Promotion through Connections: Favors or Information?

Yann Bramoullé and Kenan Huremović*

January 14, 2018

Abstract

Connections appear to be helpful in many contexts such as obtaining a job, a promotion, a grant, a loan or publishing a paper. This may be due either to favoritism or to information conveyed by connections. Attempts at identifying both effects have relied on measures of true quality, generally built from data collected long after promotion. Building on earlier work on discrimination, we propose a new method to identify favors and information from data collected at time of promotion. Under natural assumptions, we show that promotion decisions for connected candidates look more random to the econometrician, due to the information channel. We derive new identification results and show how to use heteroscedastic probit models to estimate the strength of the two effects. We apply our method to the data on academic promotions in Spain studied in Zinovyeva and Bagues (2015). We find evidence that connections both convey information and attract favors. Our results are consistent with evidence obtained from data collected five years after promotion.

Keywords: Promotion, Connections, Social Networks, Favoritism, Information.

JEL classification: C3, I23, M51.

*We thank participants in seminars and conferences and Olivier Chanel, Russell Davidson, Bernard Fortin, Habiba Djebbari, Bruno Decreuse, Mark Rosenzweig, Marc Sangnier, Christian Schluter, Adam Szeidl, Paolo Zacchia and Natalia Zinovyeva for helpful comments and suggestions. For financial support, Yann Bramoullé thanks the European Research Council (Consolidator Grant n. 616442). Bramoullé: Aix-Marseille University (Aix-Marseille School of Economics) & CNRS, yann.bramoulle@univ-amu.fr; Huremović: IMT School for Advanced Studies, kenan.huremovic@imtlucca.it.

1 Introduction

Connections appear to be helpful in many contexts such as obtaining a job, a promotion, a grant, a loan or publishing a paper.[1] Two main mechanisms help explain these wide-ranging effects. On the one hand, connections may convey information on candidates, projects and papers that helps recruiters, juries and editors make better decisions. On the other hand, decision-makers may unduly favor connected candidates, leading to worse decisions.[2] These two mechanisms have opposite welfare implications and empirical researchers have been trying to tease out the different forces behind the impacts of connections. Existing studies generally do so by building measures of candidates' "true" quality. Researchers then compare the quality of connected and unconnected successful candidates. Information effects are likely to dominate if connected successful candidates have higher quality; favors are likely to dominate if connected successful candidates have lower quality. For instance, articles published in top economics and finance journals by authors connected to editors tend to receive more citations, a sign that editors use their connections to identify better papers (Brogaard et al. (2014)). By contrast, Full Professors in Spain who were connected to members of their promotion jury publish less after promotion (Zinovyeva and Bagues (2015)), consistent with favoritism.

This empirical strategy, while widely used, carries three important limitations. First, building a measure of true quality may not be easy or feasible. Looking at researchers' publications or articles' citations requires a long enough time lag following promotion or publication. And such measures are in any case imperfect proxies of quality. Second, identification is only valid if the impact of success on measured quality is the same for connected and unconnected successful candidates, see Zinovyeva and Bagues (2015, p.283). This assumption is critical but not neces-

[1]The literature on jobs and connections is large and expanding. Recent references include Beaman and Magruder (2012), Brown et al. (2016), Hensvik and Skans (2016), Pallais and Sands (2016). On promotions, see Combes et al. (2008), Zinovyeva and Bagues (2015). On grants, see Li (2017). On loans, see Engelberg et al. (2012). On publications, see Brogaard et al. (2014), Colussi (2017), Laband and Piette (1994).

[2]Favor exchange within a group might increase the group's welfare to the detriment of society, see Bramoullé and Goyal (2016). In this paper, we focus on the immediate negative implications of favoritism.

sarily plausible, and usually cannot be tested. Third, connections may both convey information and attract favors. This empirical strategy may allow researchers to identify which effect dominates; it does not allow them to estimate their relative strengths.[3]

We develop a new method to identify why connections matter, building on earlier work on discrimination. Our method allows researchers to estimate the magnitudes of the two effects, without relying on measures of true quality. Our method exploits data collected at time of application: observable characteristics of candidates and whether they were successful.[4] It looks for indirect signs of the two effects on the relationship between candidates' observables and success.

Consider candidates applying for promotion. They are evaluated by a jury and some candidates are, by chance, connected to jury members. When connections convey information, the jury has an extra signal regarding connected candidates' ability. This signal is unobserved by the econometrician and could be positive or negative. To the econometrician, then, the promotion decision looks more random for connected candidates.[5] We show how the strength of the information channel can be recovered, under appropriate assumptions, from this *excess variance* in the latent error of connected candidates. Favors can then be recovered by estimating and comparing the *promotion thresholds* faced by connected and unconnected candidates. Favors lead to systematic biases in evaluation and the difference between promotion thresholds measures the magnitude of the underlying favors.

We develop an econometric framework based on normality. We make use of probit models with heteroscedasticity to detect and estimate excess variance. We clarify the conditions under which favors and information are identified. Identification fails to hold if both effects depend in an arbitrary way on candidates' observables

[3]In a context of grant applications, Li (2017) develops a new method to recover the respective strengths of favors and information. Her method relies on observed measures of true quality and jury evaluations.

[4]As with any analysis of the reasons behind the effect of connections, our method requires some kind of exogenous shocks on connections. In other words, we assume that the problem of identifying whether connections matter has been solved and focus on the question of identifying why they matter.

[5]A similar idea underlies Theorem 4 in Lu (2016); we discuss this relation in more detail below.

(Proposition 1). Identification holds, however, under slight restrictions on this dependence, for instance in the presence of an exclusion restriction or under linearity assumptions (Theorem 1).

We then reanalyze the data on academic promotions in Spain assembled by Zinovyeva and Bagues (2015) and containing information on all candidates for promotion to Associate and Full Professor in the Spanish academic system between 2002 and 2006. To be promoted, candidates had to pass a highly competitive nation-wide exam. They were evaluated by a jury whose members were picked at random from a pool of eligible evaluators, providing exogenous shocks on connections. The data covers six types of connection between candidates and evaluators, classified as weak or strong. From data at time of promotion, Zinovyeva and Bagues (2015) estimate the causal impacts of connections. They find positive and significant impacts of both weak and strong ties on promotion for candidates at both Associate and Full Professor level.

We investigate the reasons behind these impacts on the same data. Applying our new method, we find strong evidence of information effects associated with both weak and strong ties at Associate Professor level, when the uncertainty on candidates' academic ability is still strong. We also detect favors associated with strong ties, but not with weak ties, at that level. By contrast, we do not detect information effects at Full Professor level, when uncertainty on candidates is presumably low. We detect strong favors associated with both weak and strong ties at that level, consistent with widespread favor exchange operating in the Spanish academic system at the time. These results, obtained through our indirect method from data at time of promotion, are consistent with results obtained through quality measures from data collected five years after promotion, see Section VI.

Our analysis contributes to a growing empirical literature on connections. We develop the first empirical method to identify favors and information effects from data collected at time of promotion and apply it to analyze academic promotions in Spain. This method could be applied in many other contexts, and could notably be used to cross-validate results obtained from quality measures.

Our analysis builds on ideas first identified in the literature on discrimination. Heckman and Siegelman (1993) and Heckman (1998) clarify key implications of differences in the variances of unobservables across groups. They show that variance differences invalidate the use of standard models of binary outcomes to detect discrimination. Neumark (2012) shows how probit models with heteroscedasticity can help address this issue. He reanalyzes the data from Bertrand and Mullainathan (2004) and finds stronger evidence for race discrimination than in the original study, once differences in variance across racial groups are accounted for. We adapt and extend these ideas to the study of connections. We show that differences in variance help identify the informational content of connections, an idea consistent with Theorem 4 in Lu (2016). Lu (2016) provides a theoretical analysis of random choice under private information. He shows that better private information generates choices that look more dispersed to the econometrician. To our knowledge, we are the first to implement this insight in an empirical context. We obtain novel identification results. The first part of Theorem 1, on exclusion restrictions, extends the identification argument of Neumark (2012). The second part of Theorem 1, on linearity, shows that identification may hold even without exclusion restrictions. We provide the first empirical application of these ideas to the study of the impact of connections.

The paper proceeds as follows. The next section illustrates the identification strategy with the help of a simple model. Section III introduces the general model and establishes formal identification results. Section IV presents the data. Section V discusses key features of the empirical implementation. Section VI presents empirical results. Section VII concludes.

2 A simple model

In this Section, we introduce a simple model to explain and illustrate our identification strategy. This model is similar to models analyzed in Heckman and Siegelman (1993, Appendix 5.D), Neumark (2012) and Zinovyeva and Bagues (2015, Section

I). We develop our general model and derive formal identification results in Section III.

Candidates apply for promotion and are evaluated by a jury that makes promotion decisions. We assume that the jury grades candidates and that candidates with higher grades are promoted.[6] These grades may be affected by connections to jury members, as described below. Let a_e be the exam-specific promotion threshold: a candidate is promoted iff her grade is higher than or equal to a_e. This threshold may notably depend on the number of candidates applying for promotion in that wave and discipline.

We assume that candidate i's true ability a_i can be decomposed into three parts:

$$a_i = \beta \mathbf{x}_i + u_i + v_i \tag{1}$$

where $\mathbf{x}_i \in \mathbb{R}^m$ denotes a vector of m characteristics observed by the econometrician and the jury, u_i is unobserved by both the econometrician and the jury, and v_i is observed by the jury but not the econometrician. We assume that $E(u_i|\mathbf{x}_i) = E(v_i|\mathbf{x}_i) = 0$, which is without loss of generality in a linear context.[7] Thus, u_i and v_i represent parts of unobserved characteristics that cannot be explained by observables. Assume further that $E(u_i|v_i) = 0$ and that unobservables are normally distributed: $u_i \sim N(0, \sigma_u^2)$ and $v_i \sim N(0, \sigma_v)$. Denote by Φ the cumulative density function of a normal variable with mean 0 and variance 1. Denote by $c_i \in \{0, 1\}$ a binary variable describing whether candidate i is connected ($c_i = 1$) or unconnected ($c_i = 0$) to the jury.

Consider an unconnected candidate first. We assume that her grade is equal to the jury's expectation of her ability $E(a_i|x_i, v_i) = \beta \mathbf{x}_i + v_i$. Unconnected candidate i is then promoted if and only if $\beta \mathbf{x}_i + v_i \geq a_e$. From the econometrician's point of view, the probability of an unconnected candidate with characteristics $\mathbf{x}_i$ being

[6] We develop our approach under the assumption that the econometrician does not have data on jury evaluations.

[7] If $E(u_i|\mathbf{x}_i) = \mathbf{x}_i\beta_u$, define $\hat{u}_i = u_i - E(u_i|\mathbf{x}_i)$ and similarly for $\hat{v}_i$. Note that $E(\hat{u}_i|\mathbf{x}_i) = E(\hat{v}_i|\mathbf{x}_i) = 0$. This yields $a_i = \mathbf{x}_i(\beta + \beta_u + \beta_v) + \hat{u}_i + \hat{v}_i$, which is then equivalent to equation (1).

promoted is equal to:

$$p(y_i = 1|c_i = 0, \mathbf{x}_i) = \Phi(\frac{\beta\mathbf{x}_i - a_e}{\sigma_v}) \tag{2}$$

where $y_i = 1$ if candidate i obtains the promotion and 0 otherwise.

Next, consider a connected candidate. We make two simplifying assumptions in this Section. We assume, first, that being connected to the jury is random. In the empirical application, this holds conditional on the expected numbers of connections, see Section IV. This implies that connected and unconnected candidates have the same distributions of observables and unobservables and, in particular, $E(u_i|c_i) = E(v_i|c_i) = 0$. Second, we neglect issues related to the number and types of connections. These issues are accounted for in our general model, see Section III.

Being connected to the jury has two different effects on grades. On the one hand, the jury obtains additional information on the candidate's ability. We assume that the jury observes a noisy signal $\theta_i = u_i + \varepsilon_i$ where $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$, and updates its belief on the candidate's ability based on this additional information. On the other hand, the jury may want to favor the connected candidate. We assume that favors take the shape of a grade premium B due to connections.

A connected candidate's grade is thus equal to her expected ability $E(a_i|x_i, v_i, \theta_i) = \beta\mathbf{x}_i + E(u_i|\theta_i) + v_i$ plus the bias from favors B. Since $E(u_i|\theta_i) = \frac{\sigma_u^2}{\sigma_u^2+\sigma_\varepsilon^2}\theta_i$, connected candidate i is promoted if and only if $\beta\mathbf{x}_i + \frac{\sigma_u^2}{\sigma_u^2+\sigma_\varepsilon^2}\theta_i + v_i + B \geq a_e$. From the econometrician's point of view, the signal θ_i is included in the latent error and generates extra variance on the jury's decision. Variance of the latent error is now equal to $\sigma_v^2 + \frac{\sigma_u^4}{\sigma_u^2+\sigma_\varepsilon^2}$. *Let* $\sigma^2 = 1 + \frac{\sigma_u^4}{\sigma_v^2(\sigma_u^2+\sigma_\varepsilon^2)} > 1$ *denote the excess variance of the latent error of connected candidates compared to unconnected ones.* The probability of a connected candidate being promoted is equal to:

$$p(y_i = 1|c_i = 1, \mathbf{x}_i) = \Phi(\frac{\beta\mathbf{x}_i + B - a_e}{\sigma\sigma_v}) \tag{3}$$

Comparing equations (2) and (3), we see that information and favors have different impacts on the probability to be promoted. When a jury has better information on connected candidates, this reduces the magnitude of the impact of observable characteristics on the likelihood to be promoted. By contrast, favors lead to a shift in the effective promotion threshold, from a_e to $a_e - B$, leaving the impact of observables unchanged.

We illustrate these effects in Figure 1. The solid black curve depicts $p_u(y_i = 1|\mathbf{x}_i)$, the probability of an unconnected candidate being promoted as a function of observed ability. The dashed curve depicts the probability of a connected candidate being promoted when only information effects are involved. Note that the whole curve is less steep. The observed probability of being promoted varies less with observed ability. Formally, an increase in σ leads to a second-order stochastic dominance shift of the whole curve. This also implies that the apparent impact of connections is negative for very good candidates for which $\beta\mathbf{x}_i \geq a_e$. This apparent negative impact is due to an asymmetry in the effects of good and bad news on the candidate's chances of being promoted. While good news does not improve already good chances by much, bad news significantly reduces the chances of good candidates. On average, more accurate information conveyed by connections thus reduces the observed probability of being promoted for very good candidates.

The short-dashed curve depicts the probability of a connected candidate being promoted when only favors are involved. The curve is now translated to the left, inducing a first-order stochastic dominance shift. The shape of the whole curve is preserved. The apparent impact of connections is now positive for all candidates. Finally, the grey curve depicts $p_c(y_i = 1|\mathbf{x}_i)$ when both effects are involved.

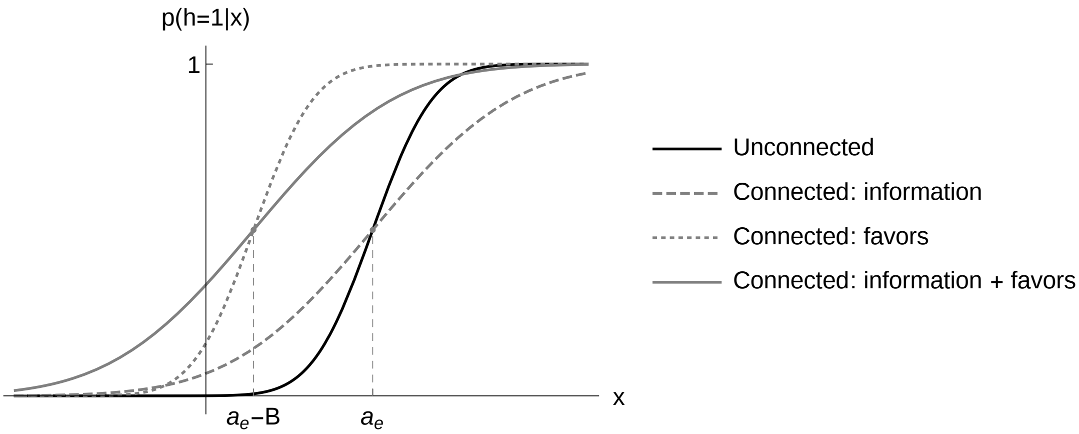


Figure 1: Effects of a connection on promotion probability

Both effects can thus be identified from data on promotion.[8] Differences in observables' impacts on promotion between the connected and the unconnected can be used to recover information effects. Differences in estimated promotion thresholds between the connected and the unconnected can then be used to recover favors. From an econometric point of view, the differential information that the jury has on connected candidates generates a form of *heteroscedasticity.* The latent error has a higher variance for the connected than for the unconnected. This property allows us to rely on standard techniques developed to analyze heteroscedasticity in probit estimations in our empirical analysis below.

To sum up, both information effects and favors can be identified from data on promotion in a simple model where the two effects are constant. We extend this model and develop our econometric framework in the next Section.

3 Identification

We now develop our general framework. We maintain the assumption that connections are random and extend the model in three directions. We incorporate baseline heteroscedasticity, varying information and varying favors. Information effects and favors may notably depend on the number and types of a candidate's connections to the jury. In line with the empirical application, we consider two types here - strong and weak ties; the framework and results directly extend to a finite number of types. Denote by n_{iS} and n_{iW} the number of strong and weak ties that candidate i has to the jury.

First, we assume that the variance of v_i may depend on i's observables $\boldsymbol{x}_i$. Thus, $v_i \sim N(0, \sigma_v^2(\mathbf{x}_i))$. In the empirical analysis, we adopt standard assumptions regarding heteroscedasticity in probit regressions, see Section IV. To state our identification results below, we only require that such baseline heteroscedasticity does not raise identification problems in standard probit estimations. More pre-

[8] Formally, identification in this model holds under the standard assumption that $\sigma_v = 1$ and is a direct consequence of Theorem 1 below.

cisely, consider unconnected candidates. We have: $p(y_i = 1|n_{iS} = n_{iW} = 0, \mathbf{x}_i) = \Phi[(\beta\mathbf{x}_i - a_e)/\sigma_v(\mathbf{x}_i)]$. We assume that β and $\sigma_v(.)$ are identified from the sample of unconnected candidates.[9]

Second, we assume that the private signal received by the jury on a connected candidate may depend on the candidate's number and types of connections and on his other observable characteristics. Denote by $\sigma \geq 1$ the excess variance in latent error generated by this signal. We now have $\sigma = \sigma(n_{iS}, n_{iW}, \mathbf{x}_i)$ where, by assumption, $\sigma(0, 0, \mathbf{x}_i) = 1$. Third, the bias from favors B may also depend on the number and types of links and on the candidate's characteristics: $B = B(n_{iS}, n_{iW}, \mathbf{x}_i)$, with $B(0, 0, \mathbf{x}_i) = 0$. While we generally expect both σ and B to be increasing in the number of connections, we do not impose it in what follows. In the end, the probability of being promoted, conditional on connections and observables, is equal to:

$$p(y_i = 1|n_{iS}, n_{iW}, \mathbf{x}_i) = \Phi[\frac{\beta\mathbf{x}_i + B(n_{iS}, n_{iW}, \mathbf{x}_i) - a_e}{\sigma_v(\mathbf{x}_i)\sigma(n_{iS}, n_{iW}, \mathbf{x}_i)}] \tag{4}$$

The simple model presented in Section II is a particular case with $\sigma_v(\mathbf{x}_i) = \sigma_v$ and $B(n_{iS}, n_{iW}, \mathbf{x}_i) = B$ and $\sigma(n_{iS}, n_{iW}, \mathbf{x}_i) = \sigma$ if $n_{iS} + n_{iW} \geq 1$.

Under which conditions is the general model identified? Note that our identification strategy will not work without some form of restriction on $B(.)$ and $\sigma(.)$. If bias B varies with observable x_i^k in a direction opposite from direct effect β_k, this leads to an apparent reduction in the impact of x_i^k on the likelihood of a connected candidate being promoted. If this happens on all observables and without further restrictions, it prevents the identification of the information effect. We next state this negative result and derive a formal proof in the Appendix.

Proposition 1

Consider model (4). Suppose that the accuracy of the signals conveyed by connections and the bias from favors depend in an arbitrary way on connections and on other observable characteristics of candidates. Then, favors and information effects cannot

[9] As is well-known, a probit model with coefficients (β, a_e) and variance $\sigma_v(\mathbf{x}_i)$ cannot be distinguished from one with coefficients $(\lambda\beta, \lambda a_e)$ and variance $\lambda\sigma_v$. We therefore adopt the classical normalization that $\sigma_v(\mathbf{0}) = 1$ in our econometric specifications.

be identified from promotion data.

We now derive our main result. We show that identification holds under mild restrictions on bias and excess variance. We consider two types of restrictions: exclusion restrictions and parametric assumptions.

Theorem 1

Consider model (4).

(Exclusion restriction). Suppose that characteristic k leaves σ and B unaffected and that $\beta_k \neq 0$. Then, the model is identified and the functions $\sigma(n_{iS}, n_{iW}, \mathbf{x}_i^{-k})$ and $B(n_{iS}, n_{iW}, \mathbf{x}_i^{-k})$ are non-parametrically identified.

(Linearity). Suppose that $\ln(\sigma(n_{iS}, n_{iW}, \mathbf{x}_i)) = \delta(n_{iS}, n_{iW})\mathbf{x}_i$ and $B(n_{iS}, n_{iW}, \mathbf{x}_i) = \gamma_0(n_{iS}, n_{iW}) + \gamma_1(n_{iS}, n_{iW})\mathbf{x}_i$, with $\gamma_0(0,0) = 0$ and $\delta(0,0) = \gamma_1(0,0) = \mathbf{0}$. Then, the model is identified and the functions $\delta(n_{iS}, n_{iW})$, $\gamma_0(n_{iS}, n_{iW})$ and $\gamma_1(n_{iS}, n_{iW})$ are non-parametrically identified.

To see why the first part of Theorem 1 holds, suppose that σ and B do not depend on x_i^k. From data on the unconnected, we can recover β_k, the direct effect of x_i^k on grade, and $\sigma_v(.)$. Take candidates with numbers of connections n_{iS} and n_{iW} and with other characteristics $\mathbf{x}_i^{-k}$. From data on these candidates, we can recover the heteroscedasticity-corrected impact of x_i^k on grade, equal to $\beta_k/\sigma(n_{iS}, n_{iW}, \mathbf{x}_i^{-k})$. If $\beta_k \neq 0$, we obtain $\sigma(n_{iS}, n_{iW}, \mathbf{x}_i^{-k})$. The bias $B(n_{iS}, n_{iW}, \mathbf{x}_i^{-k})$ can then be obtained as the difference in inferred promotion thresholds between unconnected candidates and candidates with ties n_{iS} and n_{iW} and characteristics $\mathbf{x}_i^{-k}$.

Therefore, our identification strategy works as long as one exclusion restriction is present in the model. As with instrumental variables, the excluded variable should have a direct impact on unconnected candidates' likelihood of being promoted and should not directly affect the accuracy of the signals conveyed by connections nor the bias from favors they may generate. In particular, a model where excess variance and bias from favors only depend on numbers of connections is identified. We estimate several variants of such models in the empirical analysis below.

Even without exclusion restrictions, the model can still be identified thanks to functional form assumptions. The second part of Theorem 1, proved in the Appendix, shows that this notably holds when excess variance is log-linear in observables while bias is an affine function of observables. In this case, again, dependence on connections can be arbitrary and is fully identified. To achieve non-parametric identification in practice may, of course, require a very large number of observations. In the empirical analysis below, we adopt standard parametric assumptions. All models estimated in Section VI are covered by Theorem 1.

4 Data

We apply our framework to the data on academic promotions in Spain assembled and studied by Zinovyeva and Bagues (2015). We describe the main features of the data here and refer to their study for details. From 2002 to 2006, academics in Spain seeking promotion to Associate Professor (*profesor titular*) or Full Professor (*catedrático*) first had to qualify in a national exam (*habilitaćion*). All candidates in the same discipline in a given wave were evaluated by a common jury composed of 7 members, responsible for allocating a predetermined number of positions. These exams were highly competitive and obtaining the national qualification essentially ensured promotion. A central feature of this system was that jury members were picked at random from a pool of eligible evaluators. The random draw was actually carried out by Ministry officials. The data contains information on all candidates for academic promotion during that period, their connections to eligible evaluators and to jury members, and their success or failure in the national exam.

Overall, there are $31,243$ applications for 967 exams: $17,799$ applications for 465 exams for Associate Professor (AP) positions and $13,444$ for 502 exams for Full Professor (FP) positions. We have information on candidates' demographics and academic outcomes at time of application. Observable characteristics include gender, age, whether the candidate obtained his PhD in Spain, number of publications, number of publications weighted by journal quality, number of PhD students supervised, number of PhD committees the candidate was on, and number of previous attempts at promotion. Table 1 provides descriptive statistics. Standards regarding research output may of course differ between disciplines. To analyze applications in a common framework, we follow Zinovyeva and Bagues (2015) and normalize research indicators to have mean 0 and variance 1 within exams.[10] The

Table 1: Descriptive statistics: Observables

	All	AP	FP	Eng.	H&L	Sci.	Soc. Sci.
Female	0.34	0.40	0.27	0.21	0.45	0.30	0.39
	(0.47)	(0.49)	(0.44)	(0.41)	(0.50)	(0.46)	(0.49)
Age	41.21	37.49	46.14	38.74	41.86	41.97	40.39
	(7.59)	(6.41)	(6.06)	(7.07)	(7.62)	(7.55)	(7.54)
PhD in Spain	0.78	0.83	0.70	0.83	0.77	0.76	0.78
	(0.42)	(0.37)	(0.46)	(0.38)	(0.42)	(0.43)	(0.42)
Past Experience	0.81	0.73	0.91	0.85	0.63	0.89	0.88
	(1.27)	(1.27)	(1.26)	(1.36)	(0.94)	(1.40)	(1.30)
Publications	12.84	8.12	19.09	7.76	11.45	16.99	9.22
	(18.31)	(14.06)	(21.18)	(12.88)	(11.39)	(24.10)	(11.61)
AIS	0.72	0.70	0.74	0.52	-	0.80	0.62
	(0.53)	(0.57)	(0.48)	(0.37)	-	(0.51)	(0.75)
PhD Students	1.00	0.24	2.00	0.83	0.61	1.45	0.66
	(2.11)	(0.88)	(2.75)	(1.61)	(1.63)	(2.60)	(1.58)
PhD Committees	3.61	0.88	7.23	2.40	3.04	4.81	2.67
	(6.76)	(2.55)	(8.65)	(4.42)	(5.99)	(8.21)	(4.99)
Observations	31,243	17,799	13,444	4,783	9,005	12,858	4,597

Notes: Average values of observable characteristics at time of exam. Standard deviation in parentheses. FP and AP stand for exams for Full Professor and Associate Professor positions respectively. Eng., H&L, Sci., and Soc. Sci. are abbreviations for Engineering, Humanities and Law, Sciences, and Social Sciences, which are 4 broad scientific areas in our sample. AIS is the sum of international publications weighted by corresponding Article Influence Scores. The table replicates Table 2 in Zinovyeva and Bagues (2015).

data also contain information on six types of link between candidates and evaluators. We adopt Zinovyeva and Bagues (2015)'s classification of these links in strong and weak ties.[11] A candidate is said to have strong ties to his PhD advisor(s), to his coauthors and to his colleagues. He has weak ties with members of his own PhD

[10]We also normalize age and past experience to have mean 0 within exams.

[11]The data also contains information on indirect connections, for instance when a candidate and an evaluator had a common member on their PhD committees. Zinovyeva and Bagues (2015) do not find any effect of indirect connections and we do not include them in our analysis.

committee, with members of the PhD committees of his PhD students and with his fellow PhD committee members.[12] Overall, 34.8% of candidates have at least one strong connection with a member of their jury and 20.6% have at least one weak connection. Table 2 provides further information on connections.

Table 2: Descriptive statistics: Connections

	All	AP	FP	Eng.	H&L	Sci.	Soc. Sci.
Strong connections	31.71	29.08	35.18	37.78	27.65	31.13	34.94
Advisor	3.17	2.97	3.43	4.60	3.29	2.43	3.50
Coauthor	5.44	3.26	8.32	6.10	2.84	7.44	4.24
Colleague	29.71	27.74	32.31	36.02	26.15	28.50	33.46
Weak connections	18.79	7.33	33.97	17.06	23.63	16.43	17.71
Own PhD committee member	7.08	5.31	9.43	8.05	10.22	4.39	7.48
His PhD students' committee member	4.45	0.69	9.42	4.70	4.81	4.27	3.96
Fellow PhD committee member	11.65	1.82	24.66	8.84	13.90	11.59	10.31

Notes: The percentage of candidates with at least one connection to the jury. The table replicates Table 3 in Zinovyeva and Bagues (2015).

5 Empirical Implementation

We now apply our identification strategy to the data on academic promotions in Spain. We discuss three key features of the empirical implementation: the random assignment of evaluators; the exam-specific promotion thresholds; and the specific models being estimated.

5.1 Random assignment of jury members

Our identification result, Theorem 1, relies on the assumption that the distribution of unobservables for candidates with connections (n_{iS}, n_{iW}) does not depend on (n_{iS}, n_{iW}). In the data, random assignment of jury members ensures that this holds *conditional on the expected number of connections to the jury.* That is, candidates may vary in the extent of their connections to eligible evaluators. From the number

[12] A connection which is both strong and weak is classified as strong.

Table 3: Balance tests

	AIS	Publications	PhD students	PhD committees	Past experience
	Without controls for the expected number of connections				
Strong	0.009	0.010	-0.017^{***}	-0.012^{*}	0.009
	(0.007)	(0.007)	(0.006)	(0.007)	(0.008)
Weak	0.007	0.051^{***}	0.180^{***}	0.298^{***}	0.027^{***}
	(0.007)	(0.008)	(0.010)	(0.011)	(0.007)
	Including controls for the expected number of connections				
Strong	-0.001	-0.011	0.002	-0.006	-0.004
	(0.010)	(0.011)	(0.010)	(0.010)	(0.012)
Weak	-0.005	-0.008	0.013	0.010	0.003
	(0.011)	(0.013)	(0.016)	(0.016)	(0.012)
Observations	31,243	31,243	31,243	31,243	31,243

Notes: Results of 10 regressions of observables (columns) on the number of strong and weak connections to the jury (rows). In the upper panel regressions, we do not control for the expected number of connections. The lower panel regressions include controls for the expected number of strong connections to the jury and the expected number of weak connections to the jury. OLS estimates. Standard errors clustered on the exam level are in parentheses. *p<0.1; **p<0.05; ***p<0.01.

of eligible evaluators and the numbers of weak and strong ties of a candidate to eligible evaluators, we can compute the expected numbers of connections of this candidate to the jury. Conditional on these expected numbers, actual numbers of connections are then effectively random.

We present the corresponding balance tests in Table 3.[13] Controlling for candidates' expected numbers of connections, we do not detect significant residual correlations between observable characteristics and actual number of connections. Therefore, a conditional version of Theorem 1 holds in this context. The unconnected candidates' probability of being promoted $p(y_i = 1|n_{iS} = n_{iW} = 0, En_{iS}, En_{iW}, \mathbf{x}_i)$, the excess variance due to better information $\sigma(n_{iS}, n_{iW}, En_{iS}, En_{iW}, \mathbf{x}_i)$ and the bias from favors $B(n_{iS}, n_{iW}, En_{iS}, En_{iW}, \mathbf{x}_i)$ may depend on the expected numbers

[13]To be consistent with our main regressions, we run balance tests conditioning directly on the expected numbers of connections. By contrast, Zinovyeva and Bagues (2015) control for expected connections through an extensive set of dummies, see Table 4 p.278. Incorporating these dummies raises computational issues in our non-linear setup. Results from Table 1 show that even in a simple linear formulation, actual connections are conditionally uncorrelated with observable characteristics.

of connections to the jury. Under the assumptions underlying Theorem 1, the conditional information effect and bias from favors are then identified. Note that the expected numbers of connections represent measures of social capital, built from information available to the jury. In the empirical analysis we therefore simply include them in the set of candidates' observable characteristics.

5.2 Exam-specific promotion thresholds

Our approach incorporates exam-specific promotion thresholds. This is important, since the bias from favors is identified from differences in promotion thresholds between connected and unconnected candidates. We consider two ways to account for exam-specific thresholds empirically: exam fixed effects a_e and exam grouped effects $a_e = \mathbf{a}\mathbf{z}_e$, where $\mathbf{z}_e$ is a vector of exam-level characteristics. Our first approach is to include a full set of exam fixed effects, which means that regressions then include 967 exam dummies. While exam fixed effects impose, in principle, less restrictions, they raise several problems in practice. They may not be identified for exams with small numbers of candidates, due to full predictability. They also raise computational difficulties due to the high dimensionality of the non-linear optimization problem to be solved in the estimations. Moreover in circumstances where grouped effects are appropriate, estimations based on fixed effects may be inefficient.

To address these issues, we consider exam grouped effects as in Bester and Hansen (2016). We allow promotion thresholds to depend on level, scientific area and application wave fixed effects (leading to 72 dummies in total) and on the number of candidates, the number of positions, the proportion of filled positions and the proportion of unconnected candidates. The model with exam grouped effects is of course nested in the model with exam fixed effects. In contexts where the model with fixed effects is estimable in practice, we can then test whether considering grouped effects leads to a significant loss of explanatory power.

5.3 Econometric model

In the empirical analysis, we estimate different specifications of model (4). The general model features three key ingredients: baseline heteroscedasticity $\sigma_v(\mathbf{x}_i)$, excess variance from better information $\sigma(n_{iS}, n_{iW}, \mathbf{x}_i)$ and bias from favors $B(n_{iS}, n_{iW}, \mathbf{x}_i)$. Note that the first two elements are closely related, since $\sigma_v(\mathbf{x}_i)\sigma(n_{iS}, n_{iW}, \mathbf{x}_i)$ represents the variance of the latent errror for candidates with connections n_{iS}, n_{iW} and characteristics $\mathbf{x}_i$.

We adopt a standard formulation for baseline heteroscedasticity, see Wooldridge (2010). We assume that the logarithm of the variance of v_i, the determinant of the ability of unconnected candidates observed by the jury but not by the econometrician, is a linear function of observable characteristics:

$$\sigma_v(\mathbf{x}_i) = \exp(\delta\mathbf{x}_i) \tag{5}$$

and where the constant is excluded from the x_i's. To gain in statistical and computational efficiency, we do not include all characteristics in σ_v in our preferred specification. Our estimation procedure is discussed in more detail in the Appendix.

We model information effects by building on this heteroscedasticity formulation. We consider increasingly complex specifications: (1) constant information effects $\sigma(n_{iS}, n_{iW}, \mathbf{x}_i) = \exp(\delta_c)$ if $n_{iS} + n_{iW} \geq 1$; (2) information effects depending on numbers and types of links: $\sigma(n_{iS}, n_{iW}, \mathbf{x}_i) = \exp(\delta_S n_{iS} + \delta_W n_{iW})$; and (3) information effects depending on numbers and types of links as well as other observable characteristics: $\sigma(n_{iS}, n_{iW}, \mathbf{x}_i) = \exp[(\ \delta_S\mathbf{x}_i)n_{iS} + (\delta_W\mathbf{x}_i)n_{iW}]$. Thus, each new strong tie with the jury increases latent error variance by $\exp(\delta_S)$ in formulation (2) and by $\exp(\delta_S\mathbf{x}_i)$ in formulation (3). These assumptions allow us to study the determinants of the variance of the latent error in a common, coherent framework. In addition, observe that formulation (3) can be obtained as the first element of the Taylor approximation of $\ln(\sigma(n_{iS}, n_{iW}, \mathbf{x}_i)/\sigma_v(\mathbf{x}_i))$ with respect to n_{iS}, n_{iW} and $\mathbf{x}_i$, for any function σ.

We also model increasingly complex specifications of the bias from favors: (1) constant bias: $B(n_{iS}, n_{iW}, \mathbf{x}_i) = B$ if $n_{iS} + n_{iW} \geq 1$; (2) bias depending on the numbers and types of links, linearly: $B(n_{iS}, n_{iW}, \mathbf{x}_i) = \gamma_S n_{iS} + \gamma_W n_{iW}$, or in a quadratic way: $B(n_{iS}, n_{iW}, \mathbf{x}_i) = \gamma_{1S} n_{iS} + \gamma_{2S} n_{iS}^2 + \gamma_{1W} n_{iW} + \gamma_{2W} n_{iW}^2 + \gamma_{SW} n_{iS} n_{iW}$; and (3) bias depending on connections and other observables: $B(n_{iS}, n_{iW}, \mathbf{x}_i) = (\gamma_{0S} + \gamma_S \mathbf{x}_i) n_{iS} + (\gamma_{0W} + \gamma_W \mathbf{x}_i) n_{iW} + \gamma_{2S} n_{iS}^2 + \gamma_{2W} n_{iW}^2 + \gamma_{SW} n_{iS} n_{iW}$. Quadratic terms help capture decreasing marginal impacts of additional links. For instance in the quadratic variant of formulation (2), a new strong tie with the jury increases bias by $\gamma_{1S} + \gamma_{2S}$ for an unconnected candidate and by $\gamma_{1S} + 3\gamma_{2S}$ for a candidate who already had one strong tie but no weak tie.

6 Empirical Analysis

6.1 Main results

We develop our empirical analysis in three stages. We first estimate a version of the simple model discussed in Section II, where the extent of information and favors are constant. We then account for the number and types of links, keeping both effects independent of observables. Finally, we estimate a model with full dependence on links and observables.

We first examine the impact of having at least one connection of any kind to the jury. We estimate constant favors and information effects, accounting for baseline heteroscedasticity. Denote by c_i the connection dummy: $c_i = 1$ if $n_{iS} + n_{iW} \geq 1$ and 0 otherwise. We thus estimate the following model.

$$p(y_i = 1 | \mathbf{x}_i, c_i) = \Phi[(\beta \mathbf{x}_i + B c_i - a_e) \exp[-(\delta \mathbf{x}_i + \delta_c c_i)]] \tag{6}$$

We consider grouped exam effects in our main regressions, and justify this choice in Section VI.B. Results of the estimation of Model (6) are reported in Table 4.[14]

On the whole sample, both the estimated bias from favors B and the estimated information effect δ_c are positive and statistically significant. They are also both positive and significant when estimated on promotions to Associate Professor. By

[14]For clarity, we do not report estimates of the impact of candidates' and exams' characteristics on promotion (β) and on baseline variance (δ) in the Tables.

Table 4: Binary connections: Model (6)

	(All)	(AP)	(FP)
Favors	0.179***	0.208**	0.227**
	(0.055)	(0.084)	(0.091)
Information	0.174***	0.245***	0.072
	(0.055)	(0.069)	(0.090)
Observations	31,243	17,799	13,444

Notes: All specifications include controls for the full set of observable characteristics, expected number of connections of each type, and the baseline heteroscedasticity. Heteroscedastic probit estimates. Exam grouped effects. Standard errors clustered on the exam level are in parentheses. *p<0.1; **p<0.05; ***p<0.01.

contrast, we detect favors but no information effect on promotions to Full Professor. Thus, connected candidates appear to face lower promotion thresholds at both levels and connected candidates for Associate Professor display excess variance in their latent errors. In other words, observable characteristics have lower power to explain promotion decisions in their case.

Are these effects quantitatively significant? How much do connections help? And how much does each mechanism contribute to the overall impact? To answer these questions, we compute for each candidate the predicted impact of a change in his connection status. We focus, for clarity, on unconnected candidates with at least one link to eligible evaluators. These computations could easily be replicated on other subsamples. Consider, then, unconnected candidate i. Model (6) can be used to predict how much i's probability of being promoted would change if i became connected. Denote estimated coefficients with hats. The difference in predicted promotion probabilities is equal to:

$$\begin{aligned}\Delta p_i/\Delta c_i &= p(y_i = 1|\mathbf{x}_i, c_i = 1) - p(y_i = 1|\mathbf{x}_i, c_i = 0) \\ &= \Phi[(\hat{\beta}\mathbf{x}_i + \hat{B} - \hat{a}_e)\exp[-(\hat{\delta}\mathbf{x}_i + \hat{\delta}_c)]] - \Phi[(\hat{\beta}\mathbf{x}_i - \hat{a}_e)\exp[-(\hat{\delta}\mathbf{x}_i)]]\end{aligned}$$

We can further decompose the overall impact of a change in connection status into two parts: one due to favors $[\Delta p_i/\Delta c_i]^F = \Phi[(\hat{\beta}\mathbf{x}_i + \hat{B} - \hat{a}_e)\exp[-\hat{\delta}\mathbf{x}_i]] - \Phi[(\hat{\beta}\mathbf{x}_i - \hat{a}_e)\exp[-\hat{\delta}\mathbf{x}_i]]$ and the other due to information $[\Delta p_i/\Delta c_i]^I = \Phi[(\hat{\beta}\mathbf{x}_i + \hat{B} -$

$\hat{a}_e) \exp[-(\hat{\delta}\mathbf{x}_i + \hat{\delta}_c)]] - \Phi[(\hat{\beta}\mathbf{x}_i + \hat{B} - \hat{a}_e) \exp[-\hat{\delta}\mathbf{x}_i]$.[15] Thus, $\Delta p_i/\Delta c_i = [\Delta p_i/\Delta c_i]^F + [\Delta p_i/\Delta c_i]^I$. Finally, we compute the averages of these values over all individuals in the sample.

We depict the results of these counterfactual computations in Table 5 and Figure 2. Table 5 reports averages of initial predicted promotion probability (first column), the average predicted change in promotion probability due to connection (second column), the part of this change due to information (third column) and the part due to favors (fourth column). Thus, an unconnected candidate with some link to eligible evaluators only has, on average, a 0.08 chance of being promoted, reflecting the highly competitive nature of these promotions. A chance connection to the jury leads to a relative increase in promotion probability of 80%. This relative impact is higher for candidates at the Associate Professor level (+91%) than for candidates at the Full Professor level (+76%).[16] The larger part of this effect is due to information for AP candidates (63% of the total impact). By contrast, favors is the main determinant of this impact for FP candidates (71% of the total impact). Overall, these numbers provide a quantitative picture of the impact of connections. Becoming connected to the jury almost doubles a candidate's chances of obtaining the promotion. Consistent with the estimation results, favors appear to dominate for FP candidates while the information effect dominates for AP candidates.

Figure 2 then depicts how the change in predicted promotion probability $\Delta p_i/\Delta c_i$, and its two components $[\Delta p_i/\Delta c_i]^F$ and $[\Delta p_i/\Delta c_i]^I$ vary with predicted probability $p_i = \Phi[(\hat{\beta}\mathbf{x}_i - \hat{a}_e) \exp[-(\hat{\delta}\mathbf{x}_i)]]$. We see that $[\Delta p_i/\Delta c_i]^I$ has an inverted U-shape, reaching a maximum for p_i close to 0.1 and becoming negative for high values of p_i. By contrast, $[\Delta p_i/\Delta c_i]^F$ is initially increasing over a larger range and may only decrease for high values of p_i. These qualitative patterns are consistent with Figure 1. In particular, and as discussed in Section 3, better information on candidates

[15] We assume here that the exam's promotion threshold a_e is unaffected by the change in connection status of candidate i. Alternatively, we can exploit the exam grouped effects specification to predict how the exam's promotion threshold is affected by the change in connection status of candidate i. Estimation results, available upon request, show that accounting for this impact has a negligible effect on the estimates and predicted impacts.

[16] To compute the impact of connectedness for a subsample, we rely on estimates of Model (6) for this subsample as presented in Table 4.

Table 5: Marginal effects of connections: Model (6)

	Baseline	Marginal effect		
	Predicted	Total	Information	Favors
All	0.080***	0.064***	0.035***	0.029***
	(0.002)	(0.004)	(0.011)	(0.010)
AP	0.088***	0.080***	0.050***	0.030**
	(0.002)	(0.007)	(0.014)	(0.013)
FP	0.063***	0.048***	0.014	0.034**
	(0.003)	(0.006)	(0.017)	(0.016)

Notes: Average marginal effect of being connected computed for candidates unconnected to the jury but with at least one connection to eligible evaluators. Standard errors clustered on the exam level are in parentheses. Standard errors are calculated using the delta method. *p<0.1; **p<0.05; ***p<0.01.

appears to lower the promotion probability of candidates with very good CVs. On average for these candidates, the impact of bad news dominates the impact of good news. Overall, $\Delta p_i/\Delta c_i$ displays a clear inverted U shape for AP candidates, reaching a maximum around p_i equal to 0.2, due to the key role of the information effect. By contrast, FP candidates with better observable characteristics benefit more from being connected to the jury.

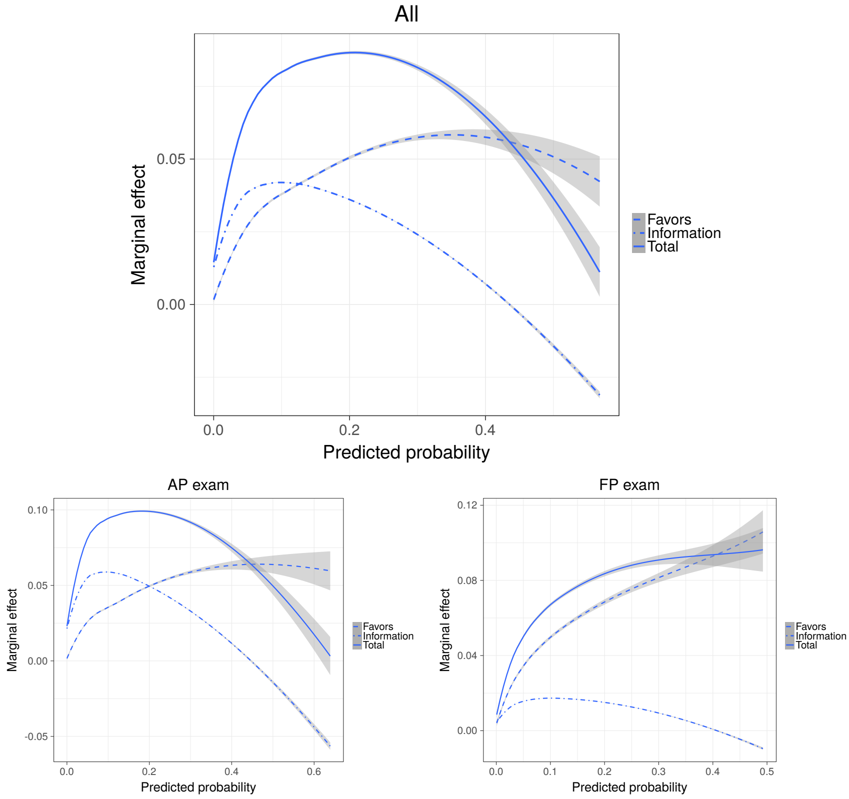
We next assume that both the bias from favors and the information effect depend on the number and types of links. We estimate a model with linear bias and log-linear variance:

$$p(y_i = 1|\mathbf{x}_i, n_{iS}, n_{iW}) = \Phi[(\beta\mathbf{x}_i + \gamma_S n_{iS} + \gamma_W n_{iW} - a_e)\exp[-(\delta\mathbf{x}_i + \delta_S n_{iS} + \delta_W n_{iW})]] \tag{7}$$

as well as a model with quadratic bias and log-linear variance:

$$\begin{aligned} p(y_i = 1|\mathbf{x}_i, n_{iS}, n_{iW}) = \Phi[(\beta\mathbf{x}_i + \gamma_{1S} n_{iS} + \gamma_{2S} n_{iS}^2 + \gamma_{1W} n_{iW} + \\ \gamma_{2W} n_{iW}^2 + \gamma_{SW} n_{iS} n_{iW} - a_e)\exp[-(\delta\mathbf{x}_i + \delta_S n_{iS} + \delta_W n_{iW})] \end{aligned} \tag{8}$$

Results are reported in Table 6. In the Left panel we report estimation results from Model (7). On the whole sample, both the favor bias and information effects from strong ties are positive and significant; they are positive but insignificant for weak ties. For Full Professor applications, we detect both favors and information effects from strong ties and, in addition, favors from weak ties. For Associate Professor ap-



Notes: Nonparametric fit using LOESS method. The grey region depicts 95% confidence intervals. Plots are constructed using estimated model (6) on subsamples indicated above each plot.

Figure 2: Marginal effects of connections: Decomposition

plications, we do not detect favors in this specification; however, we detect strongly significant and positive information effects for both strong and weak ties. Note that the effects of weak ties tend to be imprecisely estimated on the subsample of AP applications. This is due to the fact that candidates at this level have, on average, relatively few weak ties (see Table 2). In the Right panel of Table 6, we report estimation results from Model (8). Quadratic effects within the bias matter and affect overall estimation results. At the FP level, we now do not detect any information effect. By contrast, we still detect favors from both strong and weak ties. In addition, the marginal impact of an additional tie on the promotion threshold is decreasing in both cases. At the AP level, we now detect favors from strong ties and the bias is also increasing and concave in the number of ties. Information effects for both kinds of ties are positive and significant, particularly for weak ties.

Table 6: Estimation of Model (7) and Model (8)

	(All)	(AP)	(FP)	(All)	(AP)	(FP)
				Favors		
n_S	0.123***	0.071	0.120*	0.287***	0.309***	0.235***
	(0.047)	(0.070)	(0.068)	(0.047)	(0.072)	(0.061)
n_S^2				−0.051***	−0.065***	−0.036***
				(0.008)	(0.014)	(0.006)
n_W	0.038	−0.338	0.141**	0.096	−0.170	0.238***
	(0.069)	(0.230)	(0.058)	(0.073)	(0.280)	(0.067)
n_W^2				−0.026*	−0.293	−0.028***
				(0.014)	(0.240)	(0.010)
$n_S \times n_W$				0.018	0.157	−0.011
				(0.025)	(0.121)	(0.021)
				Information		
n_S	0.157***	0.240***	0.142**	0.092**	0.137***	0.095
	(0.044)	(0.055)	(0.064)	(0.040)	(0.051)	(0.061)
n_W	0.077	0.436***	−0.045	0.065	0.519***	−0.095
	(0.068)	(0.150)	(0.059)	(0.065)	(0.153)	(0.061)
Observations	31,243	17,799	13,444	31,243	17,799	13,444

Notes: Estimation of Model (7) - Left panel, and Model (8) - Right panel. All specifications include controls for the full set of observable characteristics, expected number of connections of each type, and the baseline heteroscedasticity. Heteroscedastic probit estimates. Exam grouped effects. Standard errors clustered on the exam level are in parentheses. *p<0.1; **p<0.05; ***p<0.01.

To sum up, strong connections to the jury lower the promotion threshold effectively faced by connected candidates. This impact is increasing in the number of strong ties at a decreasing rate. For applications for Full Professor, weak connections to the jury also lower the promotion threshold in a similar way. For applications for Associate Professor, we face a problem of statistical power caused by the relatively low number of weak ties. Both kinds of ties also appear to convey better information on candidates at the AP level. By contrast, we do not detect robust information effects at FP level.

We next present the outcomes of counterfactual computations on the impact of connections in Table 7, based on Model (8). We now focus on unconnected candidates who have at least one strong tie and one weak tie to eligible evaluators. For each such candidate, we compute the predicted promotion probability and the predicted increase in promotion probability caused by acquiring, through chance, one strong or weak connection to the jury. We also provide decompositions of these impacts into parts due to better information and due to favors. We then average over all candidates in the subsample. We see that one strong tie increases the promotion probability by 74% for AP candidates and by 72% for FP candidates. By contrast, one weak tie increases the promotion probability by 51% for AP candidates and by 22% for FP candidates. Thus, strong ties have higher predicted impacts than weak ties. For FP candidates, favors dominate quantitatively for both weak and strong ties. For AP candidates, favors dominate for strong ties and information effects dominate for weak ties, consistent with the estimation results.

Table 7: Marginal effects of connections: Model (8)

	Baseline	Marginal effect					
	Predicted	Total		Information		Favors	
		Strong	Weak	Strong	Weak	Strong	Weak
All	0.082***	0.060***	0.023***	0.019**	0.012	0.041***	0.011
	(0.002)	(0.004)	(0.005)	(0.008)	(0.012)	(0.008)	(0.011)
AP	0.091***	0.067***	0.046***	0.028***	0.092***	0.039***	-0.046**
	(0.002)	(0.006)	(0.012)	(0.010)	(0.023)	(0.011)	(0.018)
FP	0.068***	0.049***	0.015***	0.019	-0.017	0.030***	0.033***
	(0.003)	(0.005)	(0.005)	(0.012)	(0.011)	(0.010)	(0.010)

Notes: Average marginal effects of strong and weak connections computed for candidates unconnected to the jury but with at least one strong and one weak connection to eligible evaluators. Standard errors clustered on the exam level are in parentheses. Standard errors are calculated using the delta method. *p<0.1; **p<0.05; ***p<0.01.

These results conform to intuition. We would a priori expect strong ties to induce favors. We would also expect uncertainty on candidates' true ability to be stronger at the Associate Professor level, consistent with the stronger information effects detected at that level. The fact that weak ties generate stronger information effects is also consistent with the classic view of the role played by weak ties in information transmission, see Granovetter (1973). One finding that is, perhaps, surprising is that weak ties appear to generate favors at the Full Professor level. Note that candidates at that level have been in the academic system for a relatively long time. They have likely had many opportunities to initiate favor exchange with other academics. Overall, these findings indicate that the Spanish academic system was likely subject to generalized favoritism at the time.

These results are also consistent with - and help refine - the findings of Zinovyeva and Bagues (2015, Section IV.D.) derived from data collected 5 years after promotion. They find that research outcomes after promotion are lower for promoted candidates with strong ties, considering the whole sample and controlling for observables at time of promotion. Promoted candidates with strong ties publish less, in lower quality journals, supervise less PhD students and participate in less PhD committees. The authors state: "Our preferred interpretation of the empirical evidence is that candidates with a strong connection may have enjoyed preferential treatment, which overshadows the potential informational advantages of strong links." (p.285). By contrast, weak ties to the jury do not yield detectable differences in the research outcomes of promoted candidates. Promoted AP candidates with weak ties are more likely to eventually be promoted to full professor.

Our empirical results, obtained from data at time of promotion, are consistent with these findings. On the whole sample, we clearly detect favors resulting from strong ties. For AP candidates, we also detect information effects from weak ties. In addition, our method enables us to deepen the empirical analysis. We can detect both effects and precisely quantify their respective roles. We find, in particular, evidence of information effects from strong ties on the whole sample and no evidence of favors associated with weak ties for AP candidates.

Finally, we assume that the bias from favors and the excess variance due to better information may depend on observables. We estimate the following model:

$$\begin{aligned} p(y_i =1|\mathbf{x}_i, n_{iS}, n_{iW}) = \Phi[(\mathbf{x}_i\beta + (\gamma_{0S} + \gamma_S\mathbf{x}_i)n_{iS} + (\gamma_{0W} + \gamma_W\mathbf{x}_i)n_{iW} + \gamma_{2S}n_{iS}^2 \\ + \gamma_{2W}n_{iW}^2 + \gamma_{SW}n_{iS}n_{iW} - a_e)\exp[-(\delta\mathbf{x}_i + (\delta_S\mathbf{x}_i)n_{iS} + (\delta_W\mathbf{x}_i)n_{iW})]] \end{aligned} \tag{9}$$

We present estimation results in the Appendix, see Table A1 for AP candidates and Table A2 for FP candidates. A positive coefficient of the impact of a characteristic on bias means that favors due to connections tend to be stronger for candidates with higher values of this characteristic. Similarly, a positive coefficient on the information effect means that excess variance, and hence the quality of the extra information provided by an additional connection, is higher for these candidates. Results are rich and complex and confirm that we can detect variations in the effects of connections. For instance, younger FP candidates appear to receive less favors but to be subject to more information effect from strong ties. This is consistent with the idea that younger candidates have had less time to engage in favor exchange and to reveal their true quality through observable outcomes.

We present counterfactual computations obtained from Model (9) in Table 8. Comparing it with Table 7, we see that predicted probabilities display similar patterns across the two specifications. The average marginal impacts of gaining one

Table 8: Marginal effects of connections: Model (9)

	Baseline	Marginal effect					
	Predicted	Total		Information		Favors	
		Strong	Weak	Strong	Weak	Strong	Weak
AP	0.090***	0.078***	0.055***	0.032***	0.058**	0.047***	-0.004
	(0.003)	(0.006)	(0.012)	(0.011)	(0.028)	(0.011)	(0.026)
FP	0.066***	0.058***	0.017***	0.011	-0.011*	0.047***	0.028***
	(0.003)	(0.006)	(0.005)	(0.008)	(0.007)	(0.009)	(0.007)

Notes: Average marginal effects of strong and weak connections computed for candidates unconnected to the jury but with at least one strong and one weak connection to eligible evaluators. Standard errors clustered on the exam level are in parentheses. Standard errors are calculated using the delta method. *p<0.1; **p<0.05; ***p<0.01.

strong or weak link to the jury for unconnected candidates are comparable but slightly higher under Model (9) than under Model (8). This means that unconnected candidates have, on average, observable characteristics for which connections' impacts are slightly higher. The predicted impacts of strong ties are still higher than those of weak ties. The relative quantitative importance of the two factors is also robust to the econometric specification. Favors dominate for both strong and weak ties at the FP level and for strong ties at the AP level. Information effects dominate for weak ties at the AP level.

6.2 Robustness

In this section, we explore variations in the specification of two important features of the econometric model: exam-specific promotion thresholds and baseline variance. We first contrast estimations with exam fixed effects a_e and exam grouped effects $a_e = \mathbf{a}\mathbf{z}_e$. We compare estimation results of Model (6) under the two specifications in Table 9. The second column reports results of fixed effects estimations; the first

Table 9: Exam grouped effects vs. Exam fixed effects: Model (6)

	All		AP		FP	
	GE	FE	GE	FE	GE	FE
Favors	0.179***	0.304***	0.208**	0.074	0.227**	0.356***
	(0.055)	(0.088)	(0.084)	(0.078)	(0.091)	(0.097)
Information	0.174***	0.166***	0.245***	0.375***	0.072	0.013
	(0.055)	(0.053)	(0.069)	(0.064)	(0.090)	(0.065)
LogLik	-9965.2	-9766.6	-5847.3	-5773.2	-4064.9	-3931.7
df	98	989	61	486	61	523
LR	-	396.68	-	148.13	-	266.37
Observations	31,243	31,243	17,799	17,799	13,444	13,444

Notes: Row *df* reports degrees of freedom of the estimated model. Row *LR* reports the value of the Likelihood Ratio statistics of comparison of the unrestricted model (FE) and the restricted model (GE) in the preceding column. Heteroscedastic probit estimates. Standard errors clustered on the exam level are in parentheses. *p<0.1; **p<0.05; ***p<0.01.

column reproduces the results from Table 4. We see that the sign and statistical significance of both effects are similar for both specifications on the whole sample and on the subsample of FP candidates. On AP candidates, the information effect also has similar sign and significance. Bias from favors is positive and significant in the restricted model but positive and insignificant in the unrestricted model. Results from likelihood ratio tests show that we cannot reject the hypothesis that the grouped effects specification describes the data as well as the fixed effects specification, on each subsample as well as on the whole sample. We therefore consider grouped effects in our main regressions.

Second, we consider different specifications of baseline variance $\sigma_v(\mathbf{x}_i)$. We contrast estimations under homoscedasticity, when all individual characteristics are included, and when a subset of characteristics are included, as described in the Appendix. Results are depicted in Table 10 for Model (6). We see that the sign and statistical significance of the main effects are essentially similar for the last two specifications on the whole sample and on each subsample. Likelihood ratio tests also show that we cannot reject the hypothesis that the parsimonious specification describes the data as well as the full-fledged specification, even on subsamples. By contrast, estimates of main effects differ under homoscedasticity and the homoscedastic specification is rejected by likelihood ratio tests.[17] This confirms the importance of properly accounting for baseline heteroscedasticity. For reasons of computational and statistical efficiency, we therefore adopt the more parsimonious heteroscedasticy specification in our main regressions.

Table 10: Robustness: Baseline heteroskedasticity: Model (6)

	All			AP			FP		
	Hom.	Preferred	Full	Hom.	Preferred	Full	Hom.	Preferred	Full
Favors	0.419***	0.179***	0.186***	0.398***	0.208**	0.173*	0.412***	0.227**	0.192***
	(0.058)	(0.055)	(0.064)	(0.076)	(0.084)	(0.095)	(0.096)	(0.091)	(0.072)
Information	−0.020	0.174***	0.177***	0.055	0.245***	0.267***	−0.052	0.072	0.060
	(0.052)	(0.055)	(0.058)	(0.066)	(0.069)	(0.077)	(0.083)	(0.090)	(0.089)
LogLik	-10052.3	-9965.2	-9958.8	-5906.0	-5847.3	-5844.4	-4089.6	-4064.9	-4059.6
df	88	98	109	52	61	69	52	61	69
LR	-	174.19***	12.72	-	117.53***	5.71	-	49.42***	10.62
Observations	31,243	31,243	31,243	17,799	17,799	17,799	13,444	13,444	13,444

Notes: Row *df* reports degrees of freedom of the estimated model. Row *LR* reports the value of the Likelihood Ratio statistics of comparison of the unrestricted model and the restricted model in the preceding column. Heteroscedastic probit estimates. Standard errors clustered on the exam level are in the parenthesis. *p<0.1; **p<0.05; ***p<0.01.

[17] Table A3 in the Appendix reports similar results for Model (7).

7 Discussion and Conclusion

In this article, we propose a new method to identify the respective roles of favors and information effects in the impact of connections, building on earlier work on discrimination. Our method only exploits data collected at time of promotion and can be implemented through standard statistical software. We show that better information on connected candidates generates excess variance in latent errors. Differences in estimated variances between connected and unconnected candidates can be used to identify and quantify the information effect. Differences in estimated promotion thresholds can then be used to identify the bias due to favors. We apply our method to the data assembled and studied in Zinovyeva and Bagues (2015). Our empirical results are consistent with, and help refine, their findings obtained from data collected five years after promotion.

As with any identification strategy, our method relies on a number of assumptions. Our econometric framework is built on normality assumptions. We believe that normality is not critical, however. The fact that better private information leads to excess variance is quite general, as shown by Lu (2016). And indeed, implementing logit regressions in the empirical application leads to similar qualitative and quantitative results, see the Appendix.[18]

More importantly, we assume that information conveyed by connections is the only source of differences in latent error variances between connected and unconnected candidates. Depending on the context, however, other mechanisms could also potentially contribute to this difference. For instance, favors could have a stochastic element, perhaps as a result of debate among jury members. If the variance of this stochastic element decreases with the number of connections, our estimates would then provide a lower bound on the true information effect.[19]

Another important assumption is that the jury is risk-neutral. Risk aversion could matter in some contexts, and could lead a jury to promote a candidate with

[18]One difference is that we detect some information effects related to strong ties at the Full Professor level under some logit specifications.

[19]If the variance of this stochastic part is independent of the numbers of connections, our method and identification strategy apply without modifications.

lower expected ability in the presence of lower uncertainty on her ability. Risk aversion thus threatens identification in the whole literature on connections. For instance, evidence that connected promoted candidates publish less 5 years after promotion could conceivably be explained by risk aversion. Developing empirical methods to separately identify risk aversion, favors and information effects provides an interesting challenge for future research.

Finally, we believe that it would be interesting to combine our method with data collected after promotion. This could potentially yield more precise estimates of favors and information effects. This might also allow researchers to test critical assumptions, such as whether promotion has the same impact on subsequent outcomes for connected and unconnected candidates.

APPENDIX A

Proof of Proposition 1 A model with bias $B(.)$ and excess variance $\sigma(.)$ and an alternative model with bias $B'(.)$ and $\sigma'(.)$ yield the same conditional probability of being hired $p(y_i = 1|n_{iS}, n_{iW}, \mathbf{x}_i)$ if

$$\frac{\beta\mathbf{x}_i + B(n_{iS}, n_{iW}, \mathbf{x}_i) - a_e}{\sigma_v(\mathbf{x}_i)\sigma(n_{iS}, n_{iW}, \mathbf{x}_i)} = \frac{\beta\mathbf{x}_i + B'(n_{iS}, n_{iW}, \mathbf{x}_i) - a_e}{\sigma_v(\mathbf{x}_i)\sigma'(n_{iS}, n_{iW}, \mathbf{x}_i)}$$

Therefore, for any functions $B(.)$, $B'(.)$ and $\sigma(.)$, a model based on $B(.)$ and $\sigma(.)$ and one based on $B'(.)$ and

$$\sigma'(n_{iS}, n_{iW}, \mathbf{x}_i) = \frac{\beta\mathbf{x}_i + B'(n_{iS}, n_{iW}, \mathbf{x}_i) - a_e}{\beta\mathbf{x}_i + B(n_{iS}, n_{iW}, \mathbf{x}_i) - a_e}\sigma(n_{iS}, n_{iW}, \mathbf{x}_i)$$

have the same empirical implications. QED.

Proof of Theorem 1 Consider first the classical Probit model with heteroscedasticity:

$$p(y_i = 1|\mathbf{x}_i) = \Phi[(a + \mathbf{b}\mathbf{x}_i)\exp(-\mathbf{c}\mathbf{x}_i)]$$

Let us show that this model is identified if $a\mathbf{b} \neq 0$.[20] Identification holds if the mapping from parameters to the population distribution of outcomes is injective. Consider two sets of parameters $a, \mathbf{b}, \mathbf{c}$ and $a', \mathbf{b}', \mathbf{c}'$ such that $\forall \mathbf{x} \in \mathbb{R}^k$, $\Phi[(a + \mathbf{b}\mathbf{x})\exp(-\mathbf{c}\mathbf{x})] = \Phi[(a' + \mathbf{b}'\mathbf{x})\exp(-\mathbf{c}'\mathbf{x})]$. We must show that $a = a'$, $\mathbf{b} = \mathbf{b}'$ and $\mathbf{c} = \mathbf{c}'$.

Applying Φ^{-1} yields: $\forall \mathbf{x}$, $(a + \mathbf{b}\mathbf{x})\exp(-\mathbf{c}\mathbf{x}) = (a' + \mathbf{b}'\mathbf{x})\exp(-\mathbf{c}'\mathbf{x})$. At $\mathbf{x} = \mathbf{0}$, this yields: $a = a'$. Next, take the derivative with respect to x_k and apply at $\mathbf{x} = \mathbf{0}$. This yields $b_k - ac_k = b'_k - ac'_k$. Observe also that b_k and b'_k must have the same sign. Indeed if $x_l = 0$ when $l \neq k$, then $(a + \mathbf{b}\mathbf{x})\exp(-\mathbf{c}\mathbf{x}) = (a + b_k x_k)\exp(-c_k x_k)$. As x_k goes from $-\infty$ to $+\infty$, the sign of this expression can vary in one of three ways: it goes from negative to positive if $b_k > 0$; it goes from positive to negative if $b_k < 0$; or it stays constant if $b_k = 0$.

Assume first that $a \neq 0$ and $\mathbf{b} \neq \mathbf{0}$. Consider k such that $b_k \neq 0$, for instance $b_k > 0$. Set $x_l = 0$ except if $l \neq k$. For any x_k large enough, $a + \mathbf{b}\mathbf{x} = a + b_k x_k > 0$. Taking logs yields: $\ln(a + b_k x_k) - c_k x_k = \ln(a + b'_k x_k) - c'_k x_k$. Take the derivative with respect to x_k: $b_k/(a + b_k x_k) - c_k = b'_k/(a + b'_k x_k) - c'_k$. Take the derivative twice more: $b_k^2/(a + b_k x_k)^2 = b_k'^2/(a + b'_k x_k)^2$ and $-2b_k^3/(a + b_k x_k)^3 = -2b_k'^3/(a + b'_k x_k)^3$. Since this holds for any x_k large enough, this must hold for any x_k. At $x_k = 0$, this yields: $b_k^3 = b_k'^3$ and hence $b_k = b'_k$ and $c_k = c'_k$. If $b_k = 0$, then $b'_k = 0$ and $c_k = c'_k$.

Assume next that $\mathbf{b} = \mathbf{0}$. Then $\mathbf{b}' = 0$ and $\forall \mathbf{x}$, $a\exp(-\mathbf{c}\mathbf{x}) = a\exp(-\mathbf{c}'\mathbf{x})$ and hence $\mathbf{c} = \mathbf{c}'$. Finally, if $a = 0$ and $b_k > 0$, then for any $x_k > 0$, $\ln(b_k x_k) - c_k x_k = \ln(b'_k x_k) - c'_k x_k$ and hence $\ln(b_k) - c_k x_k = \ln(b'_k) - c'_k x_k$. This implies that $b_k = b'_k$ and $c_k = c'_k$. Thus, $\mathbf{b} = \mathbf{b}'$ and $\mathbf{c}\mathbf{x} = \mathbf{c}'\mathbf{x}$ for any $\mathbf{x}$ such that $\mathbf{b}\mathbf{x} \neq 0$, which implies that $\mathbf{c} = \mathbf{c}'$.

Observe that injectivity and identification also hold if $\mathbf{x}$ belongs to an open

[20] If $a = 0$ and $\mathbf{b} = 0$, $\forall \mathbf{x}$, $\Phi[(a + \mathbf{b}\mathbf{x})\exp(-\mathbf{c}\mathbf{x})] = 1/2$ and $\mathbf{c}$ is not identified.

set O of $\mathbb{R}^k$. The reason is that the function $\mathbf{x} \to (a + \mathbf{bx}) \exp(-\mathbf{cx})$ is analytic and that two analytic functions which are equal on an open set must be equal everywhere. Therefore, $\forall \mathbf{x} \in O$, $\Phi[(a+\mathbf{bx}) \exp(-\mathbf{cx})] = \Phi[(a' + \mathbf{b'x}) \exp(-\mathbf{c'x})] \Rightarrow$ $\forall \mathbf{x} \in \mathbb{R}^k$, $(a + \mathbf{bx}) \exp(-\mathbf{cx}) = (a' + \mathbf{b'x}) \exp(-\mathbf{c'x})$ and hence $a = a'$, $\mathbf{b} = \mathbf{b'}$ and $\mathbf{c} = \mathbf{c'}$.

Identification also holds if with some binary characteristics. Suppose that $x_i^1 \in \{0, 1\}$ and denote by $\mathbf{x}_i^{-1} \in \mathbb{R}^{k-1}$, the vector of other characteristics. Then, $p(y_i = 1|x_i^1 = 0, \mathbf{x}_i^{-1}) = \Phi[(a + \mathbf{b}^{-1}\mathbf{x}_i) \exp(-\mathbf{c}^{-1}\mathbf{x}_i)]$ yielding identification of $a, \mathbf{b}^{-1}$ and $\mathbf{c}^{-1}$. Next, $p(y_i = 1|x_i^1 = 1, \mathbf{x}_i^{-1}) = \Phi[(a + b^1 + \mathbf{b}^{-1}\mathbf{x}_i) \exp(-c^1 - \mathbf{c}^{-1}\mathbf{x}_i)]$. Rewrite $\Phi^{-1}(p) = [e^{-c^1}(a + b^1) + e^{-c^1}\mathbf{b}^{-1}\mathbf{x}_i \exp](-\mathbf{c}^{-1}\mathbf{x}_i)$. Therefore, $e^{-c^1}\mathbf{b}^{-1}$ is identified and hence c^1 is identified. Since $e^{-c^1}(a + b^1)$ is also identified, b^1 is identified.

Thus n becomes arbitrarily large, the econometrician can thus obtain consistent estimates of $a, \mathbf{b}$ and $\mathbf{c}$ if observables have full rank.

Consider, next, the following model

$$p(y_i = 1|n_{iS}, n_{iW}, \mathbf{x}_i) = \Phi[((\beta+\gamma_1(n_{iS}, n_{iW}))\mathbf{x}_i+\gamma_0(n_{iS}, n_{iW})-a_e] \exp[-(\delta+\delta(n_{iS}, n_{iW}))\mathbf{x}_i]$$

We apply the identification result on the Probit model with heteroscedascticity repeatedly. On unconnected candidates, we have: $p(y_i = 1|n_{iS} = 0, n_{iW} = 0, \mathbf{x}_i) = \Phi(\beta\mathbf{x}_i - a_e) \exp(-\delta)$ and hence a_e, β, and δ are identified. Similarly for candidates with connections n_{iS} and n_{iW}, the parameters $\gamma_0(n_{iS}, n_{iW})-a_e$, $\beta+\gamma_1(n_{iS}, n_{iW})$ and $\delta + \delta(n_{iS}, n_{iW})$ are identified. Therefore, $\gamma_0(n_{iS}, n_{iW})$, $\gamma_1(n_{iS}, n_{iW})$, and $\delta(n_{iS}, n_{iW})$ are identified. Note that to obtain consisent estimates of a_e, β, δ, $\gamma_0(n_{iS}, n_{iW})$, $\gamma_1(n_{iS}, n_{iW})$, $\delta(n_{iS}, n_{iW})$, the number of observations within exams must become arbitrarily large and observables conditional on (n_{iS}, n_{iW}) must have full rank. QED.

Preferred specification for the baseline heteroscedasticity. We first estimate model (4) on unconnected candidates, under the assumption that latent error variance is log-linear and depends on all observable characteristics. We thus estimate the following model:

$$p(y_i = 1|\mathbf{x}_i) = \Phi[(\beta\mathbf{x}_i - a_e) \exp(-\delta\mathbf{x}_i)]$$

on unconnected candidates. In our preferred specification for σ_v , we then include variables that are statistically significant in this first step as well expected numbers of connections En_{iS}, En_{iW}. We include these expected numbers given their critical role in ensuring the exogeneity of actual connections. We exclude other variables. Our preferred specification includes the following 10 observables: expected number of strong connections, expected number of weak connections, PhD students advised, AIS, age, gender, number of candidates at the exam, share of unconnected candidates at the exam, type of exam, and the indicator if the broad area is Humanities and Law. As discussed in Section VI.B. and following Davidson and MacKinnon (1984), we also test whether this restricted model indeed explains the data as well as the non-restricted model.

Additional estimation results. Results of the estimation of Model (9) for subsamples of AP candidates and FP candidates are presented in Table A1 and Table

A2 respectively. Table A3 presents the results of the robustness check with respect to the specification of the baseline heteroskedasticity in (7). Results of counterfactual computations for models (6), (8) and (9) when we use logit instead of probit are presented in Tables A4, A5, and A6 respectively.

Table A1: Estimation of Model (9): AP candidates

	Favors		Information	
	Strong	Weak	Strong	Weak
Constant	0.466**	0.587**	-	-
	(0.185)	(0.245)	-	-
n_S	−0.075***	0.020	-	-
	(0.026)	(0.128)	-	-
n_W	0.020	-0.141	-	-
	(0.128)	(0.134)	-	-
Publications	0.016	0.033	−0.061**	0.025
	(0.012)	(0.128)	(0.025)	(0.119)
PhD Committees	0.020***	-0.002	−0.017	−0.023
	(0.007)	(0.092)	(0.028)	(0.068)
AIS	−0.009	0.103	0.063**	−0.089
	(0.016)	(0.114)	(0.029)	(0.136)
PhD students	0.043**	0.086	0.078**	0.016
	(0.019)	(0.094)	(0.037)	(0.081)
Female	−0.009	0.423**	0.004	−0.452***
	(0.021)	(0.215)	(0.041)	(0.164)
PhD in Spain	−0.113	−0.650**	0.170**	0.490***
	(0.182)	(0.281)	(0.068)	(0.174)
Age	−0.004	0.001	−0.004	−0.010
	(0.002)	(0.017)	(0.004)	(0.014)
Past experience	−0.007	-0.156	0.037	0.222
	(0.024)	(0.176)	(0.025)	(0.135)
Expected strong	0.030	-0.168	−0.102***	0.060
	(0.027)	(0.189)	(0.028)	(0.150)
Expected weak	0.309	-0.415	0.180	−0.165
	(0.233)	(0.389)	(0.278)	(0.280)

Notes: Estimation of Model (9). Impact of observables (rows) on favors and the information effect from strong and weak connections (columns). All specifications include controls for the full set of observable characteristics, expected number of connections of each type, and the baseline heteroscedasticity. Heteroscedastic probit estimates. Exam grouped effects. Standard errors clustered on the exam level are in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table A2: Estimation of Model (9): FP candidates

	Favors		Information	
	Strong	Weak	Strong	Weak
Constant	0.366***	0.188***	-	-
	(0.065)	(0.058)	-	-
n_S	−0.058***	−0.044	-	-
	(0.017)	(0.035)	-	-
n_W	−0.044	−0.021	-	-
	(0.035)	(0.017)	-	-
Publications	−0.054	0.037**	0.071	−0.061**
	(0.036)	(0.019)	(0.051)	(0.028)
PhD Committees	0.007	0.038*	0.085***	0.037
	(0.007)	(0.020)	(0.027)	(0.026)
AIS	0.078**	0.140***	−0.057	−0.118***
	(0.032)	(0.034)	(0.047)	(0.037)
PhD Students	−0.008	−0.061**	0.013	0.041
	(0.010)	(0.024)	(0.020)	(0.025)
Female	0.017	0.005	0.021	0.009
	(0.025)	(0.046)	(0.047)	(0.055)
PhD in Spain	−0.009	0.064	0.085**	−0.010
	(0.013)	(0.040)	(0.039)	(0.049)
Age	0.003**	0.002	−0.012***	0.002
	(0.002)	(0.004)	(0.004)	(0.005)
Past experience	−0.001	−0.041	−0.018	0.108***
	(0.006)	(0.027)	(0.025)	(0.036)
Expected strong	0.048**	0.006	−0.069***	0.002
	(0.022)	(0.035)	(0.012)	(0.067)
Expected weak	0.019	−0.005	0.021	−0.117***
	(0.054)	(0.031)	(0.071)	(0.034)

Notes: Estimation of Model (9). Impact of observables (rows) on favors and the information effect from strong and weak connections (columns). All specifications include controls for the full set of observable characteristics, expected number of connections of each type, and the baseline heteroscedasticity. Heteroscedastic probit estimates. Exam grouped effects. Standard errors clustered on the exam level are in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table A3: Robustness: Baseline heteroscedasticity: Model (7)

	All			AP			FP		
	Hom.	Preferred	Full	Hom.	Preferred	Full	Hom.	Preferred	Full
Favors (strong)	0.221***	0.123***	0.128**	0.202***	0.071	0.055	0.236***	0.120*	0.066
	(0.037)	(0.047)	(0.051)	(0.052)	(0.070)	(0.067)	(0.051)	(0.068)	(0.061)
Favors (weak)	0.031	0.038	0.046	−0.093	−0.338	−0.369	0.085*	0.141**	0.135**
	(0.055)	(0.069)	(0.077)	(0.174)	(0.230)	(0.228)	(0.046)	(0.058)	(0.054)
Information (strong)	0.089***	0.157***	0.165***	0.124***	0.240***	0.250***	0.066	0.142**	0.182***
	(0.031)	(0.044)	(0.042)	(0.041)	(0.055)	(0.057)	(0.044)	(0.064)	(0.066)
Information (weak)	0.089**	0.077	0.074	0.280**	0.436***	0.471***	0.021	−0.045	−0.059
	(0.040)	(0.068)	(0.067)	(0.129)	(0.150)	(0.154)	(0.036)	(0.059)	(0.060)
LogLik	-10033.4	-9949.9	-9943.4	-5910.1	-5854.9	-5851.1	-4064.2	-4041.0	-4037.4
df	90	100	111	54	63	71	54	63	71
LR	-	166.91***	13.01	-	110.26***	7.70	-	46.34***	7.08
Observations	31,243	31,243	31,243	17,799	17,799	17,799	13,444	13,444	13,444

Notes: Row *df* reports degrees of freedom of the estimated model. Row *LR* reports the value of the Likelihood Ratio statistics of comparison of the unrestricted model and the restricted model in the preceding column. Heteroscedastic probit estimates. Standard errors clustered on the exam level are in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table A4: Marginal effects of connections: Model (6) with logit cdf

	Baseline	Marginal effect		
	Predicted	Total	Information	Favors
All	0.080***	0.066***	0.046***	0.020**
	(0.002)	(0.005)	(0.010)	(0.009)
AP	0.088***	0.082***	0.061***	0.021*
	(0.002)	(0.007)	(0.013)	(0.012)
FP	0.063***	0.047***	0.019	0.027**
	(0.003)	(0.006)	(0.014)	(0.013)

Notes: Average marginal effect of being connected computed for candidates unconnected to the jury but with at least one connection to eligible evaluators. Standard errors clustered on the exam level are in parentheses. Standard errors are calculated using the delta method. *p<0.1; **p<0.05; ***p<0.01.

Table A5: Marginal effects of connections: Model (8) with logit cdf

	Baseline	Marginal effect					
	Predicted	Total		Information		Favors	
		Strong	Weak	Strong	Weak	Strong	Weak
All	0.081***	0.061***	0.023***	0.026***	0.016	0.035***	0.008
	(0.002)	(0.004)	(0.005)	(0.008)	(0.011)	(0.007)	(0.010)
AP	0.090***	0.069***	0.048***	0.034***	0.091***	0.034***	-0.043**
	(0.003)	(0.006)	(0.012)	(0.010)	(0.022)	(0.010)	(0.017)
FP	0.068***	0.049***	0.014***	0.028***	-0.011	0.021**	0.026***
	(0.003)	(0.005)	(0.005)	(0.010)	(0.01)	(0.008)	(0.009)

Notes: Average marginal effects of strong and weak connections computed for candidates unconnected to the jury but with at least one strong and one weak connection to eligible evaluators. Standard errors clustered on the exam level are in parentheses. Standard errors are calculated using the delta method. *p<0.1; **p<0.05; ***p<0.01.

Table A6: Marginal effects of connections: Model (9) with logit cdf

	Baseline	Marginal effect					
	Predicted	Total		Information		Favors	
		Strong	Weak	Strong	Weak	Strong	Weak
AP	0.090***	0.078***	0.055***	0.039***	0.058**	0.040***	-0.003
	(0.003)	(0.006)	(0.012)	(0.011)	(0.028)	(0.011)	(0.026)
FP	0.066***	0.058***	0.017***	0.018**	-0.013*	0.040***	0.030***
	(0.003)	(0.006)	(0.005)	(0.008)	(0.006)	(0.009)	(0.007)

Notes: Average marginal effects of strong and weak connections computed for candidates unconnected to the jury but with at least one strong and one weak connection to eligible evaluators. Standard errors clustered on the exam level are in parentheses. Standard errors are calculated using the delta method. *p<0.1; **p<0.05; ***p<0.01.

References

Beaman, L. and Magruder, J. (2012). Who gets the job referral? evidence from a social networks experiment. *The American Economic Review*, 102(7):3574–3593.

Bertrand, M. and Mullainathan, S. (2004). Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *The American Economic Review*, 94(4):991–1013.

Bester, C. A. and Hansen, C. B. (2016). Grouped effects estimators in fixed effects models. *Journal of Econometrics*, 190(1):197–208.

Bramoullé, Y. and Goyal, S. (2016). Favoritism. *Journal of Development Economics*, 122:16–27.

Brogaard, J., Engelberg, J., and Parsons, C. A. (2014). Networks and productivity: Causal evidence from editor rotations. *Journal of Financial Economics*, 111(1):251–270.

Brown, M., Setren, E., and Topa, G. (2016). Do informal referrals lead to better matches? evidence from a firms employee referral system. *Journal of Labor Economics*, 34(1):161–209.

Colussi, T. (2017). Social ties in academia: A friend is a treasure. *Review of Economics and Statistics*, forthcoming.

Combes, P.-P., Linnemer, L., and Visser, M. (2008). Publish or peer-rich? the role of skills and networks in hiring economics professors. *Labour Economics*, 15(3):423–441.

Davidson, R. and MacKinnon, J. G. (1984). Convenient specification tests for logit and probit models. *Journal of Econometrics*, 25(3):241–262.

Engelberg, J., Gao, P., and Parsons, C. A. (2012). Friends with money. *Journal of Financial Economics*, 103(1):169–188.

Granovetter, M. S. (1973). The strength of weak ties. *American journal of sociology*, 78(6):1360–1380.

Heckman, J. (1998). Detecting discrimination. *The Journal of Economic Perspectives*, 12(2):101–116.

Heckman, J. and Siegelman, P. (1993). The urban institute audit studies: Their methods and findings. In M. Fix and R. Struyk, eds. *Clear and Convincing Evidence: Measurement of Discrimination in America*:187–258.

Hensvik, L. and Skans, O. N. (2016). Social networks, employee selection, and labor market outcomes. *Journal of Labor Economics*, 34(4):825–867.

Laband, D. N. and Piette, M. J. (1994). Favoritism versus search for good papers: Empirical evidence regarding the behavior of journal editors. *Journal of Political Economy*, 102(1):194–203.

Li, D. (2017). Expertise versus bias in evaluation: Evidence from the nih. *American Economic Journal: Applied Economics*, 9(2):60–92.

Lu, J. (2016). Random choice and private information. *Econometrica*, 84(6):1983–2027.

Neumark, D. (2012). Detecting discrimination in audit and correspondence studies. *Journal of Human Resources*, 47(4):1128–1157.

Pallais, A. and Sands, E. G. (2016). Why the referential treatment? evidence from field experiments on referrals. *Journal of Political Economy*, 124(6):1793–1828.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT Press.

Zinovyeva, N. and Bagues, M. (2015). The role of connections in academic promotions. *American Economic Journal: Applied Economics*, 7(2):264–292.