Gender effects in Dutch research funding: a statistical investigation of the Research Talent Programme 2012-2021

Albers, C.; Molen, S.J. van der

Gender effects in Dutch research funding:
A statistical investigation of the Research Talent Programme 2012–2021

Casper Albers

Heymans Institute for Psychological Research, University of Groningen, the Netherlands

Sense Jan van der Molen

Leiden Institute of Physicics, Leiden University, the Netherlands

Author Note

Correspondence concerning this article should be addressed to Casper Albers, Grote Kruisstraat 2/1, 9712TS Groningen, the Netherlands. Email: c.j.albers@rug.nl

Abstract

In 2015, the Dutch research council, NWO, took measures to combat gender bias disadvantaging female applicants in a popular three-tiered funding scheme called the Talent Programme. Using all available data for the last 10 years of applications, we study whether these measures had an effect. We find strong statistical evidence of a shift in gender effects in favour of female applicants in the first tier, called Veni. Gender differences are not found in the two other tiers, the Vidi and Vici schemes.

Keywords: gender, science funding, the Netherlands

Gender effects in Dutch research funding: A statistical investigation of the Research Talent Programme 2012–2021

Introduction

One of the main sources of research funding in the Netherlands is the Talent Programme of the Dutch Research Council, NWO. This funding scheme consists of three tiers called Veni, Vidi, and Vici, respectively, after Julius Caesar's (in)famous phrase. Veni-grants (at most 280k€) can be applied for by young scientists who are within three years of receiving their PhD-degree. Vidi-grants (at most 800k€) can be applied for by scientists up to eight years after receiving their PhD-degree, and Vici-grants (at most 1.5M€) are open to those within fifteen years of obtaining their PhD-degree. In certain situations, such as childcare responsibilities, these terms can be extended.

In this study we investigate possible gender effects in the assessment procedure of the Talent Programme. We have chosen to put the main emphasis on the Veni-scheme for a number of reasons. First and foremost, potential gender bias in the Veni system has been studied extensively in recent years. In 2015, Van der Lee and Ellemers (2015) argued that this grant scheme disadvantaged women, which led to national newspaper articles and discussion in the Dutch parliament (Bussemaker, 2015). Despite methodological criticism (Albers, 2015; Volker & Steenbeek, 2015) on the analyses that formed the basis of these discussions (Van der Lee & Ellemers, 2015), NWO decided to take several measures to combat gender bias in their funding schemes, such as introducing implicit bias training for committee members. Now that the measures taken by NWO have had considerable time to take effect, we aim to evaluate their influence. To explicitly include the possibility that some time was needed for the measures to become effective, we will not only study the gender effects in Veni awards averaged over the full time period, but also whether differences, if any, have increased or decreased over the years considered.

Other reasons to focus on the Veni grants are the following. If in this first tier gender effects occur, this automatically affects career prospects of women and men throughout their future career, e.g. due to the so-called Matthew effect (Bol, de Vaan, & van de Rijt, 2018). Furthermore, by far the highest number of grants given in the funding scheme are Veni grants, thus providing sufficient information for statistical analyses. We will analyse the publicly available data on the Vidi and Vici grants in the same way as the Veni grants, but the relatively small number of applications and grants hampers the possibility of drawing strong statistical inferences. We note that in recent years NWO has also started various calls dedicated to underrepresented groups so as to promote diversity in academia. The Talent Programme grants are not part of these calls. They are intended for all junior researchers and are thus intended to be free of (gender) effects.

In our study, we define gender effects as differences between success rates of men and women that cannot be attributed to coincidence. Gender effects include both gender bias (i.e. the effects of (unconscious) prejudice against a gender) as well as any other effects that cause systematic deviations in performance of men and women in academia.

The goal of this study is to test whether observed gender differences in the success rate of the Talent Programme grants can be attributed to coincidence or not. More precisely, we consider the following research question: 'In absence of any gender effects in quality of applications and the considerations of the assessment committee, what is the probability of finding at least the same gender difference as was found in the data of 2012-2021?'. We will answer this research question using publicly available information on the number of applications and grants, by year, gender and research domain.

Several studies have investigated (other) aspects of gender bias in Dutch academia; e.g. during the PhD-trajectory, i.e., before being eligible for a Veni-grant (Yerkes, Sonneveld, & van de Schoot, 2012), or after receiving a grant (van de Schoot, Sonneveld, & Kroon, 2012). A very recent study (Bol, de Vaan, & van de Rijt, 2022) had an objective similar to ours: to study gender effects in the NWO Talent Programme. In their case, the authors studied confidential assessment reports to find that, in the end, there is no evidence for gender effects in the final funding, although males did receive significantly better reviews. They conclude that juries tend to correct for this gender imbalance when taking the final decision to award grants. Whereas Bol et al. (2022) use data up to 2016, we also include more recent data, up to 2022. The main contribution of our study, compared to that of Bol et al. (2022), is that we focus on interactions between gender on the one hand and both year and field on the other, being interested in the question whether or not gender effects are comparable across years and fields.

To investigate our research question, we apply and compare four possible statistical models, with increasing complexity, for each of the three tiers. For the Veni tier, all models lead to the statistically significant conclusion that there is indeed a difference between the succes rates of male (lower) and female (higher) applicants overall. The models also show that this difference increases over time for all domains. For the Vidi and Vici tier, no gender differences are found.

The goal of this paper is to share and discuss the numbers and their statistics. While we hope that our work will stimulate further discussion on an explanation of the (lack of) differences found, it is outside the scope of this paper to start this debate. Hence, we refrain from interpreting the results in this present contribution.

The data

We have looked at all research grants from 2012 to the most recent grants at the moment of writing[1], restricting our attention to the publicly available data: numbers of applications and numbers of funded projects. Throughout this study, the calender year mentioned refers to the year of the funding decision, which usually is the year after the grant submission. Here, we have focused on the period from 2012 onwards. The previous period, up to 2012, had already been assessed by (Van der Lee & Ellemers, 2015). Since NWO took its measures after the latter paper appeared (in 2015), the time period chosen (2012-now) allows us to investigate the possible effects of the new policy. All data discussed here have been obtained from NWO's website[2].

For these programmes, NWO distinguishes five research fields:

- ENW: science
- TTW: applied and engineering sciences
- SGW: social sciences and humanities
- ZonMW: health research
- DO: cross-domain/interdisciplinary. (This domain has been cancelled as of 2020).

For each year and each field, we have recorded the number of submitted applications and granted applications for men and women separately. NWO publicly shares the necessary information for most but not all years, see the Supplementary Material for a detailed overview.

[1] We have included all data that were published on NWO's website until and including March 15, 2022.

[2] See NWO (2022) for the Veni data. Using the menu on the right, the data for Vidi and Vici are available. The data are also provided as Supplementary Material

The models

To model the probability of success, p_i, of a given application, we employ logistic regression (or binomial generalized linear models, McCullagh and Nelder (1989)). In these models, the expected logodds of p_i, $\log(p_i/(1-p_i))$ are predicted on the basis of a number of predictors. In our case, the success probabilities are predicted based on gender of the applicant, the field of study, and the year of application.

We distinguish four different models, of increasing complexity, based on these predictors:

1. Model 1: gender, field and year are used as additive predictors.
2. Model 2: as Model 1, but with an interaction between gender and year: the gender effect can differ per year.
3. Model 3: as Model 2, but with also an interaction between gender and field.
4. Model 4: as Model 3, but with also an interaction between year and field, i.e. all three second-order interactions.

Data for the three tiers are analyzed separately. Model fit and model parsimony are assessed through the Akaike Information Criterion.

The first model is specified by

$$\begin{aligned}\log\left(\frac{p_i}{1-p_i}\right) &= \beta_0 + \beta_M D_{M,i} + \beta_{DO} D_{DO,i} + \beta_{ENW} D_{ENW,i} + \beta_{TTW} D_{TTW,i} + \\ &\quad \beta_{ZonMw} D_{ZonMw,i} + \beta_{Year}\text{Year}_i + \varepsilon_i.\end{aligned}$$

Here, $D_{X,i}$ is used as notation for the dummy variable (also known as the Kronecker delta $\delta_{X,i}$) indicating whether person i belongs to class X (then $D_{X,i} = 1$) or not (then $D_{X,i} = 0$). A class X can stand for a research field, e.g. ENW or a gender ('M' is used as notation for male applicants, with female being the reference group for gender). The field SGW is chosen as reference field, as this field had the largest number of applications[3]. Variable 'Year' is included to measure the longitudinal effects. This variable is coded as 1 for 2012, 2 for 2013, . . ., 10 for 2021.

Subsequently, Model 2 is specified by

$$\begin{aligned}\log\left(\frac{p_i}{1-p_i}\right) &= \beta_0 + \beta_M D_{M,i} + \beta_{DO} D_{DO,i} + \beta_{ENW} D_{ENW,i} + \beta_{TTW} D_{TTW,i} + \\ &\quad \beta_{ZonMw} D_{ZonMw,i} + \beta_{Year}\text{Year}_i + \beta_{M,Year} \times \text{Year}_i \times D_{M,i} + \varepsilon_i,\end{aligned}$$

[3] Note that the choice of reference fields is arbitrary: any other choice would have yielded exactly the same predicted success rates

thus with an additional interaction term $\beta_{M,Year} \times \text{Year}_i \times D_{M,i}$. Analogously, in Model 3, interaction terms between gender and field are added, while Model 4 adds interaction terms for year and field to that.

All computations have been performed in *R* (version 4.1.2; R Core Team (2021)). The analyses of variance have been carried out using the *R* package 'car' (Fox & Weisberg, 2019).

Results

The full dataset consists of a total of 16,249 applications (6,907 from female applicants, 9,342 from male applicants). Out of these, 2,449 have been granted (1,067 for female applicants, i.e. a success rate of 15.4% ; and 1,382, for male applicants, i.e. a 14.8% success rate). There were no applicants that did not declare a gender, nor did any candidate declare a gender other than male or female. With 10,076 applicants and 1,472 funded applicants, the Veni tier is by far the largest tier. All descriptives are provided in Table 1. Note that in absolute numbers, male applicants outnumber female applicants and this gap grows with the tiers. In relative numbers, i.e. success rate, however, male applicants do not outperform female applicants, as discussed below.

As the first tier consists of 62% of all applications and 60% of all grants, we focus on this (Veni) scheme first, and in most detail. We find that all four models described predict lower success percentages for male applicants than for female applicants. Furthermore, clear differences in success rates between fields are observed, which is in line with previous studies on NWO's Veni grants (Albers, 2015; Volker & Steenbeek, 2015). To avoid the Simpson's paradox fallacy (Albers, 2015; Volker & Steenbeek, 2015), all models take field of study into account.

Table 2 displays the results of an analysis of variance on the four models, and Table 3 displays the AIC-comparisons. The latter table clearly demonstrates that inclusion of a $gender \times year$ interaction is beneficial (Model 2). Model 3, which additionally includes the four $gender \times field$ interactions, has an even lower AIC-score, indicating that the gender gap changes over time for all fields. On the other hand, the addition of the $year \times field$ terms in Model 4 provides no significant improvement to the model fit ($p = .385$), as indicated by a higher AIC-value. Thus, we will look at Model 3 in more detail, as presented in Table 4. An explanation on how to interpret the coefficients of Table 4 is given in Appendix A. In

Appendix B the R code of the analyses is provided. This, in combination with the data (Supplementary Material) will provide full results of the three other models.

Figure 1 represents the observed success probabilities and the predicted success probabilities according to Model 3 over the years considered. In this Figure, we present a graph for each field. In Figure 2 we aggregate the figures for the five domains into a single figure, using the numbers of applications per field as weights. All graphs in Figures 1 and 2 show a positive trend for grant succes rates for females and a (corresponding) negative one for males. The year at which the two lines cross varies per field. For DO, ENW and TTW the crossing takes place around 2012, where our dataset starts, whereas for SGW (around 2017) and ZonMw (around 2018), they happen later in time – although the uncertainty in these predictions is considerable. A crossing can also be observed in the aggregate predictions of Figure 2, roughly around the year 2015. As seen in Figure 1, there is considerable distance between certain observations and the corresponding predictions. This calls for some caution: whereas the model is sufficient to estimate the gender effect as a whole, it will not be sufficient for predictions for individual combinations of gender, year and field, let alone extrapolations to future years. Note that the uncertainty in the moment of crossing is also considerable, making it difficult to assess when the success rate of female applicants overtakes those of male applicants precisely. Still, this does not diminish the significant change in gender effects over time.

In Table 6 all predicted success probabilities for the Veni for all four models are listed.

In the same vein as the analyses for the first tier, the Vidi and Vici tiers are analysed. Unlike in the Veni data, for both these tiers the best performing model is Model 1, the model without any interactions of gender with one of the other variables (Table 7). Furthermore, neither in the Vidi nor in the Vici data a significant effect of gender is found (Table 8). Thus, in contrast with the Veni data, there is no evidence for any gender effect in success rate: no base rate difference, nor a change of this effect over time. The lack of significant gender effects is illustrated in Figure 3.

Discussion

Let us now revisit the research question considered, i.e. ‘In absence of any gender effects in the quality of applications and considerations of the assessment committee, then what is the probability of finding at least the gender difference as was found in the data of 2012–2021?' For the Veni tier, this p-value is found to be smaller than 0.001, i.e. there is a very significant gender difference. For the other two tiers, Vidi and Vici, no significant gender effects were found.

This does not need to imply that the assessment committees systematically disadvantage men in the Veni funding, nor that the quality of applications from men and women differ systematically. Our model is correlational and not causal. The purpose of this paper is not to find the mechanisms behind observed gender effects, nor to state whether or not they are due to gender bias, but merely to answer the question whether the observed gender effects are statistically significant. They are in the Veni data. They are not in the other tiers.

Despite their relatively high success rates in the Veni scheme, however, it does appear that more women than men leave academia before reaching the second and third tier of the Talent Programme. The fact that the percentage of female applicants clearly declines over the tiers (46% for Veni, 40% for Vidi, 33% for Vici) supports this.

One of our main results is that gender effects in the Veni tier have shifted over the years, in favour of females. It could hence be that the measures taken by NWO to combat gender effects against women - introduced after the Veni study by Van der Lee and Ellemers (2015) - have indeed been successful. However, since gender effects in the Veni's were small, or even absent, to start with (see Albers (2015); Van der Lee and Ellemers (2015); Volker and Steenbeek (2015) and Figures 1 and 2), these measures may have led to an overshoot. In a recent study, Bol et al. (2022) studied all Talent Programme data, including (confidential) scores from reviewers. These authors found that male applicants receive better reviewer scores than female applicants - indicative of gender effects in assessment. Yet, they also find evidence that external review scores were corrected for by the panels, mostly in the rebuttal phase. Furthermore, women are overrepresented at ranking positions just above the funding threshold.

Combining the conclusions by Bol et al. (2022) with our results, we hypothesize that the corrections performed by the juries may have gotten stronger over the years, yielding an overcorrection in recent times. This provokes the question what NWO can do to balance out the Veni scheme. And more generally, what policy funding agencies should have to prevent statistically relevant biases in the future. Clearly, to guarantee a proper feedback mechanism, a continuous, critical assessment of the available data over time is essential. It is our hope that this article contributes to exactly that.

CRediT authorship contribution statement

Casper Albers: Writing – original draft, Writing – review editing, Conceptualization, Methodology, Investigation, Visualization, Formal analysis. **Sense Jan van der Molen**: Writing – original draft, Writing – review editing, Conceptualization, Investigation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank Mara Yerkes and Thijs Bol for helpful feedback on the manuscript.

References

Albers, C. J. (2015). Dutch research funding, gender bias, and Simpson's paradox. *Proceedings of the National Academy of Sciences*, *112*(50), E6828–E6829.

Bol, T., de Vaan, M., & van de Rijt, A. (2022). Gender-equal funding rates conceal unequal evaluations. *Research Policy*, *51*(1), 104399. doi: https://doi.org/10.1016/j.respol.2021.104399

Bol, T., de Vaan, M., & van de Rijt, A. (2018). The matthew effect in science funding. *Proceedings of the National Academy of Sciences*, *115*(19), 4887-4890. doi: 10.1073/pnas.1719557115

Bussemaker, J. (2015). *Antwoord op vragen van het lid Yücel inzake het bericht dat 'NWO vrouwelijke wetenschappers discrimineert'.* Retrieved 2022-01-25, from `https://zoek.officielebekendmakingen.nl/ah-tk-20152016-502.html`

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). Sage.

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (2nd ed.). Chapman and Hall.

NWO. (2022). *Projecten Veni.* Retrieved 2022-01-25, from `https://www.nwo.nl/onderzoeksprogrammas/nwo-talentprogramma/\protect\discretionary{\char\hyphenchar\font}{}{}projecten-veni`

R Core Team. (2021). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Van der Lee, R., & Ellemers, N. (2015). Gender contributes to personal research funding success in The Netherlands. *Proceedings of the National Academy of Sciences*, *112*(40), 12349–12353.

van de Schoot, R., Sonneveld, H., & Kroon, A. (2012). *Mobiliteitsonderzoek vernieuwingsimpulslaureaten. rapport voor nwo [research on the mobility of innovative subsidies].* Nederlands Centrum voor de Promotieopleiding IVLOS, Universiteit Utrecht.

Volker, B., & Steenbeek, W. (2015). No evidence that gender contributes to personal research funding success in The Netherlands: A reaction to Van der Lee and Ellemers. *Proceedings of the National Academy of Sciences*, *112*(51), E7036–E7037.

Yerkes, M., Sonneveld, H., & van de Schoot, R. (2012). Genderongelijkheid in het nederlands promotiestelsel. een verkennend onderzoek. [gender inequality in the dutch phd system.

an exploratory study]. *Tijdschrift voor Genderstudies*, *3*, 6–23.

	Veni		Vidi		Vici		Total	
	Applications	Granted	Applications	Granted	Applications	Granted	Applications	Granted
Women	4,590	695	1,588	268	729	104	6,907	1,067
Men	5,486	777	2,400	411	1,456	196	9,342	1,382
Total	10,076	1472	3,988	679	2,185	300	16,249	2,449

Table 1

Numbers of applications and project fundings.

	Model 1			Model 2			Model 3			Model 4		
	df	χ^2	p-value	df	χ^2	p-value	df	χ^2	p-value	df	χ^2	p-value
Gender	1	6.573	.010	1	21.133	$< .001$	1	18,972	$< .001$	1	20.146	$< .001$
Field	4	80.256	$< .001$	4	7.185	.007	1	6.331	.012	1	.446	.504
Year	1	.982	.322	1	81.134	$< .001$	4	73.333	$< .001$	4	4.126	.389
Gender × Year	–	–	–	1	21.166	$< .001$	1	18.967	$< .001$	1	20.142	$< .001$
Gender × Field	–	–	–	–	–	–	4	16.591	.002	4	14.965	.005
Year × Field	–	–	–	–	–	–	–	–	–	4	4.173	.383

Table 2

Analysis of variance of the four models for the Veni data. The χ^2-values display the Wald test statistics, the other two columns per model the corresponding degrees of freedom and p-values.

	df	AIC
Model 1	7	478.04
Model 2	8	458.77
Model 3	12	450.24
Model 4	16	454.05

Table 3

Comparison between the four models using the Akaike Information Criterion for the Veni data.

	$\hat{\beta}$	SE	p-value
Intercept	-2.204	0.108	$< .001$
Gender: Male	0.511	0.146	$< .001$
Year	0.040	0.016	0.012
Field: DO	0.580	0.166	$< .001$
Field: ENW	0.808	0.101	$< .001$
Field: TTW	0.368	0.172	0.033
Field: ZonMw	0.046	0.123	0.712
Male × Year	-0.094	0.022	$< .001$
Male × Field: DO	-0.470	0.256	0.067
Male × Field: ENW	-0.391	0.137	0.004
Male × Field: TTW	-0.426	0.219	0.052
Male × Field: ZonMw	0.170	0.175	0.331

Table 4

Results for Model 3 for the Veni data. Field SGW is the reference field, and Female is the reference gender. Note that p-values haven't been adjusted for multiple testing (a model for each of the three tiers) yet.

		Model 1		Model 2		Model 3		Model 4	
Year	Field	Women	Men	Women	Men	Women	Men	Women	Men
2012	SGW	0.134	0.118	0.109	0.139	0.103	0.148	0.114	0.163
2013	SGW	0.133	0.117	0.113	0.132	0.107	0.142	0.115	0.152
2014	SGW	0.132	0.116	0.118	0.126	0.111	0.135	0.117	0.142
2015	SGW	0.130	0.115	0.122	0.120	0.115	0.129	0.118	0.132
2016	SGW	0.129	0.113	0.127	0.114	0.119	0.123	0.120	0.123
2017	SGW	0.128	0.112	0.131	0.108	0.123	0.117	0.121	0.114
2018	SGW	0.127	0.111	0.136	0.103	0.127	0.112	0.123	0.106
2019	SGW	0.126	0.110	0.141	0.098	0.132	0.107	0.124	0.098
2020	SGW	0.125	0.109	0.146	0.093	0.136	0.102	0.126	0.091
2021	SGW	0.123	0.108	0.152	0.089	0.141	0.097	0.127	0.084
2012	DO	0.183	0.162	0.151	0.189	0.170	0.163	0.159	0.156
2013	DO	0.181	0.160	0.156	0.181	0.176	0.156	0.168	0.151
2014	DO	0.180	0.159	0.162	0.172	0.182	0.149	0.176	0.146
2015	DO	0.178	0.157	0.168	0.165	0.188	0.142	0.185	0.142
2016	DO	0.176	0.156	0.173	0.157	0.194	0.135	0.195	0.137
2017	DO	0.175	0.154	0.180	0.150	0.200	0.129	0.204	0.133
2018	DO	0.173	0.153	0.186	0.143	0.206	0.123	0.214	0.129
2019	DO	0.172	0.152	0.192	0.136	0.213	0.118	0.225	0.125
2020	DO	0.170	0.150	0.199	0.129	0.220	0.112	0.236	0.121
2021	DO	0.169	0.149	0.206	0.123	0.227	0.107	0.247	0.117
2012	ENW	0.218	0.194	0.182	0.226	0.205	0.209	0.193	0.201
2013	ENW	0.217	0.192	0.188	0.216	0.211	0.200	0.202	0.195
2014	ENW	0.215	0.191	0.194	0.207	0.218	0.192	0.211	0.188
2015	ENW	0.213	0.189	0.201	0.198	0.225	0.184	0.220	0.181
2016	ENW	0.211	0.187	0.208	0.189	0.232	0.176	0.229	0.175
2017	ENW	0.209	0.186	0.215	0.180	0.239	0.168	0.239	0.169
2018	ENW	0.208	0.184	0.222	0.172	0.246	0.161	0.249	0.163
2019	ENW	0.206	0.183	0.229	0.164	0.254	0.153	0.260	0.157
2020	ENW	0.204	0.181	0.237	0.157	0.262	0.147	0.270	0.152
2021	ENW	0.203	0.180	0.245	0.150	0.269	0.140	0.281	0.146

Table 5

Predicted success probabilities, according to the four models [1/2]

		Model 1		Model 2		Model 3		Model 4	
Year	Field	Women	Men	Women	Men	Women	Men	Women	Men
2012	TTW	0.148	0.131	0.121	0.153	0.142	0.141	0.119	0.123
2013	TTW	0.147	0.129	0.126	0.146	0.147	0.135	0.128	0.121
2014	TTW	0.146	0.128	0.130	0.139	0.152	0.129	0.138	0.120
2015	TTW	0.144	0.127	0.135	0.133	0.158	0.123	0.148	0.118
2016	TTW	0.143	0.126	0.140	0.126	0.163	0.117	0.158	0.116
2017	TTW	0.142	0.125	0.145	0.120	0.168	0.112	0.170	0.115
2018	TTW	0.140	0.123	0.151	0.114	0.174	0.106	0.182	0.113
2019	TTW	0.139	0.122	0.156	0.109	0.180	0.101	0.194	0.112
2020	TTW	0.138	0.121	0.162	0.104	0.186	0.097	0.207	0.110
2021	TTW	0.137	0.120	0.167	0.098	0.192	0.092	0.221	0.109
2012	ZonMw	0.150	0.132	0.122	0.155	0.107	0.178	0.105	0.176
2013	ZonMw	0.148	0.130	0.127	0.147	0.111	0.170	0.109	0.169
2014	ZonMw	0.147	0.129	0.132	0.141	0.115	0.163	0.114	0.162
2015	ZonMw	0.146	0.128	0.136	0.134	0.119	0.155	0.118	0.155
2016	ZonMw	0.144	0.127	0.141	0.128	0.123	0.148	0.123	0.148
2017	ZonMw	0.143	0.126	0.147	0.121	0.128	0.142	0.128	0.142
2018	ZonMw	0.142	0.125	0.152	0.116	0.132	0.135	0.134	0.136
2019	ZonMw	0.140	0.123	0.158	0.110	0.137	0.129	0.139	0.130
2020	ZonMw	0.139	0.122	0.163	0.105	0.142	0.123	0.145	0.124
2021	ZonMw	0.138	0.121	0.169	0.099	0.147	0.117	0.150	0.119

Table 6

Predicted success probabilities, according to the four models. [2/2]

	df	AIC Vidi	AIC Vici
Model 1	7	339.10	276.71
Model 2	8	340.67	278.28
Model 3	12	347.03	284.81
Model 4	16	351.87	283.86

Table 7

Comparison between the four models using the Akaike Information Criterion for the Vidi and Vici data

	Vidi			Vici		
	df	χ^2	p-value	df	χ^2	p-value
Gender	1	.500	.480	1	1.144	.285
Field	4	44.660	$< .001$	4	9.000	.061
Year	1	.141	.708	1	4.463	.035

Table 8

ANOVA tables for the Vidi and Vici data. The p-values haven't been adjusted yet triple multiple testing.

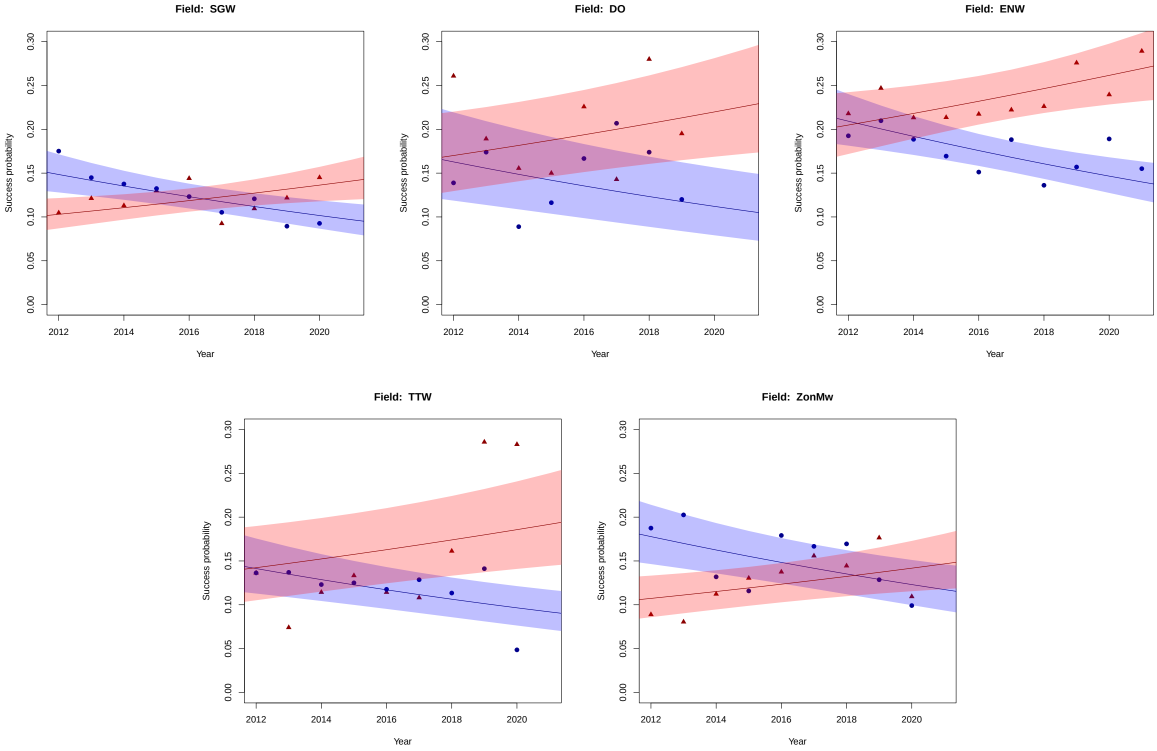

Figure 1. Observed success probabilities (triangles for women, squares for men) and predictions according to Model 3 (increasing curves for women, decreasing curves for men) for the Veni data. The shaded areas correspond to the 95% prediction intervals.

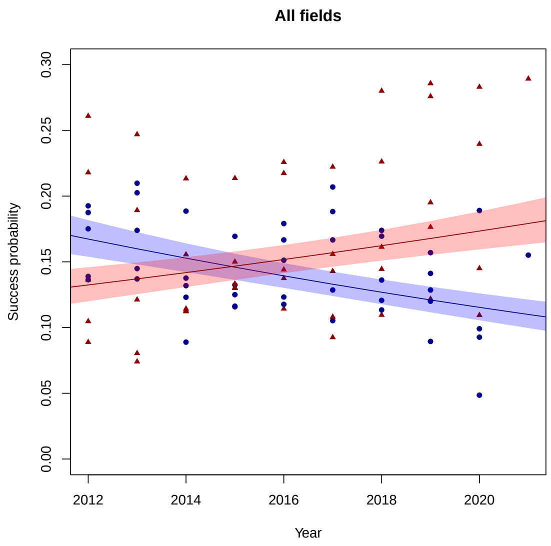

Figure 2. Observed success probabilities and predictions according to Model 3 for the Veni data, aggregated over all five fields. For an explanation of the symbols and colours, see the caption of Figure 1.

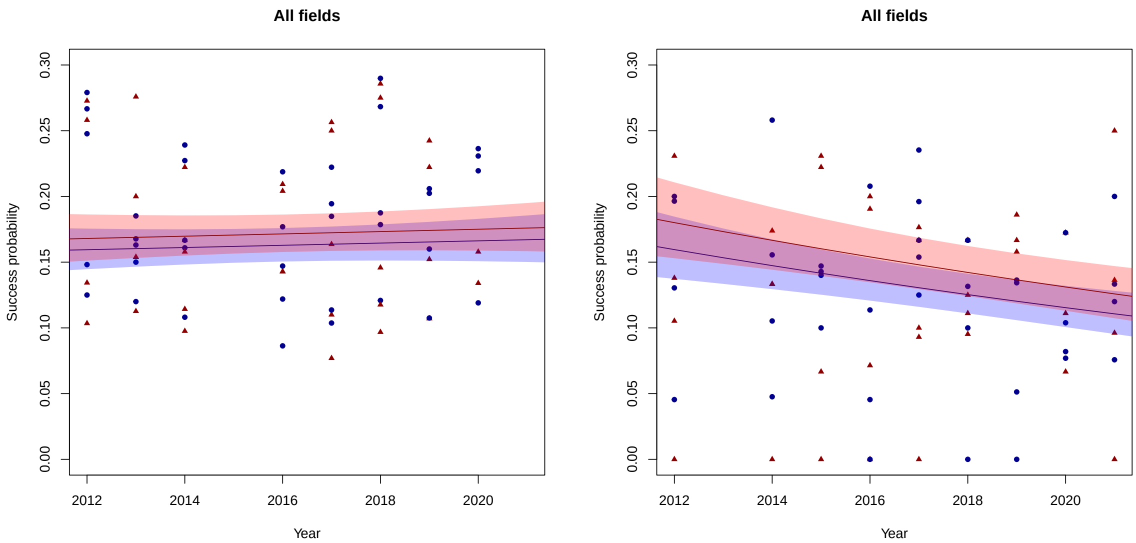

Figure 3. Observed success probabilities and predictions according to Model 1 for the Vidi (left panel) and Vici (right panel) data, aggregated over the five domains. For an explanation of the symbols and colours, see the caption of Figure 1.

Appendix A

Interpretation of the logistic regression coefficients

Although, as always in logistic models, the parameters in Table 4 cannot be directly interpreted, their sign and p-value can. To show how to use the numbers in Table 4, consider the following examples of a female and male applicant to the field ENW in 2020. For the female applicant, we have

$$\log\left(\frac{p_i}{1-p_i}\right) = -2.204 + 0.808 + 0.040 \times 9 = -1.038,$$

which corresponds to a success probability of 26.2%. For the male applicant we have

$$\log\left(\frac{p_i}{1-p_i}\right) = -2.204 + 0.512 + 0.808 + (0.040 - 0.094 - 0.391) \times 9 = -1.763,$$

corresponding to a success probability of 14.6%. For 2012, this domain had more balanced predicted success rates (20.5% for women, 20.9% for men; see Table 6).

Appendix B

Analysis code

```
venidata <- read.csv("venistats.csv", sep=";")
venidata$Field <- relevel(factor(venidata$Field), "SGW")
themodel1 <- glm(cbind(Granted, Applications - Granted) ~
    Gender + Field + Year, data = venidata,
    family = "binomial")
themodel2 <- glm(cbind(Granted, Applications - Granted) ~
    Gender*Year + Field , data = venidata,
    family = "binomial")
themodel3 <- glm(cbind(Granted, Applications - Granted) ~
    Gender*Year + Gender * Field , data = venidata,
    family = "binomial")
themodel4 <- glm(cbind(Granted, Applications - Granted) ~
    Gender*Year + Gender*Field + Year*Field, data = venidata,
    family = "binomial")

library("car")
Anova(themodel1, type = "III",test.statistic = "Wald")
Anova(themodel2, type = "III",test.statistic = "Wald")
Anova(themodel3, type = "III",test.statistic = "Wald")
Anova(themodel4, type = "III",test.statistic = "Wald")
AIC(themodel1,themodel2,themodel3,themodel4)
```