



Early career performance and its correlation with gender and publication output during doctoral education

Jonas Lindahl¹ · Cristian Colliander^{1,2} · Rickard Danell¹

Received: 7 March 2019 / Published online: 9 November 2019
© The Author(s) 2019

Abstract

Publishing in peer-reviewed journals as a part of the doctoral education is common practice in many countries. The publication output of doctoral students is increasingly used in selection processes for funding and employment in their early careers. Against the backdrop of this development, the aim of this study is to examine (1) how performance during the doctoral education affect the probability of attaining research excellence in the early career; and (2) if there is performance differences between males and females in the early career and to which degree these gender differences can be explained by performance differences during the doctoral education. The data consist of Swedish doctoral students employed at the faculty of science and technology and the faculty of medicine at a Swedish university. Our main conclusions are that (1) research performance during the doctoral education has a positive effect on attaining excellence in the early career; (2) there is an interaction between publication volume and excellence during doctoral education suggesting that a combination of quantity and quality in doctoral students' performance is indicative of future excellence; (3) there are performance differences in the early career indicating that males have a higher probability of attaining excellence than females, and; (4) this difference is partly explained by performance differences during the doctoral education.

Keywords Early career · Doctoral student · Excellence · Gender · Indicator · Bibliometrics · Performance · Prediction

Introduction

In recent decades, there has been a clear shift in Swedish research policy with an increasing emphasis on targeted initiatives to support excellent researchers and strategic research areas (Hallonsten and Silander 2012). This development is not unique for Sweden but is part of an international trend towards a more active research policy motivated by an awareness of the strategic importance of scientific knowledge production as a means to solve

✉ Jonas Lindahl
jonas.lindahl@umu.se

¹ INFORSK (Department of Sociology), Umeå University, Umeå, Sweden

² University Library, Umeå University, Umeå, Sweden

social problems, and provide a competitive advantage in a globalized economy (Hallonsten and Silander 2012; Whitley 2006).

The changing political climate—emphasizing competitiveness and performance—marked by a move towards more active research policies with targeted support for excellent researchers and excellent research environments, has also gradually extended to researchers in the early career and doctoral students (Jones 2013). An example of this change in Sweden is the development of excellence programs with a focus on identifying excellent (i.e., the most promising) young researchers and concentrating resources on this top group (Hallonsten and Hugander 2014). An international example is the ECR starting grants funded by the European Union’s Horizon 2020 Framework Programme for Research and Innovation which uses excellence as the only criterion of evaluation.

Doctoral education in Sweden, and internationally, is gradually adapting to the changing political climate. The publication output of doctoral students is increasingly assessed in accordance with standard practices of research evaluation (Lee and Boud 2009) and “the competitive ‘bar’ is rising for doctoral students in terms of both quantity and quality” (Jones 2013, p 89). Publishing in peer-reviewed journals during doctoral studies is often encouraged and sometimes required as a graduation criterion (Mason 2018). On the one hand because higher education institutions are incentivized, e.g., by national performance-based research funding systems, to encourage students to publish during their doctoral education (Mason 2018). On the other hand due to a greater awareness of the increased competition for funding and positions in the academic labor market where quantitative indicators (e.g., publication based indicators) are increasingly used in the selection processes (Brischoux and Angelier 2015; Bornmann and Williams 2017).

The increased focus on competition and performance in the early career phase have resulted in an ongoing discussion about how factors during doctoral education might affect career development and future research performance (see e.g., Sinclair et al. 2014; Frick et al. 2016). The aim of this study is to make a contribution to this discussion (see e.g., Williamson and Cable 2003; Haslam and Laham 2009; Laurance et al. 2013; Pinheiro et al. 2014; Horta and Santos 2016), by examine (1) how performance, on the one hand defined as publication volume, and on the other hand as research excellence, during the doctoral education affect the probability of attaining research excellence in the early career (i.e., the years succeeding thesis completion); and (2) if there is performance differences between males and females in the early career and to which degree these gender differences can be explained by performance differences during the doctoral education. We examine a dataset consisting of Swedish doctoral students employed at the faculty of science and technology and the faculty of medicine at a single Swedish university. The focus on excellence is motivated by (1) the “intensified quest for so-called excellence in public research systems in the Western world” (Hallonsten and Hugander 2014, p 249); and (2) the fact that excellence has not received much attention in the previous literature on the relationship between performance during doctoral education and future performance (the previous literature is more thoroughly reviewed in section Previous research and hypothesis).

Swedish doctoral education

Postgraduate studies at a Swedish University is 240 higher education credits, equivalent to 4 years of full-time studies, which includes a course component of approximately 60 higher education credits. In practice, the doctoral student usually has other tasks such as teaching or administrative duties, equivalent to 20% of a full-time job. This means that

research education usually takes around 5 years, provided full-time studies are conducted. What is relatively unusual with the Swedish university system is that all doctoral students are employed by their institutions. Most employments of doctoral students are connected to ongoing research projects and the doctoral student must be guaranteed 4 years of employment. In science and medicine, and to an increasing extent in the social sciences, the doctoral student tends to be an integral part of a research group during their education. The doctoral students usually write aggregation theses usually consisting of at least four articles published in scientific journals and a summary part in which the dissertation project as a whole is described and summarized.

When doctoral students complete their studies and the thesis is approved, it is time to take the next step in their careers. Some doctoral students will choose to refrain from continuing on a research career within the university and instead focus on the external job market. Those who are of interest in this study are the students who remain at the university. In the Swedish system, a future employer will not be able to distinguish these individuals with reference to the thesis grades; Swedish dissertations only qualify for the grade approved or not approved. Instead, they need to compete with the content and scope of the merits they collected during their education. Since Sweden basically lacks a tenure track system, there are a few permanent positions that early career researchers can compete for after they have attained their doctorate. External funding is very difficult to obtain and new PhD's will usually be employed within senior researchers projects, or compete for the more attractive postdoc positions that still exists in the Swedish system. In such a system the merits accumulated during the study period is of great importance.

Previous research and hypothesis

Questions related to the issue of whether, to which degree, and why research performance is characterized by continuity, and thus predictability, has been studied since the 1960s (Horta and Santos 2016). Two leading theories are the sacred spark theory and the theory of cumulative advantage (Allison and Stewart 1974; Bornmann and Williams 2017). The basic assumption in the sacred spark theory is that differences in scientific performance among researchers can be explained by “substantial, predetermined differences among scientists in their ability and motivation to do creative research” (Allison and Stewart 1974). The sacred spark theory does not address the doctoral education or the early career phase in particular. However, in the context of this study, the sacred spark theory would suggest that students that perform well during their doctoral education do so because they are motivated and have the inherent ability to do creative research. As the career develops for these individuals their motivation and this inherent ability persist, and they continue to perform well (Bornmann and Williams 2017).

The theory of cumulative advantage assumes that there are feedback processes in science by which advantages accumulate over time and produce patterns of increasing inequality between researchers (DiPrete and Eirich 2006). The logic of cumulative advantage suggest that the occurrence of an event increase the probability of such an event recurring in the future (Price 1976). In the reputation-and-resources model of research careers in science, processes of cumulative advantage is explained on the basis of three premises: (1) researchers compete for a limited resources; (2) it is difficult to directly observe and measure research performance; and (3) resources in science is allocated with respect to the mertonian principles of universalism, i.e., recognition in science should be based solely on

quality, and communism, i.e., resource allocation in science should maximize the overall productivity of the community (DiPrete and Eirich 2006). The purpose of resource allocation in this model is to enable future research achievements, rather than rewarding past achievements (DiPrete and Eirich 2006). Following the logic of cumulative advantage the reputation-and-resource model suggest that scientific achievements in the past enhance the probability of getting resources, and resources increase the probability to produce quality research, which in turn enhance the probability of more resources and so on (Merton 1988). In the context of this study, the theory of cumulative advantage and the reputation-and-resource model would suggest that a high performance during doctoral education would enable resources (e.g., research funding) which in turn increases the probability of continued high performance in later career phases.

The sacred spark theory and the theory of cumulative advantage provide theoretical foundations for the assumption that there are continuity between past and future research performance. The number of empirical studies examining the relationship between research performance during the doctoral education and future research performance are somewhat limited (Pineiro et al. 2014). However, there are a few. Williamson and Cable (2003) examine predictors of early career performance (i.e., publication volume) among 152 management researchers and concludes that the number of publications during the doctoral education has a positive effect on early career performance (i.e., publication volume). Laurance et al. (2013) examine the effect of publication volume during doctoral education and five other predictors on future publication volume in a sample of 203 biologists who were affiliated with institutions in North America, South America, Europe, and Australia and found a positive relationship between publication during doctoral education and publication volume later in the career. Pineiro et al. (2014) examine the effect of publishing during the doctoral education among 1310 researchers in the United States, on research performance after attaining their PhD (i.e., the average number of publications per year after Ph.D.), and found that publishing at least once during doctoral studies had a positive effect on future research performance. Two articles (Hilmer and Hilmer 2009; Buchmueller et al. 1999), focus on the field of economics and conclude that publication volume during doctoral education has a positive effect on future publication performance. A common denominator of these five studies is a focus on publication volume and the conclusion that publishing during doctoral education has a positive effect on future publication volume.

Haslam and Laham (2009) examine publication records of 60 social psychologists in the USA to determine whether performance (as operationalized by publication and the journal impact factor) during the doctoral education, can predict research performance (i.e., publication and impact) 10 years later. Haslam and Laham (2009) concludes that performance during the doctoral education has a positive effect on future research performance. Horta and Santos (2016) examined the effect of performance (i.e., publication volume) during doctoral studies on researchers' future knowledge production and impact in a dataset comprising 664 Ph.D.'s in Portugal. Horta and Santos (2016) concludes that the number of publications during doctoral studies has a positive effect on future knowledge production and impact. Both Haslam and Laham (2009) and Horta and Santos (2016) concludes that performance (i.e., publication volume) during doctoral studies has a positive effect on future impact as operationalized by a citation based indicator. Our study contributes to the literature by focusing on the relationship between both publication volume and research excellence during doctoral education, and research excellence in the later career. The focus on both the quantitative (i.e., publication volume) and qualitative dimension (i.e., excellence) of performance pre-thesis completion have, to the best of our knowledge, not received much attention in the literature. However, previous research on the predictability

of research performance with a focus on career phases after the doctoral education have found that the qualitative dimension of research performance (i.e., citation based indicators) have been predictive of future performance (see e.g. Danell 2011; Havemann and Larsen 2015; Lindahl 2018).

In this study, for hypothesis 1, we expect that doctoral students who publish more in scientific journals and also publish excellent work when preparing their thesis will have a higher probability of doing excellent research in the early career (i.e., the years following the completion of the doctoral education). Thus, we hypothesize that both quantity and quality is indicative of future excellence. The relationship between the quantity and the quality of a researchers' publication output have been investigated within the scientometric community for a long time (see e.g., Cole and Cole 1967), generally indicting a moderate to strong correlation between quantity (i.e., publication rates) and quality (i.e., citation rates). Recently a number of studies have extended this line of research to the relationship between the number of published papers and the production of excellent research (i.e., highly cited papers) with similar results (Sandström and Van Den Besselaar 2016; Larivière et al. 2016; Abramo et al. 2014). In this study we are interested in how the quantity versus quality relationships translates the context of the relationship between performance during doctoral education and future performance. By investigating the quantity versus quality question in the context of predicting future performance on the basis of performance during doctoral education, we might learn more of how scientific knowledge production during the doctoral education might affect future research performance.

Our second hypothesis revolves around gender differences during doctoral education and the early career. There are a considerable amount of empirical studies that identified gender differences in working conditions and career development for male and female researchers in various national contexts. In summary, the literature indicates that: (1) women's scientific efforts are valued lower (Wennerås and Wold 1996; Bornmann et al. 2007), (2) female researchers still have a poorer career development than their male colleagues (Xie and Shauman 2003; Ginther and Kahn 2006; Kumar 2012; Danell and Hjerm 2013), and (3) female researchers tend to publish less than their male colleagues (Cole and Zuckerman 1984; Long 1992; Xie and Shauman 1998; Prpic 2002; Fox 2005).

In the literature, the importance of informal processes is highlighted to explain the different conditions of males and females in science. Suggested explanations for the observed gender bias is often Merton's (1968, 1988) idea of a "Matthew effect" in science, as well as Rossiters (1993) idea of a "Matilda effect", suggesting that female researchers systematically receives less recognition for their work compared with male colleagues. The bias against female performance suggested by the Matthew/Matilda effect would have a negative effect on the long-term career development for women in science. Other suggested explanation is the importance of informal processes affected by gender stereotypes, e.g., role congruity theory (Eagly and Karau 2002), and the use of varying criteria in assessing men's and women's competencies and efforts (Heilman and Okimoto 2007; Phelan et al. 2008).

The effect of gender is somewhat inconclusive in the empirical literature on the relationship between research performance and other factors during the doctoral education, and future research performance. Some studies indicate that gender have an effect on future performance suggesting that males perform better (see e.g., Hilmer and Hilmer 2009; Laurance et al. 2013; Pinheiro et al. 2014; Horta and Santos 2016). Other studies indicate that gender does not have an effect, or at least not a statistically significant effect (see e.g., Buchmueller et al. 1999; Williamson and Cable 2003; Haslam and Laham 2009). Gender has mostly been used as a control variable in the previous literature. In this study we try

to improve our understanding of gender differences and performance in the early career by focusing on performance during doctoral studies as a potential explanation of gender differences in attaining research excellence during the years following thesis completion.

Thus, for hypothesis 2 we expect (a) that there is gender differences in attaining research excellence in the early career, and (b) that these gender differences can be explained by performance differences during the doctoral education. We find these hypothesis motivated by the increasing focus on competition and research performance in academia and particularly during doctoral education and in the early career. If we find gender differences in the early career that can be explained by performance difference during the doctoral education, performance would seem to progress in accordance with science as a meritocracy (Van den Besselaar and Sandström 2016) and what we would expect from theory, e.g., cumulative advantages and the reputation-and-resource model (DiPrete and Eirich 2006). To address these issues we could then focus future research on gender based performance differences during doctoral education, and policy on potential ways to improve female performance during doctoral education (Van den Besselaar and Sandström 2016). If, on the other hand, we find gender differences in performance in the early career that cannot be explained by performance differences during doctoral education, there might be some other gender bias at work and we would have to look elsewhere to understand and address these issues.

Method

Data

The data consists of 475 doctoral students who completed their studies at a Swedish university between 2003 and 2009 at the faculty of science and technology and the faculty of medicine. We performed the analyses on 310 of those who were employed or associated with the university at least 1 year between the second to the fifth year after the year of completing their thesis. Publication data was collected from DiVA, a Swedish repository for research publications, and the citation indices accessible through Web of Science. Everyone employed at or associated with the Swedish university has a personal identification code, which was used to match doctoral students with their publications. We also utilized the salary system of the Swedish university to acquire information about employment, gender, and age.

Variables

The dependent variable indicates whether an author had attained relative excellence in the early career, i.e., the second to fifth year after the year of completing their thesis. This definition of the early career is in accordance with Bazeley (2003). Generally, in research evaluation, excellence is related to the concept of quality (Hellström 2011). Excellence can be understood as a “measure of quality in the sense of equaling high or very high (‘excellent’) academic quality” (Hellström 2011, p 117) and usually refers to a specific level of quality of research output. In this study we operationalize the concept of excellence in accordance with recommended best practice and use the percentile-based citation indicator “top 10%” (Bornmann 2013, 2014). In order to operationalize the indicator for relative excellence, we identified documents that were among the top 10% cited documents in their field (i.e., subject category), taking into account document type and year. Since a document can belong

to more than one subject category in Web of Science, it may be equal to or above the 90th percentile in one category and not in others. We therefore calculated in what fraction of categories the document was equal to or above the 90th percentile.

For each author, we summed the top 10% fractions for all documents that belonged to that author. An author that was equal to or above the 90th percentile in the distribution of summed top 10% fractions was defined as being an excellent author. Excellent authors had produced a sum of top 10% fractions of at least two and constituted 10.6% of the sample.

We constructed three predictor variables to test our hypothesis. Two predictors to operationalize aspects of performance (i.e., publication volume and excellence) and one predictor to examine the effect of gender:

1. Publication volume during doctoral studies is operationalized as the number of publications indexed in the citation indices of Web of Science. Coding: *#Publications during doctoral studies*.
2. Relative excellence during doctoral studies was operationalized in the same way as the dependent variable. Coding: *Relative excellence during doctoral studies*.
3. Performance differences and gender bias was examined with a binary predictor where the value 1 represent males and 0 represent females. Coding: *Male doctoral student*.

We constructed three control variables:

(1) We expect that doctoral students who are closely integrated into their research environment will have a higher probability for future excellence. Especially within the natural and life sciences, with a more collective model for researcher education, the doctoral thesis work is an important contribution to the supervisor’s project and the doctoral student is an integral part of a larger research group, which also influences the form and content of the doctoral student’s education (Austin 2009; Becher and Trowler 2001; Delamont et al. 2000; Golde 2005; Knorr Cetina 1999; Pyhältö et al. 2009). Doctoral students within large research teams are usually more productive during and after graduate education (Platow 2012), as the actual tutoring of doctoral students is distributed among more individuals, which is important for the students’ socialization and intellectual development (Austin 2002; Fenge 2012; Lee and Boud 2009).

Collaboration and the degree of integration into the research community was operationalized with the collaborative coefficient (Ajiferuke, Burell, and Tague 1988). The collaborative coefficient is a weighted mean that incorporates the average number of authors per paper and the proportion of multi-authored papers in a single measure that can be defined as:

$$\text{Collaborative coefficient} = 1 - \frac{f_1 + (1/2)f_2 + \dots + (1/k)f_k}{N} = 1 - \frac{\sum_{j=1}^k (1/j)f_j}{N}$$

where f_j denotes the number of j -authored papers, N denotes the total number of publications, and k is the highest number of co-authors per paper of an author. Coding: *Degree of collaboration*.

(2) We expect age to be a significant predictor. In their study, Costas et al. (2010), conclude that top researchers are younger than less competitive researchers’ independent of professional category (i.e., tenured scientist, research scientist, and research professor). The authors also showed that the impact of researchers output (as measured by a citation based

indicator) decreases with age (Costas et al. 2010). Finishing doctoral studies at a more advanced age, presumably has a negative influence on the probability for future excellence, because such an individual's entry into the scientific community implies a marked deviation from the general life cycle of age-related research productivity pattern and age-creativity patterns that are visible in many research fields (Fox 1983; Jones and Weinberg 2011; Rørstad and Aksnes 2015).

Age was operationalized as the age of the doctorate at the year of the defence of the thesis. Coding: *Age at completion of doctoral studies*.

(3) There might be disciplinary differences between the doctoral students working in the natural sciences and those working in the medical and health sciences. Previous research have found differences between disciplines in science concerning factors such as publication volume and citation patterns (see e.g., Schubert and Braun 1986; Vinkler 1986), collaboration (see e.g., Newman 2001), as well as their intellectual and social organization (see e.g., Whitley 2000). Therefore, we constructed a binary control variable where the value 1 denotes a doctoral student who's thesis is classified as belonging to the natural sciences ($n=100$) and the value 0 denotes a student who's thesis is classified as belonging to the health and medical sciences ($n=210$). We conducted the classification in accordance with the first level of the classification scheme *Standard for Swedish classification of research areas 2011* (HSV 2011). Coding: *Discipline*.

Authors who attained relative excellence during doctoral studies constituted 11% of the sample. Males constituted 46% of the sample and females constituted 54%. See Table 1 for descriptive statistics for the three non-binary predictors.

Modeling considerations

We used a probit model to model the probability for future excellence in research (Greene 2012). A potential problem with our modelling approach is that our data set only allow us to observe the future performance for 310 of the 475 doctoral students. This means that the estimated probit coefficients might be affected by selection bias. To test and correct for a potential selection bias in our model we estimated a Heckman selection model (Greene 2012). The results from the Heckman selection model are presented in the section: Modelling the probability for employment at the university after completion of doctoral studies.

Another potential problem with the data set concerns the independence of the doctoral students. It is possible that the performance of an observed doctoral student is not statistically independent from the performance of other doctoral students in the same department,

Table 1 Descriptive statistics for the three non-binary predictors

Predictors	M	1st	2nd	3rd	Min	Max	Correlation coefficients	
							1.	2.
1. #Publications during doctoral studies	3.8	2	3	5	0	25		
2. Degree of collaboration	70.9	66.7	75	80.9	0	91.8	0.286	
3. Age at completion of doctoral studies	38.3	31	35	45	23	67	0.033	-0.035

M=Mean, 1st=1st quartile, 2nd=Median, 3rd=3rd quartile; 1. and 2.=Pearson's r^2 correlation coefficients

since they could be part of the same research group. Moreover, some departments for various reasons have a higher probability of producing excellent research. This problem affects inference, since correlated errors can lead to underestimated standard errors. We have, therefore, estimated robust clustered standard errors (Huber 1967; White 1980, 1982) which allow for intragroup correlation, relaxing the requirement that the doctoral students must be independent, i.e. the observations are independent across groups, but not necessarily within groups. When we calculated the robust clustered standard errors, the observations were grouped into ten research areas in accordance with the classification scheme *Standard for Swedish classification of research areas 2011* (HSV 2011). We refer to this cluster variable as *Research areas*. The dissertations are classified with the second level in this scheme, where level 1 and level 2 correspond in essence to the OECD classification scheme *Field of Research and Development*.

Results and discussion

Predicting future excellence among doctoral students

In this section the results of our analyses are presented in Table 2 which displays the estimated coefficients and Pseudo R^2 in seven probit regression models. The average marginal effects for the predictors are shown in parenthesis. The average marginal effect can be interpreted as the change in percentage points in the dependent variable for a one-unit increase in the predictor variable while controlling for the other predictors in the model. That is, we compute the marginal effect of a predictor variable for each observation as suggested by Williams (2012) while fixing the other predictors at their observed values and then taking an average across all marginal effects. We estimated seven models in order to examine our hypothesis, increase transparency of the model building and to see how the coefficients, significant levels, and the model fit change when new predictors was added. For Model 7 we provide a more in depth presentation of the predictors by means of average adjusted predictions. Average adjusted predictions can be interpreted as the probability of future excellence for individuals with a specific set of values on the predictor variables. Thus, with average adjusted predictions, we can look at the estimated probability for a range of values of a focal predictor while controlling for the other predictors in the model (i.e., we compute adjusted predictions for each observation while fixing the other predictors at their observed values in accordance with Williams (2012) and then taking an average across all predicted values).

For hypothesis 1 we expect that *doctoral students who publish more in scientific journals and also publish excellent work when preparing their thesis will have a higher probability of doing excellent research in the early career*. This hypothesis was examined with the predictors *#Publications during doctoral studies* and *Relative excellence during doctoral studies* (i.e. a sum of 10% fractions equal or larger than 2). In Model 6 (Table 2) we can see that these two predictors have positive significant effect ($P < 0.05$) on the probability of future excellence. However, as can be seen in Model 7 we also found that the doctoral students' research performance during their education, i.e. number of publications and attaining relative excellence during doctoral studies coded as top 10% articles, is complicated by a significant ($P < 0.05$) interaction effect. The interpretation of the coefficients

Table 2 Probit regression models estimating probabilities for future excellence containing probit coefficients, average marginal effects in parenthesis, and pseudo R^2

Predictors	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Male doctoral student		0.487** (0.087)	0.460** (0.081)	0.333* (0.053)	0.368* (0.057)	0.341 (0.051)	0.304 (0.043)
Age at completion of doctoral studies			-0.023* (-0.004)	-0.030** (-0.005)	-0.030** (-0.005)	-0.031** (-0.004)	-0.032** (-0.004)
#Publications during doctoral studies				0.133** (0.021)	0.124** (0.019)	0.098** (0.015)	0.045**
Degree of collaboration					0.025 (0.003)	0.026 (0.004)	0.029* (0.004)
Relative excellence during doctoral studies						0.700** (0.137)	-0.761* 0.237**
Relative excellences during doctoral studies *							
#Publications during doctoral studies							
Constant	-1.246**	-1.507**	-1.523**	-2.068**	-2.157**	-2.160**	-1.936**
Pseudo R^2	0.000	0.030	0.049	0.139	0.167	0.197	0.228

Standard errors are adjusted for 10 clusters in *Research areas*. Number of obs. = 310. Control variables “Age completing doctoral studies” and “Degree of collaboration” are mean centred. The control variable *Discipline* was not significant in any of our models and did not in essence change the coefficients or the Pseudo- R^2 in Model 7. It was therefore excluded from the analyses

* $P < 0.05$; ** $P < 0.01$

for these predictor variables must, therefore, be made with the effect of the interaction taken into account.

To interpret the interaction effect we plotted the average adjusted predictions for the number of publications during doctoral studies while distinguishing between doctoral students who had attained relative excellence and those that had not (Fig. 1; upper left). It is apparent that the effect of publication volume differs quite considerably between the groups (Fig. 1; upper left). For the doctoral students in the sample who have not attained relative excellence during their education, publication volume is a weak predictor of future excellence in research. For the doctoral students who have attained relative excellence during their education, the probability of future excellence in research increases quite rapidly for each additional publication when the total number of publications is larger than five. It should be noted that five publications is the upper quartile, so the group of doctoral students with a high probability of future relative excellence in research is rather extreme considering both their publication volume and their citedness. This result provide support for our first hypothesis and seem to suggest that a combination of quantity and quality is indicative of future excellence.

For hypothesis 2 we expect (a) that there are gender differences in attaining research excellence in the early career, and (b) that these gender differences can be explained by performance differences during the doctoral education. Hypothesis 2a and 2b was examined with the predictors *Male doctoral student* and the two performance predictors *#Publications during doctoral studies* and *Relative excellence during doctoral studies*. The coefficient for the predictor *Male doctoral student* in Model 2 (Table 2) is significant and

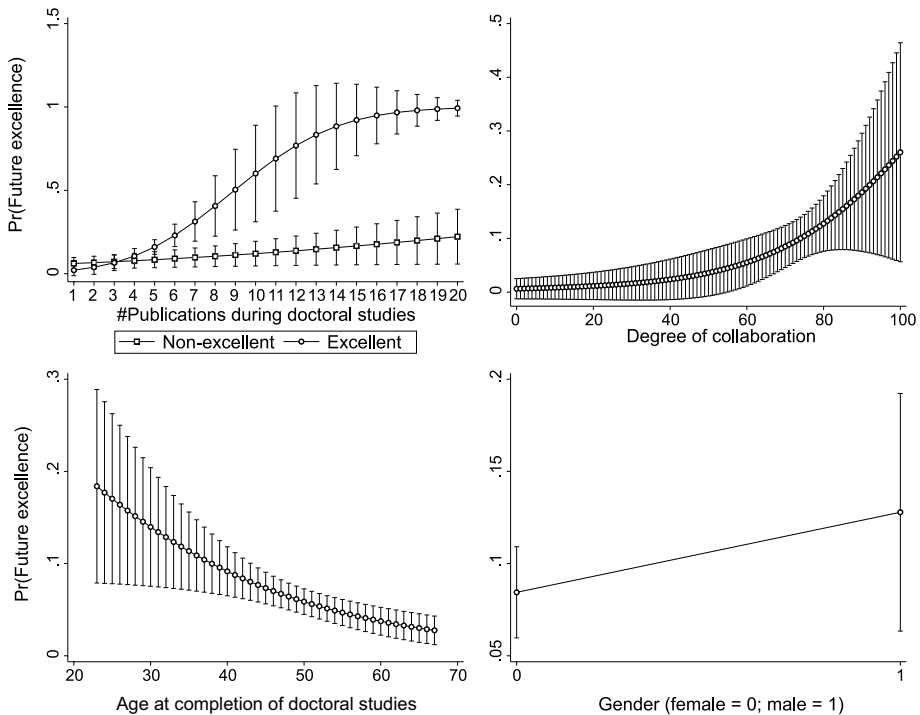


Fig. 1 Predicted margins for the predictors included in the multiple probit regression model (Model 7)

indicate that early career performance differs between genders. When the predictor *#Publications during doctoral studies* is included in Model 4 the effect of *Male doctoral student* is reduced, but still significant ($P < 0.05$). In Models 6 and 7 (Table 2), where the predictor *Relative excellence during doctoral studies* and the interaction effect is included, the effect of gender is further reduced and the effect become non-significant (i.e., $P > 0.05$). The average partial effect for the gender predictor is approximately reduced by half comparing Models 3 and 7 (Table 2).

We conclude that our results support hypothesis 2a since our model indicate that performance in terms of attaining excellence in the early career differ between females and males, where males have a higher probability of attaining excellence. Regarding hypothesis 2b we suggest some caution in interpreting the results since in Model 7 the predictor *Male doctoral student* still show a positive but non-significant effect, indicating that in our sample males have a higher probability of attaining relative excellence in the early career even when the performance predictors are controlled for. The P value change from < 0.01 in Model 2, to 0.04 in Model 4, and finally to 0.08 in Model 7. Given an alpha of 0.05, we cannot reject the null hypothesis when the P value is 0.08. However, in this case it seem reasonable to be somewhat cautious since the P values in model 4–7 are close to alpha, particularly since statistical significance testing is affected by sample size. Further, the results in the previous literature are inconclusive. Some studies indicate that gender have an effect on early career performance (see e.g., Hilmer and Hilmer 2009; Laurance et al. 2013; Pinheiro et al. 2014; Horta and Santos 2016) and some studies indicate that gender does not have an effect (see e.g., Buchmueller et al. 1999; Williamson and Cable 2003; Haslam and Laham 2009). Another reason for caution is that gender differences in performance and career development tend to be initially small and increase over time (Van den Besselaar and Sandström 2016).

Thus, concerning hypothesis 2b, we conclude that our models indicate that the observed performance difference in the early career is partly explained by performance differences during the doctoral education. However, since we can still observe a small non-significant positive effect in the gender predictor in Model 7 we want to acknowledge that there might still be gender differences/biases other than performance differences during doctoral education at work. An examination of this issue in a larger sample might generate results that are more conclusive.

We also notice that the control variables *Degree of collaboration* and *Age at completion of doctoral studies* has a significant positive and negative effect respectively, on the probability for future research excellence in Model 7. The fact that *Degree of collaboration* has a positive effect can be due to the importance of social integration of doctoral students, even though it is hard to specify exactly why this integration is of importance for a doctoral student's future research performance. It could be a combination of different factors that are embedded in this predictor, such as learning tacit knowledge, future integration into a research project or increased awareness of future research of potential interest.

In order to determine whether the interaction between the number of publications during doctoral studies and attaining relative excellence during doctoral studies is a robust pattern in our data, and not a consequence of the specific cut-off thresholds that were used to define relative excellence, we estimated five new models where we systematically changed the cut-off thresholds for the excellence predictor while keeping everything else identical to Model 7 (Table 2). These five new models consisted of two cut-off thresholds for the predictor *Relative excellence during doctoral studies* and three cut-off thresholds for the dependent variable indicating whether a doctoral student will attain future excellence. The cut-off thresholds for the predictor of attaining relative

excellence during doctoral studies was: (1) the 75th percentile in the distribution of top 10% fractions, which amounts to a top 10% fraction ≥ 1 . This group is referred to as the Top 25% doctoral students; and (2) the 90th percentile which amounts to a top 10% fraction ≥ 2 . We refer to this group as the Top 10% doctoral students.

For the dependent variable we used three different cut-off thresholds: (1) the 75th percentile (i.e., the Top 25% group); (2) the 90th percentile (i.e., the Top 10% group); and (3) the 95th percentile (i.e., the Top 5% group). Figure 2 consist of all combinations of the two cut-off thresholds for the predictor and the three cut-off threshold for the

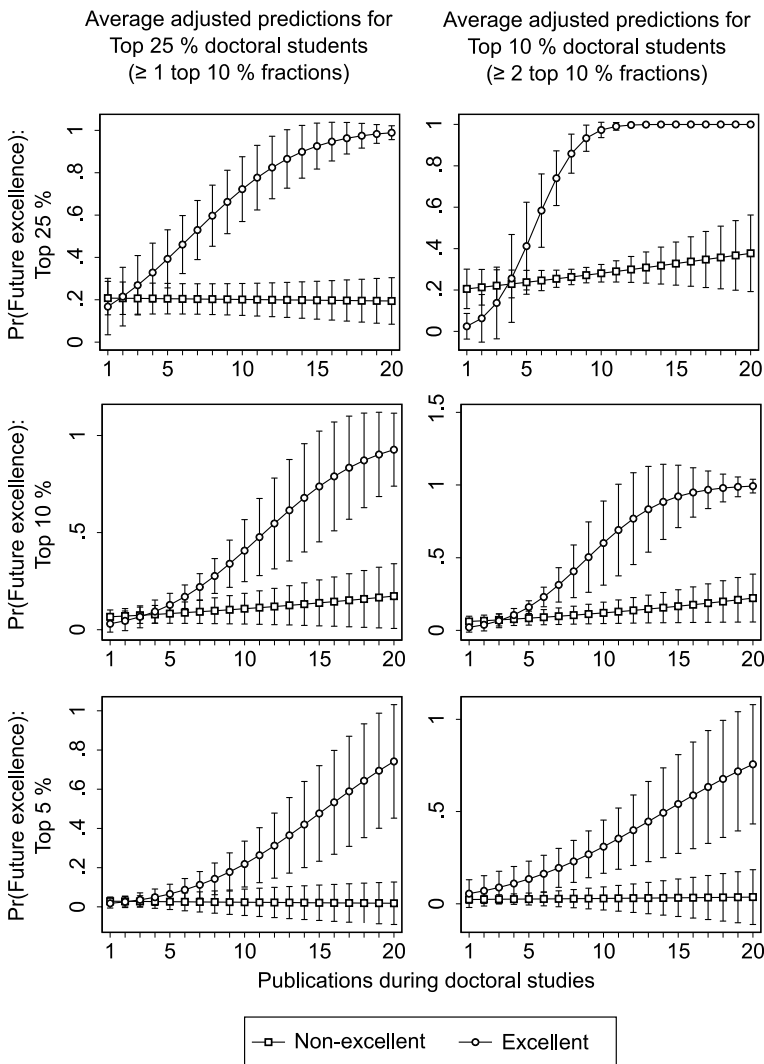


Fig. 2 Average adjusted predictions with 95% confidence intervals for the interaction between #Publications during doctoral studies and Relative excellence during doctoral studies at different cut-off thresholds

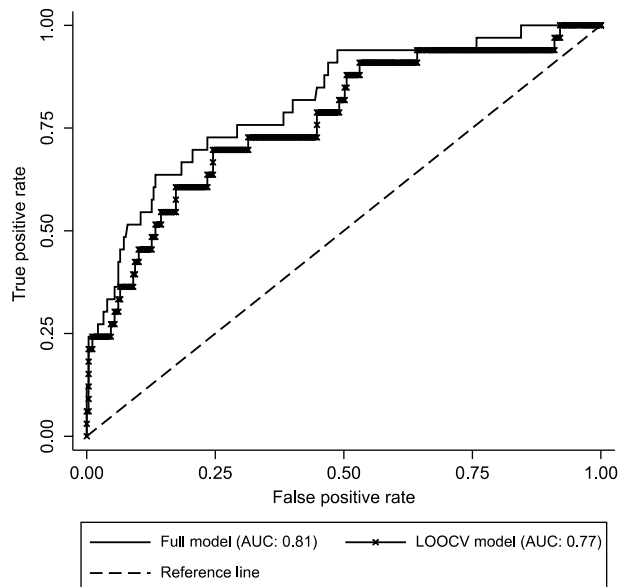
dependent variable. Overall Fig. 2 indicate that the interaction effect is a robust pattern in our data that is not dependent on the dichotomization.

Predictive value of the model

To estimate the information value of the probit model (Table 2; Model 7) in terms of predicting future excellence, we generated a ROC curve by plotting the true positive rate (i.e. the fraction of doctoral students attaining future excellence that was correctly predicted to do so) against the false positive rate (i.e. the fraction of doctoral students that did not attain future excellence but was predicted to attain future excellence) for each predicted value (Fawcett 2006). In Fig. 3, the y axis denotes the true positive rate (TPR) and the x axis denotes the false positive rate (FPR). If the ROC curve is above the diagonal line, our model performs better than expected according to a random model. If the ROC curve is below the diagonal line, our model would perform worse than a random model. If the ROC curve passes through the point (0, 1), it is a perfect classifier (Fawcett 2006). As a summary measurement of the predictive value of the model the Area Under the Curve (AUC) was calculated. The AUC is 1 when the curve passes through the (0, 1) point. If the ROC curve coincides with the diagonal line the AUC is 0.5.

In Fig. 3, two ROC curves are displayed with associated ROC areas. One ROC curve for the full model, i.e. Model 7 in Table 2, and a second ROC curve for the leave-one-out cross validation of the full model (i.e., the LOOCV model in Fig. 3). In this leave-one-out cross-validation, 310 probit models have been calculated. For each calculated model, one observation has been left out and a probit model is estimated for all remaining observations. The estimated model was then used to estimate a predicted value for the excluded observation, and this procedure was repeated for all 310 observations. We can then estimate the accuracy of our model by its ability in predicting the outcome for the excluded observation, and, as can be seen in Fig. 3, the predictive ability of the model is quite good.

Fig. 3 ROC analysis with leave one out cross validation of the model's ability to classify doctoral students according to their future performance



Modelling the probability for employment at the university after completion of doctoral studies

A potential problem with the models presented in Table 2 is that those doctoral students that are still employed at the university are selected on the variables that are used as predictors for future performances. There were 475 doctoral students completing their studies during the period of which 310 were employed at the university at least 1 year between the second and the fifth year after completing their doctoral education. However, since we only observe the future performance for a large subset of the doctoral students it is possible that the estimated coefficients are affected by selection bias. There are two types of potential selection effects in our case. First type of selection effect is that doctoral students that perform well during their educational period has a higher probability to be part of our sample, i.e. the university organization will select their future employees based on their performances, or that doctoral students that feels that they have not performed as well as their student colleagues will be less eager to choose a career in academia. Sartori (2003) notes that this non-random aspect of the sample is commonly miss-understood to be selection bias, but this process on its own does not bias estimates. Second type of selection effect, which is the most important, is that some of the “underperforming” doctoral students will continue to work at the university. If this is the case because these doctoral students has high values on an unmeasured variable that would explain their propensity to stay at the university, then we have a problematic selection bias. This selection bias will affect the constant term, i.e. the baseline probability, in one of two directions. As an example, if we have an overrepresentation in our sample of “underperforming” doctoral students that will be high performers in the future we will overestimate the size of the baseline probability, and therefore underestimate the effect of the predictors. On the other hand, if we have an overrepresentation of “underperforming” doctoral students that do not achieve excellence in the future we will underestimate the baseline probability, and therefore overestimate the effects of the predictors. With more technical terms this means that if there is a positive or negative correlation between the errors in a model that explains why the students continue to work at the university with the errors in a model that explains future performance we will have a problem with selection bias.

In order to investigate whether our model is affected by selection bias we have built a model estimating the probability for all doctoral students to stay at the university after completing their education (Table 3). In the model we have included the previously used predictors and added four predictors. The variable Non Swedish is a proxy estimating whether doctoral students are from Sweden or not. It is common that doctoral students at Swedish universities are from another country, and sometimes this is part of an exchange program financed by a third party. In such cases the doctoral student will probably not even have a work permit after completion of their studies. Even though, the indicator is a rough, and include some doctoral students that are Swedish citizens, it will catch this variability in explaining selection. The second added predictor is the year the doctoral student is completing the doctoral studies. This predictor is included since we expect the probability to get a position at the university will decrease during the period. The Swedish higher education expanded rapidly from the mid 90th until the beginning of the 00th. This was followed by an increase in the number of available positions. However, these added positions will be filled and the number of vacancies will stabilize due to normal turnover. The third variable we added is *Employment pre thesis completion* which denote the number of years a doctoral student has been employed or associated with the university during 7 years prior

Table 3 Probit regression model for probability to be employed by the university after doctoral studies

Predictors	Coefficient
#Publications during doctoral studies	0.061**
Relative excellence during doctoral studies	−0.080
Degree of collaboration	−0.008
Male doctoral student	−0.107
Age at completion of doctoral studies	0.019*
Employment pre thesis completion	0.072*
Thesis year	−0.114**
Faculty of natural science and technology	−0.570**
Non Swedish	−0.125
Constant	228.900**

Standard errors are adjusted for 10 clusters in *Research areas*. Number of obs. = 475. Pseudo $R^2 = 0.0802$. Variables “Age completing doctoral studies” and “Degree of collaboration” are mean centred

* $P < 0.05$; ** $P < 0.01$

to thesis completion. The time period of 7 years was chosen since the employment data is restricted to 7 years prior to thesis completion. Some doctoral students in our sample might have had an academic career at the university prior to their doctoral studies, e.g., as teachers or research assistants. In such cases the doctoral student might already be integrated in the workforce of a department or a research environment when they enter into the doctoral education. Such an integration might increase the probability of staying at the university for this group of doctoral students. By including the *Employment pre thesis completion* variable in our model we might catch this variability in explaining selection. The fourth added predictor is the binary control variable *Disciplinary*, where 1 denotes a doctoral student who’s thesis is classified as belonging to the natural sciences and 0 denotes a doctoral student who’s thesis is classified as belonging to the health and medical sciences. There might be disciplinary differences in the overall career advancement opportunities for the doctoral students that might affect their propensity to stay at the university. There are also differences in local career prospect between the faculty of science and technology and the faculty of medicine since MD students can have a dual organizational affiliation, i.e. both being employed by the university and the university hospital.

Age at completion of doctoral studies, *Thesis year*, *Discipline*, and *#Publications during doctoral studies* affects the probability for staying at the university (Table 3). It is notable that neither *Relative excellence during doctoral studies*, *Degree of collaboration*, nor *Non Swedish* have a statistically significant effect ($P < 0.05$). Age at completion of doctoral studies is a significant factor since it increases the probability for remaining at the university, i.e. older doctoral students has a higher probability of remaining at the university. The reason for this is not clear, but it could be the case that the older students have formed families and are not as mobile as their younger colleagues are. Other significant predictors for remaining at the university is year of thesis and what faculty the doctoral students completed their studies. The gender of the doctoral student is not a significant predictor; although, the estimated coefficient indicates that fewer male doctoral students continue their career at the university, and it is possible that the study does not have enough power to investigate this factor. It is interesting to note that *Degree of collaboration*, that was a significant predictor for future research performance, is not a significant predictor,

i.e. highly collaborative doctoral students does not have a higher probability to remain at the university.

Re-estimating our model (Table 2; Model 7) with a Heckman selection model does not change our results (Table 4). Of particular interest in the estimated Heckman model is the estimated coefficient λ which is used to calculate ρ , and ρ is the selection bias in terms of correlation between the errors. The correlation is very weak (0.019) and clearly not significant with a P value at 0.970. As a sub-sample of all those completing their doctoral studies at the university during the period the group still employed by the university can be viewed as a random sample. Obviously, this does not imply that our sample is a random sample in relation to a wider population of Swedish doctoral students (Table 4).

Conclusion

An increased focus in science on competition, performance, and excellence in the early career phase have led to an ongoing discussion about how factors and events during doctoral education might affect career development and future research performance (see e.g., Sinclair et al. 2014; Frick et al. 2016).

In this study we aim to contribute to that discussion by examining two hypothesis (1) how research performance, defined as publication volume and research excellence, during the doctoral education affect the probability of attaining research excellence in the early career (i.e., the years succeeding thesis completion); and (2) if there is performance differences between male and female in the early career and to which degree these gender differences can be explained by performance differences during the doctoral education.

Table 4 Adjusting for selection bias with a Heckman selection model does not change the estimated coefficients in Model 7 presented in Table 2

Predictors	Coefficient
Bias adjusted probit regression model	
#Publications during doctoral studies	0.045**
Relative excellence during doctoral studies	-0.761*
Degree of collaboration	0.029*
Male doctoral student	0.304
Age at completion of doctoral studies	-0.031*
Relative excellence during doctoral studies * #Publications during doctoral studies	0.237**
Constant	-1.945**
Selection model	
Age completing doctoral studies	0.022*
Employment pre thesis completion	0.077*
Thesis year	-0.119**
Faculty of natural science & technology	-0.515**
Constant	239.009**
/ λ	0.019
ρ	0.019

Control variables “Age completing doctoral studies” and “Degree of collaboration” are mean centred

* $P < 0.05$; ** $P < 0.01$

Our main results and conclusions for hypothesis 1 are (a) that performance, defined as publication volume and relative excellence, during the doctoral education has a positive effect on the probability of attaining excellence in the early career, and (b) that there is an interaction between publication volume and relative excellence during doctoral education indicating that the effect of publication volume depends on whether a doctoral student has attained excellence or not. For doctoral students who have not attained relative excellence during their education, publication volume is a weak predictor of attaining relative excellence in the early career. For the doctoral students who have attained relative excellence during their education, publication volume is a strong predictor of future excellence. However, it should be noted that the increase in probability of attaining excellence in the early career starts when the total number of publications is larger than five and this is the upper quartile. Thus, the group of the doctoral students with a high probability of future excellence in research is a rather extreme group considering both their publication volume and their citedness. These results are generally in agreement with previous research (see e.g., Hilmer and Hilmer 2009; Haslam and Laham 2009; Laurance et al. 2013; Pinheiro et al. 2014; Horta and Santos 2016). However, we want to emphasize how our study contribute to the previous literature. The main focus in previous research have been publication volume, both as predictor and dependent variables (see e.g., Williamson and Cable 2003; Laurance et al. 2013; Pinheiro et al. 2014) and/or some citation based indicator as dependent variable (Haslam and Laham 2009; Horta and Santos 2016). In this study we contribute to this literature (1) by adding relative excellence (i.e., a citation based indicator) both as a predictor and as the dependent variable, and (2) by an examination of two dimensions of performance during doctoral education, quantity and quality (i.e., publication volume and relative excellence), and how these two dimensions interact in predicting relative excellence in the early career.

Our main conclusions for hypothesis 2 are (a) that there are gender based performance differences in the early career and that males have a higher probability of attaining relative excellence than females, and (b) that this difference is partly explained by performance differences during the doctoral education. However, due to a non-significant effect in the gender predictor indicating that in our sample males have a higher probability of attaining future excellence than females even when we control for performance during the doctoral education, we want to acknowledge that there might still be gender differences/biases other than performance differences during doctoral education at work. An examination of the potential gender bias in a larger sample might generate results that are more conclusive.

A contribution of this study is that we connect gender differences during the early career (i.e., the years after thesis completion) with performance differences during the doctoral education. A future venue of research could be to examine determinants of performance differences between males and females during doctoral education and how these differences change over time.

We acknowledge that our design and measurements of performance does not distinguish between the performance of the doctoral student and the performance of e.g., potential collaborators, advisors, or the surrounding research group at large. A future venue of research could be to examine how the performance of the doctoral student is related to factors such as the performance of advisors and the surrounding research group.

Finally, our sample consisted of 475 doctoral students. Since we had observations only for those who stay at the university after the doctoral education ($n = 310$), we conducted a Heckman selection model to estimate and adjust for a potential selection bias in our analysis. From the results of the Heckman selection model we conclude that it doesn't seem to be a selection bias in our analysis.

Acknowledgements Open access funding provided by Umea University. This research was supported by the Swedish Research Council.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Abramo, G., Cicero, T., & D'Angelo, C. A. (2014). Are the authors of highly cited articles also the most productive ones? *Journal of Informetrics*, 8(1), 89–97. <https://doi.org/10.1016/j.joi.2013.10.011>.
- Ajiferuke, I., Burell, Q., & Tague, J. (1988). Collaborative coefficient: A single measure of the degree of collaboration in research. *Scientometrics*, 14(5–6), 421–433. <https://doi.org/10.1007/BF02017100>.
- Allison, P. D., & Stewart, J. A. (1974). Productivity differences among scientists: Evidence for accumulative advantage. *American Sociological Review*, 39(4), 596–606. <https://doi.org/10.2307/2094424>.
- Austin, A. E. (2002). Preparing the next generation of faculty: Graduate education as socialization to the academic career. *Journal of Higher Education*, 73(2), 94–122. <https://doi.org/10.1080/00221546.2002.11777132>.
- Austin, A. E. (2009). Cognitive apprenticeship theory and its implications for doctoral education. *International Journal of Academic Development*, 14(3), 173–183. <https://doi.org/10.1080/13601440903106494>.
- Bazeley, P. (2003). Defining 'Early Career' in research. *Higher Education*, 45(3), 257–279. <https://doi.org/10.1023/A:1022698529612>.
- Becher, T., & Trowler, P. (2001). *Academic tribes and territories*. Berkeley, CA: University of California Press.
- Bornmann, L. (2013). How to analyze percentile citation impact data meaningfully in bibliometrics: The statistical analysis of distributions, percentile rank classes, and top-cited papers. *Journal of the American Society for Information Science and Technology*, 64(3), 587–595.
- Bornmann, L. (2014). How are excellent (highly cited) papers defined in bibliometrics? A quantitative analysis of the literature. *Research Evaluation*, 23(2), 166–173.
- Bornmann, L., Mutza, R., & Daniel, H. D. (2007). Gender differences in grant peer review: A meta-analysis. *Journal of Informetrics*, 1(3), 226–238. <https://doi.org/10.1016/j.joi.2007.03.001>.
- Bornmann, L., & Williams, R. (2017). Can the journal impact factor be used as a criterion for the selection of junior researchers? A large-scale empirical study based on ResearcherID data. *Journal of Informetrics*, 11(3), 788–799. <https://doi.org/10.1016/j.joi.2017.06.001>.
- Brischoux, F., & Angelier, F. (2015). Academia's never-ending selection for productivity. *Scientometrics*, 103(1), 333–336. <https://doi.org/10.1007/s11192-015-1534-5>.
- Buchmueller, T. C., Dominitz, J., & Lee Hansen, W. (1999). Graduate training and the early career productivity of Ph.D. economists. *Economics of Education Review*, 18(1), 65–77. [https://doi.org/10.1016/S0272-7757\(98\)00019-3](https://doi.org/10.1016/S0272-7757(98)00019-3).
- Cole, S., & Cole, J. R. (1967). Scientific output and recognition: A study in the operation of the reward system in science. *American Sociological Review*, 32(3), 377–390.
- Cole, S., & Zuckerman, H. (1984). The productivity puzzle: Persistence and change in patterns of publication of men and women scientists. *Advances in Motivation and Achievement*, 2, 217–258.
- Costas, R., van Leeuwen, T. N., & Bordons, M. (2010). A bibliometric classificatory approach for the study and assessment of research performance at the individual level: The effects of age on productivity and impact. *Journal of the American Society for Information Science and Technology*, 61(8), 1564–1581. <https://doi.org/10.1002/asi.21348>.
- Danell, R. (2011). Can the Quality of Scientific Work Be Predicted Using Information on the Author's Track Record? *Journal of the American Society for Information Science and Technology*, 62(1), 50–60. <https://doi.org/10.1002/asi.21454>.
- Danell, R., & Hjerm, M. (2013). Career prospects for female university researchers have not Improved. *Scientometrics*, 94, 999–1006. <https://doi.org/10.1007/s11192-012-0840-4>.
- Delamont, S., Atkinson, P., & Parry, O. (2000). *The doctoral experience*. London: Falmer.

- DiPrete, T. A., & Eirich, G. M. (2006). Cumulative advantage as a mechanism for inequality: A review of theoretical and empirical developments. *Annual Review of Sociology*, 32(1), 271–297. <https://doi.org/10.1146/annurev.soc.32.061604.123127>.
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109(3), 573–598. <https://doi.org/10.1037/0033-295X.109.3.573>.
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>.
- Fenge, L.-A. (2012). Enhancing the doctoral journey: The role of group supervision in supporting collaborative learning and creativity. *Studies in Higher Education*, 37(4), 401–414. <https://doi.org/10.1080/03075079.2010.520697>.
- Fox, M. F. (1983). Publication productivity among scientists: A critical review. *Social Studies of Science*, 13(2), 285–305. <https://doi.org/10.1177/030631283013002005>.
- Fox, M. F. (2005). Gender, family characteristics, and publication productivity among scientists. *Social Studies of Science*, 35, 131–150. <https://doi.org/10.1177/0306312705046630>.
- Frick, B. L., Albertyn, R., Brodin, E., McKenna, S., Claesson, S., Fourie-Malherbe, M., Albertyn, R., Aitchison, C., & Bitzer, E. (2016). The role of doctoral education in early career academic development. In *Postgraduate supervision: future foci for the knowledge society* (pp. 203–219).
- Ginther, D., & Kahn, S. (2006). Does science promote women? Evidence from academia 1973–2001. In *Science and engineering careers in the United States: An analysis of markets and employment*. National Bureau of Economic Research Conference Reports. Chicago: University of Chicago Press
- Golde, C. M. (2005). The role of the department and discipline in doctoral student attrition: lessons learned from three departments. *Journal of Higher Education*, 76(6), 669–700. <https://doi.org/10.1353/jhe.2005.0039>.
- Greene, W. H. (2012). *Econometric analysis* (7th ed.). Boston: Pearson.
- Hallonsten, O., & Hugander, O. (2014). Supporting ‘future research leaders’ in Sweden: Institutional isomorphism and inadvertent funding agglomeration. *Research Evaluation*, 23(3), 249–260. <https://doi.org/10.1093/reseval/rvu009>.
- Hallonsten, O., & Silander, C. (2012). Commissioning the University of Excellence: Swedish research policy and new public research funding programmes. *Quality in Higher Education*, 18(3), 367–381. <https://doi.org/10.1080/13538322.2012.730715>.
- Haslam, N., & Laham, S. M. (2009). Ten years on: Does graduate student promise predict later scientific achievement? *Current Research in Social Psychology*, 14(10), 143–149.
- Havemann, F., & Larsen, B. (2015). Bibliometric indicators of young authors in astrophysics: Can later stars be predicted? *Scientometrics*, 102(2), 1413–1434. <https://doi.org/10.1007/s11192-014-1476-3>.
- Heilman, M. E., & Okimoto, T. G. (2007). Why are women penalized for success at male tasks? The implied communality deficit. *Journal of Applied Psychology*, 92(1), 81–92. <https://doi.org/10.1037/0021-9010.92.1.81>.
- Hellström, T. (2011). Homing in on excellence: Dimensions of appraisal in Center of Excellence program evaluations. *Evaluation*, 17(2), 117–131. <https://doi.org/10.1177/1356389011400891>.
- Hilmer, Michael J., & Hilmer, Christiana E. (2009). Fishes, ponds, and productivity: Student-advisor matching and early career publishing success for economics phds. *Economic Inquiry*, 47(2), 290–303. <https://doi.org/10.1111/j.1465-7295.2007.00108.x>.
- Horta, H., & Santos, J. (2016). The impact of publishing during Ph.D studies on career research publication, visibility, and collaborations. *Research in Higher Education*, 57(1), 28–50. <https://doi.org/10.1007/s11162-015-9380-0>.
- HSV. (2011). *Standard för svensk indelning av forskningsämnen 2011*. Stockholm: Högskoleverket (Swedish National Agency for Higher Education).
- Huber, P. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*, 1(1), 221.
- Jones, M. (2013). Issues in doctoral studies—Forty years of journal discussion: Where have we been and where are we going? *International Journal of Doctoral Studies*, 8, 083–104.
- Jones, B. F., & Weinberg, B. A. (2011). Age dynamics in scientific creativity. *Proceedings of the National Academy of Sciences*, 108(47), 18910–18914. <https://doi.org/10.1073/pnas.1102895108>.
- Knorr Cetina, K. (1999). *Epistemic cultures: How the sciences make knowledge*. Cambridge, MA: Harvard University Press.
- Kumar, N. (Ed.). (2012). *Gender and science: Studies across cultures*. Cambridge: Cambridge University Press.
- Larivière, V., Dorta-González, P., & Costas, R. (2016). How many is too many? On the relationship between research productivity and impact. *PLoS ONE*, 11(9), e0162709. <https://doi.org/10.1371/journal.pone.0162709>.

- Laurance, W., Useche, C., Laurance, S., & Bradshaw, C. (2013). Predicting publication success for biologists. *BioScience*, 63(10), 817–823. <https://doi.org/10.1525/bio.2013.63.10.9>.
- Lee, A., & Boud, D. (2009). Framing doctoral education as practice. In D. Boud & A. Lee (Eds.), *Changing practices of doctoral education* (pp. 10–25). London: Routledge.
- Lindahl, J. (2018). Predicting research excellence at the individual level: The importance of publication rate, top journal publications, and top 10% publications in the case of early career mathematicians. *Journal of Informetrics*, 12(2), 518–533. <https://doi.org/10.1016/j.joi.2018.04.002>.
- Long, J. (1992). Measures of sex differences in scientific productivity. *Social Forces*, 71(1), 159–178. <https://doi.org/10.2307/2579971>.
- Mason, S. (2018). Publications in the doctoral thesis: Challenges for doctoral candidates, supervisors, examiners and administrators. *Higher Education Research and Development*, 37(6), 1231–1244. <https://doi.org/10.1080/07294360.2018.1462307>.
- Merton, R. K. (1968). The Matthew effect in science. *Science*, 159(3810), 56–63. <https://doi.org/10.1126/science.159.3810.56>.
- Merton, R. K. (1988). The Matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property. *Isis*, 79(4), 606–623. <https://doi.org/10.1086/354848>.
- Newman, M. (2001). Scientific collaboration networks. I. Network construction and fundamental results. *Physical Review E, Statistical, Nonlinear, and Soft Matter Physics*, 64(1 Pt 2), 016131. <https://doi.org/10.1103/PhysRevE.64.016131>.
- Phelan, J. E., Moss-Racusin, C. A., & Rudman, L. A. (2008). Competent yet out in the cold: Shifting criteria for hiring reflect backlash toward agentic women. *Psychology of Women Quarterly*, 32(4), 406–413. <https://doi.org/10.1111/j.1471-6402.2008.00454.x>.
- Pinheiro, D., Melkers, J., & Youtie, J. (2014). Learning to play the game: Student publishing as an indicator of future scholarly success. *Technological Forecasting and Social Change*, 81, 56–66. <https://doi.org/10.1016/j.techfore.2012.09.008>.
- Platow, M. J. (2012). Ph.D. experience and subsequent outcomes: A look at self-perceptions of acquired graduate attributes and supervisor support. *Studies in Higher Education*, 37(1), 103–118. <https://doi.org/10.1080/03075079.2010.501104>.
- Price, D. D. S. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, 27(5), 292–306. <https://doi.org/10.1002/asi.4630270505>.
- Prpic, K. (2002). Gender and productivity differentials in science. *Scientometrics*, 55(1), 27–58. <https://doi.org/10.1023/A:1016046819457>.
- Pyhältö, K., Stubb, J., & Lonka, K. (2009). Developing scholarly communities as learning environments for doctoral students. *International Journal for Academic Development*, 14(3), 221–232. <https://doi.org/10.1080/13601440903106551>.
- Rørstad, K., & Aksnes, D. W. (2015). Publication rate expressed by age, gender and academic position—A large-scale analysis of Norwegian academic staff. *Journal of Informetrics*, 9(2), 317–333. <https://doi.org/10.1016/j.joi.2015.02.003>.
- Rossiter, M. W. (1993). The Matthew Matilda effect in science. *Social Studies of Science*, 23(2), 325–341. <https://doi.org/10.1177/030631293023002004>.
- Sandström, U., & van Den Besselaar, P. (2016). Quantity and/or quality? The importance of publishing many papers. *PLoS ONE*, 11(11), e0166149. <https://doi.org/10.1371/journal.pone.0166149>.
- Sartori, A. (2003). An estimator for some binary-outcome selection models without exclusion restrictions. *Political Analysis*, 11(2), 111–138. <https://doi.org/10.1093/pan/mpg001>.
- Schubert, A., & Braun, T. (1986). Relative indicators and relational charts for comparative assessment of publication output and citation impact. *Scientometrics*, 9(5), 281–291.
- Sinclair, J., Barnacle, R., & Cuthbert, D. (2014). How the doctorate contributes to the formation of active researchers: What the research tells us. *Studies in Higher Education*, 39(10), 1972–1986. <https://doi.org/10.1080/03075079.2013.806460>.
- Van den Besselaar, P., & Sandström, U. (2016). Gender differences in research performance and its impact on careers: A longitudinal case study. *Scientometrics*, 106(1), 143–162. <https://doi.org/10.1007/s11192-015-1775-3>.
- Vinkler, P. (1986). Evaluation of some methods for the relative assessment of scientific publications. *Scientometrics*, 10(3–4), 157–177. <https://doi.org/10.1007/BF02026039>.
- Wennerås, C., & Wold, A. (1996). Nepotism and sexism in peer review. *Nature*, 387(22), 341–343. <https://doi.org/10.1038/387341a0>.
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4), 817–838. <https://doi.org/10.2307/1912934>.

- White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*, 50(1), 1–25. <https://doi.org/10.2307/1912526>.
- Whitley, R. (2000). *The intellectual and social organization of the sciences* (2nd ed.). Oxford: Oxford University Press.
- Whitley, R. (2006). Changing governance of the public science: The consequences of establishing research evaluation systems for knowledge production in different countries and scientific fields. In R. Whitley & J. Gläser (Eds.), *The changing governance of the sciences*. Dordrecht: Springer.
- Williams, R. (2012). Using the margins command to estimate and interpret adjusted predictions and marginal effects. *Stata Journal*, 12(2), 308–331. <https://doi.org/10.1177/1536867X1201200209>.
- Williamson, I. O., & Cable, D. M. (2003). Predicting early career research productivity: The case of management faculty. *Journal of Organizational Behavior*, 24(1), 25–44. <https://doi.org/10.1002/job.178>.
- Xie, Y., & Shauman, K. A. (1998). Sex differences in research productivity: New evidence about an old puzzle. *American Sociological Review*, 63(6), 847–870. <https://doi.org/10.2307/2657505>.
- Xie, Y., & Shauman, K. A. (2003). *Women in science: Career processes and outcomes*. Cambridge, MA: Harvard University Press.