



Is there gender bias in awarding cum laude for the PhD thesis?

Peter van den Besselaar^{1,2,3} · Charlie Mom³

Received: 21 November 2023 / Accepted: 28 February 2024
© The Author(s) 2024

Abstract

In the Dutch academic system, PhD theses can be awarded with cum laude and at most 5% of all PhD graduates receive this selective distinction for their thesis. In this paper, we investigate whether there is gender bias in awarding cum laude, using data from one of the major Dutch research universities. We measure the quality of the PhD theses using bibliometric data. A main result is that the set of PhD theses receiving cum laude on average do not have a higher quality than the best theses not getting cum laude. A second main result is that, after controlling for the quality of the PhD theses, women still have a substantially lower probability to receive cum laude. These results strongly suggest that the distribution of awards suffers from gender bias, and the study creates a strong doubt about the adequacy of the procedures leading to cum laude for the PhD thesis.

Keywords Awards · Recognition · Cum laude · Gender · Bias · PhD thesis

Introduction

Within the Dutch higher education system, awarding a PhD thesis with the distinction cum laude (with honors; abbreviated as CL) does not happen often. It may differ between universities and faculties, but the overall 4.3% in our case seems a good estimate. It has been shown that men are awarded CL for their PhD thesis more frequently than women are, and this seems to be the case in all Dutch universities (De Bruin, 2018, 2019). This is not unimportant, as the CL distinction is of considerable help in the early academic career (Brouns et al., 2004), which may even have increased in the more recent period (Reschke et al., 2018).

✉ Peter van den Besselaar
p.a.a.vanden.besselaar@vu.nl
Charlie Mom
charlie@teresamom.com

¹ Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

² DZHW - German Centre for Higher Education Research and Science Studies GmbH, Berlin, Germany

³ TMC Research Amsterdam, Middenweg 203, 1098 AN Amsterdam, The Netherlands

In this paper we answer the question whether the gender difference in CL is an effect of gender bias, or alternatively whether the differences correctly reflect the quality differences of the dissertations (and the PhD candidates). Do the committee members adequately select the set of best dissertations, and does this set happen to include more male than female PhD candidates? Or is the decision not or only partly based on merit and does gender play a role: gender bias? As gender differences in CL do not fit with the generally accepted policies within Dutch higher education to promote gender equality, an explanation for the existing gender disparities is needed. To answer these questions, we collected data on PhD candidates of a research university in the Netherlands.

There are hardly studies available on gender bias in awarding CL. A recent study found that the more female members in the PhD committee, the more equal the probability to receive CL for men and women is. This homophily effect was not found for the supervisors: The gender of the PhD supervisors had no significant effect (Bol, 2023). Unfortunately, the study does not use any (proxy) measure for the quality of the thesis or of the PhD student. Although the study claims to detect gender bias in the CL distinction, it remains unclear whether this is gender *bias* or whether this is a gender *difference* that could be the effect of quality differences.

However, there are studies on other types of awards and prizes, all suggesting the existence of gender bias (Lincoln et al., 2012; Melnikoff & Valian, 2019; Silver et al., 2018). Furthermore, men also receive other signs of prestige more than women do, such as invitations for keynote speaker at scientific conferences (Casadevall & Handelsman, 2014; Dumitra et al., 2019; Johnson et al., 2016; Klein et al., 2017; Sardelis & Drew, 2016), for other invited talks (Nittrouer et al., 2017), or invitations for e.g., scientific advisory boards (Ding et al., 2012). Finally, the peer review literature has shown the many uncertainties and biases in peer and panel review processes (Chubin & Hackett, 1990; Cole et al., 1977, 1981; Cole, 1992; Lamont, 2010; Van Arensbergen et al., 2014), making it unlikely that the committee-based evaluation of PhD theses would be completely unbiased.

Gender differences versus gender bias

If there are strong gender differences in getting CL for the PhD thesis, the question is whether those differences are caused by gender bias. Bias implies a deviation of what would be merit-based decisions (Merton, 1973), and consequently, the selection of a frame of reference for measuring merit is crucial. One could prefer *distributional equality*, which would imply that the share of female PhD students that receive a CL should be equal to the share of male PhD students that get CL. In this study we use another—widely accepted—view that science is (or at least should be) a *merit-based system* and that implies that CL awards should go to the best PhD students with the best doctoral dissertations. Bias can be defined as deviation from this norm.

This of course does not immediately solve the problem, as different opinions exist about how to decide what are the best dissertations and who are the best PhD students. For example, should the most talented PhD students or the best performing ones receive the CL award? And how does one objectively measure who are the best PhD students, what are the best theses, or who are the most talented? We suggest that CL is meant as a recognition of the quality of the thesis itself, and we assume that the higher the quality, the more highly cited papers result from the thesis. This does not work in an absolute way, and one would not expect perfect correlation between the quality of the thesis and the number of highly

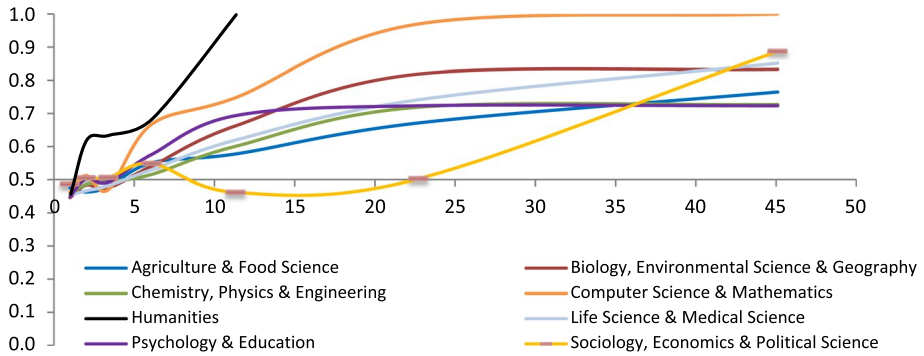


Fig. 1 Share of men (Y-axis) by productivity level (X-axis: full counts). Note that the number of researchers declines fast with the increase of productivity. *Source* Van den Besselaar and Sandström (2019)

cited papers coming out of the thesis work. However, one would expect that CL theses lead to *on average* higher bibliometric scores than very good, but not CL awarded theses. Below we will go into the details.

If we measure the quality of the PhD research by measuring the impact of the related output, the question can be answered whether the observed gender differences in receiving CL are a form of gender bias. In an earlier study, we showed that there are persisting differences between men and women researchers in terms of output and impact and, relevant for this study, these gender differences occur specifically in the top of the performance distribution (Fig. 1): among the group of high productive and high impact researchers, male researchers are strongly overrepresented (Van den Besselaar & Sandström, 2017).

Figure 1 shows only publication-counts, as high numbers of publications lead to high numbers of highly cited papers (Sandström & Van den Besselaar, 2016). This finding is relevant for the current study, as the CL recipients are expected to belong to that top segment. The fact that women less often receive CL may therefore not be bias, but the effect of male researchers being over-represented in the top-performers segment. Consequently it is crucial to include performance variables in this type of studies, something that is often omitted.

Data

The basic data were provided by the case university: Names of the PhD students and for all of them the year of the PhD degree, the faculty, CL or not, and the names of the supervisors. We received this for the 2000–2018 period and collected bibliometric data (in 2019) about the PhD students and supervisors for the period 2000–2014, in order to have useful citation data. After finishing the study, we also got data for 1997–1999 and 2019–2022. Therefore, this paper discusses data for different periods: 1997–2022, 2000–2018 and 2000–2014. For the core of the analysis the period 2000–2014 is used. The other two (longer) periods are included as these may show possible changes in awarding theses with CL over a longer period.

The data to measure the quality of the PhD thesis have been retrieved from Scopus. For a sample of 2592 researchers we collected the publication and citation data and indicators

Table 1 Available bibliometric data for the factor analysis

	PhD students	Supervisors
Collected bibliometric data (Scopus)	2592	812
No bibliometric data	714 ^a	3176

^aAll in medicine

using Scopus and SciVal, which then were controlled and cleaned, and used to retrieve and calculate a series indicators covering the selected periods.

The same was done for 812 (co-)supervisors. Table 1 summarizes the collected and cleaned bibliometric data for 3306 PhDs and their supervisors, which was a time-consuming but necessary task.

We use the following variables in the study, which are described in more detail in the next section. Here we explain why we include those.

- Bibliometric variables for the *quality of the dissertation*. These are needed to distinguish gender differences from gender bias.
- As CL-practices may differ between faculties, we use the *faculty* where the PhD is awarded as control variable.
- Dissertations generally consist of several published or submitted journal articles, and these articles are often coauthored. CL would not only reflect the quality of the work, but also that the PhD student had an outstanding and independent contribution to the work. We therefore include also a variable measuring the *level of independence*.
- The dissertation is produced in the context of the team, consisting of the PhD student and the (co)supervisors. The *team size* (between 2 and 6) and the *team composition* (male, female or mixed) can be expected to influence the level of gender stereotyping, and therefore be related to gender bias.
- Year of the PhD, in order to account for *changes over time*.

Methods

Measuring the quality of a PhD thesis

If the meritocratic system would work perfect, the quality of the PhD dissertations that received CL is higher than the quality of those that did not. As there is no independent scoring of the quality of a PhD dissertation, we need to develop a method to do this. As a proxy for the quality of dissertations we will use the papers published by the PhD student in the period of preparing the dissertation, as well as papers published in the 3 years after the year in which the student received the PhD. We assume that these papers are related to the dissertation work. Then the question of the quality of the thesis becomes another question: what is the quality of the scientific output of the PhD students in the defined period? The bibliometric toolbox SciVal of the Scopus database provides a series of indicators *at the paper level* (Elsevier, no date). We aggregated those SciVal indicators to the author level, using full and fractional counts, and absolute and relative (to the total oeuvre of the

Table 2 Bibliometrics

- (1) P: Number of full papers
- (2) Pfrac: Number of papers fractionally counted by number of authors
- (3) C: Total number of citations
- (4) Cfrac: Fractional share of total citations
- (5) C/P: Citations per publication
- (6) PP10%: Number of papers in the set of 10% highest cited papers
- (7) P10%FN: Number of papers in the set of 10% highest cited papers, field normalized
- (8) PP10%: Share of top 10% cited papers in the oeuvre
- (9) PP10%FN: Share of top 10% cited papers in the oeuvre, field normalized
- (10) P10%frac: Number of papers in the set of 10% highest cited papers, fractionally counted
- (11) PP10%FNfrac: Share of top 10% cited papers cited papers, field normalized, fractionally counted
- (12) Average FWCI: Average field weighted citation impact
- (13) Sum FWCI: Sum field weighted citation impact
- (14) SNIP: Source (journal) normalized impact per paper
- (15) SJR: SCimago Journal Rank ('impact factor')

Table 3 CL by gender

	Male	Female	Total	F/M
CL	231	129	360	
No CL	3856	3653	7509	
Total	4087	3782	7869	
Share CL	5.7%	3.4%	4.6%	0.6

Source the case university, 1997–2022, one PhD student has missing gender data

author) values. The non-field normalized and the field normalized indicators were both included. The resulting list of indicators is in Table 2:

Publication behavior and citation behavior differ between fields, and therefore the indicators are often field normalized. And as most papers are co-authored, and co-author behavior also differs between fields, many indicators correct for the number of co-authors ('fractional counting'). The indicators measure different aspects of quality and do so in different ways. For example, *P*, *Pfrac*, *P10%*, and *P10%frac* each measure some aspect of *productivity*. *P* simply measures full output, *Pfrac* corrects for the number of authors, *P10%* only takes the top cited papers into account, and *P10%frac* does this while correcting for the number of co-authors. The last two indicators also measure *impact*, as is also done by other citation-based indicators. Finally, some indicators do not measure the impact of the work of the authors, but the impact of the journals the papers are published in. These indicators are probably measuring reputation, more than impact or quality.

We only include in our analysis those research fields for which journal articles are the main research output. This excludes from this study most of the social sciences, the humanities, law and theology. Fortunately, those are also the fields with a lower gender difference in awarding CL (Table 3). Psychology and economics are included, as in those fields journal articles are dominant. Consequently, we include all dissertations in the *sciences*, *mathematics* and *computer science*, in the *earth and life sciences*, in the *medical sciences*, in *dentistry*, in *economics* and in *psychology* and *movement sciences*. Due to the

large number of medical dissertations, we decided to include all the medical CL dissertations and 50% of the dissertations that did not receive CL.¹

A next issue is the period covered by the measurement. The obvious period to take into account are the years of the PhD trajectory and a few years after the moment of awarding the PhD. But the committee may also have assessed the *potential* of the researchers, and may have based the CL decision also on *expected (future) excellence*. The ‘future’ excellence can be measured with hindsight, since the ‘future’ has already happened. In this paper we do both: using performance in the PhD period and using performance in the career.

As we measure the quality of the thesis using citation-based indicators, we restrict the analysis to the PhDs awarded the period 2000–2014.² Overall, this leads to a set of 120 CL dissertations out of 3306 dissertations, which is 3.6%.³

As publication, collaboration and citation behavior differs between the disciplines, we first *standardize* the performance indicators *by faculty* and in such a way that the mean of the performance indicators is 0 and the standard deviation is 1.

As explained in the previous section, the bibliometric indicators all measure in some way quality, and therefore we consider the various indicators *as items*, together measuring various underlying dimensions of the construct *quality*. A principal axis factoring (oblique rotation; component scores saved with regression) resulted in three dimensions, together explaining almost 80% of the variance.⁴ The resulting variables represent three distinct quality dimensions⁵:

- Dimension 1: *Relative impact* of the oeuvre of a researcher: *share* of top cited papers in the total oeuvre of the researcher, the average FWCI, and citations per publication.
- Dimension 2: *Total impact* of the researcher, in fractional count, consisting of the following indicators⁶: Fractional publications, fractional citations, fractional number of top 10% highly cited papers, sum FWCI.
- Dimension 3: *Journal impact*, average of the impact of the journals in which the papers of an author were published.

The *total impact* dimension includes the indicators measuring total output, total impact, and the sum of highly cited papers, suggesting that higher quantity produces also a higher quantity of top papers (Sandström & Van den Besselaar, 2016). We tested whether the indicators also form reliable scales, and that is the case. The Cronbach α for the three indicators are 0.911 (5 items), 0.861 (3 items), and 0.851 (2 items) respectively, and cannot be improved by leaving out one of the items.⁷ Concluding, we have a quality concept with

¹ We use a stratified sample that represents the distribution over years and over gender in the total set. Within each category, the 50% was selected randomly. The very labor-intensive part of the study was the collection of performance data from (in this case) Scopus.

² The bibliometric data were collected in 2019 using a at least 5 year citation window, which means that we can estimate citation data for all the PhDs from 2000 until 2014.

³ This is for the 2000–2014 period. Over the longer 2000–2018 period, and including all faculties, it is 4.3%.

⁴ The PAF was done on both the whole career data and the PhD period data.

⁵ The items belonging to each component are explained above.

⁶ We do not use the full-count based indicators, but when including those, they load together with the fractional-count based publication and citation indicators (component 2).

⁷ For the career data the Cronbach α 's are .911 .929 and .849 for component 1, 2 and 3 respectively.

three dimensions (total, relative, and journal impact) and the scales used to measure those dimensions are reliable.

The data are analyzed using logistic regression (the dependent variable is binary: CL or no CL). The logistic regression is first done for men and women separately, in order to test whether the independent performance variables have a similar effect for men and women. If not, we include interaction terms between gender and the performance variables in the analysis.

Then we run a logistic regression also including gender as a variable. The odds ratio for gender gives a first insight in the effect of gender (bias) on getting CL. However, the odds ratio only gives the value of the dependent variable for the independent variables at value = 0. As logistic regression is non-linear, the gender differences may be dependent on the level of the independent (here: performance) variables (Mood, 2009). In order to correct for this, we calculate the predicted probabilities (using the predicted margins in STATA 17) to display the probability to receive CL by gender for a variety of values of the performance variables.

After having done this, we did a few controls. We restrict the sample of PhD receivers to those with at least one supervisor that has at least one PhD student with CL, and run the same analyses—this because only a small subgroup of supervisors has supervised a PhD student that was awarded CL. By restricting the analysis to this set of supervisors, we may get a better picture of what variables affect the probability to get CL. Secondly, we restrict the sample to the CL awardees plus the best performing non-CL PhD receivers, to assess whether CL was indeed awarded to the best PhD students.

Independence

Good researchers are expected to be independent, and independence is often mentioned as a quality criterion for evaluating e.g. grant proposals or job applicants. However, independence does not mean ‘doing research alone’, as most research has become a team activity. Independence can be defined as having developed the ability to develop and implement an own research agenda. In earlier work, we defined independence as independence from the context where a researcher started the career: the PhD project (Van den Besselaar & Sandström, 2019; Möller et al., 2022). Several independence indicators were calculated, focusing on the relation to the PhD supervisor(s). We use as independence indicator the share of the output of a researcher that was not coauthored with any of the supervisors, as this constitutes the independently done research leading to independently written output,⁸ and it is measured twice: for the period up to the year of PhD and at the end of the available career period, which is the PhD year plus 15.⁹

Team characteristics

Doing research for a PhD thesis is done in a team with one or more supervisors. Some team characteristics may influence the CL decision such as the size of the team and the gender of the supervisor(s). The number of supervisors could be calculated from the basic data. The gender of the PhD students was available in the basic data, but that was not the case for the

⁸ Another indicator measures topical independence from the supervisors: the share of research topics in which the supervisors are not active (Van den Besselaar & Sandström, 2019; Möller et al., 2022).

⁹ Independence could not be calculated for all PhD students. Among those are 11 CL awardees.

supervisors. Relevant web pages were used to find those first names, which generally are gender specific. In cases where the first name was inconclusive, other information was used such as text about a supervisor using personal pronouns.

Field differences

The overall share of PhD thesis with the CL recognition is low, and if one looks at individual faculties it is even lower in several of those. Due to these low numbers the gender differences fluctuate heavily over time, and a difference of one or two CL awards may influence the picture considerably. When going to lower levels of aggregation this becomes an even larger problem. We therefore do the analysis on the level of the university as a whole, using faculty as control variable. In the [Annex](#) we show for each of the faculties a graphical representation of the fluctuations in both the number of CL and the gender differences. Despite the variety over time and the differences between faculties, the overall picture remains similar to one that emerges at university level.

The case

The case is a large research-intensive university in the Netherlands, which is in the top 100 in the 2023 Leiden Ranking—based on the share of top 10% cited papers.

PhD students work for several years on their PhD research project, with between one up to six supervisors and co-supervisors. The median duration of the project is 5 years.¹⁰ Almost all supervisors are full professors, whereas the co-supervisors generally are associate and sometimes assistant professors.¹¹ After the supervisors approve a dissertation, it is submitted to an examination committee that is created by the dean for the specific dissertation. The committees consist of about five members who should be specialists on the dissertation topic, and it decides whether the PhD degree will be awarded. The supervisors are not member of the committee, and the majority of members cannot come from the faculty of the supervisors. When CL is proposed (generally by the supervisors), the committee is extended with two members. If more than one member of the extended committee votes against CL, the distinction is not awarded.¹²

In the period 1997 till 2022 a total of 7870 PhD students finished their PhD thesis. Over this period there was a substantial annual growth: In the early years until 2002 there were less than 200 PhDs annually, this number increased slowly during the first decade, and peaks around 2015 and then levels out above 400 (Fig. 2). The figure shows a steep drop in the number of PhD awardees in 2022, which is probably caused by the Corona pandemic that may have influenced the time to finish the PhD project. One would therefore expect again an increase in 2023 or 2024. The decline did not work out stronger for women than for men.

¹⁰ <https://www.rathenau.nl/nl/wetenschap-cijfers/wetenschappers/van-promovendus-tot-promotie>.

¹¹ Only full professors had the right to approve the thesis, and although there are recent changes with respect to the *Ius Promovendi*. Those do not affect the analysis.

¹² The decision to award CL is reserved for the PhD committee, but the supervisor generally is the one who proposes it. For readability, we use a formally incorrect phrase: Where we write “supervisors awarding CL” one should read “supervisors who had at least one PhD student that was awarded CL by the PhD committee”. And where we write “Supervisors who never gave CL”, one should read: “Supervisors who have no PhD student that was awarded CL”.

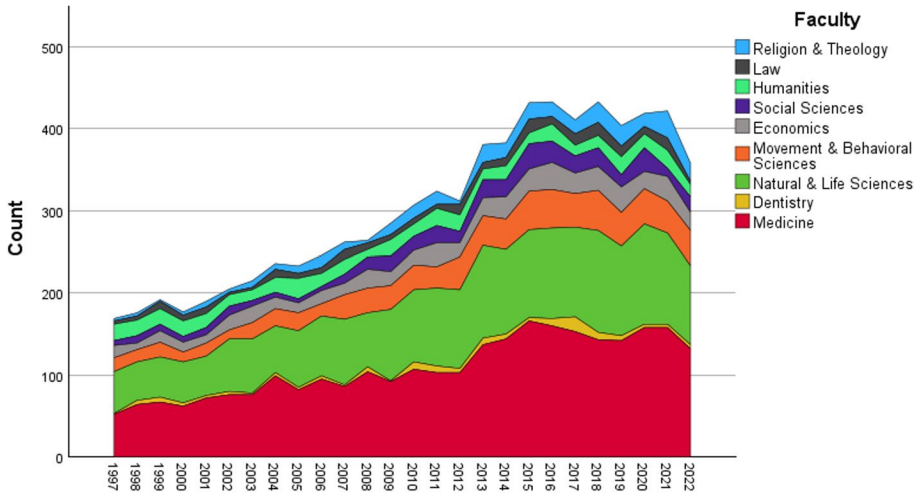


Fig. 2 Development of the number of PhD students by faculty

In total 48.1% of all the PhD students are women. Over the whole period 360 PhD students received CL, of which 35% women, implying that men have almost twice as high a probability to receive CL than women (Table 3).

The share of female PhD students increased from below 35% in 1997, rising to about 53% in 2014, and then appears stable (Table 4). The share of women receiving CL for the PhD fluctuates strongly but remains considerably lower than the share of men (Table 4).

If we distinguish between disciplines, quite some differences become visible. First of all, the numbers of PhDs per discipline differ strongly. Some 36% of all PhD degrees are in the medical school plus 2% in dentistry, and 28% are in the two science faculties. The remaining 34% are divided among the other disciplines, with behavioral and movement

Table 4 5-Year moving average share of women and percentage of men and women with CL

Year	Share women in PhD students (%)	Men with CL (%)	Women with CL (%)	Year	Share women in PhD students (%)	Men with CL (%)	Women with CL (%)
1997	35.1	3.7	3.2	2010	45.5	5.0	3.4
1998	36.2	4.2	3.1	2011	47.4	5.4	3.0
1999	37.3	5.8	2.4	2012	49.6	5.0	2.8
2000	39.0	5.2	2.2	2013	51.9	4.5	2.7
2001	40.3	6.2	2.3	2014	53.3	3.9	3.2
2002	40.8	6.3	2.2	2015	53.9	4.6	3.1
2003	41.2	6.9	2.2	2016	53.7	4.1	2.8
2004	41.3	6.3	3.6	2017	53.2	5.1	3.3
2005	41.4	7.4	3.9	2018	53.5	6.2	3.6
2006	41.8	6.7	4.4	2019	52.9	7.1	4.0
2007	42.7	6.8	4.0	2020	53.5	6.4	3.9
2008	43.3	5.8	4.4	2021	53.9	7.2	4.6
2009	44.1	5.7	3.5	2022	54.6	6.8	4.7

Table 5 CL by Faculty (1997–2022)

Faculty	Total	M	M-CL	W	W-CL	CL%	M-CL%	W-CL%	W%/M%	M% - W%
DEN	140	70	2	70	2	2.9	2.9	2.9	1.00	0.0
SBE	523	369	8	154	3	2.1	2.2	1.9	0.90	0.2
REL	341	275	25	66	4	8.5	9.1	6.1	0.67	3.0
N&E&L	2190	1382	61	808	23	3.8	4.4	2.8	0.64	1.6
HUM	451	256	25	195	12	8.2	9.8	6.2	0.63	3.6
LAW	228	108	10	120	7	7.5	9.3	5.8	0.63	3.4
M&B	775	318	24	457	21	5.8	7.5	4.6	0.61	3.0
SOC	387	166	15	221	12	7.0	9.0	5.4	0.60	3.6
MED	2833	1142	61	1691	45	3.7	5.3	2.7	0.50	2.7
Total	7869	4087	231	3782	129	4.6	5.7	3.4	0.60	2.2

DEN dentistry, *SBE* economics, *REL* theology and religion studies, *N&E&L* natural sciences (incl. computer science and mathematics, earth sciences and life sciences), *HUM* humanities, *LAW* law, *M&B* movement and behavioral sciences (incl. psychology), *SOC* social sciences, *MED* medical sciences), *M* male, *M-CL* men with CL, *W* women, *W-CL* women with CL, *CL%* % of PhDs with CL, *M-CL%* share CL, men only, *W-CL%* share CL, women only, *W%/M%* share W with CL divided by share M with CL, *M% - W%* share CL for men minus share CL for women

sciences with a share of about 10% as largest of the smaller fields (Table 5, first column for the absolute numbers).

Findings

Who receives the CL-award? Some basic findings

As reported above, about 4.3% of the PhD recipients were awarded CL, and men twice as often as women. Did this change over time? In Fig. 3, we show the development over the period 1997–2022. The figure shows for each year the total number of CL cases (the bars), and the difference between the share of men that received CL and the share of women receiving CL (the line). In 2001, only men (14 in total) received CL, and no women, leading to a very high gender difference. The different colors of the bars represent the different moments the data were collected, and the different types of data we have available.

Figure 3 shows that gender differences are fluctuating strongly, and in most years, men have a higher probability to receive CL than women. Linear and non-linear trend lines suggest some decline in the gender difference, but that is fully dependent on the outlier in 2001. In the Annex we show the same plots for each of the disciplines separately. The patterns differ between the faculties, but the important characteristics remain the same. Strong fluctuations are found and the observations in favor of men are clearly higher than the observations in favor of women. What is also visible in the Annex, that in the smaller faculties the numbers of CL receivers is very low, and that there are several years when no CL was awarded at all.

Table 6 gives a break down by field, showing that the *stronger codified fields* (the STEM fields and the mathematics-oriented fields like economics) have a low percentage of PhD theses with CL (between 1.6% and 4.3%). In these fields, the probability that men receive

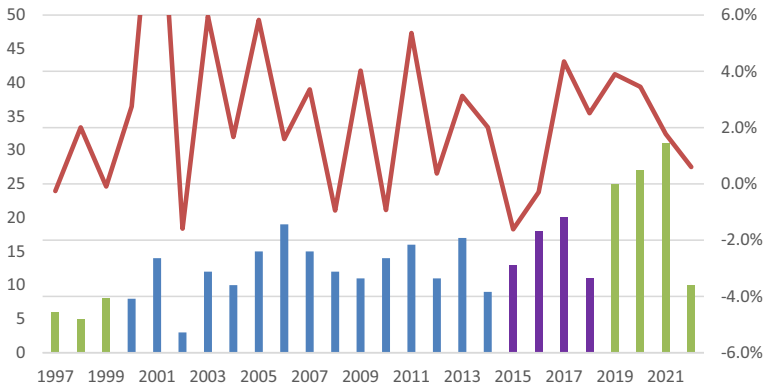


Fig. 3 CL by gender over time. Red line: % men CL - % women with CL (right Y-axis). Bars: number of CL awards by year (left Y-axis). Blue bar: data collected in 2019, including bibliometric data. Purple bar: data collected in 2019, no bibliometric data. Green bar: data collected in 2023, no bibliometric data

Table 6 CL by faculty and gender

Faculty	CL%	W/M
Business and Economics	1.6	0.00
Earth and Life Sciences	2.0	0.62
Dentistry	2.8	0.53
Medicine	3.3	0.51
Natural Sciences	4.3	0.40
Behavioral (Psychology) & Movement Sciences	6.4	0.47
Humanities	7.9	0.47
Social Sciences	7.6	0.73
Law	8.3	0.92
Theology and Religious Studies	8.9	0.77

%CL percentage of PhDs with CL, *W/M* = $W-CL\%/M-CL\%$

CL is twice as high (or more) as for women, except for Earth and Life sciences.¹³ The latter may be due to some large social sciences-oriented departments in that faculty.

Within social sciences, law, and religion studies, the share of PhD students receiving CL is substantially higher (between 7.6% and 8.9%) and the differences between men and women are relatively small. The Behavioral and Movement sciences and the Humanities faculties both have an ambiguous position. Behavioral and Movement sciences belong to the STEM group when we look at the gender differences. However, these fields have a much higher percentage of PhD students that receive CL, which is more similar to the other social sciences. The same holds for the humanities: with respect to the share of PhDs with CL, humanities are similar to the social sciences, but in terms of gender differences it belongs to the group with the STEM fields.

¹³ In the period considered here (2000–2014), Earth and Life sciences was a separate faculty, which merged in 2018 with the Natural sciences faculty.

Table 7 Supervisors by having awarded CL

	Sup	Sup%	CL	PhD ^b	CL%
Has awarded CL	462	11.60	248	2981	8.3
Has not awarded CL	3526	88.40	0	4884	0.0

2000–2018, 1 Case has missing gender data. Note that there are non CL PhDs who have supervisors from both groups, as a consequence the bottom number of PhDs contains some of the same PhDs as the top number

^bPhD students can be in both groups

Who awards CL?

Not all supervisors have supervised a PhD student that received CL; in fact, most of the supervisors (88.4%) have not (Table 7). For the group of PhD students that did have a supervisor who has awarded PhDs with CL, the probability to receive CL is 8.3%: 248 out of 2981 (Table 8). This is almost double the rate then among all supervisors included in this study.

Of all students that had at least one supervisor who at least once awarded CL, the share of female PhD students is 48.9%, compared to 46.3% for those without CL awarding supervisors. Therefore, the low share of women among CL awardees is not a distributional effect: that would only be so if those supervisors that award CL would have a rather low number of female PhD students—which is not the case.

Furthermore, those supervisors that did award CL, do that in various frequencies. Some have high percentages CL receivers, whereas others have lower percentages. Those awarding CL do so on average for 8.3% of their PhD students, which is much higher than the overall average. Of course, when you have only one or two PhD students, and one of them gets a CL, the percentages are 100% and 50%. However, there are quite a few supervisors with many PhD students who also have a high share of CL receivers. The other way around, there are also supervisors with many PhD students that never awarded CL. There are some differences in performance between the group of PhD students that had at least one supervisor that has awarded CL at least once, compared to the PhD students that has no such a supervisor, but these differences are very small. Summarizing, awarding CL is very unevenly distributed among supervisors, and the differences in success rates for male and female PhD students are not explained by the gender distribution of the PhD students among those two groups of supervisors.

Table 8 PhD students by type of supervisor

	PhD	PhD%	CL	CL%	W	M	W%
PhDs with CL granting supervisor	2981	52.0	248	8.3	1457	1523	48.9
PhDs without CL granting supervisor	2748	48.0	0	0	1272	1476	46.3

2000–2018, 1 Case has missing gender data

Table 9 CL by gender, faculty and quality (PhD period)

	<i>B</i>	SE	Wald	df	Significance	Exp(<i>B</i>)	95% CI for Exp(<i>B</i>)	
							Lower	Upper
Women vs. men	- 0.858	0.246	12.161	1	< .001	0.424	0.262	0.687
Faculty			29.196	5	< .001			
Faculty(1)	- 0.763	0.272	7.859	1	0.005	0.466	0.274	0.795
Faculty(2)	- 1.733	0.411	17.801	1	< .001	0.177	0.079	0.395
Faculty(3)	- 0.072	0.289	0.061	1	0.805	0.931	0.528	1.641
Faculty(4)	- 2.324	0.771	9.081	1	0.003	0.098	0.022	0.444
Faculty(5)	- 1.08	0.765	1.993	1	0.158	0.34	0.076	1.521
PhD year	- 0.045	0.026	3.036	1	0.081	0.956	0.91	1.006
Relative impact	- 0.096	0.122	0.627	1	0.428	0.908	0.715	1.153
Total impact	0.609	0.084	52.338	1	< .001	1.839	1.559	2.169
Journal impact	0.527	0.106	24.898	1	< .001	1.694	1.377	2.084
Independence	- 0.011	0.004	6.855	1	0.009	0.989	0.981	0.997
Team size	0.03	0.133	0.051	1	0.822	1.03	0.794	1.338
Constant	- 2.226	0.202	121.235	1	< .001	0.108		

Logistic regression; all PhD students in the selected faculties; impact indicators calculated over the *t* - 3 to *t* + 3 period

N = 2253; Nagelkerke *R*² = 0.203

Gender bias?

The independent variables are the three impact variables as well as the PhD student’s independence, team size, and gender. The faculty and the year of receiving the PhD degree were used as control variables. A logistic regression for men and women separately was used to test whether the independent variables have the same effects on the probability to get CL for men and women. This is the case, and testing interaction in the model with gender as variable did not lead to significant interaction terms. The logistic regression including gender but without interaction terms was used to test whether gender has an effect on CL (Table 9).

Two impact variables (total and journal impact) have an expected significant *positive* effect on the probability of obtaining CL, and the relative impact has *no effect*. Being *female* has a negative effect on the probability to get CL: the odds ratio for gender is 0.424 suggesting that men receive CL about twice as often compared to women after controlling for the quality of the PhD work and other covariates. This analysis unexpectedly suggests that independence has a negative effect on the chance of receiving CL. We come back to this when discussing the findings. Finally, the number of supervisors (team size) has no significant effect.¹⁴ The year of the PhD has a marginal significant negative effect, implying that in the later years the probability to receive CL slightly declines.

The odds ratio for gender is sensitive for non-linearity and measures the gender effect for independent variables at value zero. Using STATA 17, we calculated the predicted

¹⁴ The gender distribution among the supervisors also has no effect (not included in the analysis reported here).

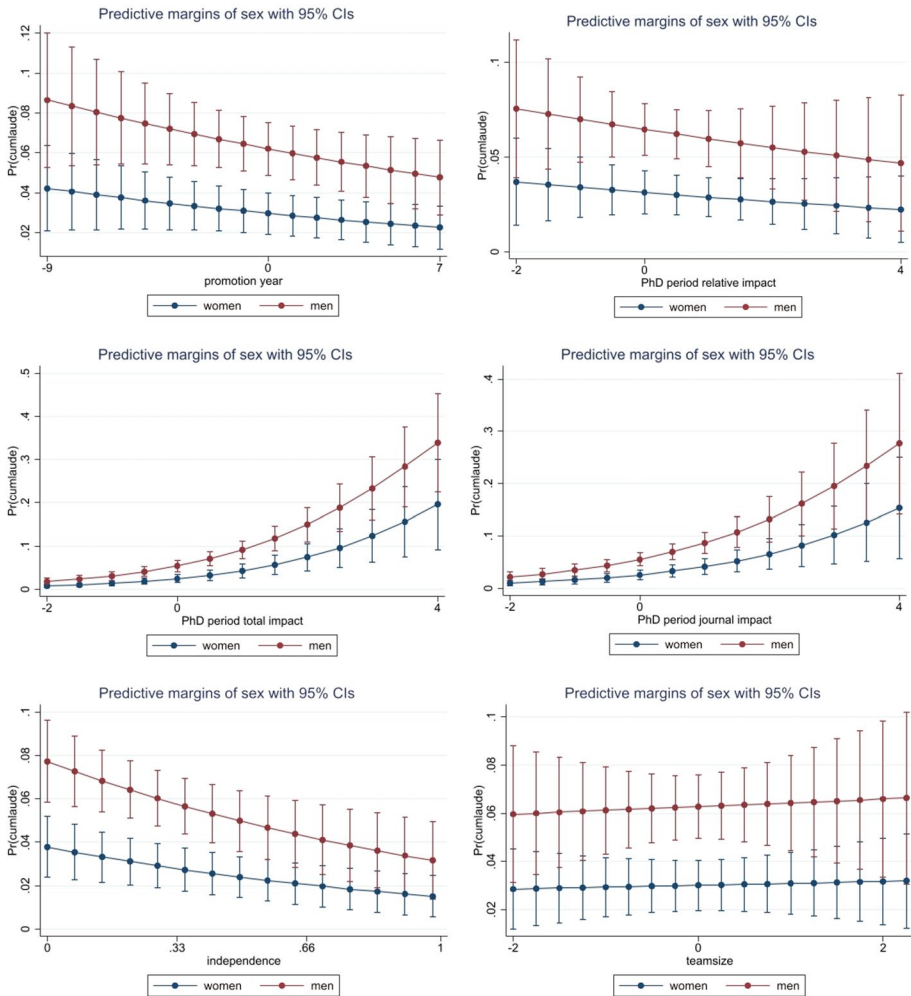


Fig. 4 Predictive margins of the probability to receive CL by sex (All cases, all impact variables are measured between $t - 3$ and $t + 3$, $t = \text{PhD year}$). (Colour figure online)

probability to receive CL for men and women *at the various levels of the independent variables* (Fig. 4). The top-left figure shows that over the years, the probability to get CL declines, but the men-women ratio remains about constant: In each year men get it about twice as often as women do—after controlling for the three impact variables. For most of the values of the *relative impact* (top-right figure), men have a significantly higher probability to get CL than women do, although this performance variable itself has no effect. The *total impact* (middle-left) and the *journal impact* (middle right) correlate strongly positive with the CL probability. The analysis shows that at the lower levels of the independent variables there is a significant gender difference, but at the high-performance levels, the 95% confidence intervals start to overlap, which is due to the low number of observations at those levels. The higher the performance level, the larger the difference between men and women, but the ratio remains constant: at all levels of the two impact variables men

Table 10 CL by gender, faculty and quality (Career)

	<i>B</i>	SE	Wald	df	Significance	Exp(<i>B</i>)	95% CI for Exp(<i>B</i>)	
							Lower	Upper
Women vs. men	- 1.027	0.245	17.616	1	< .001	0.358	0.222	0.578
Faculty			30.393	5	< .001			
Faculty(1)	- 0.653	0.267	5.993	1	0.014	0.521	0.309	0.878
Faculty(2)	- 1.7	0.401	17.983	1	< .001	0.183	0.083	0.401
Faculty(3)	0.038	0.285	0.018	1	0.893	1.039	0.595	1.815
Faculty(4)	- 2.361	0.753	9.817	1	0.002	0.094	0.022	0.413
Faculty(5)	- 1.136	0.78	2.121	1	0.145	0.321	0.07	1.481
PhD year	- 0.03	0.027	1.318	1	0.251	0.97	0.921	1.022
Relative impact	0.057	0.122	0.222	1	0.638	1.059	0.834	1.344
Total impact	0.401	0.068	34.491	1	< .001	1.493	1.306	1.707
Journal impact	0.509	0.1	26.135	1	< .001	1.664	1.369	2.023
Independence	- 0.005	0.005	1.236	1	0.266	0.995	0.985	1.004
Team size	- 0.01	0.13	0.005	1	0.941	0.99	0.768	1.278
Constant	- 2.085	0.338	37.998	1	< .001	0.124		

Logistic regression; all PhD students in the selected faculties; Impact indicators calculated over the career *N* = 2250; Nagelkerke *R*² = 0.180

have a double chance of getting CL.¹⁵ In other words, the predictive margins support the gender difference found in the logistic regression.

The higher a PhD student scores on *independence* (bottom-left), the lower the probability to receive CL, and at the same level of independence, the probability to receive CL is lower for women than for men. Finally, at the same *team size* men have a twice as high probability to receive CL than women. Some covariates affect the probability to get cum laude, but they do not affect the gender difference.

CL for exceptional talent?

One may argue that the CL recognition is not so much based on the quality of the dissertation itself, but on the talent of the PhD student recognized by the committee and on the expectation that this exceptionally talented researcher will produce exceptional work in the future. We tested that by not using the impact scores of the PhD awardees in the period of the PhD work, but over the longer career. Table 10 shows the findings, which are not very different from the previous analysis. Again, total impact and journal impact have a positive effect on the probability to receive CL, but independence, team size and the PhD year have no significant effect. The effect of gender is even stronger than in the previous analysis.

Inspecting the marginal probability plots (Fig. 5), one sees that similar to the previous analysis, the probability to get CL declines over time (top-left). Men have a significantly higher relative impact, but as the regression lines are almost horizontal, a

¹⁵ This suggest interaction, which was tested and was not statistically significant.

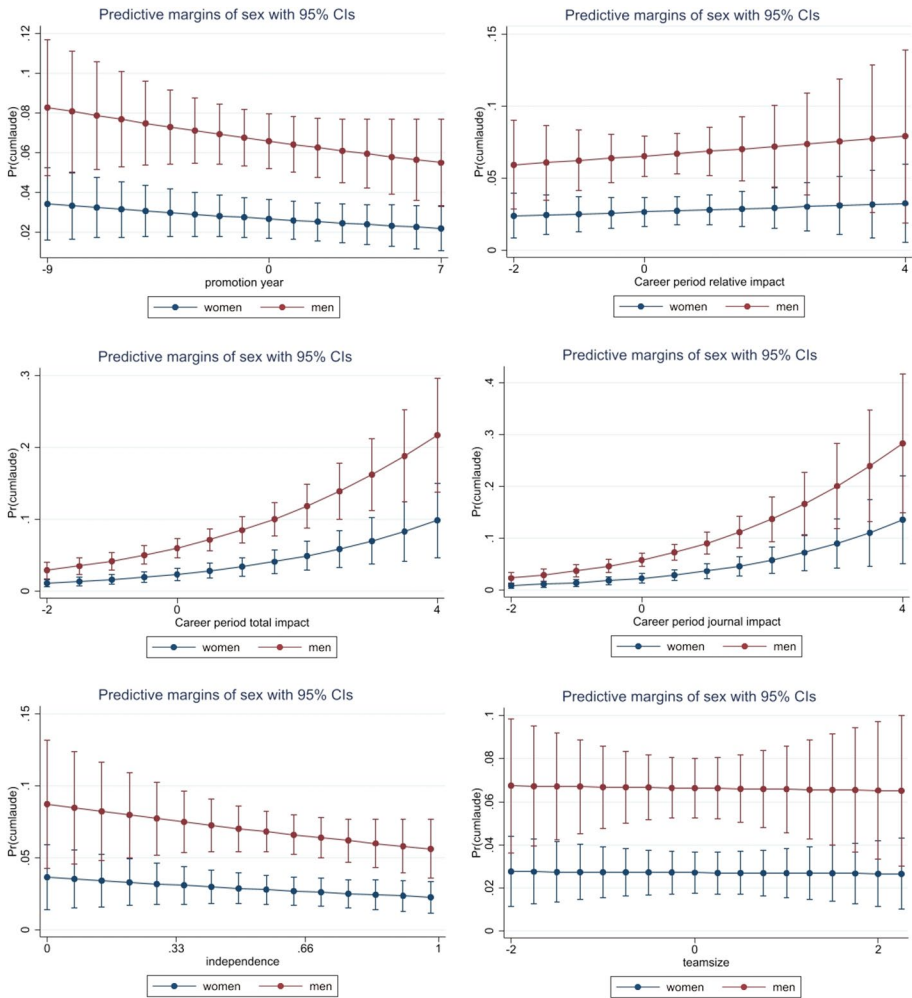


Fig. 5 Predictive margins of the probability to receive CL by sex (All cases; impact variables are measured over the career). (Colour figure online)

higher relative impact does not increase the probability to get CL (top-right). However, the estimated CL probability does increase with total impact (middle-left) and journal impact (middle right). Compared to the previous analysis (Table 9; Fig. 4) independence does not play a role (bottom-left). Although men seem to work in larger teams, team size has again no effect as also here the regression lines are horizontal (bottom-right).

Table 11 CL by gender, faculty and quality (top group—PhD period)

	<i>B</i>	SE	Wald	df	Significance	Exp(<i>B</i>)	95% CI for Exp(<i>B</i>)	
							Lower	Upper
Women vs. men	- 0.714	0.322	4.91	1	0.027	0.49	0.261	0.921
Faculty			10.911	5	0.053			
Faculty(1)	- 0.445	0.348	1.635	1	0.201	0.641	0.324	1.267
Faculty(2)	- 1.199	0.490	5.994	1	0.014	0.301	0.115	0.787
Faculty(3)	0.055	0.386	0.02	1	0.887	1.056	0.496	2.251
Faculty(4)	- 1.697	0.799	4.509	1	0.034	0.183	0.038	0.878
Faculty(5)	- 0.607	0.929	0.427	1	0.513	0.545	0.088	3.364
PhD year	0.036	0.034	1.094	1	0.296	1.037	0.969	1.109
Relative impact	- 0.911	0.191	22.72	1	< .001	0.402	0.277	0.585
Total impact	- 0.211	0.120	3.098	1	0.078	0.810	0.640	1.024
Journal impact	0.778	0.176	19.619	1	< .001	2.176	1.543	3.070
Independence	- 0.014	0.005	7.337	1	0.007	0.986	0.977	0.996
Team size	0.218	0.174	1.571	1	0.21	1.244	0.884	1.748
Constant	0.834	0.311	7.172	1	0.007	2.302		

Logistic regression; all PhD students in the top segment; Indicators calculated over the $t - 3$ to $t + 3$ period
N = 347; Nagelkerke = R^2 0.334

CL recipients versus the best non-recipients

In the previous analyses, all PhD graduates were included. As we controlled for the three impact variables, the result suggests that there should be enough high performing women that would—based on their performance—qualify for CL. Therefore, it is useful to have a closer look into the set of researchers with high-quality dissertations. That group is defined in terms of the total impact variable. As explained above, relative impact is an inadequate measure, and the journal impact is increasingly seen as an inadequate measure for assessing individual researchers (DORA; Leiden Manifesto). The threshold for the belonging to the best non recipients of CL is being in the top 10% in their faculty, resulting in a group of 238 top researchers, and together with 109 CL-recipients leads this to a sample of 347 included in the analysis.

The same logistic regression is used for this top group. In this model the relative impact variable has a significant negative effect, and total impact has a marginally significant negative effect, but journal impact has a significant positive effect on the probability to receive CL (Table 11). This suggests that the committees do focus on top journals (journal impact) instead of on individual top publications (total impact)—suggesting a flawed selection process.

Within the top group, being female lowers the probability to get CL considerably: the odds ratio is 0.49, not much different from the result of the previous analyses. Independence has also here a negative effect, and team size and the PhD year a nonsignificant positive effect.

The predictive margins plots (Fig. 6) again show that with a similar score on the independent variables, men have a higher probability than women to be awarded CL. Similar to some of the other analyses, the gender differences are not significant especially at the higher end of the distributions. The low number of observations are responsible for this.

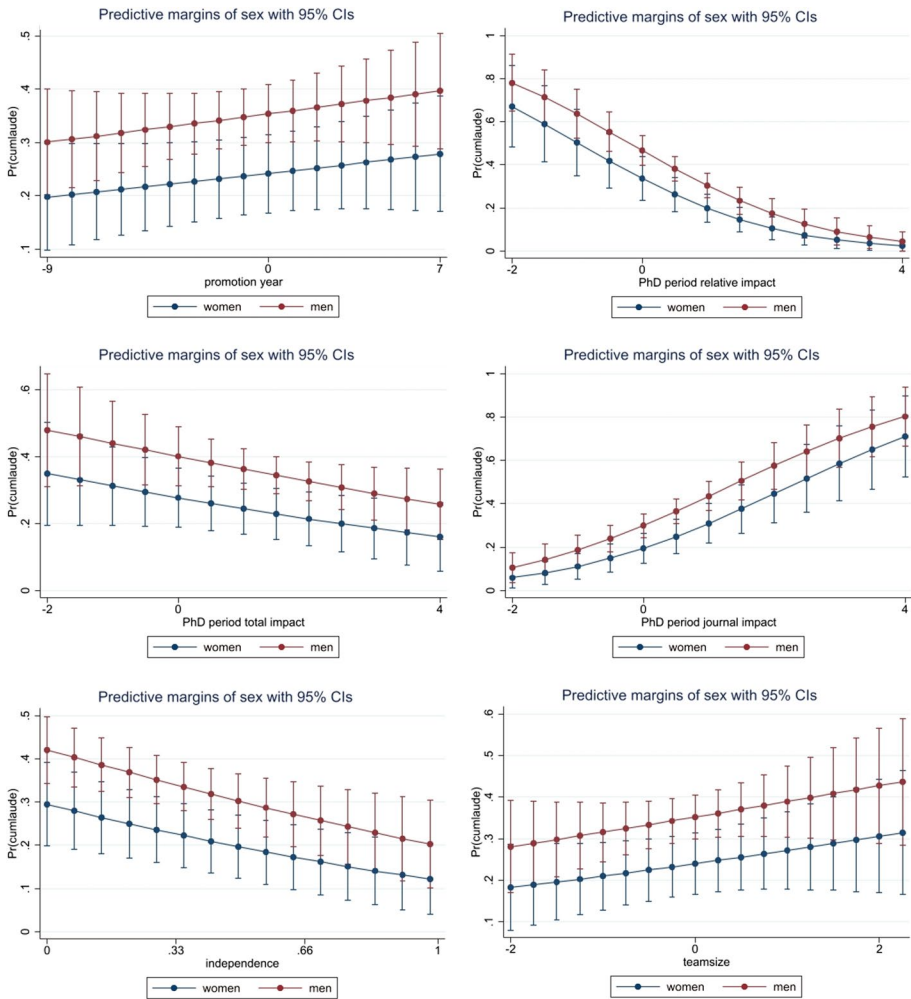


Fig. 6 Predictive margins of the probability to receive CL by sex (Top group, all impact variables are measured between $t - 3$ and $t + 3$, $t = \text{PhD year}$) (Colour figure online)

The predictive margins plots for the impact variables are interesting: The model estimates a *lower* probability to get CL for those that score high on the total impact scale, whereas the model estimates a higher probability for those that score high on the journal impact scale. It would be interesting to analyze a cohort of PhD students that graduated after the adoption of principles like those proposed by DORA and CoARA.

Conclusions

The raw data showed that men get twice as often CL than women do. In the logistic regression, we controlled for quality of the PhD thesis, for the independence of the PhD student, team size (in terms of the number of supervisors), for the faculty, and for the year

of obtaining the PhD degree. We find that gender remains having a strong and for women negative effect on the probability to receive CL. In fact, including the covariates does hardly change this effect.

Secondly, at each level of the three impact variables, the probability to get CL is about twice as high for men than for women. The ratio becomes slightly better for women at the high end of the performance distribution, but the difference remains large.

Thirdly, the higher the scientific independence of the PhD student, the lower the probability to receive CL. This holds both for men and women.

Fourthly, the PhD students that received CL do not outperform the best performing PhD students who did not get CL. On the contrary, the CL receivers are in terms of the total impact indicator outperformed by the best performing non-CL PhDs. At the same time, the findings indicate that the assessment and decision may be strongly based on the journal impact.

Fifthly, a large majority of the supervisors—also supervisors with a high number of PhD students—never awarded CL. This suggests that some of the best PhD students may not have received CL because awarding CL was not an option for their supervisors.

Sixthly, the data show that several of the CL-recipients have a very low (Scopus covered) output and impact—in the PhD period but also over the longer career. In the fields under study here, this is rather remarkable.

As CL may have strong career effects, our results taken together suggest that the current procedure may create unjustified differences.

Discussion

Although in other domains (like grant decision-making) gender bias seems to have declined or even disappeared (Van den Besselaar & Mom, 2021; Ceci et al., 2023), this does not seem to be the case in decisions about awards and other signs of recognition. This study supports the conclusion that awarding CL is strongly gender biased. After controlling for several covariates, the effect of gender remains large, and also stable over time. An explanation may be in the rather unstructured decision-making process and the lack of operational criteria, which does open the door for gender stereotypes and other sources of bias (Kahneman et al., 2021).

Most covariates showed an expected effect in the models, but that is not the case for the independence variable. Developing towards scientific independence scientist is generally seen as an important characteristic, and therefore the expectation was that the level of independence would have a positive effect on the probability of receiving CL, but the opposite was found. This hints at relational or collaboration mechanisms: closer collaboration is better for receiving a CL recognition, and in the very early career the variable may not measure independence but the collaboration and relation with the supervisors. Having only a few papers coauthored with the supervisor(s) may indicate a social distance to the supervisor(s). Our findings suggests that network effects play a role: The closer the collaboration with the supervisors, the more likely it is to receive CL for the dissertation, which indicates the existence of other sources of (e.g. network based) bias in the decision making.

The findings have a policy implication: As the CL distinction depends at least partly on gender and on the network ties to the supervisors, the decision-making process should either be formalized much stronger, in order to guarantee that all PhD students are evaluated in a similar and operationalized way. Or, if one expects that this is not possible, one may consider stopping with awarding CL for the PhD thesis all together.

This study also has several limitations. Firstly, one may dispute the way the quality of the PhD theses is measured using bibliometric indicators. However, as no other independent assessments of the theses exist, this is the best option available. But there are also positive arguments for the selected approach: an excellent PhD thesis can be expected to lead (in the fields we covered here) to impactful papers. And although a lot of randomness may affect the citations received at the individual level, at the group level one would expect that CL-dissertations would lead to higher scores on the total impact variable than the non-CL dissertations would. A second issue is the selection of the top group: Our top includes about 15% of the total population, of which about a third received the CL distinction. One may find this top group too large, and it might be reasonable to use a more selective definition. However, if we would make the top group smaller, then the difference with the CL receivers will even increase, as a more restricted group of non-CL top performers have even higher average impact scores. The CL receivers would then even more clearly underperform compared to the others in the top group. Thirdly, the explained variance (Nagelkerke pseudo R^2) found in the various regression analyses is rather low, suggesting that other factors than gender, independence and impact may be behind the CL decision. Alternatively, the decision-making could be largely random. A last issue is that we do not know when PhD trajectories exactly start, only when they end, and it is possible that CL is partly awarded for being quick (using little time) or being thorough (spending more time). Nor do we know how many hours per week the CLs work compared to the other PhD students.¹⁶

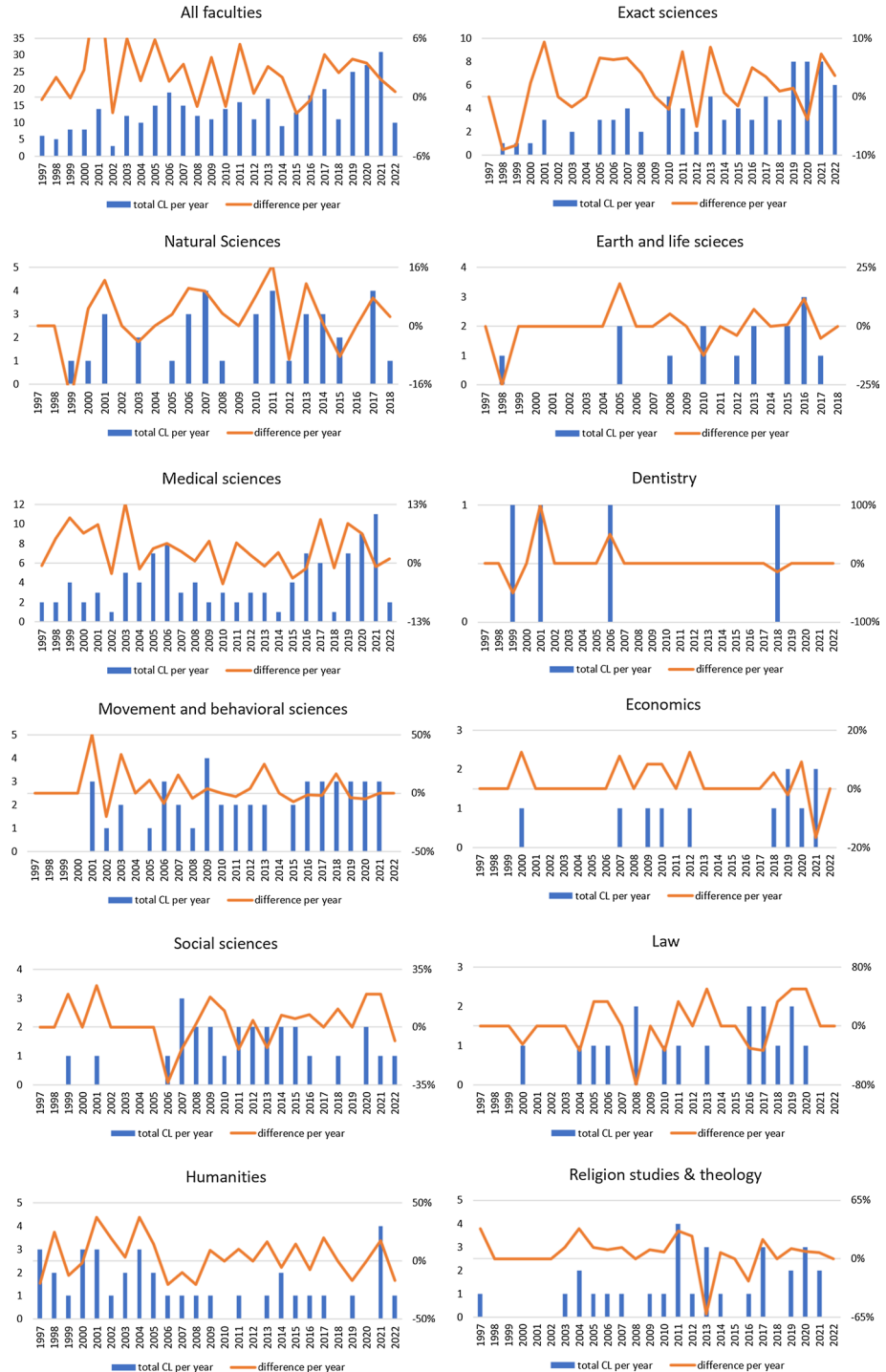
Follow up studies could investigate whether several variables not included in this report partly explain the CL decision, e.g.: (1) the personal relation between the supervisor(s) and the PhD student, (2) the personality of the PhD student, (3) personalities of the supervisors, (4) disciplinary differences between the supervisors within a team, (5) the composition of the PhD committee, and (6) the workload of supervisors (the number of simultaneously supervised PhD students, including those that do not finish successfully). Beyond those, (7) differences exists between internal PhD students with an employment contract with the university and external PhD students who do not work for the university, and (8) PhD students differ in terms of nationality. (9) Some characteristics of the supervisors and of the committee members may play a role too, such as the scientific quality, gender, and the attitude towards the CL distinction. Finally, (10) the composition of the committee may be related to bias in the CL decision such as the gender composition (as found by Bol, 2023), or the age of the committee members. Unfortunately, our data does not cover the committees.

There are two lines of research that we are exploring the moment: Firstly, as supervisors select who they propose for CL, the quality of the personal relations between the PhD student and the supervisor may play a role. The findings reported above about independence variable also point in that direction. We are analyzing the acknowledgement texts of a sample of PhD theses. The preliminary findings indicate that these texts may enable to develop indicators for the quality of the personal relation between the supervisor and the PhD student, using language as an expression of relationships, opinions, and emotions (Pennebaker et al., 2001; Tausik & Pennebaker, 2010).

Secondly, and without entering a theoretical discussion about which personality factors may play a role and why, the personality of a PhD student may also affect the probability to receive CL. In order to measure that, one could use a questionnaire including a validated personality test. This can only be done meaningfully for a recent sample of PhD students. The resulting variables, such as the social relation with the supervisor and the PhD student's personality, may have a gender dimension.

¹⁶ About 48% of all PhDs in the Netherlands have a employment contract per 2018 (<https://www.rathenau.nl/nl/wetenschap-cijfers/wetenschappers/van-promovendus-tot-promotie>).

Annex: cum laude by faculty



Orange line: % men CL - % women with CL (right Y-axis). Bars: number of CL awards by year (left Y-axis)

Acknowledgements The authors acknowledge Tijs van den Broek (Vrije Universiteit Amsterdam) for comments on an earlier version, and Harry Ganzeboom (Vrije Universiteit Amsterdam) and Luis Sanz-Méndez (CSIC, Spanish National Research Council) for methodological advice. The studied university provided the data about the PhD graduates. The report underlying this article was presented to the rector and the deans of the university in July 2021. An earlier version was presented at the ISSI conference in Bloomington, Indiana in July 2023 (Van den Besselaar & Mom, 2023). This version is substantially extended.

Funding Part of the work has been supported by the European Commission, Grant 824574, the GRANTeD Project.

Declarations

Conflict of interest The first author is a member of the Distinguished Reviewers Board of Scientometrics. No other competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bol, T. (2023). Gender inequality in 'cum laude' distinctions for PhD students. *Scientific Reports*, *13*(1), 20267.
- Brouns, M., Bosman, R., & van Lamoen, I. (2004). *Een kwestie van kwaliteit*. Rijksuniversiteit Groningen.
- Casadevall, A., & Handelsman, J. (2014). The presence of female conveners correlates with a higher proportion of female speakers at scientific symposia. *Mbio*, *5*(1), e00846-13.
- Ceci, S. J., Kahn, S., & Williams, W. M. (2023). Exploring gender bias in six key domains of academic science: An adversarial collaboration. *Psychological Science in the Public Interest*, *24*(1), 15–73.
- Chubin, D. E., & Hackett, E. J. (1990). *Peerless science: Peer review and U.S. science policy*. SUNY Press.
- Cole, S. (1992). *Making science: Between nature and society*. Harvard University Press.
- Cole, S., Cole, J. R., & Simon, G. A. (1981). Chance and consensus in peer review. *Science*, *214*(4523), 881–886.
- Cole, S., Rubin, L., & Cole, J. R. (1977). Peer review and the support of science. *Scientific American*, *237*(4), 34–41.
- De Bruin, E. (2018, October 19). *Helpt van de promovendi is vrouw. Maar cum laude krijgen ze zelden*. NRC.
- De Bruin, E. (2019, June 14). *Universiteiten bestrijden m/v-verschil bij promoties (Universities fight M/F disparity in PhD promotions)*. NRC.
- Ding, W. W., Murray, F., & Stuart, T. E. (2012). From bench to board: Gender differences in university scientists' participation in corporate scientific advisory boards. *Academy of Management Journal*, *56*(5), 1443–1464.
- Dumitra, T. C., Trepanier, M., Lee, L., Fried, G. M., Mueller, C. L., Jones, D. B., & Feldman, L. S. (2019). Gender distribution of speakers on panels at the Society of American Gastrointestinal and Endoscopic Surgeons (SAGES) annual meeting. *Surgical Endoscopy*, *34*, 4140–4147.
- Elsevier Research Intelligence. (n.d.) *Research metrics guidebook*. Elsevier.
- Johnson, C., Smith, P. K., & Wang, C. (2016). Sage on the stage: Women's representation at an academic conference. *Personality and Social Psychology Bulletin*, *43*(4), 493–507.
- Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise, a flaw in human judgment*. Little, Brown Spark.

- Klein, R. S., et al. (2017). Speaking out about gender imbalance in invited speakers improves diversity. *Nature Immunology*, 18(5), 475–478.
- Lamont, M. (2010). *How professors think; inside the curious world of academic judgment*. Harvard University Press.
- Lincoln, A. E., Pincus, S., Bandows Koster, J., & Leboy, P. S. (2012). The Matilda Effect in science: Awards and prizes in the US, 1990s and 2000s. *Social Studies of Science*, 42, 307–320.
- Melnikoff, D. E., & Valian, V. V. (2019). Gender disparities in awards to neuroscience researchers. *Archives of Scientific Psychology*, 7, 4–11.
- Merton, R. K. (1973). The normative structure of science. In R. K. Merton (Ed.), *The sociology of science* (pp. 267–278). University of Chicago Press.
- Möller, T., Mom, C., & Van den Besselaar, P. (2022). What is researcher independence and how can it be measured? In N. Robinson-Garcia, D. Torres-Salinas & W. Arroyo-Machado (Eds.), *Proceedings 26th international conference on science and technology indicators, STI 2022*, 2022.
- Mood, C. (2009). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European Sociological Review*, 26, 67–82.
- Nittrouer, C. L., Hebl, M. R., Ashburn-Nardo, L., Trump-Steele, R. C. E., & Lane, D. M. (2017). Valian VV (2017) Gender disparities in colloquium speakers at top universities. *Proceedings of the National Academy of Sciences of USA*, 115(1), 104–108.
- Pennebaker, J. W., et al. (2001). *Linguistic inquiry and word count LIWC 2001*. Lawrence Erlbaum.
- Reschke, B. P., Azoulay, P., & Stuart, T. E. (2018). Status spillovers: The effect of status conferring prizes on the allocation of attention. *Administrative Science Quarterly*, 63, 819–847.
- Sandström, U., & van den Besselaar, P. (2016). Quantity and/or Quality? The importance of publishing many papers. *PLoS ONE*, 11(11), e0166149.
- Sardelis, S., & Drew, J. A. (2016). Not “Pulling up the Ladder”: Women who organize conference symposia provide greater opportunities for women to speak at conservation conferences. *PLoS ONE*, 11(7), e0160015.
- Silver, J. K., Blauwet, C. A., Bhatnagar, S., Slocum, C. S., Tenforde, A. S., Schneider, J. C., Zafonte, R. D., Goldstein, R., Gallegos-Kearin, V., Reilly, J. M., & Mazwi, N. L. (2018). Women physicians are underrepresented in recognition awards from the Association of Academic Physiatrists. *American Journal of Physical Medicine and Rehabilitation*, 97(1), 34–40.
- Tauzik, Y. R., & Pennebaker, J. W. (2010). Psychological meaning of words. LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29, 24–54.
- Van Arensbergen, P., Van der Weijden, I., & Van den Besselaar, P. (2014). The selection of talent as a group process; a literature review on the dynamics of decision-making in grant panels. *Research Evaluation*, 23(4), 298–311.
- Van den Besselaar, P., & Mom, C. (2021). Gender bias and grant allocation—A mixed picture. Preprint, [arXiv:2205.13641](https://arxiv.org/abs/2205.13641)
- Van den Besselaar, P., & Mom, C. (2023). Gender and merit in awarding cum laude for the PhD thesis. In *Proceedings of ISSI 2023—The 19th international conference of the International Society for Scientometrics and Informetrics*, 2023 (Vol. 1, pp. 611–625).
- Van den Besselaar, P., & Sandström, U. (2017). Vicious circles of gender bias, lower positions and lower impact: Gender differences in scholarly productivity and impact. *PLoS ONE*, 12(8), e0183301.
- Van den Besselaar, P., & Sandström, U. (2019). Measuring researcher independence using bibliometric data: A proposal for a new performance indicator. *PLoS ONE*, 14(3), e0202712.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.