Charity and Favoritism in the Field:
Are Female Economists Nicer (To Each Other)?

Jason Abrevaya
Daniel S. Hamermesh

May 2010

DISCUSSION PAPER SERIES

Charity and Favoritism in the Field:
Are Female Economists Nicer
(To Each Other)?

Jason Abrevaya
University of Texas at Austin

Daniel S. Hamermesh
University of Texas at Austin,
Maastricht University, NBER and IZA

ABSTRACT

Charity and Favoritism in the Field: Are Female Economists Nicer (To Each Other)?[*]

Using a very large sample of matched author-referee pairs, we examine how the gender of referees and authors affects the former's recommendations. Relying on changing matches of authors and referees, we find no evidence of gender differences among referees in charitableness toward authors; nor do we find any effect of the interaction between the referees' and authors' gender. With substantial research showing gender differences in fairness, the results suggest that an ethos of objectivity can overcome tendencies toward same-group favoritism/opposite-group discrimination.

Corresponding author:

Daniel Hamermesh
Department of Economics
University of Texas
Austin, TX 78712
USA
E-mail: hamermes@eco.utexas.edu

I. Introduction

Discrimination is perhaps the most heavily researched topic in the field of labor economics and perhaps even among all endeavors in applied economics. Much less work has been done on differences in fairness/charitableness by individuals with different characteristics, but that too is attracting increasing attention (e.g., Andreoni and Vesterlund, 2001, and the extensive summary by Croson and Gneezy, 2009). Very little research has combined these two topics, asking whether the amount of favoritism/discrimination exhibited varies with the extent of the match between the parties (but see Parsons *et al*, 2011, Price and Wolfers, 2010, for differences by race/ethnicity, and Dillingham *et al*, 1994, for some sparse evidence on gender).

Our purpose here is to combine these two questions, focusing on differences by gender. In particular, we ask whether women are generally more or less generous than men in making up-or-down recommendations about others' work, and whether their degree of generosity is affected by the gender of those whose output they are asked to judge. While we despair of distinguishing favoritism toward one's own group from discrimination against another, at least our results provide some evidence on whether the extent of favoritism/discrimination differs by gender of the favoring/discriminating party.

Perhaps analogously to the theory of religious sects (Iannaccone, 1992), one might argue that the degree of solidarity within a group is a function of its relative size— that smaller groups will be more cohesive and more likely to favor other members of the group. We test this possibility too here, examining whether women favor other women more when they account for a smaller share of the relevant population of those making the judgments or being judged.

II. Modeling Preferences and Favoritism

We seek to model interactions between members of two groups, with within-group distinctions by gender. Call the two groups that are involved authors and referees. Members of the two groups interact on what they perceive to be a one-to-one basis, with authors seeking a judgment of their work, referees either giving their approval or not. Denote authors as A_i ($i = 1,\ldots, I$) and referees as R_j ($j = 1,\ldots, J$). Denote the gender of an author or referee by $f(A_i)$ or $f(R_j)$, equaling 1 for females and 0 for males. The utility of referee R_j when matched with author A_i is:

(1) $U(A_i , R_j) = \mu_i + \psi_j + \alpha f(A_i) + \beta f(R_j) + \lambda f(A_i)f(R_j) + \varepsilon_{ij}$.

The terms μ_i and ψ_j indicate the idiosyncratic values of author i's work and referee j's valuation of papers, respectively. Finally, there is some randomly distributed unobservable effect, ε_{ij} , that results with each author-referee match.

A paper is recommended for acceptance if:

(2) $U(A_i , R_j) > 0$,

which occurs with probability:

(3) $Pr\{\mu_i + \psi_j + \alpha f(A_i) + \beta f(R_j) + \lambda f(A_i)f(R_j) + \varepsilon_{ij} > 0\}$.

To understand the meaning of the parameters in (3), view the idiosyncratic μ_i and ψ_j as random draws, so that the composite error term is $e_{ij} = \mu_i + \psi_j + \varepsilon_{ij}$. Letting $G(.)$ denote the c.d.f. of $-e_{ij}$, the acceptance-recommendation probabilities of interest are:

Pr(accept | male referee, male author) $= G(0)$
Pr(accept | male referee, female author) $= G(\alpha)$
Pr(accept | female referee, male author) $= G(\beta)$
Pr(accept | female referee, female author) $= G(\alpha+\beta+\lambda)$.

Therefore, the effect of female authorship on a male referee's acceptance probability is $G(\alpha)$-$G(0)$ and on a female referee's acceptance probability is $G(\alpha+\beta+\lambda)$-$G(\beta)$. The relevant "difference-in-difference" effect is then $[G(\alpha+\beta+\lambda)$-$G(\beta)]$-$[G(\alpha)$-$G(0)]$.[1] This difference-in-difference will be positive (negative) if female referees are comparatively more (less) generous than male referees when matched with female authors.

The formulation of utility in (1) implicitly assumes that the referees j can identify the gender of authors i. If this were not so, then we would expect that $\lambda \equiv 0$, since, from the viewpoint of referees j, $f(A_i)$ would be noise. (Note that α could still be non-zero if paper quality differs by authors' gender.) Generalizing (1) and (3) to account for the possibility that in some cases the identity of authors is known, while in others it is not, we obtain:

(3') $\Pr\{\mu_i + \psi_j + \alpha f(A_i) + \beta f(R_j) + \lambda f(A_i)f(R_j)$

$+ \alpha'Z_{ij}f(A_i) + \beta'Z_{ij}f(R_j) + \lambda'Z_{ij}f(A_i)f(R_j) + \varepsilon_{ij} > 0\}$,

as the expanded condition for the probability of recommending acceptance, where Z_{ij} is an indicator of whether $f(A_i)$ is known to referee j. With this expanded formulation, we would expect $\lambda = 0$ and would infer whether there is within-group favoritism from the difference-in-difference effect, which for $Z_{ij}=1$ (author gender observed) is $[G(\alpha+\beta+\lambda+\alpha'+\beta'+\lambda')$-$G(\beta+\beta')]$-$[G(\alpha+\alpha')$-$G(0)]$.

III. Matching Data to the Model

In order to estimate the parameters describing charity and favoritism, we need a panel of referees and authors that is sufficiently long that individual idiosyncrasies can be accounted for through multiple observations on the same referee matched to different

---

[1]Note that, unlike in the linear regression setting, the difference-in-difference is not given directly by the interaction-variable parameter λ. See Ai and Norton (2003).

authors. Referees should also be aware of the author's gender, so that there is scope for them to indulge their preferences, if any, for their own gender. Obviously, the latter problem does not arise in laboratory work—parties' gender can, if the experimenter desires, be identified to others. The former problem generally cannot, however, be handled extensively in what are typically very short-duration laboratory experiments.

Our data set contains all the submissions to a leading field journal that were sent out to referees between 1986 and early 2008. The journal had a strict policy of double-blind refereeing, so that it took some effort for referees to discover whether they matched the author's gender. We cannot know whether or not they discovered that a match did or did not exist—whether $Z=1$ or 0. Evidence from another journal (Blank, 1991), however, suggests that even in the late 1980s referees could identify authors of half the papers they were assigned. This fact suggests that $Z = 1$ for many of these assignments.

In case the ease of identifying authors has changed, we need a proxy for Z that might indicate whether it was possible for the referee to make this discovery. Today it is trivial to discover the identity of the authors of most unpublished scholarly papers by doing an internet-based search for the paper's title. Such internet discovery was presumably far less prevalent during the early part of our sample.[2] We thus proxy Z by dividing the sample period into three parts: 1986-1994, when author gender could not be identified via the internet (but perhaps could be through working paper series, direct knowledge of the paper, etc.); 2000-2008, when authors' gender could be identified easily via the internet; and 1995-1999, when the degree of identifiability via the internet

_____

[2]Whereas 45 percent of American adults in 2000 had reported using the internet in the previous month, only 9 percent had reported doing so in 1996 (*Statistical Abstract of the United States*, various years). The internet-usage statistic is not available before 1996, but would no doubt indicate lower usage percentages.

may have been unclear. In most of our comparisons and estimation, we drop matches from this middle period.[3] To the extent that authors' gender is more likely to be known to the referee (Z=1) in the late period (2000-2008) than in the early period (1986-1994), our ability to detect statistically a gender-matching effect would be greater in the late period.

2940 initial submissions were sent to at least one referee. In the early period not all authors' names were listed in the data file, so that, as Table 1 shows, for only 70 percent of the papers were all the authors identifiable. In the later period the records were nearer to being complete. Authors' gender was completely identifiable on 80 percent of the manuscripts, with our inability to identify authors' gender due mostly to ambiguity about the gender identification of various given names. Around one-sixth of the papers for which the gender of all authors could be identified had only female authors, and one-third had at least one female. Both fractions increased between the early and late periods, significantly so for the "any female" category. Because of differences in both levels and changes of the "all female" and "any female" classifications, our analysis considers both classifications as indicators of female authorship ($f(A_i)=1$).

The journal used 6165 referees to judge these papers, but we were unable to identify the gender of 32 of them. A total of 1514 different referees judged papers during this time period. One instance of refereeing was most common, but 179 individuals judged at least ten manuscripts (and the busiest referee judged 36 papers). Of the

---

[3]Dividing the entire period into two equal-length parts and replicating our analysis does not qualitatively change any of our conclusions. Nor are our results altered qualitatively when we include all observations from 1986-2008 in the estimation.

identifiable referees, 19 percent were women, a percentage that increased significantly between the early and late periods, as Table 2 demonstrates.

The identification strategy here relies on multiple matches between a referee and a variety of authors, and multiple matches of a particular article to referees. It is thus identical to the strategy used to identify firm-worker match effects by Abowd *et al* (1999) and, in the context of discrimination, by Parsons *et al* (2011). This approach is unaffected by the identity of the (very few) editors, so long as they do not assign referees to articles based on their belief that particular referees will or will not discriminate/show favoritism that depends on the gender of the authors.

As Table 2 also shows, the matching of authors and referees was not random by gender over the entire period: Female referees were more likely than male referees to be matched with papers that had any or all female authors. This was not true during the early period. The relatively few papers that had female authors were only slightly more likely than others to be assigned a female referee; but in the late period there was more gender matching, especially of papers on which all authors were women. The difference in the extent of gender-matching with "all-female" and "any-female" papers may reflect the specialization of women in certain sub-fields. Regardless, this phenomenon justifies accounting for this non-randomness in our estimation.

Each referee was asked to rate the assigned paper on a four-point scale, with the possible ratings being: accept; accept with minor changes; accept with major changes; reject. As the first two columns of Table 3 indicate, referees recommend that roughly half of the papers be rejected. A recommendation of outright acceptance is extremely rare—most positive recommendations involve the referee offering asking for major

changes in the manuscript. The crucial thing to note in the table is the comparison by gender. In the entire sample, and in each sub-period, there is absolutely no evidence of any difference in charitableness by gender: Chi-square tests are very far from rejecting the null hypothesis that the distributions of judgments by male and female referees are the same. Ignoring possible gender differences in the quality of papers that are assigned, the evidence in Table 3 is consistent with the view that the women judging others' work were no more or less charitable than their male counterparts within any time period.[4]

Comparing responses across periods, however, suggests a slightly different implication. While the rejection rate among males rose by 2.1 percentage points between the two periods, the rejection rate among female referees rose by 11.0 percentage points, a double-difference of 8.9 percentage points (s.e. = 3.9). The women who refereed in the late period were significantly less charitable than their male counterparts, as compared to the earlier period.

This last comparison ignores the likelihood that the identity of the authors changed and the possibility that the female referees were assigned different kinds of papers in the two periods (which we saw was true along the dimension of gender matching). To examine the charitableness issue further, and to test for gender-matched favoritism, Table 4 presents estimates of the equations (3') describing refereeing decisions. In light of the fact that there are very few recommendations of outright acceptance or even acceptance with minor revisions, we define the outcome as non-reject versus reject. For both all-female and any-female papers, we interact the referee's gender

---

[4]Referees' views do matter: All 13 papers rated "accept" by at least one referee and not below "accept with minor revision" by the other(s) were eventually published in the journal. Also, there is substantial, but far from complete agreement on quality by referees. For example, if one referee rated a paper at least "accept with minor revision," 28 percent of the other referees (in two-referee cases) also rated it at least this highly.

with an indicator of the authors' gender. Also, to account for the possible changing non-observability of authors' gender, and thus to provide estimates of α', β' and λ', we include interactions with an indicator for the judgment being made in the late period.

We first estimate the equations using simple logit regressions, then using conditional logit regressions (conditional on the identity of the referee). The conditional-logit results limit variation in the outcome to differences in the interaction of a particular referee with authors of different gender, thus accounting for the referee-specific idiosyncrasies ψ_j in (3'). While we have no information on referees' characteristics (other than gender), we include a quadratic in the number of previous times a person had refereed for the journal—a measure of experience.

The results from both the logit and conditional-logit specifications in Table 4 yield the same conclusions about the possibility of gender favoritism: It is simply not evident in these data. None of the difference-in-difference estimates (the "Difference" rows near the bottom of the table) is statistically significantly different from zero. The point estimates are also quite small, ranging from 0.1 percentage points to 4.4 percentage points for the logit specifications. Even if referees in both periods could identify the gender of the authors of the papers they were assigned (so that Z=1 equally in both periods), our estimates for the separate sub-periods show that no favoritism/discrimination based on gender matching/non-matching was apparent.[5] Moreover, the triple differences, the coefficients on the interaction term of female referees with female authors and the Late period (2000-2008), are all essentially zero.

---

[5]One additional possibility is that some of the control variables, particularly experience, are related to the impacts of the gender match and its effects in the separate sub-periods. Interactions of referee experience with the referee-author match and its interaction with the sub-period identifier were essentially zero and had no effect on our conclusions about the possibility of favoritism.

The biggest change from the early (1986-1994) to the late (2000-2008) period appears in the author-gender effect upon male referees. Male referees between 1986 and 1994 were more likely to recommend rejection of female-authored papers (11.2 percentage points more likely in the first column, with a 95-percent confidence interval of [4.1, 18.2]). This difference disappears in the late period for the logit specifications. The change does not necessarily reflect anything about gender favoritism, however, as an increase in the quality of female-authored papers could also explain the difference.

One might thus argue that the relevant idiosyncrasies are manuscript- rather than referee-specific.[6] To examine this possibility, we re-estimated (3') using logits conditional on manuscript identifiers (which causes any variable that does not vary across referees on the same manuscript to drop out of the equation). Here again, the author-referee gender interaction was insignificantly different from zero in both the early and late periods. Finally, when we account for both referee and manuscript idiosyncrasies, we find no qualitative differences (although standard errors increase, as one would expect, with the increase in the number of fixed effects).[7]

These estimates do not account for the likelihood, as indicated by the across-period comparisons in Table 3, that a change in the mix of referees altered observed behavior between periods. Perhaps women who refereed during the early period were

---

[6],Clustering by manuscript or by referee has essentially no effect on the standard errors.

[7]Another potentially confounding problem is a gender difference in self-selection. Willingness to complete the assigned task may differ by gender, with women perhaps being more compliant (as they are in their propensity to complete surveys—Moore and Tarnai, 2002). Differential selectivity will only bias the results if the propensity to complete the task is related to the charitableness of doers and refusers. We cannot get at this problem, since we have no information on non-compliant referees. If men are more likely not to comply, and non-compliers are nastier, then our results are biased in favor of finding that female referees are less charitable than males. A related difficulty may arise from the selection of referees by editors. We cannot solve these problems completely; but the fact that the fraction female referees is slightly above the fraction female in the American Economic Association (Donald and Hamermesh, 2006) should allay some concerns about this issue.

inherently more favorable to female authors, but could not observe authors' gender, while female referees who entered the refereeing pool during the late period discriminated against female authors. This might have occurred because of the increase in female representation in the set of referees—possibly a reduced sense of solidarity among later cohorts of female economists. If this were correct, we would estimate $\lambda' = 0$, even though the agents' preferences exhibited favoritism in one case (the early referees) and discrimination in the other case (the new referees in the late period).

To examine this possibility, we re-estimated the equations, restricting the sample to the 295 individuals who refereed in both the early and late periods. The estimates for this reduced sample are presented in Table 5. A comparison to the results in Table 4 suggests that the apparent lack of favoritism or discrimination could not have resulted from the changing mix of referees. Even within this sub-sample there is no evidence of a significant change in behavior toward female authors when the authors' genders became more readily observable. The parameter estimates are quite similar across the two tables.

IV. Conclusions and Implications

Whereas many previous studies have found differences in altruism by gender, our examination of a unique and very large sample on author-referee outcomes in a high-stakes field environment yields no evidence of gender differences. Even accounting for the idiosyncrasies of both the judge and the judged, we still find no such differences. Moreover, we find no evidence of relative favoritism toward one's own gender by either men or women.

Female and male economists, at least in this specific setting, appear to behave similarly and in a gender-neutral manner. This might be the result of some inherent sense

of fairness, with participants feeling that exercising their prejudices is inappropriate in this particular judging activity—that "there own identity is often tied to their self-concept as experts who are able to stand above their personal interest" (Lamont, 2009, p. 9). Moreover, given the absence of an interaction of experience with gender, our results suggest either that there is no self-selection by gender attitudes, or that fairness/non-discrimination develops very early in the scholars' professional careers. Combined with previous findings, our results suggest that gender differences in fairness and favoritism are context-specific. Future research, including laboratory experiments, might examine how the extent of fairness/lack of favoritism depends on the perceived importance of the particular two-sided relationship.

REFERENCES

John Abowd, Francis Kramarz and David Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 67 (March 1999): 251-333.

Chunrong Ai and Edward C. Norton, "Interaction Terms in Logit and Probit Models," *Economics Letters*, 80 (July 2003): 123-129.

James Andreoni and Lise Vesterlund, "Which is the Fair Sex? Gender Differences in Altruism," *Quarterly Journal of Economics*, 116 (February 2001): 293-312.

Rebecca Blank, "The Effects of Double-Blind versus Single-Blind Reviewing: Experimental Evidence from the *American Economic Review*," *American Economic Review*, 81 (December 1991): 1041-1067.

Rachel Croson and Uri Gneezy, "Gender Differences in Preferences," *Journal of Economic Literature*, 47 (June 2009): 448-474.

Alan Dillingham, Marianne Ferber and Daniel Hamermesh, "Gender Discrimination by Gender: Voting in a Professional Society," *Industrial and Labor Relations Review*, 47 (July 1994): 622-33.

Stephen Donald and Daniel Hamermesh, "What Is Discrimination? Gender in the American Economic Association, 1935-2004," *American Economic Review*, 96 (September 2006): 1283-1292.

Laurence Iannaccone, "Sacrifice and Stigma: Reducing Free-riding in Cults, Communes, and Other Collectives," *Journal of Political Economy*, 100 (April 1992): 271-291.

Danna Moore and John Tarnai, "Evaluating Nonresponse in Mail Surveys," in Robert Groves, Don Dillman, John Eltinge and Roderick Little, eds., *Survey Nonresponse*. New York: Wiley, 2002.

Michèle Lamont, *How Professors Think*. Cambridge, MA: Harvard University Press, 2009.

Christopher Parsons, Johan Sulaeman, Michael Yates and Daniel Hamermesh, "Strike Three: Discrimination, Incentives and Evaluation," *American Economic Review*, 101 (2011): forthcoming.

Joseph Price and Justin Wolfers, "Racial Discrimination among NBA Referees," *Quarterly Journal of Economics*, 125 (2010): forthcoming.

Table 1. Author Characteristics, 1986-2008 (percentages)

	1986-2008	1986-1994	2000-2008	p-value for test of difference across periods
All authors identified	80.7	69.7	87.3	<0.001
All female authors, if all authors identified	16.9	15.6	18.5	0.104
Any female authors, if all authors identified	35.2	27.8	41.1	<0.001
N =	2940	1116	1095	

Table 2. Distribution of Female Referees, 1986-2008

	1986-2008	1986-1994	2000-2008	p-value for test of difference across periods
Percent Female:	18.7	14.3	22.9	<0.001
N =	6133	2347	2173	
Assignment:				
Matched with all female authors, if all authors identified	23.3	16.1	29.0	
Not matched with all female authors, if all authors identified	18.3	14.2	21.8	
p-value (test of random matching)	0.001	0.445	0.004	
Matched with any female authors, if all authors identified	21.8	15.9	25.2	
Not matched with any female authors, if all authors identified	17.7	14.0	21.6	
p-value (test of random matching)	0.001	0.323	0.067	

Table 3. Referee Recommendations by Gender, Percentages, 1986-2008

	Full sample 1986-2008		Early period 1986-1994		Late period 2000-2008	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**
Reject	51.9	51.1	45.8	49.9	56.8	52.0
Major	33.9	34.6	38.7	34.2	31.5	35.4
Minor	13.1	13.1	13.7	14.4	10.9	11.3
Accept	1.1	1.2	1.8	1.5	0.8	1.3
N =	1146	4987	336	2011	498	1675
p-value (test of difference by gender)	0.92		0.41		0.25	

Table 4. Logit and Conditional Logit Estimates of Effects on the Non-Reject Probability, Full Sample, 1986-1994, 2000-2008*

	"Female Author" = All Female Authors		"Female Author" = Any Female Authors	
	Logit	Cond. Logit	Logit	Cond. Logit
Year	0.0301	-0.0055	0.0294	-0.0060
	(0.0126)	(0.0251)	(0.0126)	(0.0251)
Late	-0.4739	-0.4039	-0.4942	-0.4452
	(0.1759)	(0.2351)	(0.1785)	(0.2381)
Experience	-0.0160	0.0875	-0.0147	0.0905
	(0.0152)	(0.0370)	(0.0153)	(0.0369)
Experience Squared/100	0.0275	-0.2115	0.0227	-0.2202
	(0.0651)	(0.1051)	(0.0652)	(0.1050)
Female Referee	0.1485		0.1077	
	(0.1264)		(0.1357)	
Female Referee*Late	-0.3907	-0.4342	-0.3896	-0.4301
	(0.1710)	(0.3134)	(0.1909)	(0.3285)
Female Author	-0.4516	-0.5849	-0.1797	-0.3198
	(0.1488)	(0.1764)	(0.1090)	(0.1319)
Female Author*Late	0.5126	0.6016	0.2995	0.4353
	(0.2040)	(0.2529)	(0.1488)	(0.1878)
Female Author* Female Referee	0.0329	0.2476	0.1771	0.0830
	(0.3701)	(0.4500)	(0.2784)	(0.3584)
Female Author* Female Referee*Late	0.0311	0.1336	-0.0510	0.1157
	(0.4555)	(0.5803)	(0.3489)	(0.4719)
Pseudo-R^2	0.0049	0.0088	0.0041	0.0071
N =	4520	3389	4520	3389

Table 4, continued

Effect of female authorship on no-reject probability, early period (1986-1994):

Female referees	-0.104 [-0.268, 0.060]	-0.084 [-0.281, 0.114]	-0.001 [-0.125, 0.124]	-0.059 [-0.220, 0.102]
Male referees	-0.112 [-0.182, -0.041]	-0.142 [-0.222, -0.063]	-0.045 [-0.098, 0.008]	-0.079 [-0.142, -0.016]
Difference	0.007 [-0.171, 0.186]	0.059 [-0.154, 0.271]	0.044 [-0.091, 0.180]	0.020 [-0.153, 0.193]

Effect of female authorship on no-reject probability, late period (2000-2008):

Female referees	0.031 [-0.079, 0.140]	0.098 [-0.057, 0.254]	0.031 [-0.070, 0.133]	0.078 [-0.055, 0.211]
Male referees	0.015 [-0.053, 0.084]	0.004 [-0.085, 0.093]	0.030 [-0.020, 0.080]	0.029 [-0.037, 0.094]
Difference	0.015 [-0.113, 0.144]	0.094 [-0.085, 0.273]	0.001 [-0.131, 0.134]	0.049 [-0.099, 0.197]

*95% confidence intervals in brackets; standard errors in parentheses.

Table 5. Logit and Conditional Logit Estimates of Effects on the Non-Reject Probability, Two-Period Referees, 1986-1994, 2000-2008*

	All Female Authors		Any Female Authors	
	Logit	Cond. Logit	Logit	Cond. Logit
Year	0.0026	0.0025	0.0027	0.0018
	(0.0193)	(0.0282)	(0.0193)	(0.0282)
Late	-0.3234	-0.4552	-0.3296	-0.4975
	(0.2412)	(0.2747)	(0.2445)	(0.2787)
Experience	0.0636	0.0771	0.0640	0.0804
	(0.0250)	(0.0404)	(0.0249)	(0.0404)
Experience Squared/100	-0.2058	-0.1836	-0.2089	-0.1912
	(0.0856)	(0.1127)	(0.0855)	(0.1125)
Female Referee	0.3291		0.3141	
	(0.1711)		(0.1836)	
Female Referee*Late	-0.7020	-0.3709	-0.7785	-0.5469
	(0.2609)	(0.32620	(0.2898)	(0.3570)
Female Author	-0.4412	-0.6111	-0.1637	-0.3108
	(0.1951)	(0.2222)	(0.1493)	(0.1687)
Female Author*Late	0.5623	0.6079	0.2729	0.4316
	(0.2860)	(0.3246)	(0.2082)	(0.2384)
Female Author* Female Referee	-0.1481	0.1469	-0.0149	-0.0510
	(0.4846)	(0.5567)	(0.3692)	(0.4384)
Female Author* Female Referee*Late	0.1691	-0.0676	0.2747	0.4815
	(0.6565)	(0.7625)	(0.5157)	(0.6264)
Pseudo-R^2	0.0078	0.0090	0.0065	0.0079
N =	2277	2102	2277	2102

*Standard errors in parentheses.