Will gender parity break the glass ceiling? Evidence from a randomized experiment[*]

Manuel F. Bagüés[†]
Universidad Carlos III

Berta Esteve-Volart[‡]
York University

October 2006

Abstract

This paper studies whether the gender composition of recruiting committees matters. We make use of the exceptional empirical evidence provided by Spanish public examinations, where the allocation of candidates to evaluating committees is random. We analyze how the chances of success of over 75,000 male and female candidates to the Judicial Corps in the period 1995-2005 were affected by the gender of their evaluators. We find that a female (male) candidate is significantly *less* likely to pass the exam whenever she is randomly assigned to a committee where the share of female (male) evaluators is relatively greater. We also find that 'gender parity'committees are fairer than male committees. The evidence found in this paper suggests that the imposition of gender parity in recruiting committees will not increase the number of women in decision-making positions; paradoxically, it will further reduce it. However, it will increase fairness.

Keywords:
JEL Classication:

1   Introduction

Legislation and policy encouraging gender parity, or gender quotas, in top political or public positions has recently been approved in some countries and is being considered in many others. In Chile, Michelle Bachelet has just appointed 50% of women to her cabinet. In Norway, since 1988 there must be a minimum of 40% of each gender in publicly appointed committees, boards, and councils. Furthermore, since January 2005, Norwegian state-owned enterprises are required to have at least 40% board representation of each gender. In 2004 the newly elected Spanish Prime Minister, Socialist José Luis Zapatero, appointed 50% of women to his cabinet, and announced that "there will be gender parity in all selection committees in the state administration, public organizations and related firms".[1] Private corporations in Spain are also receiving governmental guidelines towards greater participation of women at boards.[2]

The reason for imposing gender parity in top positions lies in the extremely low percentages of women in decision-making posts, at both the public and private sectors. Nevertheless, it is not clear that imposing gender parity in top positions is going to increase the numbers of women in other high positions. Indeed, although implicit in many discussions of gender parity policy, there is no clear evidence for the hypothesis that the lack of women in top positions is due to men discriminating against women.

How could gender parity increase the number of women in top level positions? Directly: women who get to top-level positions because of gender parity or quotas could hire more women than their (male) predecessors. Indirectly: because women at the top could become role models (if women are not getting to top positions because of social norms, having more women at the top might help change the social norm), or because women at the top can affect choices in ways that might help other women get to top positions (such as flexible working hours choices in the private sphere, or public expenditure choices that benefit women more, in the political sphere).[3]

The focus of this paper is this direct effectiveness of gender parity policy: does increasing the number of women in top positions lead to the recruitment of more women? The answer to this question is related to whether this is due to men discriminating against women: in the case that discrimination is statistical, gender parity will not be effective. Similarly, in the case that the low numbers of women at the top are due to social norms, gender parity could work only if in connection with a major cultural change as, for example, the women's liberation movement in the late 1960s and early 1970s in the United States.

A neat empirical analysis of the effects of gender parity is hard to come by, because of the endogeneity

---

[1] See the Official State Bulletin (BOE) number 57, March 8th 2005, page 8111. See, as well, the *IV Plan de Igualdad de Oportunidades entre Mujeres y Hombres, 2003-2006* (Ministry of Social Affairs, Spain).

[2] The guidelines are as follows: (1) Board members are encouraged to clarify why there might be no women at the board, and any measures taken. (2) Whenever there are any vacancies, the board is encouraged to ensure that the selection process is not biased against females. The board should also intently look for women who have the desired professional profile. (3) Firms are encouraged to include a discussion of the gender distribution of positions and report any changes. The report also points out that, due to 'old boys' network effects and hysteresis, this sort of policy is necessary in order to increase the low percentages of women at boards (*Proyecto de Código Unificado de Recomendaciones de Buen Gobierno de Sociedades Cotizadas,* January 2006).

[3] Duflo and Chattopadhyay (2004).

encountered: in most real-life cases the composition of hiring committees is likely to be related to candidates' characteristics. Thus, usually it is not possible to rule out the possibility that the number of female members in a committee might be somehow related to the average quality of females in that field. In order to avoid this problem, ideally we want to observe how committees of different characteristics evaluate the quality of candidates that have been randomly allocated.

This paper provides evidence from such an experiment. Here, we show that greater numbers of women in decision-making positions do not necessarily imply that other women will have it easier to get into decision-making positions. In fact, we find that committees where the proportion of women is greater actually make it more difficult for women candidates than for men candidates. We show this exploiting the exceptional evidence provided by the particular mechanism that is used in Spain to select members for the public sector.[4] In Spain, individuals who want to have access to any public official positions are required to pass a specific public exam at the national level. These public exams, which are also common in other countries in Europe and Latin America, typically involve an extremely large number of candidates.[5] The allocation of candidates to evaluating committees is determined by a lottery—so that the process whereby candidates and committees are mapped is based on a first random draw and, most crucially, is completely orthogonal to gender.[6]

Thus, the beauty of this evaluation process is (1) that candidates are allocated to committees randomly, which automatically eliminates concerns about omitted characteristics of candidates and other issues; (2) that the subjects and the experiment are actually taken from real life, with real-life payoffs, hence avoiding the usual *caveats* of experimental work; and (3) that the experiment is relevant because of the importance, and magnitude, of public exams.

In this paper we use information on 75,000 public exam candidates over 1995-2005 from a number of judicial public exams to explore the relationship between the gender composition of the evaluating committees and the candidates' probability of success by gender. We find that a female candidate's chances of passing the exam are 5.5% greater if she is evaluated by a committee with more male committee members than the median committee, than if she is evaluated by a committee with fewer male committee members than the median. As we show in the empirical anaylsis, these differences are statistically significant. We also find that this effect is due to both male committee members being relatively softer on female candidates, and female committee members being relatively tougher on female candidates. Our results also indicate that the decisions taken by committees with 'gender parity' are consistent with the true quality of candidates; in contrast, when most committee members are male, more female candidates of relatively low quality succeed.

The evidence in this paper has great relevance for various reasons. First, the Spanish government and top judicial authority are considering imposing gender parity in all public recruiting committees, including the committees we study here. Second, successful candidates to public exams become figures

---

[4] For an economics study of public exams in Spain see Bagüés (2005), who finds that the probability to pass a public exam is largely affected by randomness, nepotism and localism.

[5] In 2003, approximately 175,000 individuals including, of whom 131,000 were university graduates, were preparing for public exams in Spain ("Encuesta de Población Activa").

[6] Committees are mostly composed by top public officers, but also other professionals in the field. For instance, in judge exams, committees are composed by seven members of the Judicial Corps but also two non members, namely one lawyer, and one law university professor.

who deliver relevant decisions to society.[7] Furthermore, many political figures in Spain belonged to the Judicial Corps before landing important posts in the government. Thus the gender composition of evaluating committees has an important effect on the gender composition of the elite. For this reason, this 'experiment' constitutes an example of a randomized design that operates in a very important framework: the elite formation of society. Finally, recent literature shows that the composition of the government is not irrelevant to policy choices, suggesting that women might have different preferences on public expenditure than men.[8] All in all, and whatever the reason behind the effect found in this paper, imposing gender parity in public exam committees for the judiciary would not increase, but rather reduce, the number of women in the judiciary, and in turn, the government—and this could later translate into policy choices.

1.1 Gender parity effectiveness: an illustration

For the sake of illustration, we provide here an example that is consistent with the evidence in our paper. Is gender parity actually effective in raising the numbers of women in other positions? In Spain, the cabinet gender parity that was imposed by the Socialist party in 2004 (from 18.75% women in cabinet in 2002) has not been followed by higher numbers of women in other top public positions.[9] That is, newly appointed female ministers have not hired more women than their (mostly male) predecessors (see Figure 1).[10]
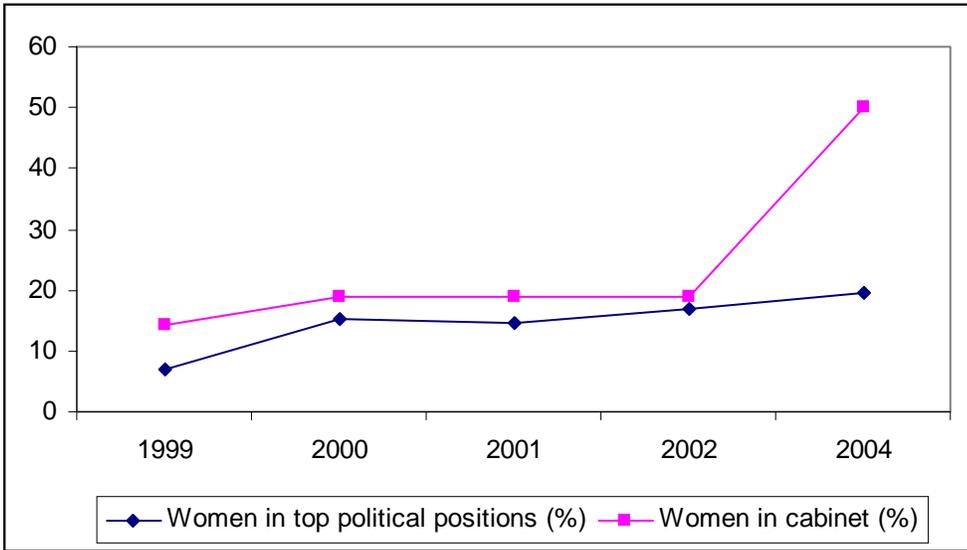

Figure 1 Women into political positions and cabinet members, Spain,1999-2004

[7] This is especially important in light of the evidence consistent with the view that judicial resolutions are not gender neutral. Indeed, even though the empirical evidence is not perfectly exogenous, male judges seem to favour female defendants (see e.g. Nagel and Hagan 1983, Parisi 1982).

[8] Pande (2003), Duflo and Chattopadhyay (2004).

[9] With 19.6% of women in 2004 compared to 16.8% in 2002 with the then ruling Popular party, an increase that is consistent with previous years.

[10] We thank Florentino Felgueroso for pointing out this fact.

2 Related literature

To our knowledge, there is no study dealing with the effectiveness of gender quotas. However, there is a large body of literature that exploits experimental data and finds empirical evidence suggesting that gender matters: Goldin and Rouse (2000) find that the adoption of a screen in the orchestras' hiring process of musicians fostered impartiality in hiring and increased the proportion of women hired, Lavy (2004) compares data on blind and non-blind scores that high school students receive on matriculation exams in their senior year in Israel, and finds gender discrimination against *male* students in each subject. Blank (1991) compares single-blind and double-blind reviewing of papers submitted to *The American Economic Review* and finds a small, insignificant effect, in that female authors fare better under double-blind reviewing.

There is also a large literature that investigates productivity differences by gender: by looking at wages (Blau and Kahn 1994), academic research of individual investigators (Veugelers 2006, Mairesse and Turner 2002), or by looking at the returns to firms by gender of their CEO (Wolfers 2006). This literature typically finds that, conditioning by a number of controls (such as tenure in the case of Veugelers 2006), men are more productive than women.

3 Background information

3.1 Gender policy

The reason for promoting gender policy —and particularly for promoting gender parity in top positions— lies in the extremely low percentages of women in decision-making posts at both the public and private sectors. In political positions, only in 12 out of 179 countries women accounted for at least 30% of parliamentary seats in 2003.[11] In Italy and France, only 3% and 4% respectively of the 50 largest companies' board directors are women.[12] In the US, women constituted only 3.4% of the top level management in a sample of firms in 1997.[13]

In the past, policy towards gender equality in the professional and public spheres seemed to focus on the so-called *equal opportunities* approach.[14] Underlying this approach was the pipeline theory, according to which women are moving their way through the pipeline and into top level jobs. Accordingly, policy was designed to encourage women's higher education and skills on the understanding that providing women with the same human capital as men would enable them to reach the top positions they seemed unable to attain.[15] The evidence for the pipeline theory is disappointing: in Norway,

---

[11] Sweden leads the list with 45% (UN Millennium indicators).

[12] According to a report from the Aspen Institute Italia (The Economist).

[13] Bertrand and Hallock (1999).

[14] In the US and other countries there has also existed affirmative action policy, involving quotas but generally not at the 50-50 level (Fryer and Loury 2005).

[15] This approach is still in use and sometimes in connection with gender parity, e.g. the French Parliament passed legislation in 2001 mandating gender parity in party lists for a variety of elections (see Frechette *et al* 2005 for an analysis of the political economy of gender quotas).

despite significant increases in the female ratio among students in higher education (about 50% since 1986), only two out of ten professors are women.[16] In the same vein, there is a prevailing view that women have started to move up into management and public positions, but once they reach a certain point, the so-called 'glass celing', they do not seem to go any further. Whatever the reason behind the existence of a glass ceiling, pessimism about the pipeline theory might explain the more recent approach, that of *equal results*: the imposition of gender parity in outcomes, such as cabinet, or board membership.[17] The motivation for imposing gender parity seems to be the rationale that once more women are in top positions in the public sphere, it should be easier for other women to get to that level—in other words, gender parity in decision-making could break the glass ceiling.[18]

A related issue arises from the heterogeneity of women and representation. Since women do not constitute a homogeneous group, the effect of imposing gender quotas might generate different effects. For instance, in less developed countries, female leaders have typically been widows or daughters of deceased male leaders (e.g. Indira Gandhi, Benazir Bhutto). If women who have easier access to politics (or public office skills) disproportionally share the same background, a gender quota will create female policy-makers with little in common with the majority of women. In this sense, a gender quota might have the effect of replacing middle-class men with high-class women.

3.2   Public examinations

The evaluation through nation-wide public exams has traditionally been used to determine the access to a variety of public positions in many countries. Public examinations were already used in Imperial China: from the start of the Sui Dynasty (in 605) to its abolition near the end of the Qing Dynasty (in 1905), the Imperial Examination System lasted continuously for 1300 years. Public examinations, which have been in place for many years to avoid corruption, are still in use in Spain as well as other countries in Europe and Latin America: in those countries, passing a public examination is required in order to obtain a permanent position in the public sector—including those in the most coveted occupations (e.g. notary, judge, or state lawyer). In this paper we use data from judicial public examinations in Spain.

To get an idea of the magnitude of these selection process, approximately 130,000 individuals with a university degree were preparing full-time for public examinations in Spain in 2003, respresenting over 20% of the population with a university degree of ages 25 to 29[19]—a very large number given the low success rate (4.9% in our data base).[20]

---

[16] Rogg (200?).

[17] The shift in policy is obvious in the Spanish case. In a recent governmental document on proposals towards gender equality, equality of opportunity is mentioned only once, while gender parity appears in the document six times (ORDEN PRE/525/2005, March 7 (*BOE, March 8, 2005*)).

[18] Gender quotas are often imposed on either of the following two stages of the selection process: the stage of finding aspirants, e.g. those willingly to be considered for nomination, or the stage of nominating the actual candidates (e.g. to be placed on the ballot by the party). In some cases, gender quotas are imposed on a third stage, the already selected or elected candidates. Here we implicitly consider gender quotas at a different stage, the *candidate evaluation* stage, but our evidence could also apply to the other quotas.

[19] According to the "Encuesta de Población Activa" there were 175,000 inviduals preparing for public examinations in 2003, of whom 130,000 were university graduates.

[20] Even though they are not listed in aggregate statistics, there are also many individuals preparing public exams on

Passing a public examination requires acquiring an enormous amount of knowledge which is often extremely specific and thus not likely to be useful for the labor market. For instance, public examinations to become a member of the Judicial and Fiscal Corps require learning by heart thousands of pages of law articles—successful candidates' average time of preparation for judge public exams is three years and 11 months,[21] which is in accord with the huge amount of material candidates must prepare.

Candidates to becoming members of the Judicial Corps must have a first degree in law. Every year, once the number of candidates is known, evaluating committees are formed. The committees are formed within a month of the publication of the candidates' names in the BOE (Official State Bulletin), and their composition is published in the BOE as well. No committee member is in two committees. In general, an evaluating committee is formed for every 500 candidates. Candidates are ranked in alphabetical order and committees are ranked numerically. Then a lottery decides the initial according to which the alphabetical list of candidates will be matched with the list of evaluating committees. For instance, in January 2004, the letter that was randomly chosen was "S", hence the first candidate in the list whose initial was "S" was matched to the first committee in the list, and so on. Thus, the process whereby candidates and committees are mapped is based on a first random draw and, most crucially, it is completely orthogonal to gender.

The rules and composition of evaluating committees differ by exam, but the committees are generally composed by both members of the Judicial Corps and non members (law professors or lawyers) (see Table 2).[22] Each committee is presided over by one member, who appoints the other members according to the rules.

Each committee evaluates a set of students, usually orally.[23] The topics that candidates are evaluated on are drawn from the set of possible topics, which are listed in the BOE beforehand. The exam lottery consists of balls numbered after the topics in the exam. Five balls are drawn, determining a particular five-question exam. A candidate receives an evaluation if he manages to answer the questions—many candidates fail to get an evaluation for this reason. The grading system of exams is as follows. At the end of each session, candidates are evaluated by committee members' ballots containing the grade for the candidate.[24] The grade ranges between zero and five per topic. For each candidate, the minimum and the maximum grade ballots are excluded. For the rest of the ballots, grades are added and divided by the number of ballots, which determines each test's mark. Candidates who do not achieve at least half the maximum grade in each test are disqualified. Candidates who do not achieve at least more than half the maximum grade in three of the five topics in each test are disqualified too. Candidates' final grades are obtained by adding their grade in each test. The committee's decision is made on a majority basis. In case of a tie, the president decides the final outcome.

a part-time basis.

[21] This is calculated for candidates who passed the exam in the years 2001 to 2004 (*Datos Estadísticos de los Jueces en Prácticas*, Consejo General del Poder Judicial).

[22] For details, see the appendix.

[23] Except for the judge and prosecutor exam for which a preliminary eliminatory multiple choice test was introduced in 2003.

[24] However, unofficial information confirms this is only used for unclear cases.

4 Description of the data

We analyze four types of Spanish public exams: exams to become (1) judge, (2) prosecutor, (3) court secretary, and (4) a joint exam to become either judge or prosecutor that has been in place since 2001 (the judge and prosecutor exams were separate until 2000).[25]

The data we use here have been compiled from Spanish official publications (BOE, selected issues) for public exams between 1995 and 2004. Our data base contains information on about 20 exams in which nearly 75,000 candidates were evaluated by 150 committees.[26] We investigate the relationship between gender composition of committees and candidates' success by gender using two types of data on public exams.

First, we use committee level data on the link between gender composition of committees and the gender of candidates. We have data on committees for exams over the period 1995 to 2004. In particular, we use data on 150 committees, for which we know members' characteristics. Usually the composition of committees that is first announced is somewhat different from the composition of committees that finally get to evaluate candidates,[27] so we use the data that correspond to the final committee composition. We do not observe the gender of all candidates, but we do observe the gender of successful candidates.

Second, we take advantage of a multiple choice test that was introduced in 2003 for the judge and prosecutor exam. Since then, the exam has consisted of three stages, all of them qualifying; the first one is a multiple choice test, the other two are two oral tests. For this, we have individual data on the 2003 and 2004 exams from over 4,000 candidates each year.

Next we discuss some descriptive statistics.

In Table 1 we show information from candidates. There are typically more female than male candidates (almost 70% for most exams). We do not have information on the total number of candidates by gender for previous years; we do though have data on the gender of successful candidates. Candidates from the capital represent about a fifth of all candidates; more than Andalusia, the most populated region in Spain, and about twice as much as Catalonia. We have also collected some information on candidate 'experience', in the sense of the number of times candidates take the public exam. Since we have the complete names and ID number of candidates we can track them over time. However, this procedure is limited to the number of years we have information for each public exam. For example, in Table 2 we put information from Bagüés (2005) using the numbers for the 2005 judge and prosecutor exam. Out of the candidates who took that exam, at least 82.3% had taken the exam once before (i.e., 2004). Similarly, 65.7% of candidates had taken the exam twice before, 48.7% had taken the exam three times before, and 23.3% had taken it five times before (i.e. 2000, 2001, 2002, 2003, and 2004). That is, almost a fourth of all candidates to the judge and prosecutor in 2005 had already taken the exam at

---

[25] We cannot use data from the court secretary exam in 1997 because the exam outcome was not published in the BOE.

[26] The set of candidates is not different from one year to another; some of them apply repeteadly.

[27] Reasons are varied, but some individuals initially appointed as committee members have other important commitments or may have moved because for job reasons.

least five times before, which goes to demonstrate the exam difficulty and candidate perseverance.

Table 2 displays committee composition by type of exam, and the number of observations that are available for them respectively. As seen from the table, each committee has to be formed by Judicial Corps members (judges, prosecutors, court secretaries, and public defenders) and non members (law professors and lawyers). Members of the Judicial Corps cannot make more than half the committee (*Ley Orgánica 6/1985, July 1*). The rules for committee composition and formation are described in more detail in the appendix.

In Figure 2, we plot the percentage of women in committees, by year, and type of exam. Typically there are more men than women in committees; in the average committee women outnumber men only in the court secretary exams in 1998 and 2002 (about 57% were women). For the other types of exams, the average percentage of women is between 20 and 30%.
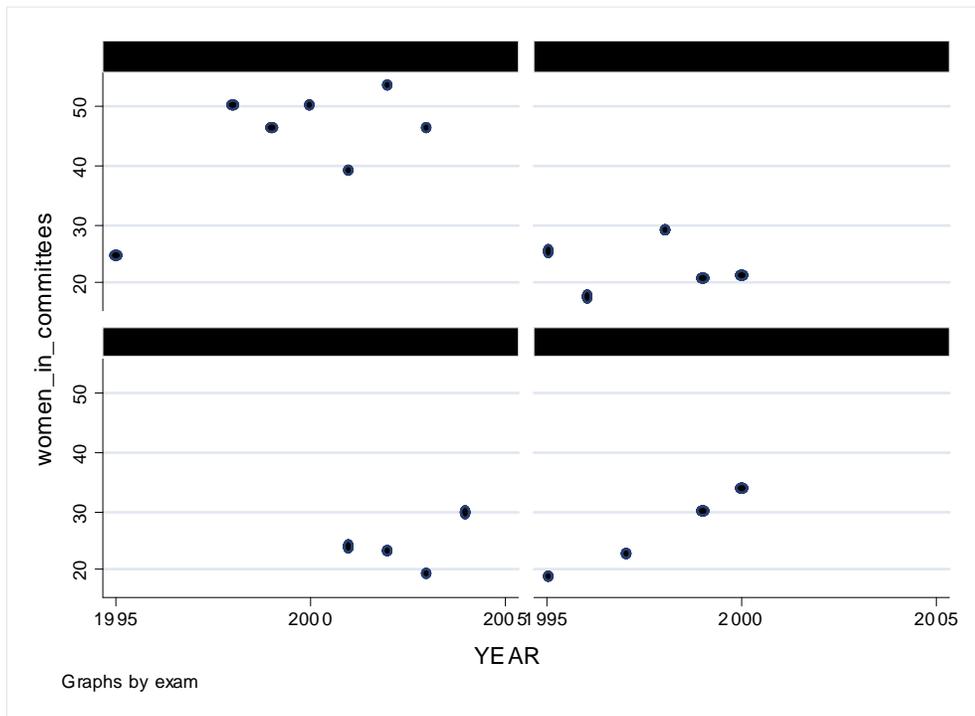
Figure 2 Women in committees (%), by type of exam

Table 3 presents some descriptive statistics on committee members' characteristics by gender. In general, female committee members are younger and have a lower relative rank than their male counterparts.

Now we compare the characteristics of committees to the characteristics of the total Judicial Corps population. We want to see if the incidence of women in committees for public exams is different than their incidence in the whole judicial population, or whether it is representative of this population. Figure 3 shows the percentage of women judges, prosecutors, and court secretaries, who were in

judicial public exam committees over 1995-2004.[28] We also have data on 1995, 2000, 2002 and 2004 for the percentage of women in judges, prosecutors, and court secretaries in Spain. In the figure, the percentage for judges in the population and committees are very similar; this is also true for court secretaries except for 2002. The prosecutor series seem more distinct: males are overrepresented in committees.
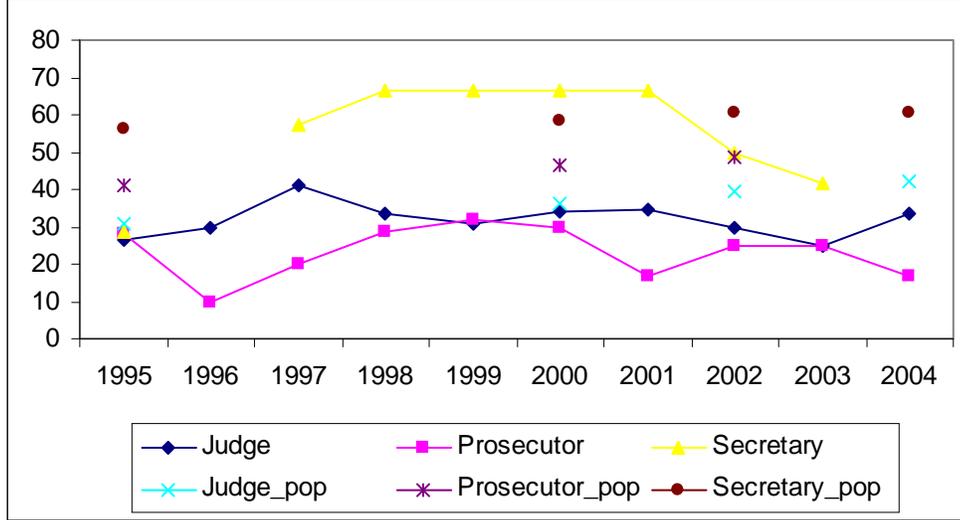

Figure 3 Women in population and in committees (%), by profession

5   Empirical analysis

In this section our goal is to estimate whether the gender composition of committees affects candidates' success according to their gender. Ideally, we would like to observe the individual vote of each committee member. However, we cannot do that, because only the committees' final aggregate decision is publicly available.

Given this limitation, our empirical strategy is twofold. In our first empirical strategy, we exploit committee composition and exam outcome for exams over the period 1995 to 2004. We use the female share in the committee as measure of the gender composition of the committee. For this, we have data on 150 committees (who evaluated about 75,000 candidates). Since, as described in the previous section, the allocation of candidates to committees is orthogonal to gender, we also know that the proportion of women evaluated by every committee is statistically similar.

We analyze the relationship between the gender composition of an evaluating committee and the candidate's probability to succeed by gender. We do this by looking at regressions of the form:

$$y_{cet} = \beta s_{cet} + \alpha_{et} + \varepsilon_{cet} \tag{1}$$

---

[28] These data exclude the committee presidents who must belong to Superior Courts of Justice and therefore constitute a different figure.

where y denotes a measure of successful candidates (e.g., female successful candidates, or male candidates), s denotes the share of women in the committee, e denotes type of exam (judge, prosecutor, court secretary, or judge and prosecutor together), t denotes year, α_{et} and is an exam and year fixed effect.

In each committee the number of positions is more or less fixed,[29] therefore we cluster standard errors at the exam level.

In Table 4 we present results from running regression (1) using the number of female successful candidates as dependent variable in the left panel, and the number of male successful candidates as dependent variable in the right panel.[30]

In column (1), pooling all committee members (regardless of their membership to the Judicial Corps), the share of women in the committee shows a negative, not significant effect on the number of sucessful female candidates; in column (4) the effect is positive, and significant, for the number of male successful candidates.

In columns (2) and (5) we distinguish between members in the committee who belong to the Judicial Corps and those do not. In column (2), the share of female Judicial Corps members is strongly negatively related with the number of women who pass the exam (significant at the 1 percent level). In column (5), we observe a positive effect on the number of successful men; this is significant at the 5 percent level. The effect for non members is not statistical significant. This is consistent with the fact that, in reality, it is only the members of the Judicial Corps (and especially the president) who are relevant to the decision.

In sum, we find that the gender composition of the committee matters. Now we want to look at information about committees in order to try and learn more about this effect. For this, we run

$$y_{cet} = \beta s_{cet} + \delta X_{cet} + \alpha_{et} + \varepsilon_{cet} \tag{2}$$

where X is a set of committee characteristics. We also disaggregate Judicial Corps members between the president and the rest of Judicial Corps members.

In columns (3) and (6) we also introduce the mean age of Judicial Corps members of the committee (and a quadratic term), their mean age of entry into the Judicial Corps, and their mean ranking.[31]

In column (3), the female share of Judicial members in the committee is still significantly negatively related with the number of successful women. The gender of the president also has a negative effect, but it is not significant. Ranking is the only committee characteristic that matters: more highly ranked members are related to more women succeeding (at the 10 percent level). In column (6), while the female share in Judicial committee members does not have a significant effect for members other than the president, more male candidates are successful if the committee president is a woman.

---

[29] See appendix for details.

[30] Since the first committee in each exam might have a different number of positions, we are also including a dummy variable that equals one for the first committee.

[31] The ranking is a measure of both quality and seniority.

In the case of successful men, we do find that age is important, if only marginally. The effect is nonlinear: more men pass the exam when evaluated by older committees, up to age 50; from then on, older committee members are associated with more women passing the exam. That is, the fact that generally male committee members are older does not explain the gender effect we find—in fact, older committee members then to be associated with more male successful candidates (there is no effect on females as seen in column (3)).

In Table 5 we run regression (1), this time with the total number of candidates who pass the exam (left panel), and the percentage of female candidates that pass the exam (right panel).

In columns (1) to (3), none of the variables seem to affect the total number of successful candidates. The reason for looking into this is that whenever some committees do not fill all positions, other committees could potentially use those vacancies if there are more satisfactory candidates than available positions.[32] The results suggest that this is not important for the gender composition of the committee.

In columns (4) to (6) we show results using the percentage of successful women. Even though these results should be taken with some caution in statistical terms, they constitute a good summary of results in Table 4. Looking at column (4), the female share of committees is related with a lower percentage of women passing the exam at the 10 percent level—regardless of whether the committee members are Judicial Corps members or not. In column (5), the share of female Judicial Corps members lowers the percentage, the effect seems to come from the president but also the other members. As in Table 5, we also find a significant nonlinear age effect, which works in the same direction as previously.[33]

In Figure 4 we show graphically the results. We classify committees with respect to the median committe's gender composition, that is we classifying committees into those where the percentage of women is lower than the median (52), and those where it is higher (58). The average percentage of female succesful candidates is higher for the former, and the difference is significant at the 5 percent level.

Taken together, the results from Tables 4 and 5 are clear: female (male) candidates are more likely to be successful when evaluated by committees in which there are more women (men).[34] However, these results do not give us precise information about the reasons behind this phenomenon. In fact, they are consistent with two hypotheses: (i) male committee members are more generous with female candidates, and (ii) female committee members are stricter with female candidates. Hypothesis (i) is consistent with positive gender discrimination working in favor of male candidates, while hypothesis (ii) is consistent with some form of statistical discrimination being suffered by female candidates. If the quality of a professional cannot be observed and it is the average quality of that professional group that she is judged by, minorities are more affected by (the quality of) new entrants than non minorities: having just a few women in the Judicial Corps means that one new female judge creates

---

[32]See appendix..

[33]The age of entry of the president, and the age of the president, do not have any effect in regression (1). They are not included in tables 4 and 5 because there are some missing observations in those series.

[34]Results also hold if we do not include the 1995 exam for court secretary, which included a multiple choice test.

a greater impact on the average quality of female judges than the effect of one new male judge on the average quality of male judges.
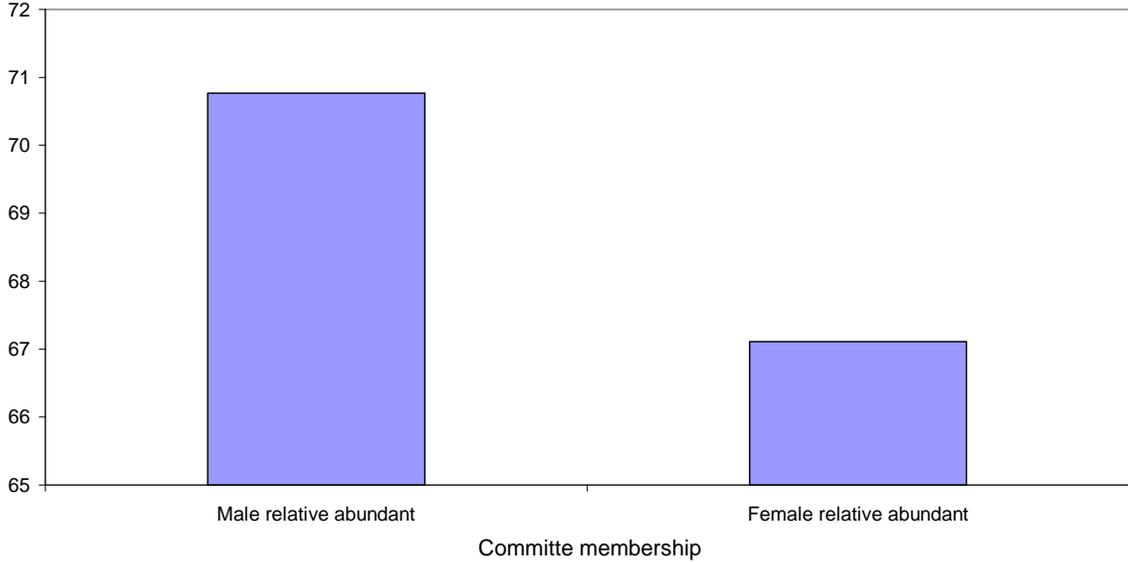

Figure 4 Female successful candidates (%), by type of committee membership

In order to try and shed light on this issue, in our second empirical strategy we are going to use our candidate-level information. We take advantage of the fact that in 2003 a multiple choice test was introduced in the judge and prosecutor exam. Since then, that particular exam consisted of three stages, all of them qualifying; the first one was a multiple choice test, and the other two were two oral tests. Using data available up to now (for the 2003, 2004 and 2005 exams) for this type of public exam, we use candidates' multiple choice test performance as a proxy of the candidates' objective quality. Even though the multiple choice test mark is arguably a noisy proxy for candidates' quality, it is nonetheless a consistent proxy. We cluster standard errors at the committee level; results are shown in Table 6.

In order to study the link between the probability of a candidate's success and the gender composition of the committee, we need to consider the sample of candidates that have successfully passed the multiple choice test—because these are the candidates who are actually evaluated by the committees. In this sample, we have information on fifteen committees. Of those, in six committees there is only one female member; in the other nine, there are more women.[35] Again, we use as measure of committee membership by gender the share of female members in the committee. For this we run:

$$y_{it} = \sigma Fcandidate_{it} + \varphi s_{it} + \eta Fcandidate * s_{it} + \phi quality_{it} + \varphi quality_{it}^2 + \gamma_t + \varepsilon_{it}, \qquad (3)$$

---

[35] There are no committees with zero women in this sample.

where t denotes year, the dependent variable is the probability that candidate i passes the exam; *Fcandidate* is a dummy variable equal to one in the case that candidate i is a woman, and *quality* denotes the multiple choice test mark.

In column (1) in Table 6, we find that female candidates have a lower probability of passing the exam if they are evaluated by a committee with a larger share of women, controlling by experience(whether the candidate has already taken the exam at least once, twice or thrice before) and provincial fixed-effects. The effect is similar in magnitude to the effect we found in Tables 4 and 5, and is statistically significant at the 5 percent level—notice that the sample here (fifteen committees) is fairly small. In columns (2) and (3) we see that the effect comes from stage three: column (2) shows that gender does not matter for stage two—however, female (male) candidates are less (more) likely to pass stage three the higher the share of females in the committee (column (3)). This effect is significant at the 1% level.

How big is this effect? Simple calculations show that, for the 2003 judge exam, a woman's chances of passing the exam are 9.4% greater if she is evaluated by a male committee (that is, a committee with only one female member) than if she is evaluated by a committee with a majority of female members.[36] This figure equals 14% and 13% for 2004 and 2005 respectively.

What can we say about hypotheses (i) and (ii)? Here we we make again use of the fact that the multiple choice test provides us with a consistent proxy for candidate quality, and assume that the mark in the test is an unbiased proxy of true quality. Using the estimated coefficients in column (1), a female candidate who is allocated to a committee in which there is a majority of women has 0.98% lower probability to succeed than would correspond to her according to her multiple choice test mark. This probability is 1.89% greater if the candidate is allocated to a committee where there is only one woman. That is, one third of the effect can be attributed to female committee members being stricter with female candidates (or less strict with male candidates), while the other two thirds of the effect can be attributed to male committee members being less strict with female candidates (or stricter with male candidates).

In sum, the evidence from the multiple choice test shows that female (male) candidates are more (less) likely to pass the public exam whenever there are fewer women in the committee that evaluates the last stage in the exam, and this seems to be due more to male evaluators being less strict with female candidates relative to male candidates.

The numbers above suggest that committees with more female members are closer to the true quality of candidates. Now, if we compare gender parity committees with committees with only men, calculations are as follows: a female candidate who is allocated to a committee in which there are no men has 2.6% higher probability to succeed than would correspond to her according to her multiple choice test mark. Interestingly, however, this probability is a meagre 0.65% lower if the candidate is allocated to a committee with gender parity, thus gender parity committees assign female candidates a probability to succeed that is much closer to the probability that would correspond to them according to their

---

[36] Committees are composed of nine members. The maximum number of female members in committees over the years we study here has been five in 2004 and 2005, and four in 2003.

multiple choice mark—that is, committees with the same number of women and men are fairer than male committees.

6 Discussion

There are several explanations consistent with the low numbers of women in top level positions in general observe in reality. First, there could exist taste discrimination, i.e. discrimination *à la* Becker,[37] where women do not get to the top because men in top positions do not hire them since that would entail a utility loss for them. Related would be the possible existence of 'old boys' network effects: the traditionally overwhelming presence of men in top positions in management and the public sphere might make it more difficult for women to get to that level.[38] Second, women could be, in average, less productive than men—that would generate statistical discrimination in which women are not hired because they are identified with the average woman.[39] For instance, Lazear and Rosen (1990) describe how statistical discrimination in the promotion process makes it difficult for women to progress up the job ladders—such discrimination would be rational and operates on the belief that women are more likely to withdraw from labor market activities than men. Gneezy *et al* (2003) find that women perform worse when the environment (here, the labor market) is competitive, and they compete against men. The story in that paper, as well as Goldin and Rouse (2003), is consistent with statistical discrimination. Third, women could not be getting into decision-making positions because of other reasons, such as parental role transmission (both female and male), identity issues (Akerlof and Kranton 2000),[40] women's lower confidence (Bengtsson *et al* 2004) or, more generally, social norms.

In the previous section, we have found that female candidates that are evaluated by relatively female committees have more chances to pass a public exam than if they are evaluated by relatively male committees. We have also found that while this effect is due to both male committee members being relatively softer on female candidates and female committee members being relatively tougher on female candidates, the former accounts for about two thirds of the effect.

In this section we suggest possibilities that could be behind this result. Basically, we distinguish an irrational choice interpretation, and some other interpretations that are consistent with rational choice theory.

One the one hand, committee members might have some sort of irrational taste. For instance, it might be that committee members taste-discriminate against candidates of the same gender.

On the other hand, the result could be consistent with rational choice theory under various interpretations: committee members might be maximizing (i) the profession's quality, (ii) their individual

---

[37] Becker (1971).

[38] Bertrand and Hallock (1999) find some evidence against the 'old boys' network hypothesis in the US. They fail to find gender discrimination once they control for individual characteristics such as experience.

[39] Phelps (1972).

[40] Their model is consistent with empirical evidence that women who work more outside the home also work more inside the home. This evidence goes against Becker's (1965) theory of comparative advantage in gender roles (Akerlof and Kranton 2000, using data from the US for 1983-1992) .

utility, or (iii) their own group quality.

First, quality maximization is consistent with our results if women think that men are of higher quality than women. Lack of confidence of women—here, evaluators—might translate into female evaluators believing that the average female candidate is of lower quality than the average male candidate. If women lack confidence for whatever reason, regardless of the fact that this belief might be wrong, we would get results consistent with the hypotheses above. What could justify such a belief? It could be due to parental role transmission or, more generally, social norms. In our particular case, from Table 3 we see that male committee members are usually more highly ranked in their profession than female committee members (this is true for all exams but prosecutor exam committee members).[41] Note that this possibility would not necessarily imply the existence of statistical discrimination, as the average grade for the multiple choice is basically the same for both groups (30% for men and 29% for women), and the variance does not vary by gender either. Indeed, as long as the belief exists, we could observe findings such as our result. The underconfidence interpretation is consistent with other experimental literature, such as Bengtsson et al (2004).

Second, another possibility is that female committee members' identity suffers if they are thought to be more generous towards their own gender by their fellow committee members. If that is the case, being generous towards the opposite gender would ensure not being identified with such a *feminist* bias and hence would imply a gain in identity and higher utility.

Similarly, female committee members might get a utility gain from anti-female bias if they think that these are 'natural' attributes of (formerly only male) committee members. Thus, in order to maximize their utility, female evaluators enhance their identity as committee members by having an anti-female bias. This could in fact compensate for their own identity loss for having an occupation that does not conform to traditional gender roles.

Third, committee members might be maximizing the prestige of the whole professional collective. Suppose that there exists uncertainty about the quality of female and male members of the Judicial Corps. That is, imagine that the average quality of members is publicly known, but the individual quality is not. In that case, committee members would want to increase the average quality of their group. Considering that the gender dimension is one identity dimension (i.e., female judges might be seen differently than male judges), both men and women in committees would be interested in increasing their respective group's average quality. This would be consistent with women being tougher on female entrants (and men being tougher on male entrants).[42]

Some, or a combination, of the explanations in these three groups could be behind our findings. However, the issue is further complicated by the fact that committee formation is not given. However, it is regulated by the rules described in the appendix. Moreover, committee composition is orthogonal to the gender of evaluated candidates. All in all we cannot say more about what might be going on within committees at this point. We cannot discard non-linearities: perhaps men in committees

---

[41] This in fact comes from the fact that female committee members are younger than male committee members, but could still sustain a wrong belief.

[42] This does not take into account that both groups would end up increasing the quality of the whole group; given uncertainty, the situation would be consistent with rational choice.

behave differently when there are women in the committee as well.[43] In truth, in the near future we shall examine a more longer-term picture: that which includes data from the time the first committee that included women was formed to evaluate in public exams.

7 Conclusions

Gender parity policy is being adopted in many countries on the basis that women are underrepresented in top level positions at both the public and the private spheres. The motivation underlying the imposition of gender parity is the existence of (taste) discrimination against women. Related is the perception that there exists a 'glass ceiling'—beyond which women cannot go any further. If women are not able to break the glass ceiling, imposing gender parity at the top level should increase hiring of other women, and therefore in turn increase the percentage of women in decision-making.

This paper uses data from Spanish public exams that provide evidence that gender parity will not increase the number of women in decision-making positions, in fact, our main finding is that gender parity would reduce that number. The analysis of Spanish exams is adequate because of the characteristics of their evaluation process: first, candidates are allocated to committees randomly, which automatically eliminates concerns about endogeneity; second, the subjects and the experiment are actually taken from real life, hence avoiding the usual *caveats* of experimental work; and finally the experiment is relevant because of the importance, and magnitude, of public exams in Spain and other countries.

We use data from candidates to the Judicial Corps between 1995 and 2004, and find that female (male) candidates are more likely to be successful when evaluated by committees in which there are more women (men). In particular, we find that a female candidate's chances of passing the exam are 5.5% greater if she is evaluated by a committee with more male committee members than the median committee, than if she is evaluated by a committee with fewer male committee members than the median. Moreover, as we show in the empirical anaylsis, these differences are statistically significant. Our evidence suggests that gender parity will not increase the numbers of women making it to the elite. In fact, in the case of the Judicial Corps, imposing gender parity in the committees will reduce the number of female members: more men in our committees would have increased the number of successful women in 3.5% (1271 women would have been hired instead of 1230).[44]

These results suggest that the reason for the low numbers of women in decision-making positions does not lie with taste discrimination. The results suggests that the reason is more complicated—either consistent with statistical discrimination in connection with women's higher likelihood to withdraw from labor market activity,[45] or consistent with social norms and identity-based gender roles.[46] The main policy lesson from this paper is thus that introducing gender parity in recruiting committees

---

[43] Dufwenberg and Muren (2005) actually find an non-linear gender effect in an experiment to test for generosity and gender committee composition.

[44] These numbers have been calculated with respect to the median committee.

[45] Lazear and Rosen (1990).

[46] Akerlof and Kranton (2000).

will not increase the numbers of women in decision-making posts. Instead, a deeper understanding of factors limiting women's participation should be gathered in order to derive adequate policy. In the Spanish case particularly, a quick glance at cabinet members resumes shows the difficulty of reconciling family and career, which brings us to Lazear and Rosen (1990).[47]

Another implication of our work is that, given our finding that evaluation is sensitive to committee compostion, it is important to introduce anonymity in the selection process by public examination, so as to make selection procedures more objective.[48]

Finally, it would be interesting to gather data from exams previous to 1995, as the inclusion of more past data should be useful in order to observe any trends and understand this phenomenon better.

8   Appendix: Committee composition and formation

The exams for judges and prosecutors were separate until 2000, but since 2001 it has been joint for both types of candidates; and rules have been common since then.

For the prosecutor exam, in place until 2000, the rules were as follows. Committees evaluating prosecutor candidates had to be composed of one prosecutor from the Superior Court of Justice (who was to preside over the committee), two prosecutors, one judge, one professor in law, one lawyer (to be chosen by the Lawyers' Corps), a public defender (affiliated with the Ministry of Justice), and one prosecutor working for the Technical Secretariat in the State's General Prosecuting Office (who was to be the committee's secretary) (*Ley 50/1981, December 30*).

The prosecutor exam consisted of two oral tests. In the first one, the candidate had to develop five topics drawn from a lottery; the five topics had be developed within 75 minutes, and the candidate could not spend more than 20 minutes in each topic. The topics were in the following fields: one in general theory of law and constitutional law, two in civil law, and two in criminal law. In the second test, candidates had to develop a topic drawn from a lottery for each of these fields: civil procedural law, criminal procedural law, administrative law, commercial law, and labor law. The candidate was allowed 30 minutes before the development of each topic, but could not have access to any law textbook or document during the exam.

The date, time, and venue for the first test was to be announced at least 20 days in advance. The details for the second test were to be determined by the first evaluating committee, but there had to be at least one month between both tests (*Orden December 18, 1996*).

For the judge exam, we have to distinguish before and after 1998. Until 1997, the committee was to be presided over by the President of the Superior Court of Justice or a Judge of the Superior Court of Justice. The rest of the members had to include one judge, one prosecutor, two law professors in two different fields, one professional lawyer, and another judge, who was to be the committee's secretary

---

[47] While the eight male cabinet members have 24 children in total, the eight female cabinet members only have five in total (*El País*, October 16, 2005).

[48] The recent introduction of a multiple choice test in one type of public exam goes in the right direction.

(*Ley Orgánica 6/1985, July 1*). From 1998 until 2000, the committee was to be presided over the President of the Superior Court of Justice or a Judge of the Superior Court of Justice. The rest of the committee members had be two judges, one prosecutor, two law professors in two different fields, one lawyer with over ten years of experience, one public defender, a court secretary of first category, and another judge, member of the technical agencies in the General Council of the Judicial Corps with a law degree, who was to be the committee's secretary. Whenever it was not possible to include law full professors, associate professors were to be considered (*Ley Orgánica 5/1997, December 4*).

In the case of exams to become judge or prosecutor, in place since 2001, committees are composed by nine members. Each committe is formed by a committee president, who must be a top member of the Judicial Corps—either a judge, in the case of oddly numbered committees, or a public prosecutor, in the case of evenly numbered committees; all from a Superior Court of Justice. The president appoints the other eight members. The composition of the other eight members in the committee must include two judges, two public prosecutors, one law professor, one public defender, and one lawyer with over ten years of professional experience. Finally, the committee secretary is a court secretary of first category.

There is the possibility that one or more committees consider that the number of satisfactory candidates within the pool of evaluated candidates is lower than the number of positions that the committee has been assigned. In that case, the vacant positions are transferred to other committees for which the number of satisfactory candidates is larger than the number of positions the committee has been assigned (*Ley Orgánica 9/2000, December 22*). That is, there is not a direct trade off between any two candidates for a specific fixed number of positions, as the number of positions is in fact somewhat variable.

Since 2003, the judge and prosecutor exams consists of three stages, all of which are qualifying. First a multiple choice test, and then two oral tests. The material covered by the first and second stages is the same. The topics in the material are published in the BOE.

Court secretary exam committees are composed by seven members. The president must be a judge, and there must be a second category prosecutor, three court secretaries, one law professor, and one lawyer with more than five years of professional experience (*Ley Orgánica 19/2003, December 23*). In 1995, the exam contained a multiple choice test.

References

[1] Akerlof, George, and Rachel E. Kranton (2000), Economics and Identity, *Quarterly Journal of Economics*, Vol. 115(3), pp. 715-33.

[2] Bagüés, Manuel F. (2005), ¿Qué determina el éxito en unas oposiciones?, FEDEA Documento de Trabajo 2005-01

[3] Becker, Gary S. (1965), A Theory of the Allocation of Time, *Economic Journal* 75, pp. 493-517.

[4] Becker, Gary S. (1971), *The Economics of Discrimination* (Chicago: The University of Chicago Press)

[5] Bengtsson, Claes, Mats Persson and Peter Willenhag (2004), Gender and Overconfidence, Institute for International Economic Studies (Stockholm University), Seminar Paper No. 730

[6] Bertrand, Marianne and Kevin F. Hallock (2001), The Gender Gap in Top Corporate Jobs, *Industrial and Labor Relations Review*, Vol. 55(1), pp. 3-21

[7] Blank, Rebecca M. (1991), The Effects of Double-Blind versus Single-Blind Reviewing: Experimental Evidence from *The American Economic Review*, *American Economic Review*, Vol. 81(5), pp. 1041-67.

[8] Blau, Francine D. and Lawrence M. Kahn (1994), Rising Wage Inequality and the U.S. Gender Gap, *American Economic Review*, Vol. 84(2), pp. 23-8.

[9] Mairesse, Jacques and Laure Turner (2006), Productivity Differences across Individuals in Public Research: An Econometric Study of French Physicists' publications and citations, *Annales d'Economie et de Statistiques*, forthcoming.

[10] Ministerio de la Presidencia, *Boletín Oficial del Estado*, selected issues

[11] Ministerio de Asuntos Sociales, *V Plan de Igualdad de Oportunidades entre Mujeres y Hombres, 2003-2006*

[12] Dahlerup, Drude (2006), Women, Quotas and Politics, London: Routledge

[13] Duflo, Esther, and Raghabendra Chattopadhyay (2004), Women as Policy Makers: Evidence from a Randomized Experiment, *Econometrica*, Vol. 72(5), pp.1409-1443

[14] Dufwenberg, Martin and Astri Muren (2005), Gender Composition in Teams, *Journal of Games and Economic Behavior*, 30

[15] Frechette, Guillaume R., Francois Maniquet and Masimo Morelli (2005), Incumbents' Interests, Electoral Rules, and Endogenous Affirmative Action Laws: Gender Bias Leads to Gender Quotas, mimeo, New York University

[16] Fryer, Roland G. and Glenn Loury (2005), Affirmative Action and Its Mythology, *Journal of Economics Perspectives*, Vol. 19(3)

[17] Gneezy, Uri, Muriel Niederle and Aldo Rustichini (2003), Performance in Competitive Environments: Gender Differences, *Quarterly Journal of Economics*, Vol. 118(3), pp. 1049 - 1074

[18] Goldin, Claudia and Cecilia Rouse (2000), Orchestrating Impartiality: The Effect of "Blind" Auditions on Female Musicians, *American Economic Review*, Vol. 90(4), pp. 715-41.

[19] Grupo Especial de Trabajo sobre Buen Gobierno (2006), *Proyecto de Código Unificado de Recomendaciones de Buen Gobierno de Sociedades Cotizadas*

[20] Kelchtermans, S. and R. Veugelers (2005), Top research productivity and its Persistence: A Survival Time Analysis for a Panel of Belgian Scientists, CEPR Discussion Paper 5415

[21] Lavy, Victor (2004), Do Gender Stereotypes Reduce Girls' Human Capital Outcomes? Evidence from a Natural Experiment, mimeo, Hebrew University of Jerusalem

[22] Lazear, Edward P. and Sherwin Rosen (1990), Male-Female Wage Differentials in Job Ladders, *Journal of Labour Economics*, Vol. 8(1)

[23] Nagel, Ilene and John Hagan (1983), Gender and Crime: Offense Patterns and Criminal Court Sanctions, in *Crime and Justice: An Annual Review of Research*, pp. 91-144, ed. by M. Torny and N. Morris, Vol. 4 (Chicago: The University of Chicago Press)

[24] Pande, Rohini (2003), Can Mandated Political Representation Provide Disadvantaged Minorities Policy Influence? Theory and Evidence from India, *American Economic Review*, Vol. 93(4), pp.1132-1151

[25] Parisi, Nicolette (1982), Are Females Treated Differently? A Review of the Theory and Evidence on Sentencing and Parole Decisions, in *Judge, Lawyer, Victim, Thief*, pp. 205-20, ed. by N. Rafter and E. Stanko (Boston: Northeastern University Press)

[26] Phelps, Edmund (1972), The Statistical Theory of Racism and Sexism, *American Economic Review*, Vol. 62, pp. 659-61.

[27] Rogg, Elisabet (200?), Gender Revolution in Science, KILDEN article

[28] Wolfers, Justin (2006), Diagnosing Discrimination: Stock Returns and CEO Gender, *Journal of the European Economic Association*, 4(2-3).

	Judge	Prosecutor		Judge & Prosecutor						Court Secretary			
Year	2000	1999	2000	2001	2002	2003	2004	2005	2006	2000	2001	2002	2003
Female (%)	68.3	69.5	69.2	68.2	68	68.6	68.2	68.7	68.8	71.9	71	71.3	72.2
Years of experience													
One			68.9	79.9	79.0	77.6	78.0	82.3	75.6		53.7	61.1	47.4
Two					59.0	60.0	59.9	65.7	62.4			36.7	29.5
Three						42.9	44.8	48.7	48.5				17.9
Four							30.7	35.3	34.9				
Five								23.4	25.3				
Six									16.7				
Number of candidates	5575	5078	4474	5374	5122	4973	4731	4084	3918	1226	1523	1680	2246

Notes: experience calculated from looking at names repeated overt the years. e.g. candidates with one year experience are candidates who are taking the exam for at least the second time.

Table 1: Descriptive statistics - candidate characteristics, by type of exam and year

	Judge	Prosecutor	Judge & Prosecutor	Court Secretary
Available years	1995.1996,1998, 1999,2000	1995,1997,1999, 2000	2001,2002,2003, 2004	1995,1998,1999, 2000,2001,2002, 2003
Average number of committees	10	9	8.25	4.43
Members per committee	9 (1995,1996), 10	8	9	7
President's occupation	Superior Court Judge	Superior Court Prosecutor	Superior Court Judge or Prosecutor[1]	Judge[3]
Other members' occupations				
Judge	$3/4^{2}$	1	2	1/0
Prosecutor	1	3	2	0/1
State lawyer	1	1	1	1/0
Lawyer	1	1	1	1
Professor	2	1	1	1
Court secretary	0/1	0	1	2/3

Notes: 1\Every odd (even) committee is presided by a judge (prosecutor), alternatively. In our data set, a prosecutor (judge) has been the committee's president 48.5% (51.5%) of the time. 2\The composition of committees for the judge and court secretary exams changed in 1998. We provide the profession composition before (first number) and after (second number) the change, where applicable. 3\From a Superior Court of Justice until 1997. For committee composition rules see the appendix.

Table 2: Committee composition, by type of exam

	Judge		Prosecutor		Judge & Prosecutor		Court Secretary	
	Female	Male	Female	Male	Female	Male	Female	Male
Age	41.3	48.8	38.7	44.8	48.3	52	40.2	45.4
	(8.67)	(11.5)	(6.71)	(11.3)	(7.95)	(12.3)	(5.23)	(10.9)
Entry age	29.4	30.4	30	29.2	29.7	29.8	28.5	31
	(4.37)	(5.81)	(5.19)	(4.31)	(5.18)	(4.96)	(3.53)	(6.39)
Relative rank	0.51	0.69	0.46	0.41	0.47	0.63	0.57	0.61
	(0.22)	(0.26)	(0.28)	(0.20)	(0.19)	(0.27)	(0.25)	(0.31)
Number of observations	110	373	78	210	69	228	104	162

Notes: standard errors in parentheses. The relative rank index has been calculated based on the member's ranking within her age group (for Judicial Corps members only). A higher number means higher in the rank. Entry age is the age at which the member entered the Judicial Corps, and therefore does not apply to non-members (i.e. professors and lawyers).

Table 3: Descriptive statistics - committee members, by type of exam

	Dependent variable:					
	Female successful candidates			Male successful candidates		
	(1)	(2)	(3)	(4)	(5)	(6)
Female share in the committee	-1.85			2.11*		
	(1.47)			(1.23)		
Female share in the committee, members of Judicial Corps only		-3.02***			1.70**	
(including the president)		(0.83)			(0.73)	
Female share in the committee, members of Judicial Corps only			-1.72**			0.70
(excluding the president)			(0.74)			(0.85)
Gender of president (=1 if female)			-2.72			3.30***
			(1.96)			(1.20)
Female share in the committee, not members of Judicial Corps		0.39	0.29		0.33	0.46
		(1.04)	(1.02)		(0.30)	(0.97)
Mean age of committee members			-0.48			0.86*
			(0.83)			(0.44)
Mean age of committee members, squared			0.004			-0.01*
			(0.01)			(0.004)
Mean entry age of committee members			0.14			-0.01
			(0.13)			(0.18)
Mean ranking of committee			8.16*			-1.66
			(4.19)			(5.03)
Exam type dummies*Year dummies	yes	yes	yes	yes	yes	yes
Number of observations	150	150	150	150	150	150

Notes: robust standard errors clustered at the exam level in parentheses. Regressions include a dummy for the first committee, which sometimes evaluates fewer candidates than other committees (not shown). Age, age of entry and ranking is only available for Judicial Corps members; entry age is the age at which the member entered the Judicial Corps.*significant at 10%; **significant at 5%; ***significant at 1%.

Table 4: Gender composition of committee and female and male successful candidates

	Dependent variable:					
	Total successful candidates			Female successful candidates(%)		
	(1)	(2)	(3)	(4)	(5)	(6)
Female share in the committee	0.26			-0.09*		
	(1.44)			(0.05)		
Female share in the committee, members of Judicial Corps only		-1.32			-0.13**	
(including the president)		(0.80)			(0.04)	
Female share in the committee, members of Judicial Corps only			-1.02			-0.08**
(excluding the president)			(0.98)			(0.03)
Gender of president (=1 if female)			0.57			-0.12***
			(2.84)			(0.02)
Female share in the committee, not members of Judicial Corps		0.72	0.74		0.03	0.02
		(0.94)	(0.99)		(0.05)	(0.04)
Mean age of committee members			0.38			-0.05**
			(0.85)			(0.02)
Mean age of committee members, squared			-0.004			0.001**
			(0.01)			(0.001)
Mean entry age of committee members			0.13			0.001
			(0.17)			(0.01)
Mean ranking of committee			6.51			0.11
			(6.80)			(0.15)
Exam type dummies*Year dummies	yes	yes	yes	yes	yes	yes
Number of observations	150	150	150	150	150	150

Notes: robust standard errors clustered at the exam level in parentheses. Regressions on the first panel include a dummy for the first committee, which sometimes evaluates fewer candidates than other commitees (not shown). Age, age of entry and ranking is only available for Judicial Corps members; entry age is the age at which members entered the Judicial Corps. *significant at 10%; **significant at 5%; ***significant at 1%.

Table 5: Gender composition of committee and total successful candidates and percentage of female successful candidates

	Candidates who passed stage one (multiple choice test)		
	Probit		
	Dependent variable =1 if success, =0 otherwise		
	Exam	Stage two only	Stage three only
	(1)	(2)	(3)
Female candidate (=1 if female)	0.026**	0.026	0.08
	(0.011)	(0.025)	(0.06)
Female share in the committee	0.05	-0.05	0.31***
	(0.04)	(0.06)	(0.07)
Female candidate*	-0.065**	0.05	-0.38***
Female share in the committee	(0.035)	(0.07)	(0.12)
Multiple choice mark	0.027**	0.05***	0.04
	(0.014)	(0.02)	(0.05)
Multiple choice mark–squared	-0.0001	-0.0002	-0.0002
	(0.0001)	(0.0001)	(0.0004)
Experience controls	yes	yes	yes
Year fixed effects	yes	yes	yes
Province fixed effects	yes	yes	yes
Predicted probability	0.07	0.22	0.38
Number of observations	5185	5251	1323

Notes: robust standard errors, clustered at the committee level in parentheses. The exam consists of three stages, all of them qualifying. The first stage is a multiple choice test. Stages two and three are both oral tests. Here we consider candidates who passed stage one and hence are evaluated by committees. *significant at 10%; **significant at 5%; ***significant at 1%.

Table 6: Probability of success by gender and gender composition of committee, 2003-2005