Kai Mühleck* and Ulrike Schwabe**

Challenging the intuition: Is a same-gender supervisor beneficial for doctoral students?***

Abstract: It continues to be a puzzle that women are disproportionally often dropping out of academic careers. Researchers and policymakers have suggested that same-gender supervisors are important for tightening this 'leaky pipeline'. Especially in subjects with a strong overrepresentation of men, it seems likely that female supervisors work as positive role models and help preventing discrimination. Anticipating this effect, female doctoral students might also prefer supervisors of the same gender.

Therefore, we ask how widespread a gender match is between doctoral student and supervisor in Germany and whether a gender match between supervisors and doctoral students is beneficial for the doctorate and for a possible scientific career thereafter. For our data we draw on the first survey of the 'German National Academics Panel Study (2018)'; to address causality concerns we apply entropy balancing for our estimations.

Our analyses confirm that both female and male doctoral students are more likely to have a supervisor of the same gender. Furthermore, results show that female supervisors have a positive effect on satisfaction with mentoring and academic self-concept for both female and male doctoral students.

Keywords: Doctoral students, gender-match, Nacaps, same-gender, supervisor, scientific career

Eine Herausforderung für die Intuition: Sind Betreuende gleichen Geschlechts für Promovierende von Vorteil?

Zusammenfassung: Es ist nach wie vor nicht gänzlich klar, warum Frauen überproportional häufig aus der akademischen Karriere ausscheiden. Wissenschaftler:innen und politische Entscheidungsträger:innen haben die Vermutung geäußert, dass

* Dr. Kai Mühleck (corresponding author), German Centre for Higher Education Research and Science Studies (DZHW), Lange Laube 12, 30159 Hannover, E-Mail: muehleck@dzhw.eu.

** Ulrike Schwabe, German Centre for Higher Education Research and Science Studies (DZHW), Lange Laube 12, 30159 Hannover, E-Mail: schwabe@dzhw.eu.

*** We thank the two guest editors and two anonymous reviewers for valuable comments on an earlier version of the manuscript. We are indebted to Frauke Peter for her particularly helpful hints on our estimation strategy. Mike Mielcarski and Lisa-Marie Steinkampf provided valuable help as student research assistants.

Betreuende gleichen Geschlechts wichtig wären, um die sogenannte „leaky pipeline“ zu schließen. Vor allem in Fächern, in denen Männer stark überrepräsentiert sind, könnten Betreuerinnen als positive Vorbilder fungieren und dazu beitragen, Diskriminierung zu verhindern. In Erwartung dieses Effekts könnten weibliche Promovierende auch weibliche Betreuende bevorzugen. Vor diesem Hintergrund fragen wir, wie verbreitet es in Deutschland ist, dass Promovierende und Betreuende das gleiche Geschlecht haben und ob dies für die Promotion und für eine mögliche anschließende wissenschaftliche Karriere vorteilhaft ist. Als Datengrundlage verwenden wir die erste Befragung der „National Academics Panel Study (2018)“. Um das Problem der Kausalität der Zusammenhänge zu adressieren, verwenden wir entropy balancing für unsere Schätzungen. Unsere Analysen bestätigen, dass sowohl weibliche als auch männliche Promovierende mit höherer Wahrscheinlichkeit Betreuende desselben Geschlechts haben. Darüber hinaus zeigen die Ergebnisse, dass Betreuerinnen sowohl bei weiblichen als auch bei männlichen Promovierenden einen positiven Effekt auf die Zufriedenheit mit der Betreuung und das akademische Selbstkonzept haben.

Stichworte: Betreuende; gleichgeschlechtlich; Nacaps; Promovierende; wissenschaftliche Karriere

1 Introduction

In the course of educational expansion, the representation of women in academia has considerably increased in recent decades, with women outnumbering men among entrants to higher education as well as among higher education graduates for most degrees, in most OECD countries (OECD 2020a, b). As a consequence, today the share of tertiary-educated women within the working-age population in the majority of OECD countries is larger than the share of tertiary-educated men (OECD 2020c). In fact, many countries have started promoting higher education among men (OECD 2019a) with a view to redressing the balance.

While women are close to having reached parity among doctoral graduates (OECD 2019b), they are underrepresented at higher levels of the academic career such as among university teachers (OECD 2020c). To some extent, this difference certainly reflects ‘historical’ gender-inequalities. However, even 10 years ago women had almost reached parity among doctoral graduates (OECD 2012) and among first-degree graduates they have now outnumbered men for at least one and a half decades (OECD 2008).

Thus, it is unlikely that persisting gender inequalities are exclusively due to student cohorts with a female majority not yet having reached these levels. Studies on countries such as Switzerland (Schubert/Engelage 2011) and Germany (Lörz/Mühleck 2019), with particularly low proportions of female professors (OECD 2020c; Konsortium Bundesbericht Wissenschaftlicher Nachwuchs 2021, Russ 2021), have corroborated that at each step of the academic career the share of women dropping

out continues to be greater than the share of men. This phenomenon has been described with the catchy metaphor of a 'leaky pipeline' in education and in science (e.g., Berryman 1983; Alper 1993). Thus, equal participation of women and men, especially in leading positions, in academia is still an important subject of higher-education research and remains on the agendas of higher-education policymakers and professionals (e.g., Cheung 2021, BMBF 2021a 2021b, Forschung und Lehre 2020).

Despite its importance, the leaky pipeline phenomenon remains a puzzle. Researchers and policymakers have suggested that same-gender supervisors are important for fostering the academic careers of women. Especially in subjects with a strong overrepresentation of men, female doctoral students could be confronted with negative stereotypes, distorted perceptions of their performance, and less academic integration, resulting in e.g., less satisfaction, lower self-esteem or even dropout. It seems likely that female supervisors lessen these negative effects and work as positive role models (Kanter 1977, Hirshfeld 2010, Solanki/Xu 2018). Anticipating this effect, female doctoral students might also prefer supervisors of the same gender. More generally, it has been supposed that supervisors show more understanding towards students of the same gender and that cooperation with them is more enjoyable (Gaule/Piacentini 2017). Thus, we intuitively assume that a same-gender supervisor is beneficial for doctoral students, be they male or female.

While such thoughts seem initially compelling, empirical evidence is mixed and differs substantially by field of study (e.g., Edmunds 2016; Gaule/Piacentini 2017; Hilmer/Hilmer 2007, Neumark/Gardecki 1998, Solanki/Xu 2018). There is a considerable body of empirical research referring to the United States but, to the best of our knowledge, there exists as yet no study for the German case. Moreover, little is known about the social mechanisms behind the association of a gender match and an academic career. Mostly, empirical research either does not or cannot tackle the question of the causality of this association (an exception is Carrell et al. 2010). Against this backdrop we ask (i) how widespread gender-matching is between doctoral students and supervisor in Germany and (ii) whether gender-matching between doctoral student and supervisor is beneficial for a successful doctorate and a possible scientific career afterwards.

The paper is structured as follows: Firstly, we give an overview of previous research. Thereafter, we present theoretical and conceptual considerations also addressing the question of causality. This is followed by a description of our database, the 'German National Academics Panel Study (Nacaps)' and our analytical strategy. We then present bivariate results and multivariate results of regressions using entropy-balancing weights. We close with a summary and discussion including limitations, future research avenues, and policy implications.

2 Previous research

Looking at the existing literature, a couple of specific focuses, imbalances, and lacks of research come to the fore. Firstly, most of the research comes from the United States and also focuses on the United States, i.e., for reviewing the state of empirical research, we did not intentionally focus on the United States, but there are hardly any studies on the effect of gender-matching supervisors that refer to other countries. Moreover, analyses focusing on doctoral students only are scarce and therefore, we also consider studies on students or graduates at bachelor's and master's levels. Finally, most studies focus on a specific field of study and rarely fields of study are compared to each other. We will use this feature of previous research and structure our brief literature review along the lines of fields of study.

Looking at STEM subjects, first, evidence from the United States clearly supports the positive effect of same-gender faculty, especially for women. Doctoral students of chemistry tended to pick same-gender advisors and both male and female students with a same-gender advisor were more productive and more likely to become professors themselves (Gaule/Piacentini 2017). These positive effects of a gender-matching advisor were greater for female doctoral students then for the males (Gaule/Piacentini 2017). Female doctoral students in STEM subjects had a higher chance of graduating if they had a female advisor or if a relatively large proportion of the faculty at their institute was female; for male doctoral students no such effect was observed (Main 2018). Female bachelor students achieved better grades in facultative math and science classes if they were taught by women (Carrell et al. 2010). Moreover, the proportion of female faculty in introductory math and science courses was found to be positively associated with female bachelor students choosing further math and science classes as well as with going for a master's in STEM subjects (Carrell et al. 2010). Female students were less active in STEM classes then their male peers and less often asked for help; this difference lowered if courses were taught by women (Solanki/Xu 2018). Grades of students were generally lower if instructors were female, but this disadvantage lessened if students were female as well (Solanki/Xu 2018). Female students had a lower subject-specific self-efficacy; female instructors did not have a significant effect on this difference (Solanki/Xu 2018). Interestingly, the studies of Solanki and Xu (2018) and Carrell et al. (2010) reported that the (relative) positive effect of female faculty on the performance of female students was particularly large for the highly-skilled, i.e., those that may be suited to an academic career.

Regarding the field of medicine, Edmunds et al. (2016) reviewed 52 studies (most of them referring to the United States) on the question of why women are less likely than men to pursue an academic career. Many studies reported that women had more problems than men in finding adequate mentors and also that women had difficulties in finding gender-matching mentors. One study, however, found that both female and male students thought that the other gender had better

mentoring (Edmunds et al. 2016) Some studies reported that female students were more likely to choose advisors of lower rank and that they valued a supportive relationship with the mentor more than the reputation of the mentor (Edmunds et al. 2016). There was also some evidence that women might have specific mentoring needs (Edmunds et al. 2016). These findings suggest that the career outlook of female doctoral students might benefit from more female advisors. However, to our knowledge no study on medicine has so far directly investigated the effect of a student-advisor gender match.

In contrast to STEM fields, studies on doctoral students in economics revealed no clear evidence for a positive effect of same-gender advisors. Neumarck and Gardecki (1998) found that female doctoral students of economics in the United States had slightly higher completion rates and graduated more quickly at institutes with higher numbers of female faculty. Numbers of female faculty, however, did not shorten the time till first job placement for women or result in higher chances of securing a first job at a PhD-granting institute. This might be due to the fact that institutes with larger numbers of female faculty were also lower tier institutions. What's more, the gender of the dissertation chair for female doctoral students had no significant effect on any of the outcome variables (Neumarck/Gardecki 1998). About 10 years after Neumarck and Gardecki's study, Hilmer and Hilmer (2007) did another study focusing on U.S. doctoral students in economics. Surprisingly, they found a positive gender-mismatch effect in the sense that female students with male advisors were more likely to attain a research-related first job then male students with male advisors. Looking at female students only, the gender of advisors had no significant effect. Generally, female students issued fewer publications then male students with male advisors. This was associated with female students being more likely to enrol in programs with less reputation and to pick dissertation advisors of lower rank (Hilmer/Hilmer 2007). Hilmer and Hilmer (2007) assume that economics is lacking female 'star-advisors' that could (additionally) push careers of female doctoral students. While this may be the case, all in all, the results of both studies suggest that the student-advisor gender match has little or no effect on the academic career outlook of doctoral students in economics.

An exception with respect to field of study is the paper of Bettinger and Long (2005) covering first-year students of colleges with a range of subjects. Using longitudinal data, they analyze the impact of having female faculty members in initial courses on additional course attendance, the overall number of credit points, and the choice of the major for female students. Overall, the results indicate some positive effects of matched gender for female students. It turns out that the effects of having female instructors for female students' outcomes vary significantly between the subjects, without a clear pattern, however. In contrast, focusing on male students in female-dominated fields, findings show strong positive effects of having male instructors for the acquisition of credit points and choice of major for male students in education. Despite the fact that this study is not restricted to

certain subjects, it also provides a sophisticated estimation strategy. Based on the argument that selection into initial courses is far from random, an instrumental variable approach is applied to deal with endogeneity. The term-specific variation in the likelihood of female-taught courses functions as a valid instrument to capture selection into courses based on students' gender preferences.

All in all, the literature shows that advisors of the same gender are generally preferred. There is some evidence supporting the claim that female advisors have positive effects on the study results and career prospects of female students. Evidence further indicates that a gender match has positive effects, generally with possibly greater effects for women than for men. However, results differ quite strongly across subjects, e.g., contrasting between economics and STEM. Moreover, other socio-demographic characteristics, e.g., race or ethnicity (Alston et al. 2017; Riegle-Crumb et al. 2020), seem also to be relevant.

By summarizing the state of the empirical literature, we see research gaps in a couple of aspects; the *first aspect* is the restricted geographical scope. Empirical evidence almost exclusively refers to the U.S.-American context. This raises the question of whether the U.S. results can be generalized towards other countries. To the best of our knowledge there exists as yet no empirical study on German higher education in general or on German doctoral students in particular (for secondary education in Germany, see Helbig 2012, Neugebauer et al. 2011). However, Germany seems to be an interesting case. The share of women among professors is relatively low compared to other European countries (European Commission 2021) and this continues to be raised as a pressing challenge on the political agenda (most recently e.g., Konsortium Bundesbericht Wissenschaftlicher Nachwuchs 2021). At the same time, Germany addresses this issue with policy measures and, which, in fact, do seem to contribute to recent improvements (Löther 2019).

The *second aspect* refers to the selected set of subjects analyzed, such as STEM fields. Often, only one subject is covered, and thus, the question of generalization of findings to other subject/discipline-specific contexts also arises. Referring to differences among subjects, it is still an open empirical question, how single subjects differ from the 'average effect' of gender matching across all subjects.

Such analyses require large-scale surveys with a sufficient number of observations, the *third aspect* that we have detected. Existing evidence is often based on administrative data (e.g., Bettinger/Long 2005; Carrel et al. 2010; Gaule/Piacentini 2017; Hilmer/Hilmer 2007; Neumarck/Gardecki 1998), on smaller local surveys (e.g., Riegle-Crumb et al. 2019) or qualitative data (e.g., Alston et al. 2017; Hirshfield 2010). While administrative data usually provide sufficient samples sizes, they lack subjective evaluations like motivation to obtain a doctoral degree, relationship to supervisor or satisfaction with mentoring during doctoral studies. Since suitable data at the national level are already rare, international comparisons are currently not possible at all.

A *fourth aspect* is the obvious problem of endogeneity. Students selecting into a mentoring relationship with an advisor of a specific gender is not a random assignment. Thus, factors like e.g., goals or personality traits may drive this self-selection process and may at the same time influence the academic career trajectories. Only a few studies referred to in the literature review explicitly address the issue of gender matching being an endogenous variable and that 'selection into treatment' may be due to unobserved variables. However, this is highly relevant for causal reasoning. By design, experimental studies are a good way to fully debilitate the (self-)selection concerns but, of course, such experiments would be hard to accomplish and probably immoral. Only one study that we are aware of made use of a natural experiment; Carrell and colleagues (2010) conducted research on students at the U.S. Air Force Academy from graduating classes 2001 to 2008. In this institute, students are randomly assigned to professors in required core courses. Since all faculty members use the same syllabus and test scores, equivalence in teaching has been ensured. The findings indicate only small effects of professor's gender for male students' performance, but, substantial effects for female students, especially in math and science. Since students can usually not be randomly assigned into courses or to supervisors, ex-post estimation approaches are necessary when using survey data. As described above, Bettinger and Long (2005) used an instrumental variable approach to meet the objection of selection on unobservables. As described below, we will use entropy balancing to account for systematic differences between students with and without a gender-matching advisor.

Finally, the *fifth aspect* concerns theoretical considerations. Many studies start from the assumption that a gender match would have a positive effect on educational or academic careers. While such a correlation hypothesis seems intuitively compelling, without some theoretical considerations it is unclear why a gender match should have such an impact, i.e., which social mechanisms are at stake. A more sophisticated theoretical framework could strongly contribute to strengthening our understanding of the social mechanisms and also help in a causal interpretation of findings. However, there is no established theoretical framework telling us why a gender match should impact educational or academic careers. Developing such a framework clearly goes beyond the scope of this paper, but we will in the next step present theoretical considerations leading to several testable hypotheses.

3 Theoretical considerations and hypotheses

Following our intuition, gender matching is beneficial for doctoral students: Doctoral students and supervisors of the same gender might get along with each other better, leading to stronger, more trustful and enjoyable relationships. Same-gender advisors may better understand gender-specific problems such as combining family responsibilities with doctoral studies. But besides intuitive reasoning, why should that really be the case? Why should gender matching lead to positive student outcomes during doctoral studies?

One line of reasoning is the 'theory of proportions' (Kanter 1977a, 1977b) and, related to that, the 'identity threat' (Hirshfield 2010). Both argue that it is the number of females compared to their male peers within a given context, 'skewed groups' (85:15 ratio of majority to minority; Kanter 1977a), that lead to stressful and challenging work environments. According to Kanter, members of small minorities (so-called 'tokens') stick out, are confronted with negative stereotypes by the majority, and are exposed to negative expectations that they would need to disprove. An example related to the topic of this paper could be female doctoral students in subjects where women form small minorities, e.g., in engineering. As a consequence of the social mechanisms described, performance of female doctoral students would be perceived more critically by the majority group. Moreover, the majority tends to maintain borders between groups, i.e., women would not be included in scientific networks to the same extent as men. In sum, this would lead to lower motivation and productivity among female doctoral students and poorer academic career prospects. A core assumption is that doctoral students anticipate this situation and thus choose a supervisor belonging to their own minority group, i.e., a female supervisor, in our example. A female supervisor could mitigate the 'identity threat' and the discrimination that goes with it. Obviously, this argumentation only holds if women form a small minority group. Consequently, the effect of a gender match would strongly depend on the share of females and males, both for doctoral candidates as well as for supervisors, in each respective field of study. Therefore, building on the tokenism theory, we would expect the positive effect of a gender match for women in male-dominated fields to be larger.

Another social psychological explanation, leading to similar conclusions is the 'identity-based motivation theory' by Oyserman (2007, 2009; for an application of this theory to gender matching see Solanki/Xu 2018). The core argument is that during higher education in general and the doctorate in particular students develop their academic identities, which help them to act and react in the academic world. During this process of identity-building, advisors, mentors and supervisors serve as important *role models*. If these role models have the same socio-demographic characteristics as the student, e.g., socioeconomic background, gender, ethnicity or race, it is much easier for the student to establish an analogue identity, which would then be in line with an academic career. Such a congruent academic identity leads to higher motivation, better academic performance, and developing a resilient personality to overcome difficulties in the academic system. In other words: Same-gender supervisors may be better suited to serve as role models, thus giving encouragement to same-gender students and being examples for how to pursue an academic career in the field—as a woman or as a man. Again, one might suspect that role models are especially relevant in environments where the specific role is less-established and few examples exist, i.e., female role models of being a professor could be more relevant in male-dominated fields.

A different argument could be made using theories of social networks. McPherson and colleagues (2001) argue that networks are often built following the 'homophily principle', i.e., ties between sociodemographically similar people are formed more often and are more stable. This would mean that female doctoral students could make better use of the network of a female supervisor and male doctoral students could make better use of the network of a male supervisor and therefore a gender match would again be beneficial. The homophily in academic networks might lead to the reproduction of gender pattern in the science system.

Network theory identifies factors which make networks more beneficial for their members, amongst others "the size of the network […] and resources of the tie" (Forret 2006: 151). If so, the social tie in a female-dominated (or rather, less male-dominated) field of study would be more beneficial than in a male-dominated subject. It would grant access to a larger network, and, in a field with a more balanced gender composition, female professors are more likely to have already reached outstanding positions associated with especially high resources.

While these explanations are based on coherent theoretical models, the literature additionally provides assumptions that do not belong to any parent theoretical framework. These arguments either refer to concrete and gender-specific behavior or to differences in productivity between men and women in academia.

With respect to behavior, male and female supervisors may differ in their specific mentoring styles, and, in turn, male and female doctoral students may differ in their specific mentoring needs (Gaule/Piacentini 2017). Supervisors of the respective gender might show more understanding for these gender-specific needs in mentoring and, e.g., support reconciling work and family life (Bettinger/Long 2005; Etzkowitz et al. 1994). These challenges may take on different scope and forms for male and female students, depending on gendered family roles (Lörz/Mühleck 2019).

Quite generally and intuitively comprehensible, one could assume that cooperation between mentors and students of the same gender could be more pleasant (Gaule/Piacentini 2017) which would ease work, add to motivation, and could thus promote the satisfaction and success of doctoral students.

With respect to productivity, male and female supervisors may differ in scientific reputation, productivity in terms of research output as well as in their status within organizations, e.g., being dean of a faculty (Etzkowitz et al. 1994; Gaule/Piacentini 2017; Hilmer/Hilmer 2007; Jaksztat 2017). Due to seniority, it seems likely that male advisors, on average, have a higher reputation, more resources and larger networks. If doctoral students do prefer advisors of the same gender, this would lead to differences in access to academic resources being dependent on the supervisor's gender. A gender match might thus have different consequences for the career prospects of male or female doctoral students, if, on average, a male supervisor

could grant access to more resources that are relevant for advancing an academic career.

A last argument, from an economic perspective, is directed to the supply of supervisors within doctoral subjects. Gaule and Piacentini (2017) argue that one reason for the surprisingly persistent gender gap at higher levels of the academic career is the overrepresentation of male doctoral advisors, specifically in fields like science and engineering. In this view, the underrepresentation of women in faculty positions may perpetuate itself through a lower availability of same-gender mentors for young female researchers. Likewise, an overrepresentation of women in faculty positions in specific fields could start reproducing itself through a lower availability of same-gender mentors for young male researchers.

Based on these theoretical considerations, we derive the following five hypotheses.

Firstly, several theoretical arguments lead us to expect a general preference among doctoral students for gender-matching supervisors. Female students in male-dominated fields could prefer female supervisors to avoid tokenism in male-dominated subjects. They might more generally tend to choose female supervisors to learn from a role model. Doctoral students of both genders might prefer supervisors of the same gender due to expecting this to be a more pleasant working relationship.

When testing these theoretical assumptions, we face the problem that we don't know whether the students have chosen their supervisors, or the supervisors have chosen their students. Our data unfortunately tells us relatively little about the process of how students and supervisors have selected each other. The form of doctorate is likely to influence this; looking at the different forms, we argue that student preferences do have a certain impact, even though the strength of the impact may vary.

The most prevalent form of doctorate in Germany (accounting for close to half of the students in our sample) is that of doctoral students being employed as researchers. In such cases, the supervisor and the superior are often (not always) one and the same, and therefore the supervisor has chosen the student by hiring them. At the same time, the student has decided to apply for the job or at the very least to accept the job offer. In contrast, doctoral students in structured programs or freely pursuing their doctorates (together these two forms account for slightly less than half of the students in our sample) often take the initiative and approach the professor of their choice, asking to be supervised; professors usually accept such a request. At the same time, supervisors may have encouraged being approached. Finally, a smaller share of doctoral students in Germany has scholarships without pursuing a structured program. They could have approached supervisors on their own initiative, or the supervisor could have encouraged the student to apply for a scholarship. In sum, forming a couple made up of doctoral student and supervisor is sometimes driven by the preferences of the student and sometimes rather by the

preferences of the supervisor. However, it seems very unlikely that the preferences of students have no effect. Therefore, if students tend to prefer supervisors of the same gender, this would, *ceteris paribus*, lead to a higher prevalence of gender-matching combinations between students and supervisors.

Clearly the prevalence of gender-matching combinations may also be driven by opportunity. Thus, it is important to check whether gender-matching combinations are still more likely if we distinguish between fields of study.

While we cannot test whether doctoral students prefer supervisors of the same gender, students are arguably always involved in the choice, which would, based on the theoretical arguments above, lead us to expect that the share of doctoral students with a supervisor of the same gender is disproportionally higher, i.e., female doctoral students would have a larger share of female supervisors than the overall share of female supervisors and male doctoral students would have a larger share of male supervisors than the overall share of male supervisors.

Hypothesis 1a: The share of doctoral students with a supervisor of the same gender is disproportionally higher.

Kanter's theory of proportions leads to a more specific hypothesis in this regard.

Hypothesis 1b: The share of female doctoral students with a supervisor of the same gender is disproportionally higher especially in male-dominated subjects.

Secondly, we expect that a same-gender supervisor has positive effects on a students' doctorates in various respects, i.e., that students are generally more satisfied with mentoring, that they build more academic self-esteem or self-efficacy, and that they are more optimistic about their academic career prospects after graduation.

Hypothesis 2: Doctoral students with a gender-matching supervisor are more satisfied with supervision.

Hypothesis 3: Doctoral students with a gender-matching supervisor believe more strongly in their own research abilities.

Hypothesis 4a: Doctoral students with a gender-matching supervisor are more optimistic about their career prospects in academia.

As described above, male supervisors may provide access to larger networks and resources, may have a higher reputation in the scientific community or may be more productive (e.g., due to age). This leads to a hypothesis 4b which, in contrast to hypothesis 4a, assumes that the gender-match effect differs between male and female doctoral students.

Hypothesis 4b: Male doctoral students with a gender-matching supervisor are more optimistic about their career prospects in academia and female doctoral students with a gender-matching supervisor are more pessimistic about their career prospects in academia.

Finally, the share of female doctoral candidates and also the share of female supervisors substantially vary across subjects. Tokenism theory and the identity-based motivation theory suggest that a gender match is especially beneficial for women in male-dominated subjects. According to tokenism theory we would expect discrimination against women especially in male-dominated fields, and a female supervisor could limit such discrimination. Moreover, she could serve as a role model, which is more important in fields where such role models are rare. In contrast, applying network theory, one could argue that social ties to a larger network with a wealth of resources are more beneficial than social ties to a small network with fewer resources. Thus, a female supervisor in a field offering a larger network of other female professors could be more beneficial to students' careers than a female supervisor in a male-dominated field. This reasoning leads to two conflicting hypotheses on subject-specific differences in the effect of a gender match for female doctoral students:

Hypothesis 5a: Female doctoral students in fields of study with a relatively *low* proportion of women benefit more strongly from a gender-matching supervisor.

Hypothesis 5b: Female doctoral students in fields of study with a relatively *high* proportion of women benefit more strongly from a gender-matching supervisor.

4 Data and methods

4.1 Data and measures

We use data from the 'German National Academics Panel Study (Nacaps)' on a recent cohort of doctoral candidates that were registered for doctoral studies in December 2018 at German higher education institutions (Briedis et al. 2020, Briedis et al. 2022).[1] The data of this initial cohort 2018 comprises all doctoral subjects, different forms of doctorate—e.g., being employed at a university or a research institution, getting a grant—and different stages, from just registered through to almost finished. Within the entire study design, this cohort is an exception as it presents a cross-section of all doctoral candidates registered as of 1st December 2018. A follow-up cohort was interviewed two years later; the Nacaps cohort 2020, however, only considers those doctoral candidates that had been newly registered in the interceding two years. Generally, Nacaps is designed as a multi-cohort panel study including multiple measurement points for each respondent in a given cohort (for more details see Briedis et al. 2022).

1 The scientific use file of Nacaps 2018, first wave, is available via the Research Data Centre of the German Centre for Higher Education Research and Science Studies (FDZ-DZHW): Adrian, D., Ambrasat, J., Briedis, K., Friedrich, C., Fuchs, A., Geils, M., Kovalova, I., Lange, J., Lietz, A., Martens, B., Redeke, S., Ruß, U., Sarcletti, A., Schwabe, U., Seifert, M., Siegel, M., Teichmann, C., Tesch, J., de Vogel, S. & Wegner, A. (2020). National Academics Panel Study (Nacaps) 2018. Datenerhebung: 2019. Version: 1.0.0. Datenpaketzugangsweg: On-Site-SUF. Hannover: FDZ-DZHW. Datenkuratierung: Weber, A., Birkelbach, R., Hoffstätter, U. & Daniel, A. https://doi.org/10.21249/DZHW:nac2018:1.0.0.

We only consider the first wave of Nacaps cohort 2018, because of the overall high number of observations—information on more than 20,000 respondents—which allows field-specific analyses (for details see Tables 3 and 4). As has been highlighted in the section on previous work, findings quite strongly differ by field of study. Together with the overall lack of empirical evidence for Germany, we therefore see the strongest contribution of our study to be in describing the phenomenon in as much detail as possible by also taking issues of self-selection into supervisor relationships and doctoral contexts into account.

Moreover, detailed analyses on data quality, for the representation side in particular, are available for this first wave of Nacaps cohort 2018 (Briedis et al. 2022). By design, Nacaps is a complete enumeration of registered doctoral candidates at all German higher education institutions that are legally allowed to award doctoral degrees. In order to be comparable with official statistics, the date of reference for sampling is 1st December of the corresponding year (HstatG § 5). In practice, however, there exists no official register for doctoral candidates in Germany. Thus, higher education institutions function as important gatekeepers for field access by contacting the target population. One result of these conditions in Germany is that coverage bias due to non-participation in the study can occur on two levels: The level of higher education institutions (comparable to primary sampling unit) and the level of doctoral candidates (comparable to secondary sampling unit). On the level of higher education institutions, larger higher education institutions are more likely to participate in the Nacaps study, whereas higher education institutions in East Germany and special types like church-sponsored higher education institutions and colleges of the arts are less likely to participate (for more details see Briedis et al. 2022). This coverage bias on the primary sampling unit does, however, not affect our analyses as long as respondents' gender and their field of study do not systematically vary from the entire population (secondary sampling unit). Indeed, comparisons with official statistics for registered doctoral candidates provided by the Federal Statistical Office indicate no systematic bias by gender and field of study due to unit-nonresponse at the level of doctoral candidates (Briedis et al. 2022, Vollmar 2019). Thus, we argue that results based on Nacaps can largely be generalized to the German population of doctoral students, although a complete enumeration as proposed by design has not been realized. Beyond this, Nacaps is unique as it provides current information on the situation of doctoral students in Germany.

For our analyses, we exclude from the entire sample those respondents stating that they have dropped out of doctoral studies at the time of the interview.[2] However, we have included those who reported only a temporary interruption. Most important for our purpose, we have information in the dataset on students' and main advisors' genders, so that we can model 'gender match' for each respondent. After

2 For purposes of transparency, our replication files can be found here: https://doi.org/10.21249/DZHW:muehleck2023:1.0.0.

listwise deletion on analytical variables, we ended up with a sample for our main analyses of 15,350 respondents from 53 German higher education institutions.[3]

Dependent Variables

We focus on three outcome variables as indicators for success during doctoral studies: (1) satisfaction with mentoring, (2) belief in one's own research abilities and (3) career prospects of obtaining a postdoc position after completing the doctorate. In this way, we cover different dimensions: Doctoral students' satisfaction with the supervision can be assumed to be strongly related to the overall satisfaction with the doctorate and thus with motivation to successfully complete the doctorate. Academic self-efficacy seems to be another important ingredient for an academic career as it is the belief that one holds the necessary abilities and talents. The perceived career prospects in academia, finally, can be assumed to be another important factor for motivating the successful candidate as they measure the belief of being able to further pursue an academic career after the doctorate. From a theoretical perspective all these three outcome variables can be assumed to be positively influenced by a gender-matching supervisor.

Table 1 shows the measurement as well as the means and standard deviations (SD) for our three outcomes.

On average, doctoral students are rather satisfied than dissatisfied with the supervision of their supervisors. The mean value of 3.73 is clearly above the neutral value of 3 and therefore on the positive side of the scale but also clearly below the value of 5 which would indicate being very satisfied.[4] With respect to group differences, some interesting results can be reported. Please note that all group differences in Table 1 are highly significant. First, male doctoral students are slightly more satisfied with mentoring than their female peers. The difference is far from dramatic but still highly significant. When comparing students with and without a gender-matching supervisor, we observe that, as expected, doctoral students with a gender-matching supervisor are more satisfied with supervision. A similar pattern emerges for the belief in one's own research abilities. Female doctoral students are less confident about their research abilities and, likewise, doctoral students with a supervisor of a different gender have slightly lower academic self-efficacy. As we will see below, women are more likely to have a supervisor of a different gender. Female doctoral students are more skeptical regarding their chances of becoming a postdoc then their male peers. For this dependent variable the gender differences

3 This way of handling missing data results in a reduced analytical sample; about 5,800 cases out of 21,100 are excluded from the entire analyses. However, we expect no systematic bias in results due to this procedure.

4 It might be that those doctoral candidates being less satisfied with their supervision or their situation during doctoral studies in general have not taken part in the survey at all. However, we cannot provide empirical evidence for this selectivity due to unit nonresponse on the level of doctoral candidates.

Table 1: Measurement and descriptive results for outcome variables

Dependent variable		Measurement and descriptives			
Satisfaction with mentoring		"How satisfied are you ... with the supervision of your PhD/ doctorate by your supervisor?", 5-point Likert scale: (1) "not at all satisfied", ..., (5) "very satisfied"			
	All	Male	Female	Gender-match	No gender-match
Mean	3.73	3.77	3.68	3.76	3.67
SD	1.18	1.16	1.20	1.17	1.20
Belief in own research abilities		"I have the necessary skills for a job in academia.", 5-point Likert scale: (1) "not at all certain", ..., (5) "very certain"			
	All	Male	Female	Gender-match	No gender-match
Mean	3.67	3.77	3.60	3.70	3.60
SD	1.07	1.02	1.11	1.05	1.09
Career prospects of obtaining a postdoc position		"How easy would it be for you personally to get ... a post-doc position in academia?", 10-point Likert scale: (1) "very difficult", ..., (10) "very easy"			
	All	Male	Female	Gender-match	No gender-match
Mean	4.90	5.21	4.57	5.00	4.67
SD	2.85	2.85	2.81	2.82	2.79

Source: Nacaps 2018, first wave. Own calculations. N = 15,350.
Note: All reported differences between groups are significant at p<0.001.

are somewhat stronger than for the other two variables (also taking into account the different scale). And, again confirming the familiar pattern, we find that students with a gender-matching supervisor evaluate their chances more optimistically.

All in all, these descriptive results show, that (1) women score less well on all three outcome variables, i.e., they could be among the factors explaining why women are more likely to drop out of an academic career. (2) Doctoral candidates with a gender-match score better on all three outcome variables, suggesting that this might indeed be a way to foster the academic career prospects of female doctoral candidates. Below we will test whether the multivariate models confirm this first descriptive impression.

Core Independent Variable

Our core independent variable is a dummy variable for a gender match indicating whether doctoral candidates' gender equals supervisors' gender. Following the Nacaps-specific concept of 'main supervisor', for male Ph.D. students this dummy

equals 1 if their (main) supervisor is a man, respectively for female Ph.D. candidates, if their (main) supervisor is a woman. As Nacaps data provides detailed information on up to three different doctoral supervisors and advisors, we defined 'gender match' based on the answer to the question 'Who is your main supervisor?'[5] According to the instruction in the questionnaire, this means the person who supervises your work in everyday life most intensively. This is not necessarily the same person who officially supervises the doctorate (in the sense of first supervisor respectively first reviewer of doctoral thesis). Reflecting typical German doctoral studies, for 63 percent of our analytical sample, however, the self-reported main supervisor equals the first reviewer of the thesis.

4.2 Analytical strategy

We are interested in the effect of a gender match between doctoral candidate and (main) supervisor on success during doctoral studies. Identifying this effect is complicated by the fact that assignment into matched or unmatched gender relationships during doctoral studies is *not random*. Students with and without a gender match may have differed systematically in characteristics relevant for our outcome variables *prior* to (self-)selection into gender-matched supervisory relationships. Claiming causality in 'simple' regression models based on cross-sectional data might therefore be misleading. We neither know all factors that account for (self-)selection into gender-matched supervisory relationships nor have we measured all factors that could be relevant.

However, entropy balancing offers a way to at least partially account for *pre-treatment* differences in the treatment and the control group also using cross-sectional data. Entropy balancing is a reweighting method for balanced samples (Hainmueller 2012, Hainmueller/Xu 2013). We are interested in the 'average treatment effect (ATE)' for doctoral candidates with the same gender as their (main) supervisor on success during doctoral studies. Thus, our treatment variable, gender match, is binary. Following the entropy balancing approach, we design a synthetic control group, those whose gender is not matched, on the basis of a wide range of observables that are in the data. Based on these observed characteristics, the control group is weighted with the purpose of being comparable to the treatment group. For designing the control group, we use all available information that captures differences between both groups *before* registering as doctoral students (see Table 2). To account for ascribed and further socio-demographic characteristics, we control for age, migration and social background, stable relationship, partner's education and employment as well as children. Further, we include self-rated health (Carstensen 2020, GESIS 2015) and personality traits such as Big Five (Schupp/

5 Further details on this specific question can be found here: https://metadata.fdz.dzhw.eu/en/questions/que-nac2018-ins1-B30.1?page=1&size=10&type=surveys&version=1.0.0. Last accessed: 21.3.2022.

Gerlitz 2014), general self-efficacy (Beierlein et al. 2012) and locus of control (Kovaleva et al. 2012). And finally, we consider grade-point average of higher education degree that allows for doctoral studies, doctoral subject (differentiating between STEM, biology, medicine, social sciences and arts),[6] form of doctorate (differentiating between employment at higher education institution or research institute, structured doctoral program and grant or free/external doctorate), reasons for obtaining a doctoral certificate, reasons that the respective higher education institution has been chosen and the desired characteristics of a job after completing the doctorate (Roach/Sauermann 2010).

As suggested by the literature, we use exactly this information as additional control variables in order to increase the precision of coefficient estimates in our subsequent analyses (Oster 2019). To be transparent on our estimation approach, we present the results for four different estimation strategies for each outcome variable in the appendix: (1) 'naïve' regression coefficient without balancing and without controls, (2) with control variables, but without balancing weights, (3) with balancing weights, but without control variables, and finally (4) with control variables and with balancing weights. From a methodological perspective, these comparisons of different estimation strategies give interesting insights into the deviation of point estimates by neglecting important factors as well as the quality of entropy balancing. As a rule of thumb, the entropy balancing has been successful; the closer point estimates are by comparing models with and without control variables (Oster 2019). As we strongly believe in providing the 'best' results using the fourth estimation strategy, combining entropy balancing with control variables, we only present these results for our three outcome variables in the main text.

As our outcome variables are measured on symmetric Likert scales with 5 or 11 points respectively, we run linear regression models. To test our theoretical hypotheses (compare chapter 3), we are mainly interested in two coefficients: (1) the direct effect of gender-matching on success during doctoral studies (ATE), and (2) the interaction of gender-matching with gender, and thus heterogeneity of effects. For each outcome, we report results for the whole sample in a first step as well as for subject-specific analyses in a second step. Results for the relations of interest are presented as coefficient plots (Jann 2014).[7]

With our analytical approach, we account for selectivity into treatment for the purpose of causal reasoning. However, we cannot completely refute the objection of selection by unobserved characteristics even considering a wide set of covariates.

6 When defining groups for doctoral subject, we considered the share of female doctoral candidates as well as the number of supervisors. We have separated biology from the other sciences that are combined with the other STEM fields. We did so due to the strong difference in the gender composition of biology as opposed to the other sciences and as we suspect that the gender composition in a field of study moderates how the gender match impacts on outcome variables.

7 Regression tables are provided in appendix 2.

One might think of other factors influencing a gender match as well as outcome variables that we have not measured, e.g., supervisor's reputation within the scientific community. However, our coefficient estimates are closer to the 'true causal effect' than are 'simple' regression results (also compare Figures A1-A3 in the appendix). Moreover, choosing balancing variables has forced us to think about control variables in more sophisticated way.

Table 2: Variables used to balance the control group

Variable	Measurement/Operationalization
Ascribed characteristics	
Respondent is female	Binary, yes = 1, no = 0
Match of gender of respondent and supervisor	Binary, yes = 1, no = 0
Age	Continous, age in years
Father's level of education	Categorical, 3 categories (higher education degree, doctorate; reference category: no higher education degree)
Mother's level of education	Categorical, 3 categories (higher education degree, doctorate; reference category: no higher education degree)
Respondent born outside Germany	Binary, yes = 1, no = 0
Father born outside Germany	Binary, yes = 1, no = 0
Mother born outside Germany	Binary, yes = 1, no = 0
Characteristics of doctorate	
Doctoral subject	Categorical, 6 categories (arts and humanities, biology, medicine, stem, others; reference category: social sciences)
Form of doctorate	Categorical, 3 categories (program and scholarship, free/external; reference category: appointment)
Grade point average at master's level	Continous, according to the German grading system: 1.0 – 4.0
Socio-demografic characteristics	
Children	Binary, yes = 1, no = 0
Partner/Stable relationship	Binary, yes = 1, no = 0
Partner's level of education	Categorical, 3 categories (no or occupational training, doctorate; reference category: higher education degree)
Partner's employment status	Categorical, 4 categories (part-time or other employment status, training or parental leave, not employed; reference category: full-time employment)
Partner not employed in academia	Binary, yes = 1, no = 0

Variable	Measurement/Operationalization
	Health and personality traits
Health (self-rated)	Categorical, 5 categories (ranging from "very bad" to "very good")
Risk-taking	Categorical, 7 categories (ranging from "not at all willing to take risks" to "very willing to take risks")
Locus of control	Factor variable, 2 factors (internal and external)
Self-efficacy Personality traits as Big Five	Factor variable, 5 factors (extraversion, neuroticism, openness, conscientiousness, agreeableness)
	Individual attitudes
Goals for doctorate	Binary, yes = 1, no = 0 (9 items: interest, contribution to scientific progress, common in discipline, social environment's expectations, nothing else came along, work in academia permanently, solving societal problems, reputation, career prospects outside academia)
Importance of job characteristics after doctorate	Binary, yes = 1, no = 0 (11 items: managerial responsibility, compatibility of work and family, availability of resources, opportunities for advancement, societal recognition, job security, societal benefits of work, salary level, autonomy in decision-making, working in a team, intellectual challenge)
Reasons to choose higher education institution	Binary, yes = 1, no = 0 (7 items: location, good research conditions, supervisor, university's reputation, attractive services for doctoral candidates, just came about that way, others)

5 Results

5.1 How widespread is a gender match between doctoral students and supervisors?

To begin with, our data show that a gender match between student and supervisor is more prevalent than a non-gender-match (see Table 3). Generally, the share of male supervisors among all supervisors is an astounding 75 percent. Accordingly, only a quarter of all supervisors are female. For male doctoral students the share of male supervisors is even larger and at 82 percent. In contrast, the share of female supervisors is disproportionally larger among female doctoral students and reaches one third. As argued before, this may indicate a preference of doctoral students for gender-matching supervisors. Note, however, that we cannot test to what extent this result is driven by preferences of students or by preferences of supervisors.

Table 3: Proportions of gender-matching between doctoral students and supervisors (Absolute and relative numbers)

		Supervisor		
		Male	Female	**Total**
Doctoral Student	Male	6,745 (82.36) *(58.52)*	1,445 (17.64) *(37.76)*	8,190 (100.00) *(53.35)*
	Female	4,778 (66.73) *(41.48)*	2,382 (33.27) *(62.24)*	7,160 (100.00) *(46.65)*
	Total	11,523 (75.07) *(100.00)*	3,827 (24.93) *(100.00)*	15,350 (100.00) *(100.00)*

Pearson chi2(1) = 498,3151 Pr = 0,000.

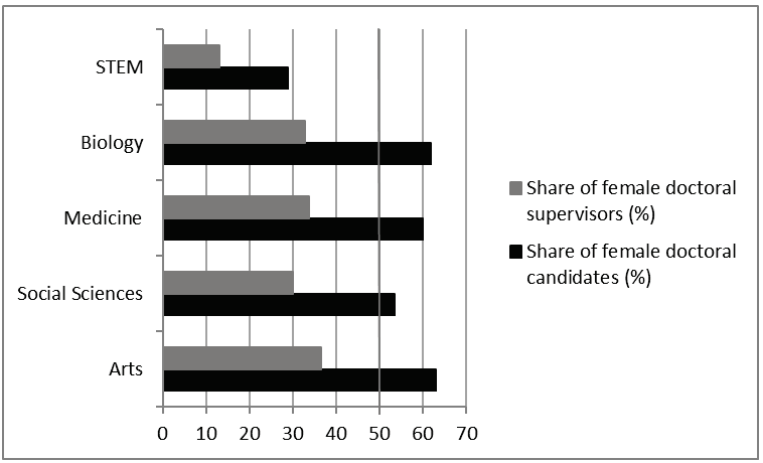
Source: Nacaps 2018, first wave. Own calculations. N = 15,350.
Note: Row percentages in parentheses, *column percentages* in parentheses and italics.

Due to the generally larger share of male supervisors, female doctoral candidates are much less likely to be matched in terms of supervisors' gender than are male doctoral candidates. Figure 1 shows another interesting result: While women are strongly underrepresented among supervisors, the gender ratio among doctoral candidates almost reaches parity (47 percent females and 53 percent males).

Of course, the relatively larger share of female doctoral students with female supervisors—or male doctoral students with male supervisors respectively—could also be due to differences in the gender composition of supervisors across subjects, i.e., due to opportunities rather than preferences. Therefore, in the next step, we look at subject differences (Figure 1, Table 4).

Not surprisingly, the share of female doctoral candidates and also female supervisors differs substantially across subjects. As displayed in Figure 1, both shares are lowest in STEM fields, and highest in biology, medicine and arts. For the latter subjects, the proportion of women among doctoral students is 60 percent or more, thus clearly crossing the line that indicates gender parity. In contrast, even in subjects with a comparatively high proportion of female supervisors, the share is far from reflecting gender parity. Biology differs considerably in the gender composition of both students and supervisors. Therefore, we look at biology separately from the other STEM fields, which are more homogeneous in this respect.

Figure 1: Proportion of female doctoral students and female supervisors across subjects

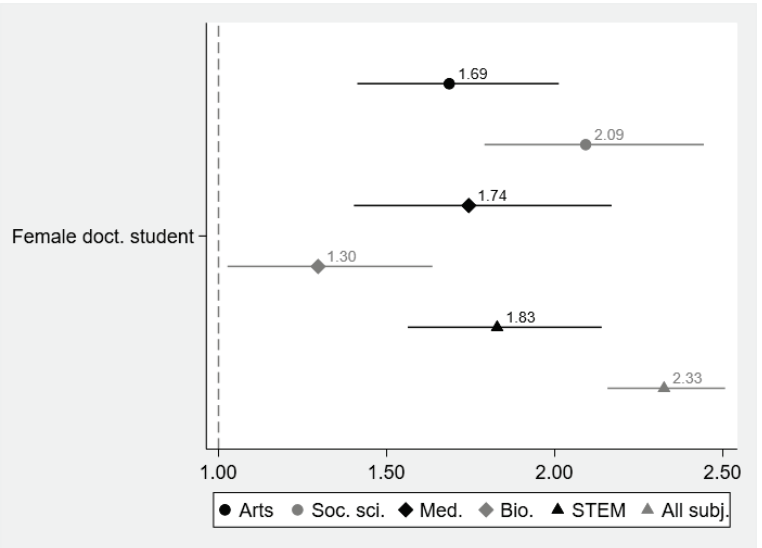


Source: Nacaps 2018, first wave. Own calculations. N = 15,350.

From a perspective of demand and supply, it could be that the observed tendency of male students having male supervisors and female students having female supervisors is primarily driven by the supply of supervisors of the respective gender in the different fields. But the results of Table 4 show that the pattern observed in Table 3 also holds across subjects. In all fields of study, the share of male doctoral students with a male supervisor exceeds the overall proportion of male supervisors and likewise the share of female doctoral students with a female supervisor exceeds the overall proportion of female supervisors. However, there are slight differences across subjects and the overall pattern is somewhat mitigated when taking on a subject-specific perspective.

To check how the pattern varies, we compared the chance of attaining a female supervisor for female and male doctoral students by running bivariate logistic regressions with gender of the doctoral student as explanatory variable. Figure 2 shows the odds ratios for all doctoral students and by subject. All odds ratios are above 1 and statistically significant, i.e., the chances of female doctoral students having a female supervisor are greater than those for male doctoral students. Generally, the chances of a female doctoral student having a female supervisor are 2.33 times higher than the chances of a male doctoral student having a female supervisor. The odds ratios vary to some extent across subjects with social sciences showing the largest odds ratio (2.09) and biology the lowest (1.30). The difference between these two subjects is statistically significant but the other differences between subjects are not.

Figure 2: Odds ratio of having a female supervisor by field of study (Coefficient plots from logistic regressions for being a female doctoral student)



Source: Nacaps 2018, first wave. Own calculations. N = 15,333.
Note: Regression results are available on request.

Note that, due to the different 'supply' of female or male supervisors across subjects, the proportions of doctoral students with a gender-matching supervisor vary (Table 4). In STEM fields, for example, nearly 90 percent of male doctoral candidates are matched; however, only about one fifth of female doctoral candidates are matched. Biology is the exception among the natural sciences; for more than 70 percent of the male doctoral students and about one third of the female doctoral students, the gender of the supervisor equals that of the doctoral candidate. The figures for medicine and social sciences are quite similar. With more than 40 percent matched female doctoral candidates, arts has the highest share of female doctoral students with a gender-matching supervisor, obviously due to the highest share of female professors.

Summing up, male doctoral students are more likely to have a supervisor of the same gender while female doctoral students are more likely to have a supervisor of a different gender. But considering the overall gender distribution of supervisors, the likelihood of having a supervisor of the same gender is disproportionally higher for both male and female candidates. This can be shown in an overall perspective and also, with minor differences between subjects, in subject-specific perspective. Thus, hypothesis 1a is confirmed with recent data for Germany. Hypothesis 1b suggested a specifically strong overrepresentation of a gender match for female candidates in

male-dominated subjects. The STEM fields would be an example of a male-dominated field. Biology would be an example of a natural science with a relatively large proportion of women among supervisors and candidates. While indeed the odds ratio in Figure 2 is relatively small for biology and differs significantly from the social sciences it does not differ significantly from STEM. In fact, the odds ratio for the STEM fields does not differ significantly from any other field. Thus, our results do not confirm hypothesis 1b.

Table 4: Proportions of gender-matching between doctoral students and supervisors across different subjects (Absolute and relative numbers)

STEM

		Supervisor		
		Male	Female	**Total**
Doctoral Student	Male	3,850 (89.10) *(73.98)*	471 (10.90) *(60.85)*	4,321 (100.00) *(72.28)*
	Female	1,354 (81.71) *(26.02)*	303 (18.29) *(39.15)*	1,657 (100.00) *(27.72)*
	Total	5,204 (87.05) *(100.00)*	774 (12.95) *(100.00)*	5,978 (100.00) *(100.00)*

Pearson chi2(1) = 57,9669 Pr = 0,000.

Source: Nacaps 2018, first wave. Own calculations. N = 5,978.
Note: Row percentages in parentheses, *column percentages* in parentheses and italics.

Biology

		Supervisor		
		Male	Female	**Total**
Doctoral Student	Male	403 (72.09) *(41.42)*	156 (27.91) *(35.29)*	559 (100.00) *(35.51)*
	Female	570 (66.59) *(58.58)*	286 (33.41) *(64.71)*	856 (100.00) *(60.49)*
	Total	973 (68.76) *(100.00)*	441 (31.24) *(100.00)*	1,415 (100.00) *(100.00)*

Pearson chi2(1) = 4,7698 Pr = 0,029.

Source: Nacaps 2018, first wave Own calculations. N = 1,415.
Note: Row percentages in parentheses, *column percentages* in parentheses and italics.

Medicine

		Supervisor		
		Male	Female	**Total**
Doctoral Student	Male	497 (74.51) *(54.31)*	170 (27.91) *(32.20)*	667 (100.00) *(41.05)*
	Female	600 (62.57) *(54.69)*	359 (33.41) *(67.80)*	959 (100.00) *(58.95)*
	Total	1,097 (67.47) *(100.00)*	529 (31.24) *(100.00)*	1,626 (100.00) *(100.00)*

Pearson chi2(1) = 25,8536 Pr = 0,000.

Source: Nacaps 2018, first wave. Own calculations. N = 1,626.
Note: Row percentages in parentheses, *column percentages* in parentheses and italics.

Social Sciences

		Supervisor		
		Male	Female	**Total**
Doctoral Student	Male	1,197 (78.08) *(51.58)*	336 (21.92) *(33.98)*	1,533 (100.00) *(46.50)*
	Female	1,110 (63.00) *(48.15)*	652 (37.00) *(66.02)*	1,762 (100.00) *(53.50)*
	Total	2,307 (70.02) *(100.00)*	988 (29.98) *(100.00)*	3,295 (100.00) *(100.00)*

Pearson chi2(1) = 88,8636 Pr = 0,000.

Source: Nacaps 2018, first wave. Own calculations. N = 3,295.
Note: Row percentages in parentheses, *column percentages* in parentheses and italics.

Arts

		Supervisor		
		Male	Female	**Total**
Doctoral Student	Male	644 (71.24) *(41.62)*	260 (28.76) *(29.71)*	904 (100.00) *(37.31)*
	Female	903 (59.49) *(58.38)*	615 (40.51) *(70.29)*	1,518 (100.00) *(62.69)*
	Total	1,547 (63.87) *(100.00)*	875 (36.13) *(100.00)*	2,422 (100.00) *(100.00)*

Pearson chi2(1) = 4,7698 Pr = 0,029.

Source: Nacaps 2018, first wave. Own calculations. N = 2,422.
Note: Row percentages in parentheses, *column percentages* in parentheses and italics.

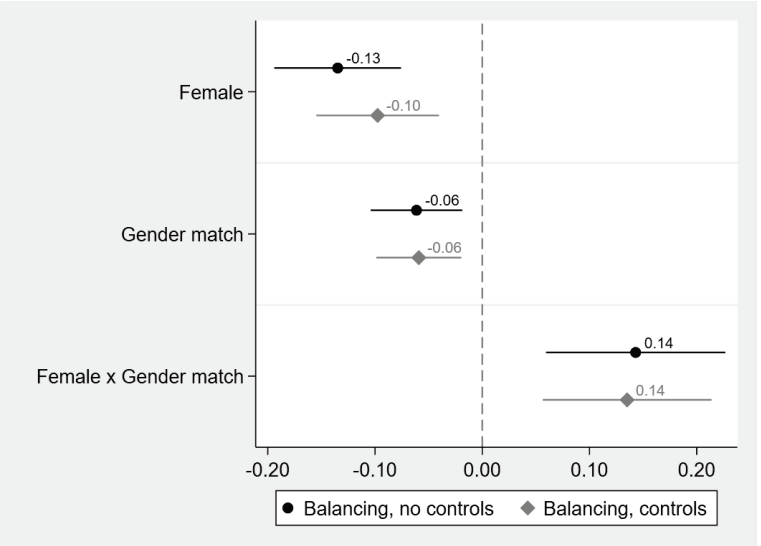
5.2 Is gender-matching beneficial for doctoral studies?

We now turn to the analytic modeling of our three dependent variables. If a gender match of students and supervisors has a positive effect on these dependent variables, as theory suggests, this would indicate that academic careers of women, or men, would benefit from a gender match.

5.2.1 Satisfaction with mentoring

Firstly, we look at the effect of a gender match on satisfaction with mentoring (Figure 3). The conditional main effect of being female is negative, i.e., compared to their male counterparts female doctoral candidates are less satisfied with mentoring. The main effect of gender matching is also negative. Note, that due to the interaction term and male students being the reference group, this is the effect of a gender match for male doctoral students. In other words, male doctoral students with a male supervisor are less satisfied with mentoring than are male doctoral students with a female supervisor. In contrast, female doctoral students with a gender match, i.e., with a female supervisor, are more satisfied than their female peers with a male supervisor, as shown by the positive interaction effect. The size of the positive effect of a female supervisor almost exactly compensates the generally lower level of satisfaction among female doctoral students.

Figure 3: Satisfaction with mentoring (Coefficient plots from linear regression)

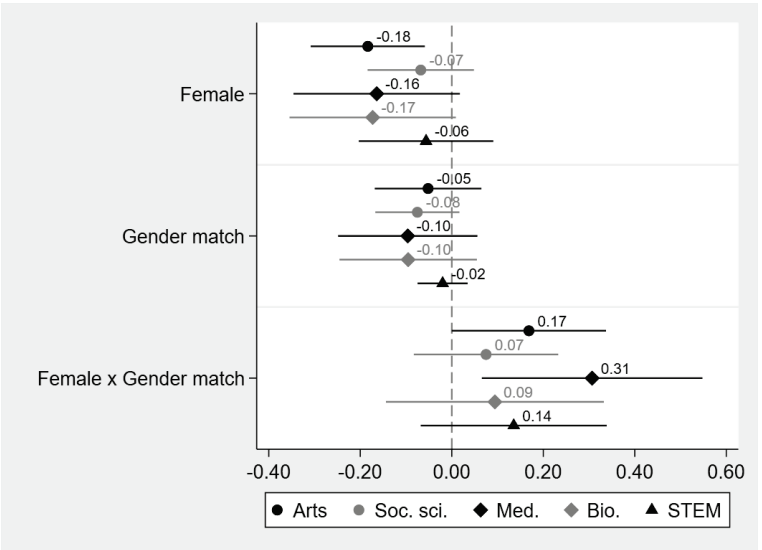


Source: Nacaps 2018, first wave. Own calculations. N = 15,333.
Note: Plot of regression coefficients for main effects of gender, gender match and interaction effect of models without control variables and with control variables. For control variables see Table 2 above. Table with full regression coefficients is provided in the appendix 2 in Table A1, models 1a and 1b.

Reconsidering our second hypothesis on a positive effect of gender-matching supervisors, results are therefore mixed. For female students, the gender match indeed has a positive effect on satisfaction; however, this is not the case for male students. This means that doctoral students with female supervisors are generally more satisfied with mentoring, irrespective of their own gender, even though this positive effect seems to be somewhat stronger in absolute terms for female students.

As a quality check, we compare the coefficients for regression models with and without control variables. We find coefficient estimates to be very similar. This is what we expect when applying entropy-balancing weights and may also be taken as a sign that the entropy balancing works well (Oster 2019). With controls, confidence intervals are slightly smaller.

Figure 4: Satisfaction with mentoring – by subject (Coefficient plots from linear regression)



Source: Nacaps 2018, first wave. Own calculations. N = 15,333.
Note: Plot of regression coefficients for main effects of gender, gender match and interaction effect of models for five groups of subjects. Model specification: with entropy balancing and controls. For control variables see Table 2 above. Table with full regression coefficients is provided in the appendix 2 in Table A2.

In Figure A1 in the appendix 1, we compare coefficients with and without entropy balancing yielding a methodologically interesting result: While point estimates do not differ strongly, confidence intervals are clearly smaller when applying entropy balancing weights. The latter lead to more efficient estimates and in fact, without the entropy balancing we would not have accepted the coefficient of the gender-match dummy as statistically significant.

Looking at subject-specific differences, the picture becomes less clear (see Figure 4). Except for the conditional main effect of female doctoral students in arts and the interaction effect between female doctoral candidates in medicine and with a gender match, all coefficient estimates are statistically insignificant. Considerably larger confidence intervals indicate uncertainty in estimation, even though numbers of respondents are not particularly small, ranging from 1,415 in biology to 5,978 in the STEM fields. Above, we formulated two conflicting subject-specific expectations. Hypothesis 5a suggested that the positive effect of a gender match would be particularly strong in male-dominated fields while hypotheses 5b suggested a particularly strong positive effect in fields with relatively low proportions of men. Our results confirm neither hypothesis 5a nor hypothesis 5b. Rather, for doctoral

students in Germany the subject as a context does not seem to make a major difference for the effect of a gender-matching supervisor on satisfaction with mentoring.

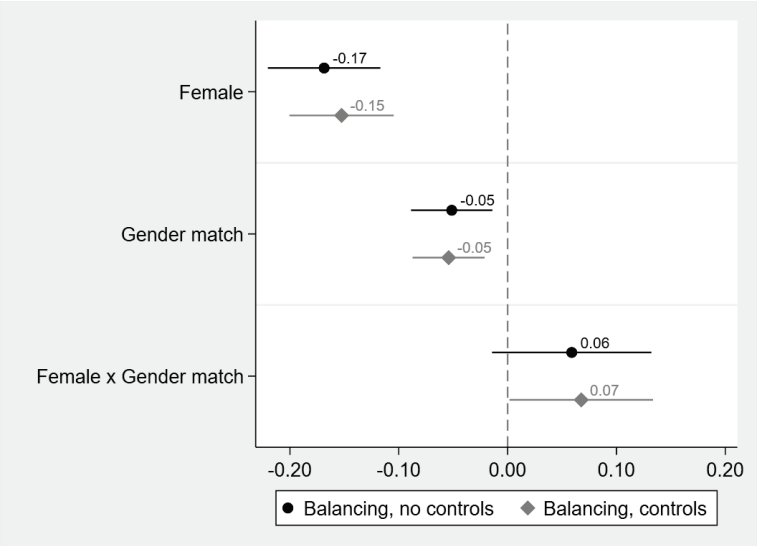
5.2.2 Belief in own research abilities

The belief in one's own research abilities is likely to be an important resource for successfully traveling the sometimes rocky road of a doctorate and an academic career in general. As Figure 5 shows, female doctoral students are significantly less well equipped with this resource and are more skeptical about their research abilities than their male peers. Does a gender-matching supervisor help to boost academic self-efficacy?

Results in Figure 5 resemble the pattern already observed for satisfaction with mentoring. There is no general positive effect of a gender match between students and supervisors. The main effect is negative, i.e., male doctoral students with a male supervisor believe somewhat less in their research abilities. For female students, though, we observe a positive interaction effect. With respect to hypothesis 3 the result is therefore mixed again and depends on the gender of doctoral students. A gender match helps only if the student is female. Putting it differently, female supervisors strengthen the academic self-efficacy of their doctoral students as compared to male supervisors. This effect does not fully compensate the lower academic self-efficacy of female doctoral students but helps to mitigate it.

Comparing estimates with and without controls we again find point estimates and confidence intervals to be quite similar. However, estimates are slightly more efficient with control variables and reveal a statistically significant interaction effect. Figure A2 in the appendix 1 provides the results for models without the entropy-balancing weights. Again, it is interesting to see that we would have overlooked several statistically significant point estimates without the entropy balancing.

Figure 5: Belief in own research abilities (Coefficient plots from linear regression)

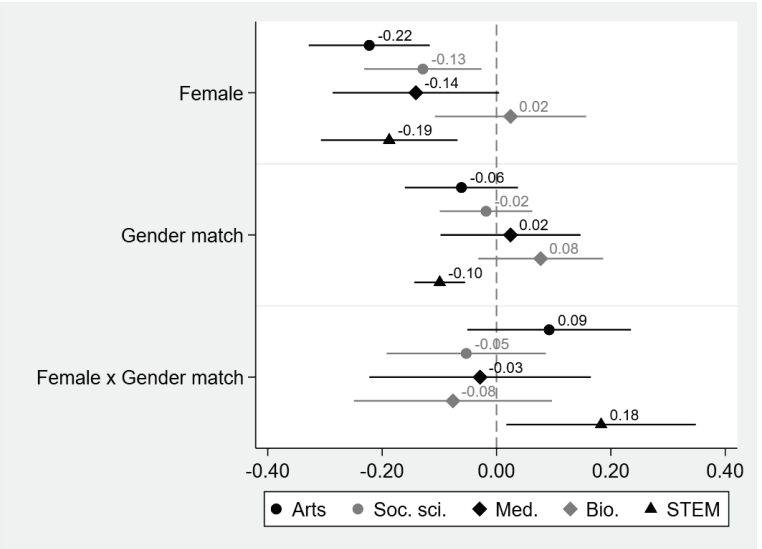


Source: Nacaps 2018, first wave. Own calculations. N = 15,280.
Note: Plot of regression coefficients for main effects of gender, gender match and interaction effect of models without control variables and with control variables. For control variables see Table 2 above. Table with full regression coefficients is provided in the appendix 2 in Table A1, models 2a and 2b.

Findings of subject-specific models show a more complex picture (see Figure 6). For three groups of subjects (arts, social sciences, and STEM) we find that female doctoral students have significantly lower levels of academic self-efficacy—as in the overall analysis. For the main effect of the gender match and the interaction, significant effects are only observed for the largest field of study, i.e., STEM. For the latter, beliefs in own research abilities are negatively affected by a gender matching, i.e., male doctoral students are less confident in their research abilities if supervised by a male mentor. This negative effect turns into the opposite if female doctoral candidates are supervised by women in STEM fields, which is in line with previous findings (Bettinger/Long 2005). For all other subjects, neither the main effects of a gender match nor the interaction terms are statistically significant. Thus the overall picture seems to be dominated by the pattern to be observed for the STEM fields. The pattern for arts is very similar, even though the main effect of a gender match and the interaction effect are not statistically significant with the given statistical power. Remarkably, these results do not support theoretical considerations about the share of female doctoral students as a relevant context condition as similar patterns are observed for the subject groups with the lowest and with the highest shares of female doctoral students and supervisors. The relatively

large positive interaction-effect in the STEM fields could be seen as supporting hypothesis 5, that suggested a relatively strong effect for male-dominated fields. But as coefficients of the different subjects overlap, the results support neither hypothesis 5a nor hypothesis 5b.

Figure 6: Belief in own research abilities – by subject (Coefficient plots from linear regression)



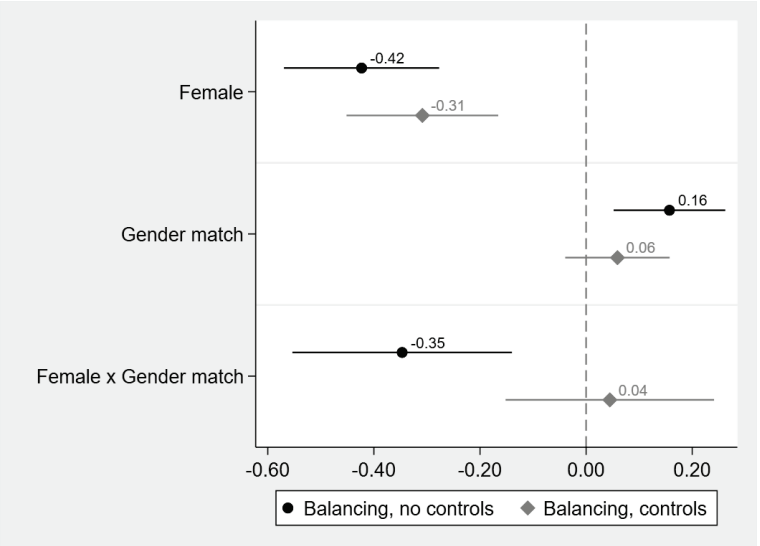
Source: Nacaps 2018, first wave. Own calculations. N = 15,280.
Note: Plot of regression coefficients for main effects of gender, gender match and interaction effect of models for five groups of subjects. Model specification: with entropy balancing and controls. For control variables see Table 2 above. Table with full regression coefficients is provided in the appendix 2 in Table A3.

5.2.3 Prospects for postdoc position

With respect to the perceived career prospects, we first need to acknowledge that results differ for the models with and without controls (see Figure A3 in the appendix 1). In either case, compared to their male peers, women are less optimistic about their chances of obtaining a post-doc position in academia. However, when applying controls, neither the main effect of a gender match in general nor the interaction effect significantly affects the perceived prospects for a postdoc position.

Interestingly enough, the results without controls seem to suggest the obverse gender match and interaction effect as for the satisfaction with mentoring and academic self-efficacy, i.e., a generally positive effect of a *male* supervisor for doctoral students of both genders. However, with controls, both effects are insignificant and thus we need to reject hypothesis 4a and hypothesis 4b.

Figure 7: Prospects for postdoc position (Coefficient plots from linear regression)

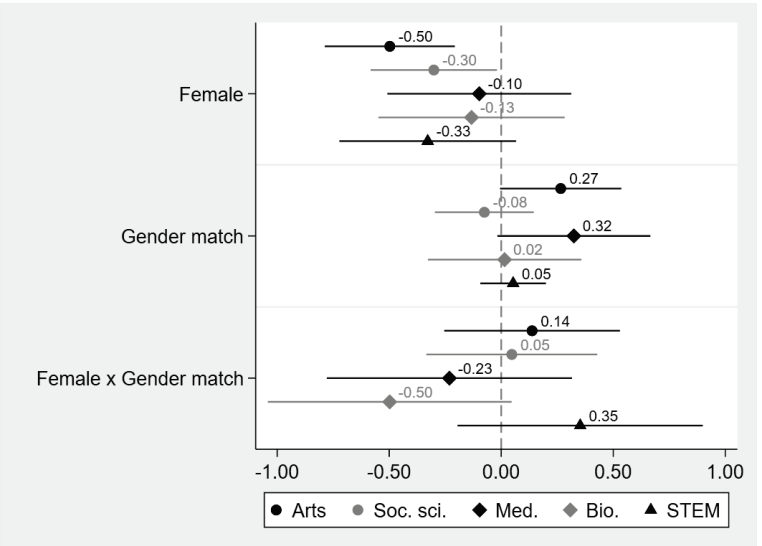


Source: Nacaps 2018, first wave. Own calculations. N = 14,915.
Note: Plot of regression coefficients for main effects of gender, gender match and interaction effect of models without control variables and with control variables. For control variables see Table 2 above. Table with full regression coefficients is provided in the appendix 2 in Table A1, models 3a and 3b.

As noted, we found relatively large differences in results with and without control variables (see Figure A3 in the appendix 1). The main gender effect and specifically the interaction term become insignificant when the control variables are included in the model. By stepwise regressions it was found that the interaction term becomes insignificant when the subjects are controlled for. With a good entropy-balancing model such differences between models with and without controls should not occur. While we must acknowledge that with the data at hand there is little we could do to improve the model, this may hint at weaknesses of the entropy-balancing model with regard to prospects for a postdoc position as dependent variable, i.e., results for this dependent variable should be interpreted with caution.

Subject-specific analyses show almost no significant effects (see Figure 8): As with the overall results, female doctoral students in arts and social sciences are less optimistic regarding their academic outlook. In line with the results for all subjects together, none of the conditional main effects of a gender match or of the interaction effects is statistically significant. Thus, again our results confirm neither hypothesis 5a nor hypothesis 5b.

Figure 8: Prospects for postdoc position – by subject (Coefficient plots from linear regression)



Source: Nacaps 2018, first wave. Own calculations. N = 14,915.
Note: Plot of regression coefficients for main effects of gender, gender match and interaction effect of models for five groups of subjects. Model specification: with entropy balancing and controls. For control variables see Table 2 above. Table with full regression coefficients is provided in the appendix 2 in Table A4.

6 Summary and discussion

In light of the 'leaky pipeline' phenomenon in the German science system, our contribution investigates (i) how widespread a gender match between doctoral student and supervisor is in Germany and (ii) whether a gender match of doctoral student and supervisor is beneficial for the doctorate and academic career prospects thereafter. To answer our two research questions, we draw on recent data from the 'German National Academics Panel Study (Nacaps)'.

Firstly, our analyses confirm a clear prevalence of gender-matching combinations between doctoral students and supervisors for both genders. This prevalence can be observed across all subject groups and is in line with previous findings mainly from the United States. Interestingly, even in subjects with a comparatively high proportion of female supervisors, the share is far from reflecting gender parity. Based on tokenism theory we suspected an especially strong overrepresentation of gender matches for female doctoral students in male-dominated fields; in such fields of study, female doctoral students could be exposed to discrimination particularly strongly and seek to find a female supervisor to avoid this. However, this hypothesis is not confirmed.

Secondly, results show that female supervisors have the expected positive effect on satisfaction with mentoring and academic self-concept for female doctoral students. This result was suggested by theory and it seems intuitive that supervisors of the same gender are somewhat beneficial. Surprisingly and challenging to our intuition, female supervisors have this positive effect on male doctoral students as well. To some extent, the effect therefore seems to be rooted in the supervisors' gender rather than in the match between doctoral students' and supervisors' gender. Thus, our hypotheses 2 and 3 on the beneficial effect of a gender match are only confirmed for female doctoral students but not for their male peers.

Thirdly, we find no significant effect of a gender match regarding the perceived prospects for a postdoc position. Thus, our results confirm neither hypothesis 4a regarding a general positive effect of a gender-match nor hypothesis 4b regarding a negative effect for women.

Fourthly, no clear pattern can be identified with respect to differences between doctoral subjects. Applying tokenism theory and the identity-based motivation theory we suspected a specifically strong beneficial effect of the gender match in male-dominated fields, such as STEM (hypothesis 5a). Considering arguments of network theory, in contrast, it seems plausible to expect specifically strong beneficial effects of the gender match in fields with relatively high proportions of women. In other words, we assumed the proportion of women in the field to be an important moderating context variable. But coefficients differed by subjects only very rarely. An exception that could be mentioned is that for STEM fields we do find a significant positive effect of the gender match on academic self-efficacy but not for the other fields of study. This might indicate that the mechanisms suggested by tokenism theory and the identity-based motivation theory are at work but again the gender-match effect for women does not differ significantly across subjects. All in all, our results therefore confirm neither the systematic differences between fields of study suggested by tokenism theory nor the systematic differences between fields of study suggested by network theory. This finding may be somewhat unsatisfactory, but it also fits with the results for bachelor students in Ohio (Bettinger/Long 2005).

Finally, from a methodological point of view it is interesting that by applying entropy-balancing weights we arrive at more accurate and thus statistically significant estimates which would otherwise have been overlooked (see Figure A1 in the appendix). Our estimation strategy helps in dealing with the endogeneity problem and strengthens the claim made in the reviewed literature of interpreting findings in a causal way. However, we cannot be sure whether we fully solved this obvious endogeneity problem with our entropy-balancing model. There may be heterogeneities between treatment and control group that are not observed and therefore cannot be controlled for. The Nacaps data provides a huge set of observed characteristics (see Table 2). This leads us to be fairly confident about our results

and their interpretation. As mentioned above, however, results on prospects for a postdoc position should be treated with some caution.

To the best of our knowledge, our contribution provides results for doctoral students in Germany for the first time. It uses recent available data and applies a sophisticated estimation strategy. Still a couple of limitations should be mentioned. These limitations offer potential for future research.

First of all, our data contains only doctoral students at an early stage of their academic careers. Even though our outcome variables are directed to further academic careers, we do not know who stays in academia after graduation from doctoral studies and which of those graduates will finally go on to a successful academic career. To answer these and similar questions for long-term effects of a gender-matching supervisor relationship during doctoral studies, we need longitudinal data capturing a time span of several years. Future waves of Nacaps offer an opportunity for longitudinal analyses.

With respect to theoretical explanations, *secondly*, the findings partly conflict with our assumptions and probably also with our intuition. Our results suggest that effects on the outcome variables are rather driven by the supervisor's gender than the gender match between doctoral students and their supervisors. Ultimately, the core question of *why* same-gender supervisors are beneficial for academic careers still remains open. For identifying the social mechanisms behind the gender-match effect (or the supervisor-gender effect), we need more information on supervisors than just gender. For example, to test whether male supervisors provide better access to influential academic networks, as proposed in hypothesis 4b, we need appropriate measures for network size and density or supervisor's reputation within the scientific community. As a forecast, some of these indicators are measured in subsequent waves of Nacaps.

A *third* point is directed to alternative estimation strategies. Instead of using entropy-balancing as a reweighting method to build a synthetic control group, one could think of matching procedures on the individual level like Coarsened Exact Matching (CEM, Blackwell et al. 2009, Iacus et al. 2012) or propensity score matching (Caliendo/Kopeinig 2008, Gangl 2010) to build statistical twins. However, as Hainmüller (2012) shows, entropy-balancing is not only easier to apply than propensity score matching and similar techniques but also yields better results. Generally, the problem with selection on unobservable variables is by design not solved with either of these estimation strategies.

Concerning possible implications of our findings for higher education policies, we would like to highlight that despite all limitations we have clear indications that 'women are helping women', as Hilmer and Hilmer (2007) had put it; i.e., policies striving to bring more women into leading academic positions and thus to further boost the prospects of women in academic careers seem to be on the right track.

Doctoral students with female advisors are more satisfied with mentoring and have are more confident in their academic abilities.

Interestingly enough, male doctoral students also seem to benefit from female supervisors. We are not fully sure how to interpret this finding. It could be that women differ in their mentoring intensity and style which could lead to more satisfaction and academic self-esteem among doctoral students. To some extent these findings seem to confirm gender stereotypes of more 'caring' female supervisors. While we cannot exclude that this is the case, there are alternative interpretations, e.g., in all likelihood, female supervisors are on average younger and at an earlier stage in their academic careers than male supervisors. This could impact on mentoring intensity and style as well, in that younger professors, whose doctoral studies were completed relatively recently, might better understand and be more open to the needs of doctoral students. Moreover, they might have more available time to care about their doctoral students and lower 'opportunity costs' due to having fewer doctoral students and fewer other obligations (and opportunities) in which to invest their time. These alternative explanations are linked to the question, who chooses whom? Are students choosing supervisors or are supervisors choosing students and what are the reasons for such decisions? In this sense, the gender match could also be an interesting outcome variable to be investigated.

References

Alper, Joe (1993): The pipeline is leaking women all the way along. *Science (New York, N.Y.)* 260(5106): 409–411. https://doi.org/10.1126/science.260.5106.409

Alston, Gelada Drew, Guy, Breonte S. & Campbell, Dean C. (2017): Ready for the professoriate? The influence of mentoring on career development for black male graduate students in STEM. *Journal of African American Males in Education* 8(4): 45–66. https://www.proquest.com/scholarly-journals/ready-professoriate-influence-mentoring-on-career/docview/2667266322/se-2

Beierlein, Constanze, Kovaleva, Anastassiya, Kemper, Christoph J. & Rammstedt, Beatrice (2012): *Ein Messinstrument zur Erfassung subjektiver Kompetenzerwartungen: Allgemeine Selbstwirksamkeit Kurzskala (ASKU)* (GESIS Working Papers No. 17).

Berryman, Sue (1983): *Who Will Do Science? Trends, and Their Causes in Minority and Female Representation among Holders of Advanced Degrees in Science and Mathematics. A Special Report.* Rockefeller Foundation.

Bettinger, Eric P., & Long, Bridget Terry (2005): Understanding teacher quality: do faculty serve as role models? The impact of instructor Gender on female students. *American Economic Review* 95(2): 152–157. https://doi.org/10.1257/000282805774670149

Blackwell, Matthew, Iacus, Stefano, King, Gary & Porro, Giuseppe (2009): Cem: Coarsened exact matching in Stata. *The Stata Journal: Promoting Communications on Statistics and Stata* 9(4): 524–546. https://doi.org/10.1177/1536867X0900900402

Briedis, Kolja, Lietz, Almuth, Ruß, Uwe, Schwabe, Ulrike, Weber, Anne, Birkelbach, Robert & Hoffstätter, Ute (2020): *Nacaps 2018: Daten-und Methodenbericht zur National Academics Panel Study 2018 (1. Befragungswelle – Promovierende).* DZHW.

Briedis, Kolja, Hoffstätter, Ute & Schwabe, Ulrike (2022). The German National Academics Panel Study (Nacaps): A New Longitudinal Survey on Early Researchers' Careers and Life Courses. *Soziale Welt* 73(4): 740–763. https://doi.org/10.5771/0038607320224740

Bundesministerium für Bildung und Forschung (2021a): *Karliczek: Mehr Frauen in Spitzenpositionen von Wissenschaft und Forschung. Bundesforschungsministerin Anja Karliczek zum Internationalen Tag der Frauen und Mädchen in der Wissenschaft*. Pressemitteilung 27/2021. https://www.bmbf.de/bmbf/shareddocs/pressemitteilungen/de/karliczek-mehr-frauen-in-spitz-von-wissenschaft-und-forschung.html

Bundesministerium für Bildung und Forschung (2021b): *Karliczek: „Wir brauchen mehr weibliche Rollenvorbilder in Wissenschaft und Forschung".: Bundesforschungsministerin Anja Karliczek zum Weltfrauentag 2021.* Pressemitteilung 42/2021. https://www.bmbf.de/bmbf/shareddocs/pressemitteilungen/de/karliczek-wir-brauchen-mehr-we--in-wissenschaft-und-forschung.html

Caliendo, Marco & Kopeinig, Sabine (2005): Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys* 22(1): 31–72. https://doi.org/10.2139/ssrn.721907

Carell, Scott E., Page, Marianne E. & West, James E. (2010): Sex and science: How professor gender perpetuates the gender gap. *The Quarterly Journal of Economics* 125(3): 1101–1144. https://doi.org/10.1162/qjec.2010.125.3.1101

Carstensen, Johann (2020): Die Messung von Gesundheit, in P. Kriwy & M. Jungbauer-Gans (Hrsg.), *Handbuch Gesundheitssoziologie*. Springer Fachmedien Wiesbaden, S. 51 – 70

Cheung, Fanny M. (2021): The "state" of women's leadership in higher education. *International Briefs for Higher Education Leaders* 9: 5–8.

Edmunds, Laurel D., Ovseiko, Pavel V., Shepperd, Sasha, Greenhalgh, Trisha, Frith, Peggy, Roberts, Nia W., Pololi, Linda H. & Buchan, Alastair M. (2016): Why do women choose or reject careers in academic medicine? A narrative review of empirical evidence. *The Lancet* 388(10062): 2948–2958. https://doi.org/10.1016/S0140-6736(15)01091-0

Etzkowitz, Henry, Kemelgor, Carol, Neuschatz, Michael, Uzzi, Brian & Alonzo, Joseph (1994): Paradox of critical mass for women in science. *Science* 266(5182): 51–54. https://www.science.org/doi/10.1126/science.7939644

European Commission (2021): *She figures 2021: gender in research and innovation: statistics and indicators*. Publications Office of the European Union. https://data.europa.eu/doi/10.2777/06090 2021

Forret, Monica L. (2006). Impact of social networks on the advancement of women and racial/ethnic minority groups, in M. F. Karsten (Hrsg.), *Gender, ethnicity, and race in the workplace*. Praeger Publishers, S. 149 – 166

Forschung und Lehre (2020): *"Leaky Pipeline" prägt weiterhin die Wissenschaft. Der Frauenanteil in der Wissenschaft nimmt mit jeder Karrierestufe ab. Fortschritte in Sachen Chancengleichheit sind träge, zeigt eine Analyse.* https://www.forschung-und-lehre.de/karriere/leaky-pipeline-praegt-weiterhin-die-wissenschaft-3277/

Gangl, Markus (2010): Nichtparametrische Schätzung kausaler Effekte mittels Matchingverfahren, in Wolf, C., Best, H. (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse.* Wiesbaden: VS Verlag für Sozialwissenschaften, S. 931 – 961.

Gaule, Patrick & Piacentini, Mario (2017): *An advisor like me? Advisor gender and post-graduate careers in science* (IZA Discussion Papers No. 10828). Bonn. Institute of Labor Economics (IZA). http://hdl.handle.net/10419/170812

GESIS – Leibniz-Institut für Sozialwissenschaften (2015): *Allgemeine Bevölkerungsumfrage der Sozialwissenschaften ALLBUS 2014 (ZA5420 Fragebogen)*. Köln. GESIS Datenarchiv.

Hainmüller, Jens (2012): Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis* 20(1): 25–46. https://doi.org/10.1093/pan/mpr025

Hainmüller, Jens & Xu, Yiqing (2013): ebalance: A Stata package for entropy balancing. *Journal of Statistical Software* 54(7): 1–18. http://hdl.handle.net/1721.1/89819

Helbig, Marcel (2012). Boys don't benefit from male teachers in their reading and mathematical competencies. Empirical evidence from 21 countries of the OECD and EU. *British Journal of Sociology of Education*, 33(5), 661–677.

Hilmer, Christiana & Hilmer, Michael (2007): Women helping women, men helping women? Same-gender mentoring, initial job placements, and early career publishing success for economics PhDs. *American Economic Review* 97(2): 422–426. https://doi.org/10.1257/aer.97.2.422

Hirshfeld, Laura E. (2010): "She won't make me feel dumb": Identity threat in a male-dominated discipline. *International Journal of Gender, Science and Technology* 2(1): 6–24. https://genderandset.open.ac.uk/index.php/genderandset/article/view/60

Iacus, Stefano M., King, Gary, & Porro, Giuseppe (2012): Causal inference without balance checking: Coarsened exact matching. *Political Analysis* 20(1): 1–24. https://doi.org/10.1093/pan/mpr013

Jaksztat, Steffen (2017): Geschlecht und wissenschaftliche Produktivität. *Zeitschrift Für Soziologie* 46(5): 347–361. https://doi.org/10.1515/zfsoz-2017-1019

Jann, Ben (2014). Plotting regression coefficients and other estimates. *The Stata Journal: Promoting Communications on Statistics and Stata* 14(4): 708–737. https://doi.org/10.1177/1536867X1401400402

Kanter, Rosabeth Moss (1977a): *Men and women of the corporation*. Basic Books.

Kanter, Rosabeth Moss (1977b): Some effects of proportions on group life: Skewed sex ratios and responses to token women". *American Journal of Sociology* 82(5): 965–990.

Konsortium Bundesbericht Wissenschaftlicher Nachwuchs (2021): *Bundesbericht Wissenschaftlicher Nachwuchs 2021: Statistische Daten und Forschungsbefunde zu Promovierenden und Promovierten in Deutschland* (1st ed.). wbv Media. https://doi.org/10.3278/6004603aw

Kovaleva, Anastassiya, Beierlein, Constanze, Kemper, Christoph J. & Rammsted, Beatrice (2012): *Eine Kurzskala zur Messung von Kontrollüberzeugung:: Die Skala Internale-Externale-Kontrollüberzeugung-4 (IE-4)*. GESIS-Working Papers GESIS-Working Papers: Vol. 19. GESIS – Leibniz-Institut für Sozialwissenschaften. https://www.gesis.org/fileadmin/kurzskalen/working_papers/IE4_Workingpaper.pdf

Lörz, Markus & Mühleck, Kai (2019): Gender differences in higher education from a life course perspective: transitions and social inequality between enrolment and first post-doc position. *Higher Education* 77(3): 381–402. https://doi.org/10.1007/s10734-018-0273-y

Löther, Andrea (2019): Is It working? An impact evaluation of the german "Women Professors Program". *Social Sciences* 8(116): 1–18. https://doi.org/10.3390/socsci8040116

Main, Joyce B. (2018): Kanter's theory of proportions: Organizational demography and PhD completion in science and engineering departments. *Research in Higher Education* 59(8): 1059–1073. https://doi.org/10.1007/s11162-018-9499-x

McPherson, Miller, Smith-Lovin, Lynn & Cook, James M. (2001): Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27: 415–444. https://doi.org/10.3410/f.725356294.793504070

Neugebauer, Martin, Helbig, Marcel & Landmann, Andreas (2011): Unmasking the myth of the same-sex teacher advantage. *European Sociological Review* 27(5): 669–689. https://doi.org/10.25656/01:7955

Neumark, David & Gardecki, Rosella (1998): Women helping women? Role model and mentoring effects on female Ph.D. students in economics. *The Journal of Human Resources* 33(1): 220–246.

OECD (2008): *Education at a Glance 2008: OECD Indicators*. Paris: OECD Publishing. https://doi.org/10.1787/23fafc19-en

OECD (2019a): Who is expected to enter tertiary education?, in OECD (Hrsg.), *Education at a Glance 2019: OECD Indicators.* Paris: OECD Publishing. S. 192 – 206. https://doi.org/10.1787/888933978113

OECD (2019b): What are the characteristics and outcomes of doctoral graduates?, in OECD (Hrsg.), *Education at a Glance 2019: OECD Indicators.* Paris: OECD Publishing. S. 246 – 263. https://doi.org/10.1787/888933978645

OECD (2020a): Who is expected to enter tertiary education?, in: OECD (Hrsg.), *Education at a Glance 2020: OECD Indicators.* Paris: OECD Publishing. S. 198 – 212. https://doi.org/10.1787/888934163724

OECD (2020b): Who is expected to graduate from tertiary education?, in: OECD (Hrsg.), *Education at a Glance 2020: OECD Indicators*. Paris: OECD Publishing. S. 214–225. https://doi.org/10.1787/888934163876

OECD (2020c): To what level have adults studied?, in: OECD (Hrsg.), *Education at a Glance 2020: OECD Indicators*. Paris: OECD Publishing. S. 38 – 52. https://doi.org/10.1787/888934161843

OECD (2020d): Who are the teachers?, in: OECD (Hrsg.), *Education at a Glance 2020: OECD Indicators.* Paris: OECD Publishing. S. 428–439. https://doi.org/10.1787/888934165985

Oster, Emily (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics* 37(2): 187–204. https://doi.org/10.1080/07350015.2016.1227711

Oyserman, Daphna. (2007): Social identity and self-regulation, in: A. Kruglanski & T. Higgins (Hrsg.), *Handbook of Social Psychology*. New York: Guilford Press. S. 432 – 552.

Oyserman, Daphna (2009): Identity-based motivation: Implications for action-readiness, procedural-readiness, and consumer behavior. *Journal of Consumer Psychology* 19(3): 250–260. https://doi.org/10.1016/j.jcps.2009.05.008

Riegle-Crumb, Catherine, Peng, Menglu & Russo-Tait, Tatiane (2020): Committed to STEM? Examining factors that predict occupational commitment among asian and white female students completing STEM U.S. postsecondary programs. *Sex Roles* 82(1–2): 102–116. https://doi.org/10.1007/s11199-019-01038-8

Roach, Michael & Sauermann, Henry (2010): A taste for science? PhD scientists' academic orientation and self-selection into research careers in industry. *Research Policy* 39(3): 422–434. https://doi.org/10.1016/j.respol.2010.01.004

Ruß, Uwe (2021): Frauen in der Wissenschaft: Zu Entwicklung und Anteil von Frauen in verschiedenen Stadien der wissenschaftlichen Karriere in Baden-Württemberg. *Statistisches Monatsheft Baden-Württemberg 1/2021*: 1–13.

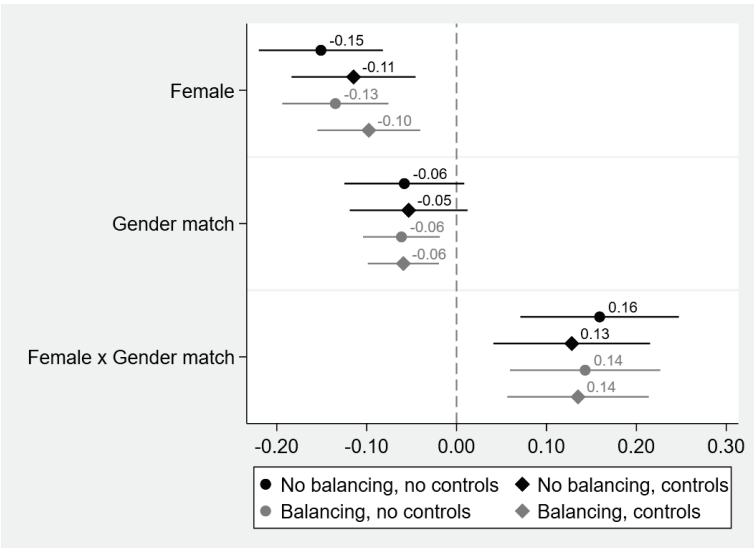
Schubert, Frank & Engelage, Sonja (2011): Wie undicht ist die Pipeline? Wissenschaftskarrieren von promovierten Frauen. *KZfSS Kölner Zeitschrift Für Soziologie Und Sozialpsychologie* 63(3): 431–457. https://doi.org/10.1007/s11577-011-0144-3

Schupp, Jürgen & Gerlitz, Jean-Yves (2014): *Big Five Inventory-SOEP (BFI-S): Zusammenstellung sozialwissenschaftlicher Items und Skalen.* https://doi.org/10.6102/zis54

Solanki, Sabrina M. & Di Xu (2018): Looking beyond academic performance: The influence of instructor gender on student motivation in STEM fields. *American Educational Research Journal* 55(4): 801–835. https://doi.org/10.3102/0002831218759034

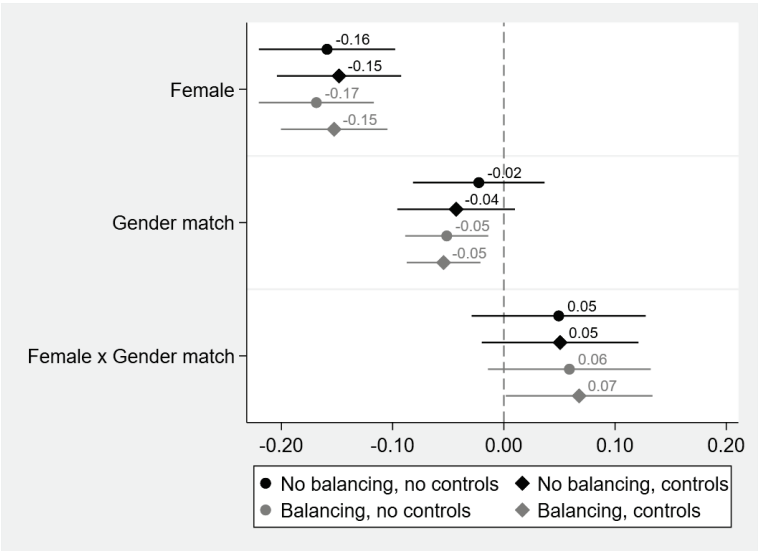
Appendix 1: Comparing estimation strategies

Figure A1: Satisfaction with mentoring – by estimation strategy



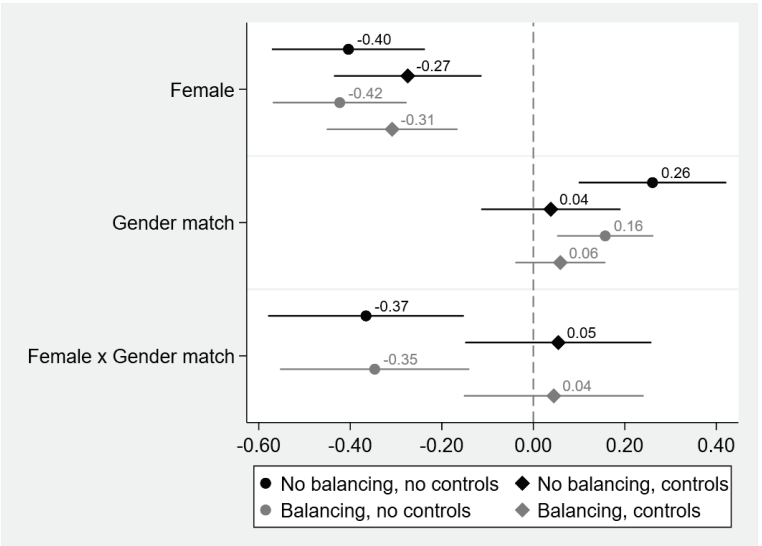
Source: Nacaps 2018, first wave. Own calculations. N = 15,333.
Note: Plot of regression coefficients for main effects of gender, gender match and interaction effect of model without entropy balancing weights and without control variables, without entropy balancing weights and with control variables, with entropy balancing weights and without control variables and with entropy balancing weights and with control variables. For control variables see Table 2 above.

Figure A2: Belief in own research abilities – by estimation strategy



Source: Nacaps 2018, first wave. Own calculations. N = 15,280.
Note: Plot of regression coefficients for main effects of gender, gender match and interaction effect of model without entropy balancing weights and without control variables, without entropy balancing weights and with control variables, with entropy balancing weights and without control variables and with entropy balancing weights and with control variables. For control variables see Table 2 above.

Figure A3: Prospects for postdoc position – by estimation strategy



Source: Nacaps 2018, first wave. Own calculations. N = 14,915.
Note: Plot of regression coefficients for main effects of gender, gender match and interaction effect of model without entropy-balancing weights and without control variables, without entropy-balancing weights and with control variables, with entropy-balancing weights and without control variables and with entropy-balancing weights and with control variables. For control variables see Table 2 above.

Appendix 2: Full regression models

Table A1: Satisfaction with mentoring, belief in own research abilities and prospects for postdoc position (Unstandardized coefficients from linear regression models)

	Satisfaction with mentoring		Belief in own research abilities		Prospects for post-doc position	
	M1a: Entropy balancing without controls	M1b: Entropy balancing with controls	M2a: Entropy balancing without controls	M2b: Entropy balancing with controls	M3a: Entropy balancing without controls	M3b: Entropy balancing with controls
Female (ref.: Male)	-0.13***	-0.10***	-0.17***	-0.15***	-0.42***	-0.31***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.07)	(0.07)
Gender match (ref.: No gender match)	-0.06**	-0.06**	-0.05**	-0.05**	0.16**	0.06
	(0.02)	(0.02)	(0.02)	(0.02)	(0.05)	(0.05)
Interaction: Female x gender match	0.14***	0.14***	0.06	0.07*	-0.35**	0.04
	(0.04)	(0.04)	(0.04)	(0.03)	(0.11)	(0.10)
Ascribed and socio-demographic characteristics and characteristics of doctorate						
Age		-0.01***		0.00*		-0.03***
		(0.00)		(0.00)		(0.01)
Father: Higher education degree (ref.: Father: No higher education degree)		-0.07***		-0.04*		0.08
		(0.02)		(0.02)		(0.05)
Father: Doctoral degree		-0.05		-0.03		0.10
		(0.03)		(0.03)		(0.08)
Mother: Higher Education degree (ref.: Mother: No higher education degree)		0.03		-0.02		0.13*
		(0.02)		(0.02)		(0.05)
Mother: Doctoral degree		0.04		0.04		0.00
		(0.05)		(0.04)		(0.12)
Born abroad (ref.: Born in Germany)		0.08		0.03		-0.08
		(0.04)		(0.04)		(0.11)
Father: Born abroad (ref.: Born in Germany)		-0.00		0.01		0.19
		(0.04)		(0.03)		(0.10)
Mother: Born abroad (ref.: Born in Germany)		-0.03		0.05		-0.12
		(0.04)		(0.03)		(0.10)
Arts & humanities (ref.: Social and behavioral sciences)		0.10**		0.28***		-0.85***
		(0.03)		(0.03)		(0.08)
Biology (ref.: Social and behavioral sciences)		-0.03		0.28***		0.94***
		(0.04)		(0.03)		(0.10)

	Satisfaction with mentoring		Belief in own research abilities		Prospects for post-doc position	
	M1a: Entropy balancing without controls	M1b: Entropy balancing with controls	M2a: Entropy balancing without controls	M2b: Entropy balancing with controls	M3a: Entropy balancing without controls	M3b: Entropy balancing with controls
Medicine (ref.: Social and behavioral sciences)		0.07 (0.04)		-0.08** (0.03)		0.73*** (0.09)
STEM (ref.: Social and behavioral sciences)		-0.07** (0.03)		0.14*** (0.02)		0.74*** (0.06)
Other subjects (ref.: Social and behavioral sciences)		0.06 (0.05)		0.09* (0.04)		0.65*** (0.13)
Program/scholarship (ref.: Appointment		0.07*** (0.02)		-0.10*** (0.02)		-0.02 (0.05)
'Free' doctorate (ref.: Appointment)		-0.06* (0.03)		-0.37*** (0.02)		-0.80*** (0.07)
Final grade HE degree		-0.00 (0.02)		-0.21*** (0.02)		-0.58*** (0.05)
Child/children (ref.: No child/children)		0.05 (0.03)		0.01 (0.02)		0.01 (0.07)
Partner (ref.: No partner)		-0.18*** (0.03)		-0.01 (0.03)		0.26** (0.08)
Partner: Vocational training (ref.: Partner with higher education degree)		0.00 (0.03)		0.07** (0.02)		-0.06 (0.07)
Partner: Doctoral degree (ref.: Partner with higher education degree)		0.03 (0.03)		0.08** (0.03)		0.12 (0.09)
Partner: Part-time employed (ref.: Partner full-time employed)		0.09*** (0.03)		0.08*** (0.02)		-0.23*** (0.07)
Partner: In training or parental leave (ref.: Partner full-time employed)		0.03 (0.03)		-0.02 (0.03)		0.05 (0.08)
Partner: Not employed (ref.: Partner full-time employed)		0.20*** (0.04)		0.03 (0.03)		-0.23* (0.10)
Partner: Not in academia (ref.: Partner in academia)		0.07* (0.03)		-0.07** (0.02)		-0.19** (0.07)

	Satisfaction with mentoring		Belief in own research abilities		Prospects for post-doc position	
	M1a: Entropy balancing without controls	M1b: Entropy balancing with controls	M2a: Entropy balancing without controls	M2b: Entropy balancing with controls	M3a: Entropy balancing without controls	M3b: Entropy balancing with controls
Health and personality traits						
Health	0.08***		-0.04***		-0.04	
	(0.01)		(0.01)		(0.03)	
Big5: Extraversion	-0.02		-0.04***		0.00	
	(0.01)		(0.01)		(0.03)	
Big5: Neuroticism	-0.06***		-0.10***		-0.18***	
	(0.01)		(0.01)		(0.03)	
Big5: Openness	-0.01		0.09***		0.12***	
	(0.01)		(0.01)		(0.03)	
Big5: Conscientiousness	0.01		0.17***		0.01	
	(0.01)		(0.01)		(0.03)	
Big5: Agreeableness	0.01		-0.04***		-0.06*	
	(0.01)		(0.01)		(0.03)	
Risk-taking	-0.02*		-0.01		0.01	
	(0.01)		(0.01)		(0.02)	
Control beliefs	0.16***		-0.00		0.21***	
	(0.01)		(0.01)		(0.03)	
Self-efficacy	0.04**		0.15***		0.22***	
	(0.01)		(0.01)		(0.03)	
Individual attitudes						
Interested in the issue	0.10***		0.02*		-0.06*	
	(0.01)		(0.01)		(0.03)	
Contribution to scientific progress	0.05***		0.13***		0.11***	
	(0.01)		(0.01)		(0.03)	
Common in my discipline	0.02**		0.05***		0.19***	
	(0.01)		(0.01)		(0.02)	
Personal environment expects it	-0.01		-0.02**		0.05*	
	(0.01)		(0.01)		(0.02)	
Nothing else came about	-0.02*		0.01		-0.04	
	(0.01)		(0.01)		(0.02)	
Contribute to solving societal problems	-0.01		0.05***		0.07***	
	(0.01)		(0.01)		(0.02)	

	Satisfaction with mentoring		Belief in own research abilities		Prospects for post-doc position	
	M1a: Entropy balancing without controls	M1b: Entropy balancing with controls	M2a: Entropy balancing without controls	M2b: Entropy balancing with controls	M3a: Entropy balancing without controls	M3b: Entropy balancing with controls
Increase my reputation		0.00		-0.01		-0.01
		(0.01)		(0.01)		(0.02)
Improve career opportunities outside academia		-0.02**		-0.01*		-0.07***
		(0.01)		(0.01)		(0.02)
Managerial responsibility		-0.04***		-0.01		-0.04
		(0.01)		(0.01)		(0.02)
Compatibility of work and family		0.04***		0.01		0.05*
		(0.01)		(0.01)		(0.03)
Availability of resources		0.02		0.01		0.01
		(0.01)		(0.01)		(0.03)
Good opportunities for advancement		-0.02		0.04***		0.01
		(0.01)		(0.01)		(0.03)
Societal recognition		-0.01		-0.04***		0.03
		(0.01)		(0.01)		(0.02)
Job security		0.02		0.03**		-0.06*
		(0.01)		(0.01)		(0.03)
Societal benefits of work		-0.01		-0.04***		-0.07**
		(0.01)		(0.01)		(0.02)
Salary level		0.03*		0.01		0.01
		(0.01)		(0.01)		(0.03)
Autonomy in decision-making		0.01		0.06***		0.11***
		(0.01)		(0.01)		(0.03)
Working in a team		-0.03**		-0.03***		-0.01
		(0.01)		(0.01)		(0.02)
Intellectual challenge		0.06***		0.08***		0.06
		(0.01)		(0.01)		(0.03)
Location		-0.05*		-0.05**		-0.09
		(0.02)		(0.02)		(0.05)
Good research conditions in my discipline		0.15***		0.03		0.17**
		(0.02)		(0.02)		(0.05)
Supervisor		0.62***		-0.00		0.07
		(0.02)		(0.02)		(0.06)

	Satisfaction with mentoring		Belief in own research abilities		Prospects for post-doc position	
	M1a: Entropy balancing without controls	M1b: Entropy balancing with controls	M2a: Entropy balancing without controls	M2b: Entropy balancing with controls	M3a: Entropy balancing without controls	M3b: Entropy balancing with controls
Good reputation of the university		-0.04 (0.02)		-0.00 (0.02)		0.04 (0.06)
Attractive services for doctoral candidates		0.09* (0.04)		-0.00 (0.03)		0.01 (0.10)
It just came about that way		0.11*** (0.03)		-0.02 (0.03)		-0.14 (0.08)
Other reasons		-0.01 (0.03)		0.04 (0.02)		-0.12 (0.06)
Constant	3.80*** (0.02)	2.59*** (0.14)	3.82*** (0.01)	2.83*** (0.12)	5.09*** (0.04)	5.50*** (0.35)
N	15333	15333	15280	15280	14915	14915
R^2	0.001	0.154	0.004	0.227	0.009	0.146

Source: Nacaps 2018, first wave. Own calculations. Standard errors in parentheses.
Level of significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A2: Satisfaction with mentoring – by subject (Unstandardized coefficients from linear regression models)

	Satisfaction with mentoring				
	Art & humanities	Social & behavioral sciences	Medicine	Biology	STEM
Female (ref.: Male)	-0.18**	-0.07	-0.16	-0.17	-0.06
	(0.06)	(0.06)	(0.09)	(0.09)	(0.07)
Gender match (ref.: No gender match)	-0.05	-0.08	-0.10	-0.10	-0.02
	(0.06)	(0.05)	(0.08)	(0.08)	(0.03)
Interaction: Female x gender match	0.17	0.07	0.31*	0.09	0.14
	(0.09)	(0.08)	(0.12)	(0.12)	(0.10)
Ascribed and socio-demographic characteristics and characteristics of doctorate					
Age	0.00	-0.00	-0.01*	-0.02*	-0.03***
	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)
Father: Higher education degree (ref.: Father: No higher education degree)	-0.11*	-0.19***	0.05	-0.14*	-0.06*
	(0.05)	(0.05)	(0.08)	(0.07)	(0.03)
Father: Doctoral degree (ref.: Father: No higher education degree)	-0.13	-0.11	-0.09	0.07	-0.03
	(0.08)	(0.07)	(0.09)	(0.12)	(0.05)
Mother: Higher education degree (ref.: Mother: No higher education degree)	-0.03	0.04	-0.05	0.11	0.14***
	(0.05)	(0.05)	(0.07)	(0.07)	(0.03)
Mother: Doctoral degree (ref.: Mother: No higher education degree)	-0.21	0.21*	-0.00	0.01	0.12
	(0.11)	(0.10)	(0.12)	(0.16)	(0.08)
Born abroad (ref.: Born in Germany)	0.16	0.36***	-0.03	-0.18	-0.01
	(0.10)	(0.10)	(0.14)	(0.15)	(0.07)
Father: Born abroad (ref.: Born in Germany)	-0.18*	-0.13	0.11	0.32*	-0.00
	(0.09)	(0.08)	(0.13)	(0.13)	(0.07)
Mother: Born abroad (ref.: Born in Germany)	-0.01	-0.07	-0.20	-0.19	0.05
	(0.09)	(0.08)	(0.13)	(0.15)	(0.07)
Program/scholarship (ref.: Appointment)	0.08	-0.01	-0.13	0.11	0.11***
	(0.05)	(0.05)	(0.08)	(0.06)	(0.03)

	Satisfaction with mentoring				
	Art & humanities	Social & behavioral sciences	Medicine	Biology	STEM
'Free' doctorate (ref.: Appointment)	-0.03 (0.06)	-0.06 (0.05)	-0.07 (0.08)	0.02 (0.13)	-0.15** (0.06)
Final grade HE degree	0.01 (0.06)	-0.00 (0.04)	0.08 (0.06)	0.08 (0.08)	0.00 (0.04)
Child/children (ref.: No child/children)	-0.08 (0.06)	0.12* (0.06)	0.05 (0.10)	0.23 (0.12)	0.14** (0.05)
Partner (ref.: No partner)	-0.23** (0.09)	-0.34*** (0.08)	-0.40*** (0.11)	-0.08 (0.10)	-0.05 (0.05)
Partner: Vocational training (ref.: Partner with higher education degree)	-0.06 (0.07)	-0.06 (0.06)	0.03 (0.09)	0.11 (0.09)	0.06 (0.04)
Partner: Doctoral degree (ref.: Partner with higher education degree)	0.06 (0.08)	0.04 (0.08)	0.22* (0.10)	0.02 (0.10)	0.04 (0.06)
Partner: Part-time employed (ref.: Partner full-time employed)	0.07 (0.06)	0.15** (0.06)	0.21* (0.10)	0.04 (0.10)	-0.02 (0.04)
Partner: In training or parental leave (ref.: Partner full-time employed)	0.10 (0.09)	-0.04 (0.07)	0.17 (0.11)	-0.11 (0.13)	-0.03 (0.05)
Partner: Not employed (ref.: Partner full-time employed)	0.20 (0.11)	0.38*** (0.10)	0.35* (0.14)	-0.09 (0.13)	0.17** (0.06)
Partner: Not in academia (ref.: Partner in academia)	0.11 (0.07)	0.22*** (0.06)	0.05 (0.09)	-0.13 (0.09)	-0.02 (0.04)

	Satisfaction with mentoring				
	Art & humanities	Social & behavioral sciences	Medicine	Biology	STEM
Health and personality traits					
Health	0.08**	0.07*	0.07	0.11**	0.07***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.02)
Big5: Extraversion	-0.05	-0.05	0.11*	-0.08	0.01
	(0.03)	(0.03)	(0.04)	(0.04)	(0.02)
Big5: Neuroticism	-0.01	-0.06*	0.03	-0.11**	-0.10***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.02)
Big5: Openness	0.04	0.02	-0.02	-0.08	-0.04
	(0.03)	(0.03)	(0.04)	(0.05)	(0.02)
Big5: Conscientiousness	0.10**	0.00	-0.03	-0.03	0.00
	(0.03)	(0.03)	(0.05)	(0.05)	(0.02)
Big5: Agreeableness	-0.01	0.03	-0.00	0.02	0.01
	(0.03)	(0.02)	(0.04)	(0.04)	(0.02)
Risk-taking	0.02	-0.04*	0.00	-0.05	-0.02
	(0.02)	(0.02)	(0.03)	(0.03)	(0.01)
Control beliefs	0.18***	0.21***	0.16***	0.12**	0.11***
	(0.03)	(0.03)	(0.05)	(0.05)	(0.02)
Self-efficacy	0.01	0.01	-0.02	0.07	0.06**
	(0.03)	(0.03)	(0.04)	(0.04)	(0.02)
Individual attitudes					
Interested in the issue	0.03	0.12***	0.06	0.13**	0.11***
	(0.03)	(0.02)	(0.04)	(0.04)	(0.02)
Contribution to scientific progress	0.09***	0.05*	0.00	0.03	0.07***
	(0.03)	(0.02)	(0.04)	(0.04)	(0.02)
Common in my discipline	-0.02	0.02	0.08**	0.02	0.02
	(0.02)	(0.02)	(0.03)	(0.03)	(0.01)
Personal environment expects it	0.02	-0.04	-0.03	0.00	-0.02
	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)
Nothing else came about	-0.05*	0.01	-0.10**	0.00	0.00
	(0.02)	(0.02)	(0.03)	(0.03)	(0.01)
Contribute to solving societal problems	0.01	-0.02	0.01	0.01	-0.00
	(0.02)	(0.02)	(0.03)	(0.03)	(0.01)
Increase my reputation	-0.02	-0.01	-0.04	-0.01	0.02
	(0.02)	(0.02)	(0.03)	(0.03)	(0.01)

	Satisfaction with mentoring				
	Art & humanities	Social & behavioral sciences	Medicine	Biology	STEM
Improve career opportunities outside	-0.03 (0.02)	0.00 (0.02)	-0.04 (0.03)	-0.04 (0.03)	-0.03* (0.01)
Managerial responsibility	-0.06** (0.02)	-0.03 (0.02)	-0.09** (0.03)	-0.04 (0.03)	-0.03* (0.01)
Compatibility of work and family	0.05* (0.02)	0.03 (0.02)	0.06 (0.04)	-0.02 (0.04)	0.02 (0.02)
Availability of resources	0.03 (0.03)	0.01 (0.02)	0.06 (0.04)	-0.05 (0.04)	0.02 (0.02)
Good opportunities for advancement	0.01 (0.03)	-0.07* (0.03)	-0.02 (0.04)	0.06 (0.05)	0.00 (0.02)
Societal recognition	0.05* (0.03)	-0.02 (0.02)	-0.04 (0.04)	-0.05 (0.04)	-0.01 (0.02)
Job security	-0.01 (0.03)	0.01 (0.02)	0.00 (0.04)	0.03 (0.04)	0.04* (0.02)
Societal benefits of work	-0.03 (0.03)	-0.02 (0.02)	0.01 (0.04)	0.02 (0.04)	-0.01 (0.02)
Salary level	-0.00 (0.03)	0.06* (0.03)	0.01 (0.04)	0.00 (0.04)	0.00 (0.02)
Autonomy in decision-making	0.00 (0.03)	0.06* (0.03)	0.09* (0.04)	0.04 (0.04)	-0.02 (0.02)
Working in a team	0.01 (0.02)	-0.03 (0.02)	-0.08* (0.04)	-0.04 (0.04)	-0.03* (0.02)
Intellectual challenge	-0.03 (0.03)	0.05 (0.03)	0.04 (0.05)	0.04 (0.05)	0.10*** (0.02)
Location	-0.03 (0.05)	-0.04 (0.04)	-0.14 (0.08)	0.06 (0.07)	-0.07* (0.03)
Good research conditions in my discipline	0.21*** (0.05)	0.19*** (0.05)	0.16 (0.09)	0.08 (0.08)	0.10** (0.03)
Supervisor	0.56*** (0.07)	0.63*** (0.05)	0.49*** (0.08)	0.62*** (0.07)	0.70*** (0.03)
Good reputation of the university	-0.04 (0.06)	-0.00 (0.05)	-0.17 (0.10)	-0.08 (0.08)	-0.00 (0.04)
Attractive services for doctoral candidates	0.09 (0.09)	0.17* (0.08)	0.01 (0.15)	-0.12 (0.12)	0.01 (0.07)

	Satisfaction with mentoring				
	Art & humanities	Social & behavioral sciences	Medicine	Biology	STEM
It just came about that way	-0.00 (0.10)	0.14 (0.08)	-0.04 (0.11)	0.19 (0.11)	0.19*** (0.05)
Other reasons	0.06 (0.06)	-0.01 (0.05)	-0.17 (0.09)	0.06 (0.09)	-0.02 (0.04)
Constant	2.79*** (0.34)	2.66*** (0.31)	3.58*** (0.47)	3.18*** (0.55)	2.68*** (0.24)
N	2420	3286	1625	1415	5973
R^2	0.173	0.170	0.133	0.168	0.195

Source: Nacaps 2018, first wave. Own calculations. Standard errors in parentheses.
Level of significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A3: Belief in own research abilities – by subject (Unstandardized coefficients from linear regression models)

	Belief in own research abilities				
	Art & humanities	Social & behavioral sciences	Medicine	Biology	STEM
Female (Ref.: male)	-0.22***	-0.13*	-0.14	0.02	-0.19**
	(0.05)	(0.05)	(0.07)	(0.07)	(0.06)
Gender match (ref.: No gender match)	-0.06	-0.02	0.02	0.08	-0.10***
	(0.05)	(0.04)	(0.06)	(0.06)	(0.02)
Interaction: Female x gender match	0.09	-0.05	-0.03	-0.08	0.18*
	(0.07)	(0.07)	(0.10)	(0.09)	(0.08)
Ascribed and socio-demographic characteristics and characteristics of doctorate					
Age	0.00	-0.00	-0.00	0.02*	0.00
	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)
Father: Higher education degree (ref.: Father: No higher education degree)	0.02	-0.10*	-0.13*	-0.13**	-0.01
	(0.04)	(0.04)	(0.06)	(0.05)	(0.03)
Father: Doctoral degree (ref.: Father: No higher education degree)	0.05	-0.05	-0.11	-0.09	-0.00
	(0.07)	(0.06)	(0.07)	(0.09)	(0.04)
Mother: Higher education degree (ref.: Mother: No higher education degree)	-0.02	0.04	-0.04	0.10	-0.04
	(0.04)	(0.04)	(0.06)	(0.05)	(0.03)
Mother: Doctoral degree (ref.: Mother: No higher education degree)	0.02	0.18*	-0.13	0.11	0.04
	(0.09)	(0.08)	(0.10)	(0.11)	(0.07)
Born abroad (ref.: Born in Germany)	0.03	0.16	0.04	-0.11	-0.01
	(0.09)	(0.09)	(0.11)	(0.11)	(0.05)
Father: Born abroad (ref.: Born in Germany)	-0.02	-0.06	-0.08	0.15	0.04
	(0.08)	(0.07)	(0.11)	(0.10)	(0.06)
Mother: Born abroad (ref.: Born in Germany)	0.03	0.06	0.11	0.11	0.03
	(0.08)	(0.07)	(0.11)	(0.11)	(0.05)
Program/scholarship (ref.: Appointment)	-0.12**	-0.14***	-0.39***	-0.12*	-0.03
	(0.04)	(0.04)	(0.06)	(0.05)	(0.02)

	Belief in own research abilities				
	Art & humanities	Social & behavioral sciences	Medicine	Biology	STEM
'Free' doctorate (ref.: Appointment)	-0.35*** (0.05)	-0.40*** (0.05)	-0.56*** (0.06)	-0.19* (0.09)	-0.27*** (0.05)
Final grade HE degree	-0.21*** (0.05)	-0.31*** (0.04)	-0.18*** (0.04)	-0.18** (0.06)	-0.14*** (0.03)
Child/children (ref.: No child/children)	-0.09 (0.05)	0.13* (0.05)	0.04 (0.08)	-0.01 (0.09)	0.01 (0.04)
Partner (ref. No Partner)	-0.13 (0.07)	0.15* (0.07)	-0.04 (0.09)	0.02 (0.08)	-0.03 (0.04)
Partner: Vocational training (ref.: Partner with higher education degree)	0.04 (0.06)	0.15** (0.05)	0.04 (0.07)	0.03 (0.07)	0.13*** (0.03)
Partner: Doctoral degree[j] (ref.: Partner with higher education degree)	0.13 (0.07)	-0.06 (0.07)	0.19* (0.08)	0.17* (0.08)	0.07 (0.05)
Partner: Part-time employed (ref.: Partner full-time employed)	0.07 (0.05)	0.01 (0.05)	0.07 (0.08)	0.04 (0.07)	0.10** (0.04)
Partner: In training or parental leave (ref.: Partner full-time employed)	0.00 (0.07)	-0.14* (0.06)	0.05 (0.09)	0.07 (0.09)	0.04 (0.04)
Partner: Not employed (ref.: Partner full-time employed)	0.20* (0.09)	-0.06 (0.08)	0.00 (0.11)	-0.02 (0.09)	0.01 (0.05)
Partner: Not in academia (ref.: Partner in academia)	0.11 (0.06)	-0.14* (0.06)	-0.09 (0.08)	-0.02 (0.07)	-0.16*** (0.04)

	Belief in own research abilities				
	Art & humanities	Social & behavioral sciences	Medicine	Biology	STEM
Health and personality traits					
Health	-0.02	-0.12***	-0.05	-0.02	-0.01
	(0.03)	(0.02)	(0.04)	(0.03)	(0.02)
Big5: Extraversion	-0.05	0.02	-0.10**	-0.04	-0.04*
	(0.03)	(0.02)	(0.03)	(0.03)	(0.02)
Big5: Neuroticism	-0.10***	-0.06**	-0.06	-0.13***	-0.09***
	(0.03)	(0.02)	(0.04)	(0.03)	(0.02)
Big5: Openness	0.12***	0.08**	0.15***	0.01	0.09***
	(0.03)	(0.02)	(0.03)	(0.03)	(0.02)
Big5: Conscientiousness	0.24***	0.15***	0.04	0.11**	0.18***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.02)
Big5: Agreeableness	-0.06**	-0.02	-0.01	0.03	-0.05***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.01)
Risk-taking	-0.05***	0.00	-0.04	-0.00	0.02
	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)
Control beliefs	-0.04	0.03	-0.08*	0.02	0.00
	(0.03)	(0.03)	(0.04)	(0.03)	(0.02)
Self-efficacy	0.08***	0.16***	0.19***	0.17***	0.16***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)
Individual attitudes					
Interested in the issue	-0.02	-0.01	0.04	-0.04	0.03*
	(0.03)	(0.02)	(0.03)	(0.03)	(0.01)
Contribution to scientific progress	0.19***	0.17***	0.10***	0.09**	0.11***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.01)
Common in my discipline	0.07***	0.07***	0.03	0.02	0.04***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)
Personal environment expects it	-0.01	-0.07***	-0.05*	-0.05*	0.03*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)
Nothing else came about	0.02	0.03	0.02	-0.03	-0.01
	(0.02)	(0.02)	(0.03)	(0.02)	(0.01)
Contribute to solving societal problems	0.03	0.03	0.09***	0.01	0.06***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)
Increase my reputation	0.01	-0.06**	0.02	-0.00	-0.03*
	(0.02)	(0.02)	(0.03)	(0.02)	(0.01)

	Belief in own research abilities				
	Art & humanities	Social & behavioral sciences	Medicine	Biology	STEM
Improve career opportunities outside academia	-0.05*** (0.01)	0.00 (0.02)	0.05* (0.02)	-0.03 (0.02)	-0.01 (0.01)
Managerial responsibility	0.03 (0.02)	-0.03 (0.02)	0.02 (0.03)	-0.01 (0.02)	-0.01 (0.01)
Compatibility of work and family	0.03 (0.02)	-0.03 (0.02)	0.03 (0.03)	0.02 (0.03)	0.03* (0.01)
Availability of resources	-0.02 (0.02)	0.05* (0.02)	0.05 (0.03)	0.04 (0.03)	0.01 (0.02)
Good opportunities for advancement	0.04 (0.02)	0.01 (0.02)	0.11*** (0.03)	0.03 (0.03)	0.05** (0.02)
Societal recognition	-0.06** (0.02)	-0.01 (0.02)	-0.13*** (0.03)	-0.03 (0.03)	-0.03** (0.01)
Job security	0.03 (0.02)	0.04* (0.02)	-0.04 (0.03)	-0.02 (0.03)	0.04** (0.01)
Societal benefits of work	-0.07*** (0.02)	-0.02 (0.02)	-0.02 (0.03)	-0.06* (0.03)	-0.05*** (0.01)
Salary level	0.05 (0.02)	0.01 (0.02)	-0.01 (0.03)	-0.02 (0.03)	-0.00 (0.02)
Autonomy in decision-making	0.10*** (0.03)	0.08** (0.03)	0.12*** (0.03)	0.11*** (0.03)	0.06*** (0.02)
Working in a team	-0.03 (0.02)	-0.04* (0.02)	-0.07* (0.03)	-0.03 (0.03)	-0.02 (0.01)
Intellectual challenge	0.12*** (0.03)	0.10*** (0.03)	0.00 (0.04)	0.13*** (0.04)	0.02 (0.02)
Location	-0.03 (0.04)	-0.08* (0.04)	-0.07 (0.06)	-0.02 (0.05)	-0.05 (0.03)
Good research conditions in my discipline	-0.01 (0.05)	-0.09 (0.05)	0.16* (0.07)	0.05 (0.06)	0.05 (0.03)
Supervisor	0.04 (0.06)	0.03 (0.05)	0.06 (0.06)	-0.04 (0.05)	-0.09** (0.03)
Good reputation of the university	-0.01 (0.05)	-0.05 (0.05)	-0.13 (0.08)	0.04 (0.06)	0.06* (0.03)
Attractive services for doctoral candidates	0.06 (0.08)	-0.04 (0.07)	0.04 (0.12)	0.17 (0.09)	-0.04 (0.05)

	Belief in own research abilities				
	Art & humanities	Social & behavioral sciences	Medicine	Biology	STEM
It just came about that way	-0.04 (0.09)	-0.04 (0.07)	-0.13 (0.09)	0.01 (0.08)	0.00 (0.04)
Other reasons	0.06 (0.05)	0.07 (0.05)	-0.10 (0.07)	-0.01 (0.06)	0.09* (0.03)
Constant	2.77*** (0.28)	3.65*** (0.27)	3.24*** (0.38)	3.13*** (0.40)	2.90*** (0.19)
N	2412	3269	1613	1413	5962
R^2	0.266	0.215	0.335	0.233	0.200

Source: Nacaps 2018, first wave. Own calculations. Standard errors in parentheses.
Level of significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A4: Prospects for postdoc position – by subject (Unstandardised coefficients from linear regression models)

	Prospects for postdoc position				
	Art & humanities	Social & behavioral sciences	Medicine	Biology	STEM
Female (ref.: Male)	-0.50***	-0.30*	-0.10	-0.13	-0.33
	(0.15)	(0.14)	(0.21)	(0.21)	(0.20)
Gender match (ref.: No gender match)	0.27	-0.08	0.32	0.02	0.05
	(0.14)	(0.11)	(0.17)	(0.17)	(0.07)
Interaction: Female x gender match	0.14	0.05	-0.23	-0.50	0.35
	(0.20)	(0.19)	(0.28)	(0.28)	(0.28)
Ascribed and socio-demographic characteristics and characteristics of doctorate					
Age	-0.03**	-0.03*	-0.04**	0.01	-0.06***
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)
Father: Higher education degree (ref.: Father: No higher education degree)	-0.16	0.29**	-0.23	-0.27	0.26**
	(0.12)	(0.11)	(0.17)	(0.16)	(0.09)
Father: Doctoral degree (ref.: Father: No higher education degree)	0.01	0.10	-0.06	0.12	0.25
	(0.19)	(0.16)	(0.20)	(0.27)	(0.13)
Mother: Higher education degree (ref.: Mother: No higher education degree)	0.18	0.11	-0.02	0.40*	0.17
	(0.12)	(0.11)	(0.16)	(0.17)	(0.09)
Mother: Doctoral degree (ref.: Mother: No higher education degree)	0.47	0.26	0.15	-0.12	-0.33
	(0.25)	(0.23)	(0.28)	(0.36)	(0.22)
Born abroad[e] (ref.: Born in Germany)	0.80***	0.14	-1.18***	-0.23	-0.40*
	(0.24)	(0.23)	(0.31)	(0.35)	(0.18)
Father: Born abroad (ref.: Born in Germany)	-0.41*	0.03	0.08	0.69*	0.24
	(0.21)	(0.20)	(0.29)	(0.30)	(0.18)
Mother: Born abroad[e] (ref.: Born in Germany)	0.25	-0.22	0.34	-0.39	-0.10
	(0.22)	(0.21)	(0.30)	(0.33)	(0.18)
Program/scholarship (ref.: Appointment)	-0.47***	-0.22*	-0.54**	0.05	0.18*
	(0.12)	(0.11)	(0.18)	(0.15)	(0.08)

	Prospects for postdoc position				
	Art & humanities	Social & behavioral sciences	Medicine	Biology	STEM
'Free' doctorate (ref.: Appointment)	-1.15*** (0.13)	-0.75*** (0.13)	-0.85*** (0.18)	-1.27*** (0.30)	-0.63*** (0.15)
Final grade HE degree	-0.18 (0.13)	-0.70*** (0.11)	-0.25* (0.13)	-0.44* (0.18)	-0.92*** (0.10)
Child/children	-0.16 (0.14)	-0.03 (0.14)	-0.31 (0.22)	-0.38 (0.28)	0.24 (0.13)
Partner (ref.: No Partner)	0.13 (0.20)	-0.12 (0.19)	-0.20 (0.26)	0.39 (0.24)	0.48*** (0.14)
Partner: Vocational training (ref.: Partner with higher education degree)	0.02 (0.17)	0.04 (0.14)	0.14 (0.20)	-0.33 (0.22)	-0.28* (0.11)
Partner: Doctoral degree (ref.: Partner with higher education degree)	0.12 (0.20)	-0.08 (0.18)	0.05 (0.23)	0.01 (0.24)	0.37* (0.16)
Partner: Part-time employed (ref.: Partner full-time employed)	-0.06 (0.15)	0.01 (0.13)	-0.29 (0.22)	-0.07 (0.22)	-0.52*** (0.12)
Partner: In training or parental leave (ref.: Partner full-time employed)	0.27 (0.21)	-0.21 (0.16)	0.16 (0.25)	0.11 (0.29)	0.23 (0.13)
Partner: Not employed (ref.: Partner full-time employed)	-0.16 (0.25)	0.21 (0.23)	0.27 (0.32)	-0.38 (0.29)	-0.33* (0.17)
Partner: Not in academia (ref.: Partner in academia)	-0.16 (0.17)	0.14 (0.16)	0.20 (0.22)	-0.08 (0.21)	-0.39** (0.12)

	Prospects for postdoc position				
	Art & humanities	Social & behavioral sciences	Medicine	Biology	STEM
Health and personality traits					
Health	-0.04	-0.05	-0.06	0.11	-0.07
	(0.07)	(0.06)	(0.10)	(0.09)	(0.05)
Big5: Extraversion	-0.08	0.04	0.04	0.09	0.03
	(0.07)	(0.07)	(0.10)	(0.09)	(0.05)
Big5: Neuroticism	-0.17*	-0.14*	-0.09	-0.10	-0.18***
	(0.07)	(0.06)	(0.10)	(0.10)	(0.05)
Big5: Openness	0.05	0.09	0.11	-0.17	0.21***
	(0.08)	(0.07)	(0.10)	(0.10)	(0.06)
Big5: Conscientiousness	-0.13	0.02	0.26*	-0.03	-0.02
	(0.08)	(0.07)	(0.10)	(0.11)	(0.06)
Big5: Agreeableness	-0.14*	-0.05	-0.24**	0.03	0.01
	(0.06)	(0.06)	(0.09)	(0.09)	(0.04)
Risk-taking	0.03	0.05	-0.00	-0.06	-0.02
	(0.04)	(0.04)	(0.06)	(0.06)	(0.03)
Control beliefs	0.29***	0.33***	0.14	-0.03	0.11
	(0.07)	(0.07)	(0.11)	(0.10)	(0.06)
Self-efficacy	0.09	0.17*	0.23*	0.50***	0.28***
	(0.07)	(0.07)	(0.10)	(0.10)	(0.05)
Individual attitudes					
Interested in the issue	-0.07	-0.17**	0.15	0.13	-0.09*
	(0.07)	(0.06)	(0.08)	(0.09)	(0.04)
Contribution to scientific progress	0.19**	0.20***	0.11	0.02	0.01
	(0.06)	(0.06)	(0.08)	(0.10)	(0.05)
Common in my discipline	0.07	0.09	0.14*	0.22***	0.24***
	(0.04)	(0.05)	(0.06)	(0.06)	(0.03)
Personal environment expects it	0.18**	0.05	0.09	-0.07	-0.03
	(0.05)	(0.05)	(0.06)	(0.07)	(0.04)
Nothing else came about	-0.20***	0.03	0.08	0.02	-0.08*
	(0.05)	(0.05)	(0.07)	(0.06)	(0.04)
Contribute to solving societal problems	0.09	-0.01	-0.01	-0.08	0.14***
	(0.05)	(0.05)	(0.07)	(0.07)	(0.04)
Increase my reputation	-0.08	-0.10	0.04	0.09	0.01
	(0.05)	(0.05)	(0.07)	(0.07)	(0.04)

	Prospects for postdoc position				
	Art & humanities	Social & behavioral sciences	Medicine	Biology	STEM
Improve career opportunities outside academia	-0.10* (0.04)	-0.07 (0.04)	0.06 (0.06)	-0.01 (0.07)	-0.13*** (0.03)
Managerial responsibility	-0.06 (0.05)	-0.10 (0.05)	0.04 (0.07)	0.11 (0.07)	-0.03 (0.04)
Compatibility of work and family	0.08 (0.06)	0.10 (0.06)	-0.20* (0.09)	-0.00 (0.08)	0.02 (0.05)
Availability of resources	-0.07 (0.06)	0.01 (0.06)	0.14 (0.10)	0.21* (0.10)	0.00 (0.05)
Good opportunities for advancement	0.10 (0.07)	0.01 (0.07)	-0.16 (0.10)	-0.08 (0.10)	0.07 (0.05)
societal recognition	0.08 (0.06)	0.01 (0.05)	0.01 (0.08)	-0.12 (0.08)	0.06 (0.04)
Job security	-0.21** (0.06)	-0.07 (0.05)	-0.37*** (0.09)	-0.29** (0.10)	0.09* (0.04)
Societal benefits of work	-0.07 (0.06)	-0.02 (0.05)	0.18* (0.08)	0.08 (0.08)	-0.13*** (0.04)
Salary level	-0.03 (0.07)	0.08 (0.07)	0.08 (0.09)	0.08 (0.10)	-0.00 (0.05)
Autonomy in decision-making	0.09 (0.07)	0.07 (0.07)	0.04 (0.10)	0.30** (0.10)	0.10 (0.05)
Working in a team	0.03 (0.05)	-0.10 (0.05)	-0.08 (0.08)	-0.03 (0.08)	0.01 (0.04)
Intellectual challenge	-0.03 (0.08)	0.08 (0.07)	0.01 (0.10)	0.12 (0.11)	0.05 (0.05)
Location	-0.10 (0.11)	-0.19 (0.11)	-0.43* (0.17)	-0.31 (0.16)	-0.00 (0.09)
Good research conditions in my discipline	0.20 (0.13)	0.19 (0.12)	0.76*** (0.20)	0.29 (0.17)	0.01 (0.09)
Supervisor	-0.01 (0.15)	0.10 (0.13)	-0.14 (0.17)	-0.41* (0.16)	0.21* (0.09)
Good reputation of the university	0.09 (0.15)	0.17 (0.13)	0.09 (0.21)	0.13 (0.19)	-0.25* (0.10)
Attractive services for doctoral candidates	-0.01 (0.21)	0.02 (0.19)	0.61 (0.35)	-0.43 (0.28)	0.09 (0.18)

	Prospects for postdoc position				
	Art & humanities	Social & behavioral sciences	Medicine	Biology	STEM
It just came about that way	0.10 (0.23)	0.12 (0.18)	-0.46 (0.24)	-0.44 (0.25)	-0.22 (0.14)
Other reasons	0.15 (0.14)	-0.16 (0.13)	-0.33 (0.20)	-0.25 (0.20)	-0.29** (0.11)
Constant	5.16*** (0.78)	6.25*** (0.74)	7.34*** (1.06)	3.49** (1.25)	7.61*** (0.63)
N	2376	3194	1555	1393	5809
R^2	0.153	0.093	0.165	0.164	0.108

Source: Nacaps 2018, first wave. Own calculations. Standard errors in parentheses.
Level of significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.