I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13372

Evidence for a Two-Women Quota in University Departments across Disciplines

Lena Janys

JUNE 2020

DISCUSSION PAPER SERIES

IZA DP No. 13372

Evidence for a Two-Women Quota in University Departments across Disciplines

Lena Janys
University of Bonn, Hausdorff Center of Mathematics and IZA

JUNE 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Evidence for a Two-Women Quota in University Departments across Disciplines*

Increasing diversity in higher tiers of occupations is a strongly debated topic and subject to legislation and reform in professional organizations in many countries. I use a novel method for detecting implicit quotas in workplaces, college admissions or birth patterns, relying exclusively on the distribution of different demographic types across different workplace locations, colleges or families. I apply this method to current employment of female professors at German universities across 50 different disciplines and show that the distribution of women, given the average number of women in the respective field, is unlikely to result from a random allocation of women across departments and more likely to stem from an implicit quota of one or two women on the department level. I also show that a large part of the variation in the share of women across STEM and non-STEM disciplines could be explained by a strict two-women quota on the department level. These findings have important implications for the potential effectiveness of policies aiming at reducing underrepresentation, as well has providing evidence how stakeholders perceive and evaluate diversity.

JEL Classification: J71, C15

Keywords: diversity, tokenism, gender, simulation methods

Corresponding author:
Lena Janys
University of Bonn
IFS-Institute of Finance and Statistics
Adenauerallee 24-42
53113 Bonn
Germany

E-mail: ljanys@uni-bonn.de

* I thank Guido Friebel, Daniela Grunow, Joachim Freyberger, Pia Pinger, Hans-Martin von Gaudecker, Christina Felfe, Amelie Schiprowski, Astrid Kunze, Shelly Lundberg, Katie Coffman and Lingwei Wu, as well as participants of the workshop "The Determinants of Careers for Women and Men: Institutions, Professions, Organizations and Households" for helpful comments and suggestions. The beta-version of the R-package to calculate all quantities is available on my GitHub-page https://github.com/LJanys/Implicit-quota-package.

1 Introduction

Increasing the representation of women and other underrepresented minorities in high-level professions, such as managers, board members, politicians or academics has been a declared goal of both political institutions, such as the Europe 2020 Strategy[1], as well as professional organizations (e.g. in economics Auriol et al. (2020)) and banks (e.g. Goldman and Sachs).[2]

There exists an active literature analyzing *explicit* quotas.[3] However, evidence on implicit quotas is extremely sparse, with the notable exception of Chang et al. (2019), who study female representation in boards. Explicit quotas are typically aimed at increasing representation with an official target, e.g. 1/3 of board seats should be filled by a female candidate. Implicit quotas, however, are not well understood. Implicit quotas can stem from discrimination on an institutional level (as was demonstrated in the case of Students for fair admissions v. Harvard college (2019) and Arcidiacono et al. (2020)[4]) or because of (numerically unspecified) pressure from funding agencies, administrations or the public. Analyzing implicit quotas can be difficult because of the problems outlined in Arcidiacono et al. (2020): First, in every single hiring or admission decision, there are many factors that can be correlated with the observable characteristic of interest and therefore analyzing individual decisions requires a wealth of individual-level (and possibly unobservable) information to mitigate the omitted variable bias. Second, because the quota itself is not known, it is a question of "where to look": the probability of hiring/admitting a person from an underrepresented minority is not *uniformly* lower given the same observable characteristics, but depends on others from the same group already in the pool. The potential spill-over effects of implicit quotas might also be different than for explicit quotas, which means that results from that literature might not carry over to implicit quotas that originate from unspecified "diversity pressure". It is possible, that in the absence of explicit quotas, entities such as university departments that are under pressure from funding agencies and the administration, hire enough women to conform to discipline standards, but are reluctant to go beyond that. In that scenario, unspecified pressure leads to the preservation of the status quo after an initial push, but with a sub-optimal distribution. This is compatible with concerns that progress in creating diversity is "stalled" or sluggish (as Lundberg and Stearns (2019) document for the academic economics profession in the US).

This paper is based on the idea that observing the same general hiring/admission circumstances in

[1]The Europe 2020 Strategy proposes a law that would increase female representation on boards of listed companies to 40 percent , see the reference number of the document SOC/471-EESC-2012-1734.

[2]"Starting on July 1st in the U.S. and Europe, we are not going to take a company public unless there is at least one diverse board candidate, with a focus on women," Solomon[CEO of Goldman and Sachs] said on CNBCs "Squawk Box" from the World Economic Forum in Davos, Switzerland.

[3]Explicit quotas have, for example, been analyzed in the context of company boards or election lists, see for example Bertrand et al. (2018) evaluating a Norwegian board reform and Maida and Weber (2019) on an Italian board reform and O'Connell (2018) use a mandated 1/3 gender quota on election list in India, to analyze whether this increases participation in legislatures.

[4]American students of Asian descent sued Harvard college for discriminatory admission policies, i.e. they claimed that Harvard college used implicit quotas to artificially limit the number of students from this demographic group

different locations can give us information about quantity restrictions (e.g. discrimination or implicit quotas). With only aggregate information on the minority status I can show the presence of implicit quotas. Using a bootstrap method, I can show that the observed implicit quota is unlikely due to chance but rather structural in nature. Research on the underrepresentation of certain demographic groups has typically focused on explaining the share of the group p. However, none of these explanations can be used to explain the distribution of the members of the underrepresented group across different sub-units.

The German university system is particularly well suited for analyzing implicit quotas: it is a big, rather homogeneous and modern university system, in which most universities offer a large range of study disciplines (rather than specializing in certain disciplines). Employment conditions across universities for professors are comparable nation-wide and there exists good-quality administrative data containing information on all university professors, including rank, gender and discipline.

The existence of implicit quotas has broad implications for the analysis of the reasons for underrepresentation, as well as for the effectiveness of policy measures that are aimed at improving diversity in organizations and professions. The existence of implicit quotas also points to some interesting implications for the behavioral mechanisms that determine how people perceive and evaluate diversity in organizations.

The tool I propose can be applied to different contexts in which groups are underrepresented, if hiring/appointment decisions are aggregated in sub-experiments. Examples include admissions to elite colleges, elections of local government entities or even comparing groups of employees at comparable levels across a firm's different locations. My method makes no assumptions about the underlying hiring process and reasons why women might be underrepresented and takes the overall share of the underrepresented minority in a profession or admissions pool as given. Studying direct discrimination, i.e., quantity restriction in one direction or another is difficult, because each hiring decision is multidimensional and therefore difficult to analyze[5]. However, given that there are multiple universities that offer the same discipline, we can see every *department* as a binomial experiment. Taking the overall share of women as given, I can examine the variation in the number of women *across* departments but within the same discipline.

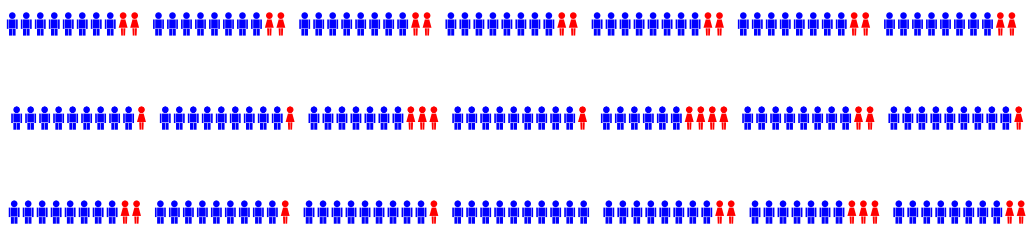
My results show that (1) the current allocation of women across departments is highly improbable in a world where candidates are drawn "gender blind" from a binomial distribution with the share of female professors p. Specifically, across disciplines there are "too few" departments with no female professors and "too many" with one or two female professors.[6] (2) These deviations from the expected value are not driven by individual disciplines. (3) I show in bootstrap simulations that the results are highly unlikely to be due to chance. (4) I find no evidence that fields that are supposedly more accepting of women (as proxied by

[5]In a different context Bharadwaj et al. (2014) discuss these challenges for the case of the "son-stopping rule".

[6]Chang et al. (2019) coin the term "twokenism" for this type implicit quota

the overall share of women in the discipline) are different in their implicit quotas and can show that the distribution of female shares across *disciplines* could be well-explained by a two-women quota per department in all disciplines.

Example 1. *Consider three different scenarios with seven sub-units (for example, university departments). The red icons symbolize female faculty. In all three cases the probability of hiring a woman is $p = 0.2$; however, the distribution of women in the first scenario seems artificially uniform and is thus less likely to be associated with a random allocation of women across departments and more likely to be the result of manipulation*[7]*. I formalize this notion Section (2).*



In the presence of implicit quotas, it is doubtful whether analyzing observable productivity indicators is necessarily informative because: (1) match quality might be worse for women, as the probability of being hired at a particular institution relies on factors outside their control (i.e., is there a woman already?). (2) if women are more likely to be hired when there are no women but less likely to be hired when there are already two, then the characteristics of the third woman in a department will be very different from the characteristics of the first one and analyses such as Huang et al. (2020) who analyze productivity patterns of scientists in different disciplines over time, will mask important heterogeneity.[8]

In my data application I can distinguish between the distribution being centered around the *mean* and being centered around the *mode* (due to the variability of department sizes), which can give us additional insights into how people perceive diversity. If implicit quotas are the result of a shift in social pressure (either by the university administration or funding agencies), it is unclear whether the "acceptable number" of women will increase over time, or whether there will be stagnation, since there is no explicit target. More worryingly, there is evidence that implicit quotas can be self-reinforcing: in a recent experimental paper Paryavi et al. (2019) analyze how descriptive norms in gender composition are acted upon by men and find that descriptive norms in their set-up do not lead to prescriptive norms and can even lead to backlash if male "employers" are informed that others have hired more women. The presence of implicit quotas can

[7] The latter two scenarios were drawn from a binomial distribution with $p = 0.2$ in R with *set.seed(400)*.

[8] In fact, one of their key findings, that research productivity differences have increased over time could be at least partially explained by an increase in quality heterogeneity.

also be informative about how employers, the public or other stakeholders perceive diversity. Additionally, implicit quotas can tell us how random processes are most likely perceived by the public, in the sense that large deviations from the mode are more likely to be interpreted as manipulations than small ones, even though tail events such as zero or seven women are relatively likely to occur as long as n is large enough.

Examining the distribution for evidence of implicit quotas is important for three reasons: First, if for a profession or discipline s the probability of hiring a woman p_s is a result of an implicit quota rather than a market solution, trying to explain p_s with either personal attributes of women or with implicit biases will not be informative, because any correlations that are found would be spurious to some degree.[9] Second, evidence for quantity discrimination has implications for the policy recommendations resulting from the analyses outlined above: proposed policies such as improved access to childcare, etc., can only ever affect p (in the best case), but will not be effective in the face of direct discrimination on the institutional level.

Direct discrimination in the form of implicit quotas, gender differences in personal attributes, preferences and implicit biases are not mutually exclusive and the presence of one does not preclude the presence of another. There is even a plausible link between biases and direct discrimination; as Goldin (2014) argues, if men perceive women as less qualified and care about the prestige of their profession, they may fear that a woman's entering a profession signals a change in the prestige of the profession and therefore might block entry to otherwise qualified women. Third, analyzing attributes of women and men already *within* the profession under direct discrimination will not yield a meaningful comparison as this type of discrimination may affect match quality for women: if the probability of being hired is high when there are no women and low when there already are women means that women (on average) will be matched to departments that have no women, not to the departments where they may be most productive, and there will be considerable heterogeneity in scholarly output, depending on how many women were already in the department during hiring.[10]

1.1 Previous Literature

The lack of women at higher tiers of professions, including academia, is often described as a "leaky pipeline" (for a first usage see Alper and Gibbons (1993)), referring to the fact that at each stage of the academic ladder (undergraduate studies, graduate studies, postdoc, tenure track professor) women leave academia.

[9] Consider the extreme case that every department has a strict two women quota: then there will be variation in the share of women across disciplines and these differences will presumably be correlated with something idiosyncratic to the discipline, but the correlation with the share of women will be entirely due to omitted variable bias.

[10] Lutter and Schröder (2016) analyze CV's for Sociology professors and find that women have fewer publications when they receive their first professorship, Madison and Fahlman (2020) analyze publication and citation patterns in Sweden and find that men had more publications and citations in both medicine and in the social sciences, and conclude that women are not held to a higher scholarly standard. Most recently Huang et al. (2020) analyze productivity patterns of male and female researchers using longitudinal data covering different disciplines.

This phenomenon and its causes have been discussed for a long time[11] and an interesting puzzle has emerged: Williams and Ceci (2015) show in a large experiment that using randomized CV's, women are *more* likely to be hired than men in a hypothetical university hiring situation; however, the number of female professors (especially on a senior level) has been rising only slowly, if at all. For example, Lundberg and Stearns (2019) write that "More recently, this growth in female representation in the economics discipline has stalled. The share of female assistant professors and PhD students has remained roughly constant since the mid-2000s". Many possible explanations for the current underrepresentation of women are being investigated in many different disciplines from Economics to Psychology. This research can broadly be categorized into three strands (1) inherent or socialized differences in abilities or preferences, (2) structural discrimination in the form of conscious or unconscious biases and (3) institutional barriers such as lack of childcare or within-family division of labor and lack of female role models. What these explanations have in common is that they all affect the overall probability p of hiring a woman for a particular job, which is in expectation equal to the overall share of women in any given (level of a) profession. None of these explanations, however, affect the *distribution* of women across different sub-units, such as different locations of the same firm or, in the case of this application, the distribution of female university professors across departments within the same discipline. As an example, consider the case of implicit bias, where a women's qualifications are systematically undervalued by 20%. We can interpret this as for any objective *qualification=1*, women are only credited with 0.8: this would *on average* lead to an underrepresentation of women in the entire profession (because the probability of hiring a women is reduced by the biased perception of quality). There is no reason that once the hiring decision is made, this will systematically bias the *perception* of quality in the next female hire. As a result, any evidence that the distribution of women is non-random can only be explained by discrimination, sorting or a combination of the two.

2 Method

We can formalize the method described above, by thinking over each employment decision $x \in X$ as a separate Bernoulli experiment with associated success probability $X \sim \mathcal{B}(p)$ where $x = 1$ means hiring a woman. This success probability corresponds to the average share of women in the profession or discipline. In that sense, every employment decision is like rolling a die with the outcome "woman is hired" occurring with probability p[12]. This allows us to think about the distribution of female faculty within disciplines but

[11] For example in the paper"Women in science: why so few?" Rossi (1965) and "Women in academe?" Graham (1970).

[12] One obvious objection is that hiring decisions are not truly independent, because once a female faculty member is hired, she is no longer in the hiring pool. However, it is sufficient that the own dependence is not too binding (as would be the case if there were only very few women in the pool, which is not the case in this application). Generally, a woman being hired at one institution, does not preclude another institution from recruiting her. Additionally, unlike Chang et al. (2019) I do not impose that the random process allocates the precise number of women in each discipline, which accounts for the fact that departments

across departments as a Binomial experiment: We denote the vector of realizations for one discipline s as a random vector $\mathbf{x}_s \sim \mathcal{B}(1, k_s, p_s)$, where k is a vector of length "number of departments", containing the sizes of the departments. We can now estimate the $\hat{p}_s$ by calculating the sample average over all departments. To see whether the *distribution* of women across different departments (within a discipline) is consistent with a random allocation, we can analyze the frequency f_s^z of women within each discipline across departments. f_s^z denotes here the frequency, i.e. if $z = 1$, then f_s^1 is the number of departments (within discipline s) that have exactly one female faculty member, f_s^2 is the number of departments that have exactly two female faculty members, etc.

The general method is therefore the following:

1. For each discipline, calculate the number of departments with exactly z women, for $z = 0, ..., 10$..
2. Generate a vector with the real department sizes as the individual k_s.
3. Simulate a distribution of female faculty within each discipline from a binomial distribution $\hat{p}_s$ with $sim = 1000$ simulation draws and average over the the simulated frequencies.
4. Compare the actual frequencies f_s^z of female faculty with the simulated number of female faculty as described above.

Example 2. *Suppose that for discipline s there are 20 different departments of different sizes, such that the first five elements of k_s might be $k = \begin{pmatrix} 20 & 23 & 30 & 34 & 32 & 36 & \dots \end{pmatrix}$. Suppose that the share of women overall is $p = 0.2$. One realization (drawn from a binomial) could be, $\mathbf{x}_s = \begin{pmatrix} 6 & 6 & 10 & 9 & 11 & \dots \end{pmatrix}$, i.e., in department 1, there are 20 faculty members and six of those are women, in department 2 there are 23 faculty members of which also 6 are women, etc. If we repeat this for different disciplines s with (potentially) different p_s, k_s, then the excess number of women shares should average out over the different disciplines. If the empirically observed distribution does not average out and/or if the share of departments which deviate in a positive or negative direction is significantly larger than 0.5, this can be interpreted as evidence against a random allocation and for manipulation.*

Testable predictions If the distribution of female faculty across departments is random and depends only on the overall probability p_s, we can make two testable predictions about the deviations of the observed frequencies f_s^z from their expected value:

could higher outside the existing pool.

1. for each z,

$$d_f^z = \sum_{s=1}^{S} f_s^z - \mathbb{E}\{f_s^z\} = 0, \tag{1}$$

i.e. the sum of the deviations over all disciplines should average out to zero.

2.

$$\mathbb{E}\left\{\frac{\left(\sum_{s=1}^{S} I(d_f^z > 0)\right)}{S}\right\} = 0.5, \text{ or equivalently } \mathbb{E}\left\{\frac{\left(\sum_{s=1}^{S} I(d_f^z < 0)\right)}{S}\right\} = 0.5, \tag{2}$$

where $I(\cdot)$ is an indicator function. This prediction tells us that the share of disciplines that have a positive/negative deviation from the expected value should be equal to 0.5.

3 Data

3.1 Data sources

I examine the distribution of female faculty in all German universities using administrative data from the German Federal Statistical Office (DeStatis) in 2015. This data contains information on all non-administrative employees at German universities.[13] For professors, it contains, among other things, information on the discipline, their gender, their salary category, their university, the year of their first appointment as a professor and their year of birth.

3.2 Definitions

University I only consider universities and not universities of applied sciences, because these do not overlap in their hiring pool and have different populations of students and working conditions.[14]

Faculty I restrict the sample to professors according to the administrative classification, both tenured and untenured, including assistant- associate and full professor level, but not substitute professors. I do not condition on civil servant status or full time employment, although neither of these would change the sample in a significant way.

[13] "Personalstatistik: Wissenschaftliches Personal"

[14] Note that the official statistics that are published on the official materials by DeStatis do not contain information on university status, so I coded status as a university myself. This was unproblematic, since the real names of the universities are included in the data and status is unambiguous, but it does mean that some officially published statistics are not directly comparable.

Discipline A discipline is defined according to the administrative three-digit classification of the statistical office.[15] The list of disciplines is given in Table (2) below. Comparing this to the official code book, there are some disciplines where the minimum threshhold for a department (see below) were not met and thus were not included.

Department There is no official designation of a department in the administrative statistics; therefore, I define a department as all faculty that are employed within a discipline at the same university. This leaves some room for error, as it is conceivable that, for example, a chemistry professor is employed in the physics department and thus in a different organizational unit. However, I argue that the hiring pool, and thus the p_s is more similar *within* discipline than within department in that case. To count as a department, it must have at least three professors.

The final sample contains $N = 1737$ departments across 50 disciplines.

4 Application

4.1 Descriptives

The upper panel in Figure (1) shows the mean share of female professors by discipline. The horizontal line marks 0.5, emphasizing the fact that there is no discipline where female professors are in the majority (STEM disciplines are highlighted). Perhaps unsurprisingly, STEM disciplines populate the lower end of the female share distribution, while humanities have a more balanced gender ratio. Of those professors employed in 2015, 50% of male employed professors were appointed in 2004 or after, while 50% of currently employed female professors were appointed in 2008 or after.

Average department size The bottom panel in Figure (1) shows the average department size, ordered by discipline. STEM disciplines are highlighted. Contrasting this with the top panel from Figure (1) another thing becomes apparent: department sizes in the humanities, where the average share of women is much higher, are much smaller than in STEM disciplines, with the social sciences in between. This is broadly consistent across countries, e.g. in 2010 the average physics department employed 29.2 full-time faculty vs. history, which had an average department size of 16.5 in 2007 among PhD granting institutions.[16]

[15] "Schlüsselverzeichnisse fur die Personalstatistik", which is also used to define disciplines used for allocating third party funding at the German Research Foundation (DFG).

[16] see White et al. (2012) and Townsend (2010)

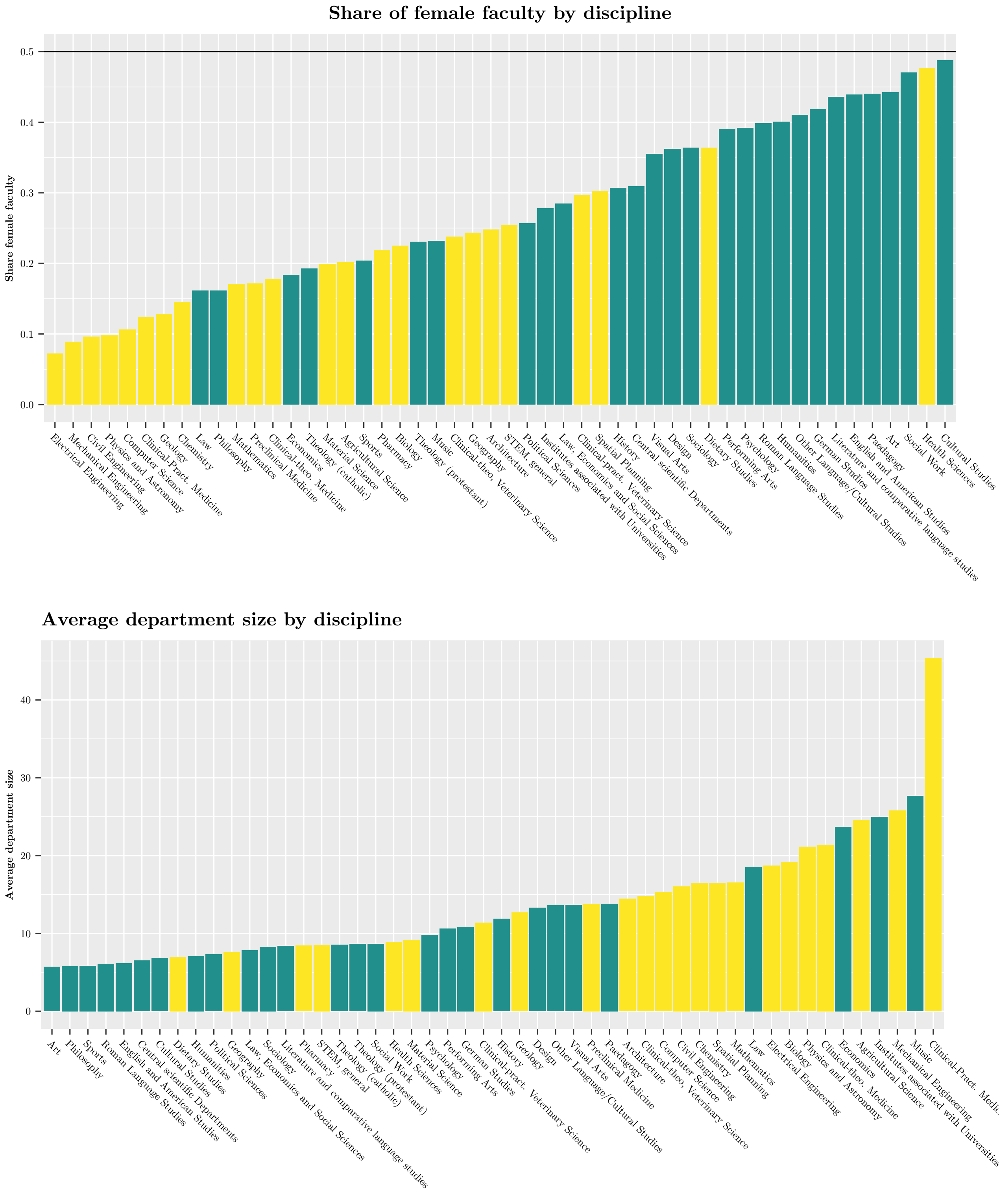


Figure 1: **Top Panel**: Share of female faculty by subject, ordered, yellow bars indicate STEM fields. The correlation between the average size of the departments and the share of female faculty is -0.51 (with 95% confidence interval (-0.69 -0.28), different from zero with p-value= 0.00013), indicating that the smaller the average department, the larger the share of women within the subject. **Bottom panel**: Average department size by discipline, ordered, yellow bars indicate STEM fields. The correlation between the average size of the departments and the share of female faculty is -0.51 (with 95% confidence interval (-0.69 -0.28), different from zero with p-value= 0.00013), indicating that the smaller the average department, the larger the share of women within the subject.

4.2 Results

In the following I present the analysis on the the deviations from the expected value as outlined above, both overall and by discipline.

4.3 Main results

The top panel of Figure (2) shows the average deviation of the number of departments with exactly z women, for $z = 0, 1, ..., 10$, summed over all disciplines. Negative deviations are highlighted in purple. According to Prediction (1), the expected deviation is equal to zero. However, there are, on average, 0.4 too few departments with exactly zero women, as well as too few departments with three, four, five and seven women. On the other hand, there are too many departments with exactly one or two women. To see to what extent this is driven by some extreme outliers in one or two disciplines, in the top panel of Figure (3) I weight the deviations by the absolute value deviation of the share of departments with a positive or negative deviation from the expected value. It is zero if either the overall deviation is close to zero, and/or the share of departments with either a positive or a negative deviation from zero is close to 0.5. According to Prediction (2), we expect about half of the disciplines to be above and half to be below the expected value, meaning that any systematic deviation in terms of the sign of the deviation points to a structural issue in academia as a whole. In the case of zero women, the share of departments with a negative deviation from expected value is 0.72, which means I weight the estimate 0.42 by $|0.5 - 0.72| = 0.22$ etc. The results do not differ qualitatively, however I consider this measure to be more robust against outliers and therefore more indicative of systemic issues because is combines both predictions in one measure.

4.4 Uncertainty

In order to quantify the uncertainty surrounding the aggregate estimates in the upper panels of Figure (2 and 3) I perform a bootstrap procedure in which I first draw a random value from a binomial distribution, with the p_s and k_s supplied from the data and treat the resulting deviations as the real data, repeating the steps outlined in Section (2). I repeat this procedure 100 times to obtain a 90% interval for each individual observed deviation. The results are displayed in the bottom panel of Figures (2) and (3) for the unweighted estimates and the weighted estimates, respectively. The 90 % interval is displayed as blue triangles, along with the individual bootstrapped observations (in grey diamonds) and the actual observations (purple dots). For both the unweighted and the weighted estimates, the actual observations for the deviations from zero, two and three women departments lie outside the bootstrapped 90% interval. In the weighted version the actual observations lie outside the bootstrapped data altogether, indicating that this distribution of deviations is highly unlikely to be the result of a random allocation. It also highlights that the weighted measure is a more meaningful metric, since random data may have large individual, discipline- specific deviations, but it is less likely that these occur over many disciplines in the same draw.

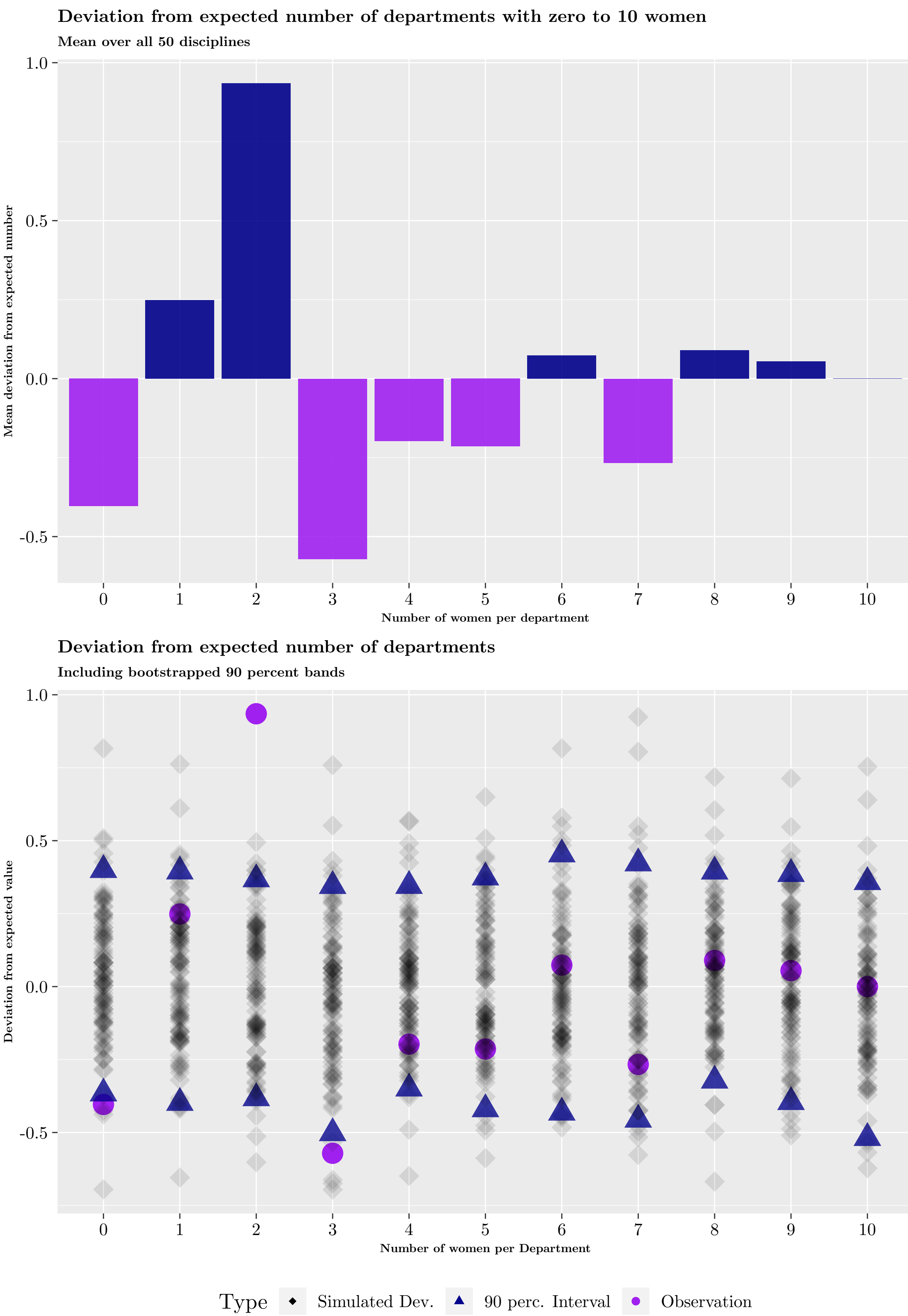


Figure 2: **Top panel**: Deviation from expected value summed over all disciplines. **Bottom panel**: Deviation from expected value summed over all disciplines, including bootstrapped 90% intervals.

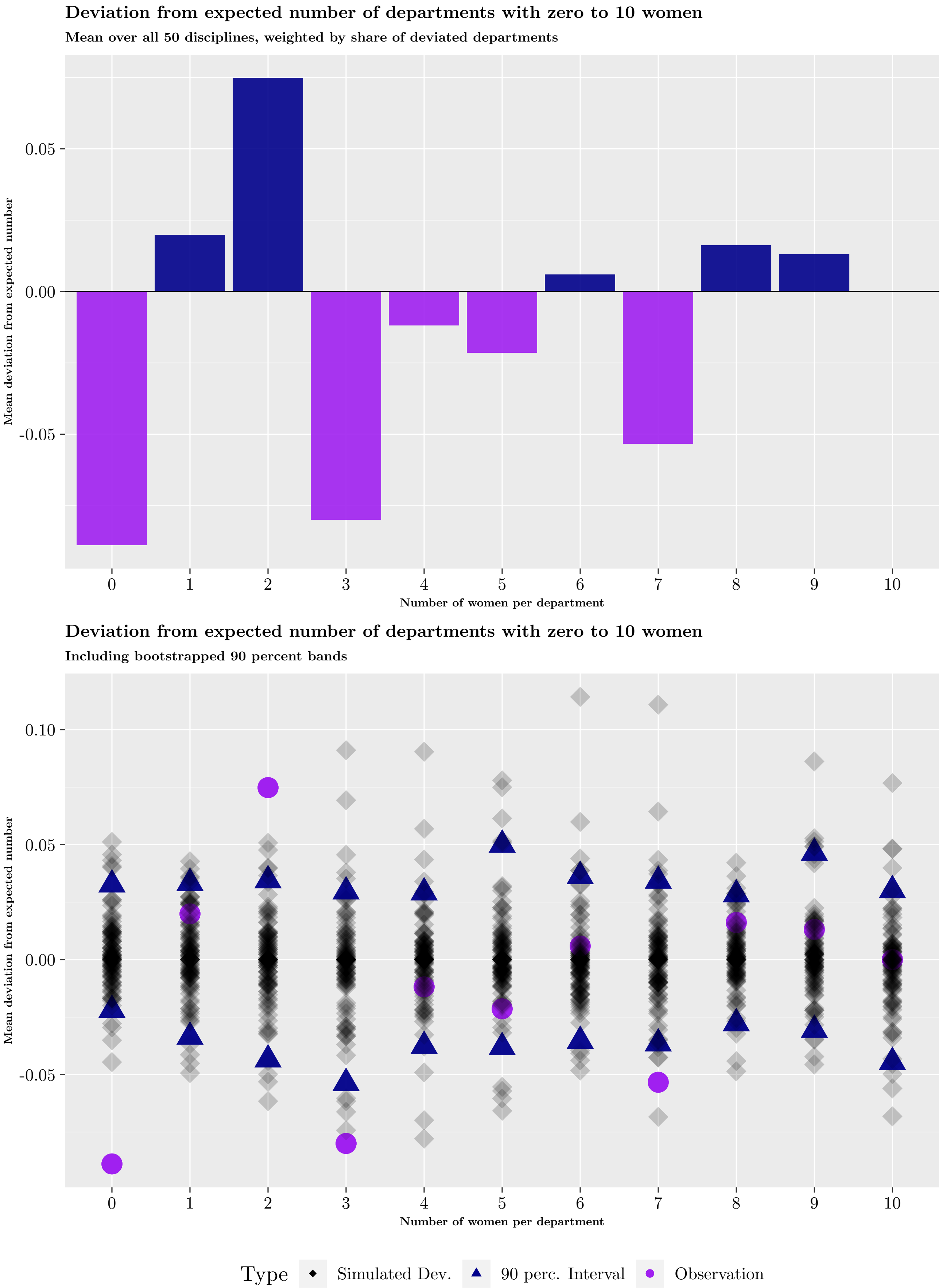


Figure 3: **Top panel**: Deviation from expected value times the deviation from 0.5 (in absolute terms). **Bottom panel**: Deviation from expected value times the deviation from 0.5 (in absolute terms), including bootstrapped 90% intervals.

4.5 Results by discipline

From the top left to the bottom-right panel, Figure (4) shows the deviation from the expected value of exactly zero–four women departments by discipline (economics is highlighted for reference) the vertical line indicates 0.5. We would expect about 25 disciplines with negative deviations and 25 disciplines with positive deviations (Prediction (1)) and that the mass below and above zero should be roughly equal (Prediction (2)). It is clear, however, that there are both too many disciplines with fewer than expected zero women departments and that the mass below zero is larger than the mass above zero (as seen in the main results in section 4.3). The opposite is true for exactly one and two women departments (top-panel right and bottom panel left): there are too many departments with exactly one or two women and the mass of departments is unequally distributed and this pattern is then again reversed when considering three women per department.

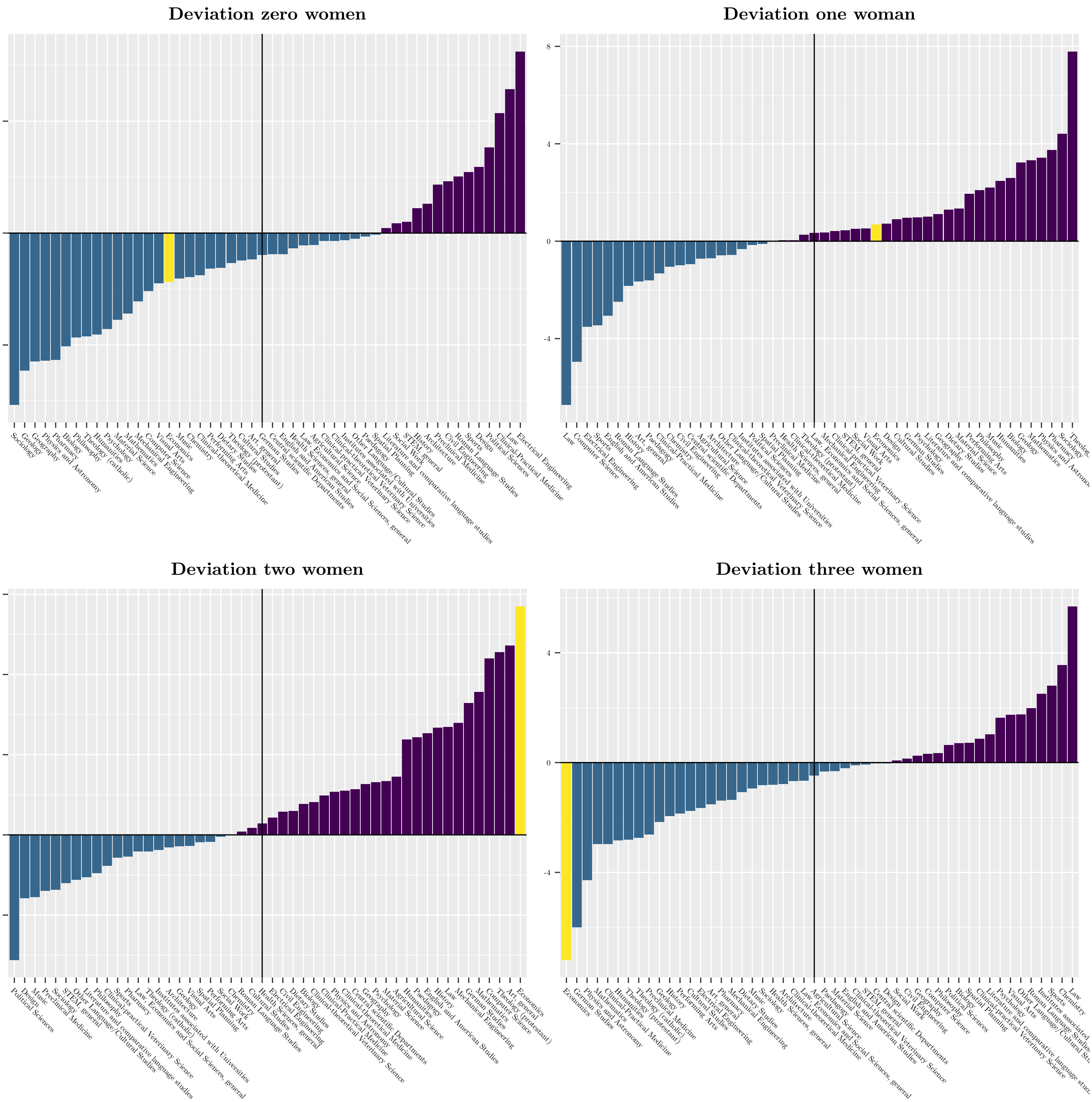


Figure 4: From top left to bottom right: The deviation from zero to three women departments, by discipline. Economics is highlighted for reference.

4.6 Effects of a fictional quota

We might hypothesize that subjects that have a lower share of female professors provide a more hostile environment and therefore are more likely to discriminate against women. Visually inspecting Figure (4) does not, however, reveal a clear pattern, and the correlation ρ_z between having a negative deviation from the expected number of zero women departments and the overall share of women is $\rho_0 = -0.035$ with a p-value of 0.86 and for $\rho_3 = 0.07$ with a p-value of 0.62. This result might seem puzzling; however, re-examining Figure (1), it becomes clear that implementing a quota would have heterogenous effects on the share of women within discipline, depending on average department size. Most of the non-STEM fields (with the notable exceptions economics and law) which have the largest share of female faculty have average department sizes below ten professors: this implies that a two women quota would mechanically lead to a much higher female share in these disciplines than in mathematics or economics, where the average department size is much larger. While I cannot explain the entirety of the variation in female shares with a quota and department sizes, we cannot rule out that non-STEM disciplines in the humanities have a larger share of female professors because an implicit quota has a heterogenous effect depending on department sizes, rather than discipline specific characteristics (such as how math-intensive these disciplines are). To illustrate this point, consider Figure (5): here I simulate the hypothetical shares of women within subjects that would result from taking the real department sizes k_s and imposing a two women quota per department, as shown for an example in table (1) for Economics.

Subject	Size	Fe. Faculty
Econ	6	0
Econ	16	2
Econ	12	2
Econ	7	1
Econ	23	5
⋮	⋮	⋮

Code	Size	FE Faculty Quota
Econ	6	2
Econ	16	2
Econ	12	2
Econ	7	2
Econ	23	2
⋮	⋮	⋮

Table 1: Left are the department size and the real number of female faculty, the right panel shows the actual number replaced by the shares.

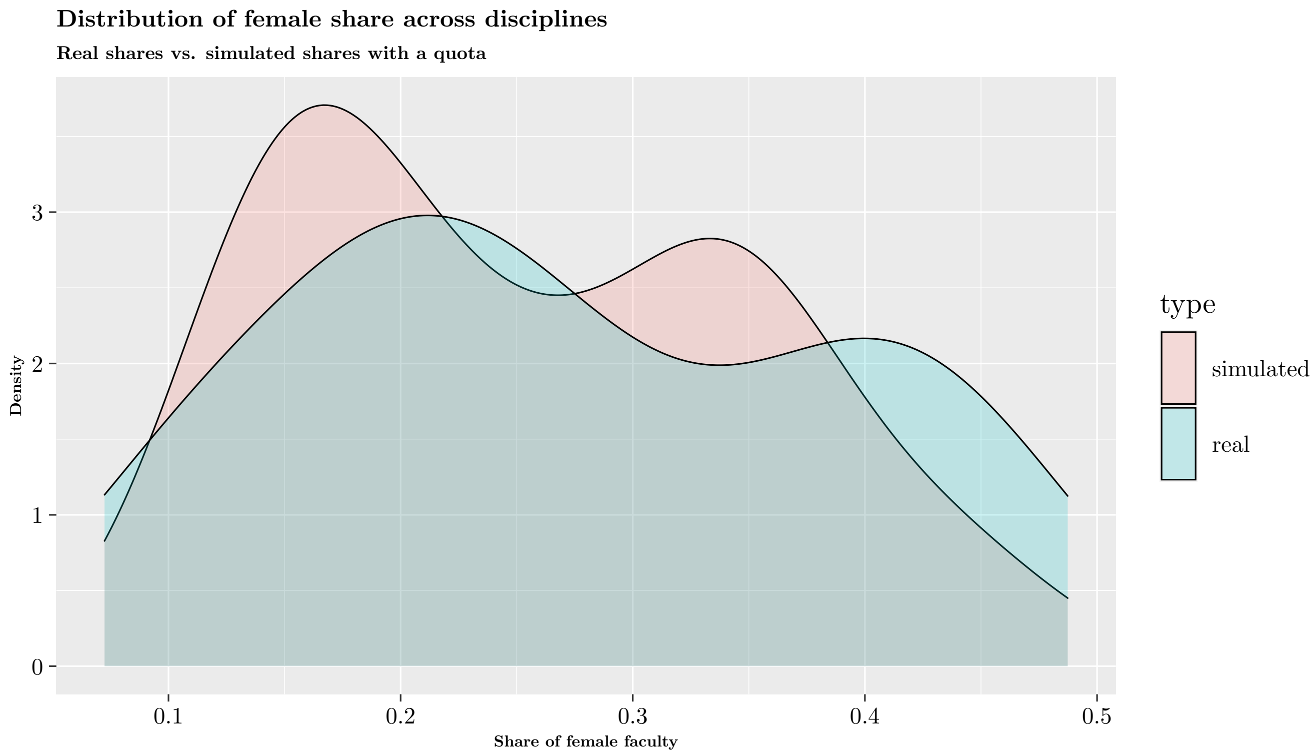


Figure 5: Density plot for the distribution of real and simulated shares of female faculty across disciplines.

A two women quota per department would lead to a mean female share of $\mu_f^s = 0.254$ overall as opposed to the real share $\mu_f = 0.257$ and can reproduce the bimodal distribution of the female share distribution, although the peaks are slightly shifted to the left. This exercise is not intended to state that this type of quota is actually implemented, but that a certain adherence to "field standard" hiring may lead to a distribution of female shares that is very similar to the one we observe. Interestingly, Ceci et al. (2014) show that the transition from graduate programs to assistant professorships shows more pipeline leakage in the fields in which women are already very prevalent (psychology, life science, social science) than in the math-intensive fields, which would be consistent with the results presented here.

5 Summary and Discussion

Analyzing employment patterns of female professors across departments and disciplines allows me to show that there exists a one- to two-women quota on the department level, which cannot be explained by traditional explanations for the underrepresentation of women among academics at the professor level. I also show that this type of implicit quota can potentially explain part of the gap in the female professor share among disciplines. The fact that I can distinguish between the implicit quota's being centered around the mean (i.e., around the relative share of female professors in a discipline, which is not the case) and a specific discrete frequency, poses interesting questions that I have not seen addressed before, namely how humans perceive structures that have the "correct level" of diversity.

It is important to emphasize that the observed patterns described above only allow me to make statements about whether or not the observed distribution can be attributed to a random allocation: we cannot say whether the "real" p (without the implicit quota) is larger, smaller or exactly the same. My results are agreeable to all options, so analyzing these patterns in a dynamic framework is needed to make more precise statements. However, the analysis above is sufficient to show that the hiring is not "gender blind" and therefore ex-post analysis of performance measures and researcher productivity should be conducted only with this caveat in mind. Going forward, it would be useful if studies that attempt this type of analysis show ex-ante that their data represent a "market solution" or address the potential (observable) quality heterogeneity within underrepresented groups directly.

Code	Subject
10	Humanities
20	Theology (protestant)
30	Theology (catholic)
40	Philosophy
50	History
80	Literature and comparative language studies
100	German Studies
110	English and American Studies
120	Roman Language Studies
140	Other Language/Cultural Studies
160	Cultural Studies
200	Sports
220	Law, Economics and Social Sciences, general
230	Political Sciences
235	Sociology
240	Social Work
250	Law
290	Economics
315	Psychology
320	Pedagogy
330	STEM, general
340	Mathematics
360	Physics and Astronomy
370	Chemistry
390	Pharmacy
400	Biology
410	Geology
420	Geography
445	Health Sciences, general
450	Preclinical Medicine
470	Clinical-theoretical Medicine
490	Clinical-Practical Medicine
560	Clinical-theoretical Veterinary Science
580	Clinical-practical Veterinary Science
620	Agricultural Science
650	Dietary Studies
690	Mechanical Engineering
710	Electrical Engineering
730	Architecture
740	Spatial Planning
750	Civil Engineering
765	Computer Science
770	Material Science
780	Art, general
790	Visual Arts
800	Design
820	Performing Arts
830	Music
920	Central scientific Departments
960	Institutes associated with Universities

Table 2: List of the included subjects, coding and name, according to the "Schlusselverzeichnisse fur die Personalstatistik Stand: 2017", translation: own.

References

Alper, J. and A. Gibbons (1993): "The pipeline is leaking women all the way along," *Science*, 260, 409–412.

Arcidiacono, P., J. Kinsler, and T. Ransom (2020): "Asian American Discrimination in Harvard Admissions," *NBER Working Paper*.

Auriol, E., G. Friebel, and S. Wilhelm (2020): "Women in European Economics," *Women in Economics, edited by S. Lundberg*, 26–30.

Bertrand, M., S. E. Black, S. Jensen, and A. Lleras-Muney (2018): "Breaking the glass ceiling? The effect of board quotas on female labour market outcomes in Norway," *The Review of Economic Studies*, 86, 191–239.

Bharadwaj, P., G. B. Dahl, and K. Sheth (2014): "Gender discrimination in the family," *The economics of the family: How the household affects markets and economic growth*, 2, 237.

Ceci, S. J., D. K. Ginther, S. Kahn, and W. M. Williams (2014): "Women in academic science: A changing landscape," *Psychological Science in the Public Interest*, 15, 75–141.

Chang, E. H., K. L. Milkman, D. Chugh, and M. Akinola (2019): "Diversity thresholds: How social norms, visibility, and scrutiny relate to group composition," *Academy of Management Journal*, 62, 144–171.

Goldin, C. (2014): "A pollution theory of discrimination: male and female differences in occupations and earnings," in *Human capital in history: The American record*, University of Chicago Press, 313–348.

Graham, P. A. (1970): "Women in academe," *Science*, 169, 1284–1290.

Huang, J., A. J. Gates, R. Sinatra, and A.-L. Barabási (2020): "Historical comparison of gender inequality in scientific careers across countries and disciplines," *Proceedings of the National Academy of Sciences*, 117, 4609–4616.

Lundberg, S. and J. Stearns (2019): "Women in economics: Stalled progress," *Journal of Economic Perspectives*, 33, 3–22.

Lutter, M. and M. Schröder (2016): "Who becomes a tenured professor, and why? Panel data evidence from German sociology, 1980–2013," *Research Policy*, 45, 999–1013.

Madison, G. and P. Fahlman (2020): "Sex differences in the number of scientific publications and citations when attaining the rank of professor in Sweden," *Studies in Higher Education*, 1–22.

Maida, A. and A. Weber (2019): "Female Leadership and Gender Gap within Firms: Evidence from an Italian Board Reform," *IZA Discussion Paper*.

O'Connell, S. D. (2018): "Can quotas increase the supply of candidates for higher-level positions? Evidence from local government in India," *Review of Economics and Statistics*.

Paryavi, M., I. Bohnet, and A. van Geen (2019): "Descriptive norms and gender diversity: Reactance from men," 2.

Rossi, A. S. (1965): "Women in Science: Why So Few?: Social and psychological influences restrict women's choice and pursuit of careers in science," *Science*, 148, 1196–1202.

Students for fair admissions v. Harvard college, H. (2019): "Case 1:14-cv-14176-ADB [2020] (UNITED STATES DISTRICT COURT DISTRICT OF MASSACHUSETTS), STUDENTS FOR FAIR ADMISSIONS, INC., v. PRESIDENT AND FELLOWS OF HARVARD COLLEGE (HARVARD CORPORATION)." Report 672.

Townsend, R. B. (2010): "Putting Academic History in Context: A Survey of Humanities Departments," .

White, S., R. Ivie, and A. Ephraim (2012): "Number of Physics Faculty Members: Results from the 2010 Survey of Physics Degree-Granting Departments," *American Institute of Physics*.

Williams, W. M. and S. J. Ceci (2015): "National hiring experiments reveal 2:1 faculty preference for women on STEM tenure track," *Proceedings of the National Academy of Sciences*, 112, 5360–5365.