The main empirical problem is that some fields can only be studied at a regular university, for example biology. For that reason we see a very unequal distribution in the Sciences (more than 80% graduated from a regular university), which makes a comparison problematic as the comparison group would be quite small. Therefore it was decided to reduce the sample to university graduates only. People who hold two degrees are included as long as the second degree is from the same STEM-field as the first one. Otherwise these people were removed. Note that this does not concern consecutive degrees (for example bachelor-master) as no master students at all are included in the DZHW sample. Lastly, people with more than 20 semesters were also removed because it is unclear if these represent the average life-courses. All in all, there are 1,218 graduates left for analyses from the STEM-fields.

Variables

The first and crucial step of the analysis is to group students into the four STEM-fields. A review of the literature shows that no generally accepted scheme exists, which might lead to bias. This problem can only be countered with transparency. Firstly, to select any STEM-field from all fields available, a definition by Chen (2009, 2) was used which stays close colloquial meaning: "...mathematics; natural sciences (including physical sciences and biological/agricultural sciences); engineering/engineering technologies; and computer/information sciences." Table Table 1 (page 19) lists the available fields of studies and how they are sorted into the four categories. It should be noted that the fields were already pre-grouped by the DZHW due to privacy-issues. Especially problematic seemed the distinction between Technology and Engineering, as here similarities seem high. As a rule of thumb, only fields with the term "Engineer" or machine construction were grouped into Engineering, all other as Technology. This is a limitation of the study, which results from a weakly developed conceptual background. Yet the distinction is kept for comparison with many other papers that employ this concept.

The three main outcome variables are hourly wages including bonuses, job-mismatch and risks of ever reporting an episode of unemployment. Hourly wages were calculated for the last reported income in wave two, which includes any bonuses. This variable was logarithmized to ease statistical inference and make the distribution more symmetric. People with very high incomes (more than 60€ per hour, which is above the 99. percentile) were deleted (9 cases). The job-mismatch indicator was created from three pentatonic rating-scale items that target whether the currently held position in wave two is adequate for the degree earned or not (alpha=0.80). Higher values stood for a worse match between degree and occupation. This variable was logarithmized as well. To indicate unemployment information from the longitudinal episode-files was used. Respondents should report any relevant episodes after graduation, for example being employed, doing internships, being unemployed or being on parental leave. When a person ever reported an episode of unemployment, he or she received the value 1, otherwise 0. Using the length of unemployment spells as a dependent variable turned out to be problematic as the distribution of values was extremely skewed and most people have none or very short unemployment spells.

Possession of skills was measured by using 24 different items with pentatonic rating scales. Graduates should indicate for each dimension the amount of skills available in wave one. The items comprise general and specific skills, language knowledge, intercultural skills or business related items, like time-management abilities or working independently. To further summarize the skills and avoid multicollinearity issues a principal-component analysis (PCA) is run to determine basic dimensions. The results show five components with an eigenvalue larger than 1 which jointly explain more than 51% of the total variance of all 24 items. To ease interpretation, components were rotated using the *promax* algorithm. This algorithm is an oblique rotation method, which seems to be more adequate in the own case as it allows factors to be correlated. The Kaiser-Meyer-Olkin measure is larger than 0.89, which is an excellent result and underlines that the variables are suited for PCA (Kaiser and Rice 1974, 112). The resulting rotated factor-loadings are shown in Table Table 2 (page 21). For a robustness check, original items and generated factors were included for comparison in an extra model. The results are similar, therefore it was decided to use the generated factors in further analyses.[1]

To make an assessment of self-selection possible, variables were selected as controls that might influence both the chance to enter a certain study and the outcomes in later life simultaneously.

1 Diagnostic tables are available upon request from the author due to space restrictions.

These variables are gender, age, federal state of university eligibility degree and federal state of university (both grouped into four categories), type of university eligibility school (*Gymnasium* VS all other), grade of university eligibility degree (GPA[2]), whether a study includes a mandatory internship, whether someone worked before the study, whether someone had a vocational training before study, type of degree (*Diplom* VS *Staatsexamen* (state examination) VS Bachelor), academic background (1 when at least one parent holds a tertiary degree, 0 otherwise) and the self-rated importance of labour market outcomes when choosing the field of study (likert scaling with 5 categories). Furthermore the status of a Ph.D.-project was controlled. It was distinguished whether a graduate never started or canceled a Ph.D. (0), still was in a Ph.D.-program (1) or already finished the Ph.D. successfully (2). This seemed to be important as the probability to enter a Ph.D.-program is correlated with field of study and especially financial outcomes are often affected by this until the project is finished. As the current analysis is not interested in assessing the effect of having a Ph.D., it was decided to control for it. This means that final results are independent of the Ph.D.-status. If long-term outcomes, say 10 years after graduation, were available, this would not be necessary as (almost) all graduates would have finished their Ph.D. projects and the overall effects of field of study could be estimated including the effect and propensity of entering such a program.

2 In Germany, the GPA can range from 1.0 to 4.0 and lower values indicate higher performances.

Methods

To assess the effect of field of study on outcomes, multiple linear regressions (OLS) are used for the outcomes “hourly wages” and “job-mismatch” and binary logistic regressions for the dependent variable “unemployment”. To receive detailed results, all fields of study are compared to all others, which results in a binary-contrasts-design. The challenge is to ensure that two fields are actually comparable with respect to all relevant variables. For example, if there is no overlap of the fields with respect to occupation in the labour market, a comparison would be impossible. Therefore it has to be guaranteed for each binary contrast that only graduates are included who have an occupation that is also taken by people in the comparison group. Again, this underlines the need for the six-contrasts design: The overlap of the population for all four fields is much smaller than the overlap of only two fields. To operationalize occupation the scheme of *Klassifikation der Berufe 2010* (KldB2010) was used, which is comparable to the ISCO-08. The system has been used at the two-digits classification which allows to group occupations together without being too specific. To calculate the overlap, for each occupation the number of graduates from the two respective fields in comparison has been checked. When there were less than ten people combined in an occupation or when the percentage of one field was higher than 95%, these occupations (and the graduates working there), were not used in the model. By doing this an acceptable number of cases could be retrieved for each binary contrast.

Descriptive results

Table Table 3 (page 22) lists most central dependent and independent variables as well as F-values, which are based on one-way ANOVAs to test for overall group differences. One can conclude from this that graduates of different fields do differ drastically with respect to some covariates. Graduates of the Science show quite a high interest in pursuing a Ph.D., which is much less popular in the other fields. Interestingly, they also show the highest percentage of females and the second lowest interest in importance of the labour market. The higher the value of this item, the less the interest in direct applicability of the field of study for a job. We can use this as a proxy to see how market-driven the students are and accordingly whether they rather study for personal interests and development.

The comparison of the factors created in the PCA is also quite interesting as the systematic differences become obvious. Especially remarkable is the very low value of factor two (leadership skills) for graduates of Science. It can be noted that fields differ with respect to skills, yet it is unclear why these differences arise. The first possibility is that they are caused by the different curricula but it could also be due to self-selection when it comes to choosing a field of study. A third possibility is that these skills are predominantly acquired in the time after graduation, as about 1.5 years lie between graduation and skill measurement. Although this seems problematic to disentangle, the models will account for this.

With respect to the outcomes we see that Science shows the lowest income on average, but also the best matches in the job. The unemployment risks are quite similar for all fields except Mathematics, which displays a lower risk. To summarize it, differences between the fields are clearly visible, yet it is unclear whether these are caused by the fields or by self-selection. The differences in covariates underline that students are already different when choosing a field of study, therefore statistical methods are required to disentangle these aspects.

Binary contrast models

As there are three outcomes, six field contrasts and three models for each, which would result in 54 extensive tables, the results are displayed in a highly compressed format which only shows coefficients and standard errors for the contrast-variable and the number of cases used (for the outcome unemployment, average marginal effects will be shown).[3] The first model-block (M1) only includes the contrast-variable (for example Sciences VS Technology, where the first field is always coded as 0 and the second as 1) and all sociodemographic controls. Furthermore only those cases are used that show an overlap with respect to occupation. The second block (M2) adds the five skill-factors as explaining variables. The third block (M3) furthermore adds occupation of the graduates as a last control variable. The idea behind this design is as follows: if the contrast-variable in model one is significantly different from zero one can conclude that this difference is not due to any background-characteristics. If the p-values become smaller in models two or three (indicated by asterisks), this reveals that the supposed mechanisms are actually able to “explain away” why outcomes are different. If p-values do not change much one knows that supposed mechanisms are probably no mechanisms at all in reality. By assessing differences in coefficients and significances, mechanisms can be tested for each contrast.

How to read the tables should be demonstrated for the contrast Sciences VS Technology and the outcome wages (page 23). In the first model there is a highly significant difference of 0.278, telling that graduates of Technology earn more on average. The second model, which additionally controls for skills, shows that the effect becomes even slightly stronger. The third model furthermore

3 More detailed tables are available upon request.

includes occupation as a control, meaning that graduates in the same positions are compared. Now the coefficient becomes smaller (0.185), yet the significance level stays identical. This highlights that occupation is a mechanism that partially explains how outcomes are linked to field of study. As the general contrast is still significant, it can be concluded from this that graduates of Technology earn more than those of the Sciences even if they work in the same jobs. Field of study therefore has an additional independent effect that cannot be explained by the model or by theory so far. The general interpretation of Table Table 4 (page 23) shows that graduates from Sciences do earn less in contrast to Technology, Engineering and Mathematics, regardless of the model. There are no significant differences for any other contrast.

When regarding the second outcome job-mismatch (Table Table 5, page 24), it becomes clear that there are no differences at all between the fields for any model. Therefore one can conclude that field of study does not predict job-mismatches in the labour market for any field-contrast, except Science-Engineering in Model 3, which seems to be an outlier here.

Finally the results for outcome risk of unemployment are introduced (Table Table 6, page 25). It should be noted that the mechanism proposed (occupation) is somewhat problematic for this dependent variable because only people are included that found a job and therefore entered a certain occupation. Accordingly, it is unclear whether the occupation held at the time of the survey does have any influence on the risk of being unemployed the time before. Summarized, it seems best to interpret only blocks one and two, which still allows controlling for skills as a mechanism. The first model shows that graduates of Engineering and Mathematics have significantly smaller risks of ever being unemployed than graduates of the Sciences. These differences remain after adding skills in model 2. Therefore one can conclude that neither background characteristics nor skills of graduates do explain field-differences completely and that there is an independent and significant field-effect. This effect seems quite substantial as, for example, graduates of Engineering have a 16 percentage points lower probability of ever being unemployed than graduates of the Sciences.

Discussion

Summary of principal findings

The descriptive statistics show that there are differences between fields of study with respect to outcome, which were especially large when hourly wages were considered. Furthermore, self-selection effects were observed, which was also revealed by the descriptive findings. Differences are especially large with respect to the percentage of female students, type of degree or whether a vocational training was done before entering the study. Therefore one can conclude that self-selection is present, which paradoxically often does not account for observed differences. For example, GPA was lowest and therefore best for graduates of the Sciences, yet wages were lowest for this group, which seems surprising, as better grades are associated with better outcomes in general.

To expand these findings, binary contrast models were run. The results are that graduates of the Sciences earn less than graduates of any other field, even in the full model (M3) which also controls for occupation and available skills. In conclusion, graduates of the Sciences earn less than other STEM-graduates, even when they work in the same positions in the labour market. Similar findings can be made with regard to risks of unemployment, yet there is no significant difference for the contrast Science-Technology. What we can learn from this is that Science seems to be clearly worse off than other fields as this is in general the only contrast that shows significantly poorer outcomes. These findings are somewhat surprising in the light of the theories proposed before. Firstly, signaling theory alone cannot explain these findings as a wide range of control variables are introduced as well as the actual skills of graduates. When field of study is only useful to select the ablest people it is unclear why the Sciences have worse outcomes in general, especially, as graduates of this field show the best GPA scores on average. Therefore it can be concluded that signaling-theory is not appropriate in the STEM context. This was anticipated as graduates should be similar on average, as all STEM-fields have similar requirements for students. Continuing with the training-costs-model, we learn that the findings do not strengthen the model. One would expect that skills are an important factor which explains general training costs, as these should be related to skills held by graduates. Yet it seems that even after controlling for them the differences persist. It is unclear why employers should estimate higher training costs for graduates with the same skills but a different field of study according to the theory. One could argue that, in reality, skills are hidden and employers do not have access to the kind of information the analyses take into account and use field of study as a proxy instead. Yet this seems improbable as, over time, employers should update their beliefs about the skills of graduates as long as they employ people from different fields. As long as field of study does influence skills, a rational employer should be able to assess the available skills quite precisely.

The general problems of the model deepen after introducing occupation which highlights that graduates of the Sciences earn less than others even if they work in the same jobs. This is a violation of the model assumptions as graduates should compete for positions in the labour market. When a position is found, wages should not depend on the field of study any longer. The general conclusion is that the model seems not appropriate to explain the findings. Lastly, it seems that Becker's human-capital-theory still describes the results best. According to his notion, one must assume that field of study has an independent influence on the productivity of graduates which is not measurable directly and not related to any background information. This explains why graduates in the same jobs with different fields of study show differences in wages, as the more productive workers have higher incomes on average.

Limitations

Before summarizing the findings, the limitation of the study should be discussed. First, it is possible that controlling for skills does not affect the results because not all skills are measured by the variables. This is certainly true to some extent, as not every single skill is included. Furthermore it is not clear which kind of skills one should consider in the first place as technically the number of them might be infinite. Yet one could argue that employing 24 different items, which are reduced to five major factors, should capture a large amount of the general variance. The findings seem too strong and robust so that minor, not included parts, could explain all differences.

Second, another problem is that skills are not measured objectively and directly after graduation, but at a time when most graduates already entered the labour market. Therefore, it might be possible that skills are also affected by the work-experience (reverse causality). These are valid claims and can only be answered finally by collecting more data. It would be highly desirable to measure skills before starting the study and directly after graduation to estimate the causal influence of the field on skills. Still there persists the problem of how to measure skills and which theoretical constructs are applicable and realistic in the process of data collection. Finally it should be noted that, after controlling for position, skills are of less importance as skills should lead to a certain position, as wages are linked to positions and not skills. Therefore the problems that are coupled to controlling for skills should be less relevant in the final model.

Third, the coding of occupations held by graduates is problematic as it is unclear whether the same code means automatically exactly the same position or not. This is a general problem which is also related to the number of cases. When the coding is too general, quite different jobs are mixed together in the same categories, which would clearly weaken the effect of field and make them less comparable. Yet, a very strict coding will decrease the number of cases in each occupation, meaning that there are no comparisons possibly in the end. Therefore the solution of the paper is pragmatically, yet could be improved in the future by collecting more data. When the number of graduates is much higher, a finer raster will be applicable.

Conclusion

The main message of the study is that more detailed analyses will be needed in the future to disentangle the shown results and analyse them in a causal fashion. More information about graduates as well as more participants would be highly desirable to test further mechanisms and reduce margins of error. Furthermore, the extremely problematic outcomes for graduates of the Sciences should be studied in detail, as these could influence economic growth and scientific progress. When structural factors are accountable for these unsatisfying results it would be central to reduce these troublesome factors to make careers in Science more attractive. Lower incomes and a higher risk of unemployment are quite detrimental factors that could easily deter potential students. Analysing and solving these issues might clearly affect the shortage of experts in these fields which will be beneficial for economic growth and prosperity.

References

Anger, Christina, Oliver Koppel, and Axel Plünnecke. 2018. "MINT-Frühjahrsreport 2018. Gutachten für BDA, BDI, MINT Zukunft schaffen und Gesamtmetall." https://www.iwkoeln.de/fileadmin/user_upload/Studien/Gutachten/PDF/2018/MINT-Fr%C3%BChjahrsreport_2018neu.pdf

Apolinarski, Beate, and Tasso Brandt. 2018. “Ausländische Studierende in Deutschland 2016. Ergebnisse der Befragung bildungsaus-ländischer Studierender im Rahmen der 21. Sozialerhebung des Deutschen Studentenwerks durchgeführt vom Deutschen Zentrum für Hochschul- und Wissenschaftsforschung.” Berlin: Bundesministerium für Bildung und Forschung (BMBF).

Arrow, Kenneth J. 1973. "Higher education as a filter. "Journal of Public Economics 2 (3):193-216.

Baillet, Florence, Andreas Franken, and Anne Weber. 2005. "DZHW-Absolventenpanel 2005. Daten-und Methodenbericht."

Chen, Xianglei. 2009. "Students Who Study Science, Technology, Engineering, and Mathematics (STEM) in Postsecondary Education. Stats in Brief. NCES 2009-161." National Center for Education Statistics.

Falk, Susanne. 2010. "Gleicher Lohn bei gleicher Qualifikation? Eine Analyse der Einstiegsgehälter von Absolventinnen und Absolventen der MINT Fächer " Beiträge zur Hochschulforschung 32 (4):23.

Glebbeek, A, NWHS WIM, and R Schakelaar. 1989. "The Labor-Market Position Of Dutch Sociologists-An Investigation Guided By A Theoretical-Model." Netherlands Journal of Social Sciences 25 (2):57-74.

Green, Francis, and Golo Henseke. 2017. “Graduates and ‘graduate jobs’ in Europe: a picture of growth and diversification.” Centre for Global Higher Education working paper series. Working Paper Nr. 25.

Hango, Darcy William. 2013. Gender differences in science, technology, engineering, mathematics and computer science (STEM) programs at university: Statistics Canada= Statistique Canada.

Hernán, Miguel A. "The C-Word: Scientific Euphemisms Do Not Improve Causal Inference From Observational Data." AJPH 108 (5):3.

Kaiser, Henry F, and John Rice. 1974. "Little jiffy, mark IV." Educational and psychological measurement 34 (1):111-7.

Kaufman, Alan S, and Elizabeth O Lichtenberger. 2005. Assessing adolescent and adult intelligence: John Wiley & Sons.

Kirkeboen, Lars J, Edwin Leuven, and Magne Mogstad. 2016. "Field of study, earnings, and self-selection." The Quarterly Journal of Economics 131 (3):1057-111.

Klein, Markus. 2010. "Mechanisms for the Effect of Field of Study on the Transition from Higher Education to Work."

Morgan, James, and Martin David. 1963. "Education and income." The Quarterly Journal of Economics 77 (3):423-37.

Pearl, Judea. 2000. Causality : models, reasoning, and inference. Cambridge, U.K. ; New York: Cambridge University Press.

Schramm, Michael, and Christian Kerst. 2009. "Berufseinmündung und Erwerbstätigkeit in den Ingenieur- und Naturwissenschaften." HIS Projektbericht.

Spence, Michael. 1978. "Job market signaling." In Uncertainty in Economics, 281-306. Elsevier.

Stanley, Becker Gary. 1983. Human capital: A theoretical and empirical analysis, with special reference to education: University of Chicago Press.

Statistisches Bundesamt. 2018a. "Zahlen und Fakten Studierende." Accessed May 18. https://www.destatis.de/DE/ZahlenFakten/Indikatoren/LangeReihen/Bildung/lrbil01.html.

Statistisches Bundesamt. 2018b. "Studierende: Deutschland, Semester, Nationalität, Geschlecht, Studienfach" Accessed August 25. https://www-genesis.destatis.de/genesis/online/logon?sequenz=tabelleErgebnis&selectionname=21311-0003.

Thurow, Lester C. 1975. Generating inequality: Basic books.

———. 1979. "A job competition model." Unemployment and inflation:17-32.

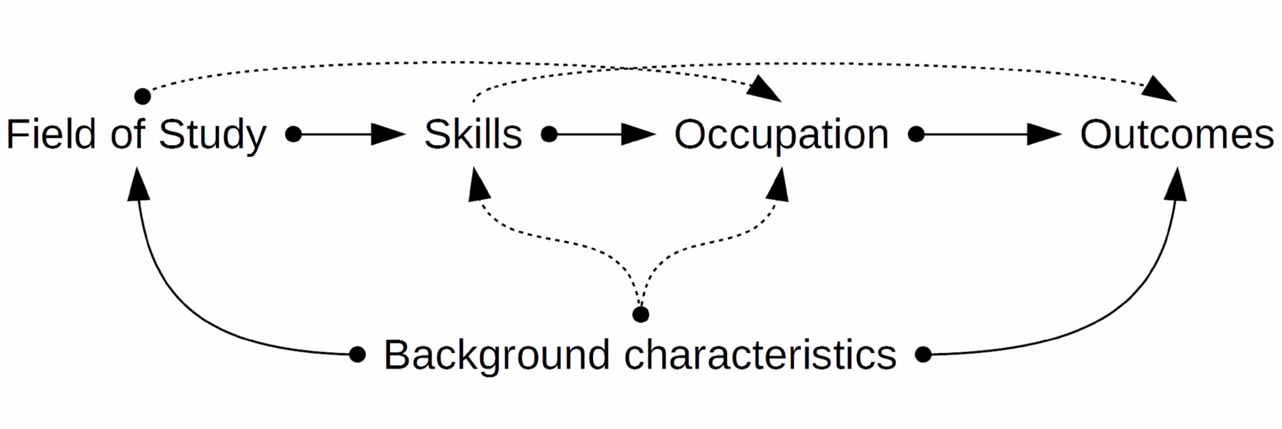
Tables

Table 1: Categorization of STEM

Category	Name of the Study
Sciences	Physics, astronomy; chemistry; pharmaceutics; biology; geosciences; health sciences
Technology	Land conservation; agricultural sciences; forestry; nutrition sciences; electrotechnology; traffic and transportation, nautical sciences; area planning; survey and mapping; mining industry, metallurgy
Engineering	Engineering (general); machine construction, process engineering; civil engineer
Mathematics	Mathematics, natural sciences (general); statistics; computer sciences

Source: DZHW Graduate Panel 2005/06. German names were not translated to keep the comparison to the pre-defined groups by the DZHW. Translations of groups by the author.

Figure 1: Theoretical framework



Source: own design.

Table 2: Principal-component-analysis

Original Item	**General skills**	**Leadership skills**	**Self-organization**	**Intercultural skills**	**Specific skills**
Special Knowledge	0.101	0.111	-0.164	**-0.226**	**0.545**
General Knowledge	**0.289**	-0.084	-0.027	0.007	0.179
Scientific Methods	0.089	-0.151	0.023	0.008	**0.459**
Foreign Languages	0.022	-0.108	-0.068	**0.602**	-0.032
Communication skills	**-0.242**	0.163	0.146	0.180	0.189
Negotiation skills	-0.134	**0.462**	-0.022	-0.039	0.122
Organization skills	-0.081	0.003	**0.476**	0.028	-0.055
IT Knowledge	**0.262**	-0.071	-0.076	**0.414**	-0.086
Flexibility	0.147	0.008	**0.251**	0.166	-0.101
Written skills	-0.072	-0.018	-0.052	**0.336**	**0.324**
Oral skills	-0.155	0.153	-0.010	0.145	**0.423**
Closing knowledge gaps	0.145	-0.100	**0.226**	-0.059	**0.256**
Leadership skills	0.036	**0.458**	-0.012	-0.085	-0.006
Economic knowledge	**0.218**	**0.449**	**-0.417**	0.034	0.036
Cooperation skills	0.023	0.050	0.188	0.189	-0.037
Time management	0.012	-0.040	**0.528**	-0.150	-0.081
Knowledge application	**0.387**	0.044	0.032	-0.020	0.113
Interdisciplinary thinking	**0.387**	0.116	-0.053	0.057	-0.046
Intercultural knowledge	0.065	0.130	-0.043	**0.394**	-0.147
Independent working skills	**0.214**	-0.097	**0.265**	-0.019	0.087
Taking responsibility	0.056	**0.228**	**0.265**	-0.076	-0.035
Conflict solving	0.055	0.422	0.044	-0.096	-0.040
Problem solving	**0.319**	0.152	0.071	-0.021	0.006
Analytical skills	**0.433**	-0.007	-0.071	0.099	0.054

Source: DZHW Graduate Panel 2005/06. Absolute factor loadings larger than 0.20 are highlighted.

Table 3: Descriptive statistics

	Science	Technology	Engineering	Mathematics	F-Value
Female %	0.669	0.479	0.276	0.345	37.51
	(0.471)	(0.501)	(0.448)	(0.476)	
Degree: Diplom %	0.649	0.751	0.939	0.591	30.25
	(0.478)	(0.433)	(0.240)	(0.493)	
Academic background %	0.620	0.607	0.658	0.640	0.83
	(0.486)	(0.489)	(0.476)	(0.481)	
Worked before study %	0.205	0.280	0.204	0.222	1.71
	(0.404)	(0.450)	(0.404)	(0.416)	
Vocational Training before study %	0.094	0.233	0.158	0.103	9.86
	(0.293)	(0.424)	(0.366)	(0.305)	
Never started or canceled Ph.D. %	0.276	0.739	0.689	0.665	65.10
	(0.448)	(0.440)	(0.464)	(0.473)	
GPA (Abiturnote)	1.882	2.272	2.168	1.962	24.72
	(0.597)	(0.581)	(0.631)	(0.551)	
Importance of labour market	3.396	3.401	3.092	2.901	12.35
	(1.085)	(1.224)	(1.155)	(1.165)	
General skills	-0.190	-0.182	0.112	0.464	13.01
	(1.526)	(1.580)	(1.461)	(1.571)	
Leadership skills	-0.237	0.206	0.006	0.145	6.14
	(1.782)	(1.701)	(1.720)	(1.680)	
Self-organization	0.313	0.045	-0.317	-0.142	4.74
	(1.700)	(1.731)	(1.791)	(1.631)	
Intercultural skills	-0.115	-0.007	0.021	0.365	7.18
	(1.487)	(1.443)	(1.582)	(1.385)	
Specific skills	0.173	-0.066	-0.121	0.012	0.97
	(1.337)	(1.469)	(1.367)	(1378)	
Log Hourly Wage Wave 2	2.847	2.991	3.202	3.091	34.45
	(0.477)	(0.408)	(0.297)	(0.379)	
Log Job Mismatch Wave 2	0.423	0.596	0.491	0.516	7.80
	(0.415)	(0.464)	(0.386)	(0.452)	
Ever unemployed %	0.373	0.385	0.357	0.281	2.81
	(0.484)	(0.488)	(0.480)	(0.450)	
N	308	257	196	203	

Source: DZHW Graduate Panel 2005/06, own calculations. Standard errors in parentheses. F-values based on one-way ANOVAS.

Table 4: All binary contrasts, outcome: logged wages

Model		S-T	S-E	S-M	T-E	T-M	E-M
M1	Coef.	0.278***	0.346***	0.317***	0.098*	0.019	-0.059
	SE	0.052	0.049	0.056	0.039	0.041	0.064
	N	421	395	325	362	348	223
M2	Coef.	0.280***	0.339***	0.292***	0.075#	-0.006	-0.095
	SE	0.052	0.051	0.059	0.040	0.043	0.067
	N	421	395	325	362	348	223
M3	Coef.	0.185***	0.266**	0.298***	0.066	0.002	-0.082
	SE	0.055	0.055	0.059	0.040	0.050	0.066
	N	421	395	325	362	348	223

Source: DZHW Graduate Panel 2005/06, own calculations.

$^{\#}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: All binary contrasts, outcome: logged mismatch

Model		S-T	S-E	S-M	T-E	T-M	E-M
M1	Coef.	0,051	-0,065	-0,007	-0,056	-0,020	0,029
	SE	0,051	0,049	0,054	0,046	0,055	0,079
	N	470	432	356	392	376	238
M2	Coef.	0,036	-0,070	-0,013	-0,042	-0,003	0,046
	SE	0,052	0,051	0,058	0,047	0,056	0,082
	N	470	432	356	392	376	238
M3	Coef.	-0,040	-0,132*	-0,054	-0,014	-0,032	0,023
	SE	0,054	0,057	0,059	0,050	0,064	0,083
	N	470	432	356	392	376	238

Source: DZHW Graduate Panel 2005/06, own calculations.

$^{\#}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: All binary contrasts, outcome: Risk of every being unemployed

Model		S-T	S-E	S-M	T-E	T-M	E-M
M1	AME	-0,076	-0,172***	-0,129*	-0,016	-0,039	-0,071
	SE	0,054	0,053	0,057	0,052	0,054	0,079
	N	476	436	361	393	377	238
M2	AME	-0,068	-0,160**	-0,141*	-0,019	-0,046	-0,016
	SE	0,055	0,056	0,059	0,052	0,055	0,084
	N	476	436	361	393	377	238
M3	AME	-0,073	-0,127$^{\#}$	-0,128*	0,046	-0,080	-0,031
	SE	0,060	0,066	0,063	0,057	0,065	0,084
	N	476	436	361	393	377	238

Source: DZHW Graduate Panel 2005/06, own calculations.

$^{\#}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$