

NBER WORKING PAPER SERIES

WHAT IS DISCRIMINATION?  
GENDER IN THE AMERICAN ECONOMIC ASSOCIATION

Stephen Donald  
Daniel S. Hamermesh

Working Paper 10684  
<http://www.nber.org/papers/w10684>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
August 2004

We thank John Siegfried and the staff of the American Economic Association, who are in no way complicit in any of the conclusions reached here, for providing some of the material used in this study, and Francine Blau, Stephen Bronars, Charlie Brown, Ronald Ehrenberg, James Poterba, and participants in seminars at several universities, for comments. Mark Pocock provided expert programming assistance. The views expressed herein are those of the author(s) and not necessarily those of the National Bureau of Economic Research.

©2004 by Stephen Donald and Daniel S. Hamermesh. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

What is Discrimination? Gender in the American Economic Association  
Stephen Donald and Daniel S. Hamermesh  
NBER Working Paper No. 10684  
August 2004  
JEL No. J71, A11

**ABSTRACT**

Measuring market discrimination is extremely difficult except in the increasingly rare case where physical output measures allow direct measurement of productivity. We illustrate this point with evidence on elections to offices of the American Economic Association. Using a new technique to infer the determinants of the chances of observing a particular outcome when there are  $K$  choices out of  $N$  possibilities, we find that female candidates have a much better than random chance of victory. This advantage can be interpreted either as reverse discrimination or as reflecting voters' beliefs that women are more productive than observationally identical men in this activity. If the former this finding could be explained by the behavior of an unchanging median voter whose gender preferences were not satisfied by the suppliers of candidates for office; but there was a clear structural change in voting behavior in the mid-1970s. The results suggest that it is not generally possible to claim that differences in rewards for different groups measure the extent of discrimination or even its direction.

Stephen Donald  
Department of Economics  
University of Texas  
donald@eco.utexas.edu

Daniel S. Hamermesh  
Department of Economics  
University of Texas  
Austin, TX 78712-1173  
and NBER  
hamermes@eco.utexas.edu

## **I. Introduction**

Nearly 50 years ago Gary Becker (1957) set out the definition of discrimination used by economists today: A premium required to work with, buy from or employ a member of the group one discriminates against when that person is, except for group membership, identical to other individuals who are not discriminated against. The supply of members of that group and the distribution of preferences among the majority generate the observed outcome of market discrimination. This concept has proven incredibly powerful and has generated an immense empirical literature designed to measure the extent of market discrimination (see, e.g., Blau and Kahn, 2000, on gender discrimination).

The difficulty with the implementation of this concept is that it requires the econometrician who attempts to measure market discrimination to be sure that the members of the group experiencing discrimination are identical in all relevant ways to members of the majority. In the case of discrimination in economic rewards in a world where the relevant output is physical and easily measured (perhaps sports—see Kahn, 1991) this may be possible; but in most activities in industrialized societies the output is not physical, and it is impossible to adjust fully for differences in the productivity of members of the minority and majority. This argument does not mean that there is no value in comparing rewards after adjusting for as many factors as we can observe. It simply suggests that the measure of discrimination need not reflect underlying tastes—the original concept of discrimination—but may instead reflect a mixture of tastes and unobservable differences in productivity.

We illustrate this proposition with a particularly stark example of apparent gender discrimination—the election of officers of the American Economic Association. We examine the determinants of the outcomes of the 92 contested elections that occurred from 1959 through 2004, relating them to the candidates' ascriptive characteristics and one measure that arguably indicates the candidates' productivity. The results suggest the difficulty of identifying not only the extent of discrimination, but also its causes and even its direction.

## II. Data and Institutions

Beginning with the officers whose terms started in 1935, the American Economic Association has sent its members slates of four nominees for each of two positions as vice-president, and four nominees for each of two positions on the Executive Committee.<sup>1</sup> Those elected have some (consultative) decision-making power over the affairs of the Association and hold offices that many might view as prestigious. Each Association member receives a ballot in the fall before the year that the candidates would take office (Year  $t$ ) and can vote for up to two candidates for each of these positions. Lists of candidates and winners of these four-person elections beginning with 1935 form the data set used here. Because of difficulties in obtaining some of the measures (and the impossibility of obtaining information on our productivity measure for elections for office before 1959), the formal analysis concentrates on elections beginning with that year, although we also present some information for the 1935-1958 elections.

The outcome is winning or losing an election. As identifiers of the candidates' characteristics we include: 1) Gender—whether the candidate is female or male. 2) Honorable—whether the candidate held or holds a position in government that carries with it the honorific “honorable.” We include this measure to examine whether the publicity attached to such positions, or perhaps the recognition that they convey of the candidate's competence, affects his/her electoral chances. 3) Affiliation. This includes measures of whether the candidate is affiliated with a “Top 5” institution (Harvard, MIT, Princeton, Chicago or Stanford), and whether he/she is not an academic. 4) Race—whether or not the candidate is an African-American. 5) Field—whether or not the candidate is a theorist or econometrician. This distinction is clearly impressionistic, so that any results on this measure must be interpreted carefully. 6) Distinction—whether the candidate is a future Nobel Prize winner. This measure is less relevant for elections during the last decade of our sample, given the likely lags between recognition by the local (American) profession and by the Swedish Nobel Committee. All but the last of these characteristics have been readily available to the

---

<sup>1</sup>Until the late 1930s the position of President of the Association was also contested. Since no women were nominated for that office during those years, we ignore elections for President.

voters, as the Association has been enclosing an information sheet with a brief vita along with the ballot, and has even included pictures since at least the early 1970s.

We base the measure of scholarly productivity on the number of citations each candidate has received to his/her work. Among candidates for office in the elections from 1968-2004 we found citations in Year  $t-2$  (the most recent complete calendar year in which impressions on the voters could have been made). For the elections for office 1959-1967 we sum citations in Years  $t-4$ ,  $t-3$  and  $t-2$ .<sup>2</sup> In all cases we calculate the candidate's share of citations among the nominees for the particular offices. Thus if all candidates in a four-person election were identical along this dimension, each would obtain a value of 0.25 for this measure.

### **III. An Initial Look at Outcomes**

While the ascriptive measures that we use are not orthogonal to each other, and while each may be correlated with our productivity measure, it is nonetheless interesting to examine how candidates' unconditional chances of electoral victory differ by their characteristics. The upper part of Table 1 presents statistics—the means and their associated standard errors—describing the shares of candidates with each characteristic, their success probability, and the average share of citations within each category. The most striking feature is that for all but the first two characteristics listed the success probability does not differ significantly from 0.50. There is no evidence from the mean outcomes that being at a “Top 5” institution, being outside academe, being African-American, being a theorist or econometrician or being a future Nobel Prize winner has a significant effect on the likelihood of victory in these elections.

Only two characteristics—gender, and having held or currently holding a high-level government position—have significant impacts on the likelihood of winning. 70 percent of “honorable” candidates are elected, significantly different from 50 percent ( $t = 2.84$ ,  $p < .05$ ). 74 percent of female candidates emerged victorious from their elections, also significantly different from 50 percent ( $t = 3.68$ ,  $p < .01$ ). It is worth

---

<sup>2</sup>Using several years' citations in the early years of the sample is necessary to reduce the sampling error resulting from the relative paucity of journals catalogued in those years. The citation counts are from the on-line *Social Science Citation Index* and include all self-citations and citations to the author regardless of his/her order in the authorship. This database has citations for individual years beginning only with 1955.

noting that the average female candidate and the average “honorable” candidate have far below the average fraction of citations, 0.25, in any given election.

The bottom part of Table 1 answers the question whether scholarly productivity—really scholarly impact—matters in these elections. It clearly does: Each successively lower quartile of candidates by citation share has a successively lower probability of electoral victory. Moreover, the probabilities are nearly symmetric around 0.50: The chance that a candidate in the top quartile of citation shares wins an election differs from 50 percent by about the same (statistically significant) amount as does that of a candidate in the bottom quartile, and similarly for candidates in the second and third quartiles of the distribution of the shares of citations.

#### **IV. Estimating a Model of the Determinants of Electoral Success**

##### *A. A Feel for the Multivariate Relationships*

The standard approach to studying the binary outcome, electoral victory, is to estimate something like a probit or a logit. For reference purposes, in Columns (1) and (2) of Table 2 we present the results of estimating two probits. The estimates corroborate and even strengthen the inferences from the descriptive statistics presented in Table 1. A candidate’s share of citations has a significant positive effect on his/her electoral chances, as do gender and having held or currently holding a high-level government position. Moreover, the implied t-statistics on the coefficients of the variables Female and Honorable are far larger than the t-statistics testing the hypotheses that their raw means differ from 0.50. This is not surprising, given that the share of citations raises the probability of election and is less than 0.25 in each of these groups. These three measures alone produce significant effects on the probability of election. In line with inferences from the descriptive statistics, none of the other five variables significantly affects electoral probabilities, although theorists and econometricians apparently do suffer some electoral disadvantage, while faculty at Top 5 institutions do reap some additional advantage.

*B. Inferring the Impacts on Electoral Chances—A Multinomial Multiple-Response Estimator<sup>3</sup>*

The estimates presented in the previous Subsection may be informative, but they are incorrect. Given the institutional arrangements governing the elections, the probabilities that each candidate will be elected must sum to two in each of the ninety-two four-person elections. While the probit correctly constrains each individual candidate's chances of being elected to be on the open unit interval, it cannot impose this more basic institutional restriction. The econometric problem—estimating the determinants of a multiple of positives from among a fixed set of choices—does not appear to have addressed before (although a somewhat related econometric issue was modeled by Bloom and Cavanaugh, 1986).

Let the underlying desirability of candidate  $j$  in election  $i$  be:

$$(1) \quad y_{ji}^* = x_{ji}\beta + \varepsilon_{ji}.$$

Within each four-person election there are  ${}_4C_2 = 6$  possible sets of winners. (Since we do not see the votes, but only the identities of the winners and losers, the problem is one of combinations, not permutations.) Let the indicator for the pair of candidates that won the election be  $z_i\{j, l\}$  for  $j \neq l$ , where  $z_i\{j, l\} = z_i\{l, j\}$ . Then the contribution of election  $i$  to the likelihood function is:

$$(2) \quad L_i = \prod_{j=1}^4 \prod_{l>j}^4 P(z_i\{j, l\} = 1 \mid x)^{z_i\{j, l\}},$$

where:

$$\sum_{j=1}^4 \sum_{l>j}^4 z_i\{j, l\} = 1$$

The issue is one of calculating the probabilities  $P(z_i\{j, l\} = 1 \mid x)$ . Assume that the  $\varepsilon_{ji}$  are independent random variables. Arbitrarily ordering the observations so that candidates 1 and 2 win the four-person election, for a general distribution of the error terms:

$$(3) \quad P(z_i\{1,2\} = 1 \mid x) = P(y_{1i}^* > y_{3i}^*, y_{1i}^* > y_{4i}^*, y_{2i}^* > y_{3i}^*, y_{2i}^* > y_{4i}^* \mid x_i) \\ = P(\{y_{1i}^* > y_{wi}^*\} \cap \{y_{2i}^* > y_{wi}^*\} \mid x_i),$$

---

<sup>3</sup>We use this term based on Wooldridge's (2002) usage for the standard case where only one of out N possible choices can be made.

where  $y_{wi}^* = \max\{y_{3i}^*, y_{4i}^*\}$ . Noting that probabilities can be written as expectations of indicator functions (1(·)), and substituting from (1), we can rewrite (3) as:

$$\begin{aligned}
 (4) \quad P(z_i\{1,2\} = 1 \mid x) &= E(1(\{y_{1i}^* > y_{wi}^*\}) \cdot 1(\{y_{2i}^* > y_{wi}^*\}) \mid x_i), \\
 &= E(E(1(\{\varepsilon_{1i} > y_{wi}^* - x_{1i}\beta\}) \mid y_{wi}^*) \cdot E(1(\{\varepsilon_{2i} > y_{wi}^* - x_{2i}\beta\}) \mid y_{wi}^*) \mid x_i) . \\
 &= E((1-H(y_{wi}^* - x_{1i}\beta))(1-H(y_{wi}^* - x_{2i}\beta)) \mid x_i),
 \end{aligned}$$

where H is the cumulative distribution function of  $\varepsilon_{ji}$ . The particular estimator depends on the assumptions made about the nature of the distribution of the  $\varepsilon_{ji}$ .

The technique we have developed here to model multinomial multiple responses would appear applicable to elections in which there are more than one winner and more candidates than winners.<sup>4</sup> Indeed, the general technique is applicable to estimating the determinants of responses in any case in which there are a fixed number  $K > 1$  of slots that must be filled from among a fixed number  $N > K$  of choices. This would include admissions to many educational institutions (e.g., the U.S. military academies, many medical schools), some scholarship competitions, quinnella bets on horse races, and the three medals awarded among the eight finalists in many Olympic track and swimming events.

We assume that the errors  $\varepsilon_{ji}$  in (1) are independent  $N(0,1)$ , generating a probit-type estimator. The specific form is derived in the Appendix, along with that for a logit-type estimator.<sup>5</sup> The results of estimating the determinants of the electoral outcomes are presented in Columns (3) and (4) of Table 2. They are qualitatively like those using the incorrect simple probits—the impacts of each independent variable on the desirability index,  $y_{ji}^*$ , do not change much compared to the similar specifications in Columns (1) and (2). The standard errors, however, are somewhat larger, and some of the coefficients that had approached statistical significance (e.g., Top 5 School) no longer do. With this correct estimation procedure again only

---

<sup>4</sup>There is a substantial literature on elections in multimember electoral districts. The theoretical literature has examined voting patterns given preferences (Gerber *et al*, 1998), while the empirical literature has focused on characteristics of winners without formally examining the elections' determinants (e.g., Niemi *et al*, 1985).

<sup>5</sup>The maximum-likelihood estimates under the assumptions about the error terms that generate this estimator are qualitatively the same as those for the probit-type estimator.

the three variables, share of citations, Female and Honorable, significantly increase a candidate's electoral chances.<sup>6</sup>

## V. A Focus on Gender Discrimination

The means in Table 1 and the estimates in Table 2 show a clear electoral advantage to female candidates. The inference about women's advantage in these elections is similar to that found by Dillingham *et al* (1994) for election to office in a much smaller association of economists. Female voters in that group had a nearly lexicographic preference for female candidates, while male voters slightly preferred female candidates but were easily swayed by candidates' other measurable qualifications. In elections to confer an honorific in another association of economists, however, female candidates were treated identically to males with objectively identical qualifications (Hamermesh and Schmidt, 2003). These results make it worth delving further into the treatment of female candidates, as it can illuminate issues of discrimination/favoritism more generally.

What might constitute gender discrimination in this case? One might argue that, conditional on being on the ballot, disparate electoral chances of otherwise identical candidates, of the sort that we observe above, indicate the presence of disparate treatment by the electorate. If one assumes that the set of potential winners is the entire Association, then gender discrimination might be implied by differential representation among the winners by gender compared to the Association, conditioning on gender differences in the "productivity" of the profession's members. While we cannot impose this latter condition, below we do examine how levels and changes in the supply of candidates and winners by gender relate to the gender mix of the profession.

---

<sup>6</sup>One possible omission is the simultaneous presence of candidates from the same institution, which describes 7 percent of the sample. An indicator for this occurrence had a negative, albeit not quite statistically significant effect, on the probability of electoral victory. Its impacts on the estimated coefficients in Column (3) were tiny, raising each slightly and slightly lowering their standard errors. A measure of the candidates' years since the Ph.D. degree was quite insignificant statistically (and negative) and had a tiny influence on the other coefficients; similarly insignificant results were obtained for an indicator of whether the candidate had previously received the John Bates Clark Medal.

### *A. Unobservable Characteristics*

One might be tempted to argue that the results demonstrate that women are the benefactors of reverse discrimination by the Association's electorate. That would be true if academic impact, as captured by the share of citations among the four candidates in any given election, were monotonically related to productivity in office and were uncorrelated with other productivity-increasing characteristics. There is no reason to infer that this is true. One can readily adduce other qualifications, such as organizational ability, willingness to accomplish tasks on time, and ability to interact productively with colleagues in reaching decisions, that might be possessed in greater amounts by female candidates and confer an advantage on them in the eyes of the voters. Psychological research shows that women score higher in personality inventories on such characteristics as Restraint, Friendliness and Personal Relations (Guilford *et al*, 1976, p. 108; Butcher and Pancheri, 1976, pp. 24-225). Without information on these characteristics, which cannot be obtained in this sample, we cannot infer whether the outcomes of these elections reflect the nondiscriminatory result that the most "productive" candidates win or instead indicate reverse gender discrimination. This is a standard problem of unobservable productive characteristics, a problem that pervades the literature on discrimination (although it is infrequently acknowledged).

### *B. Inferring the Preferences of an Unchanging Median Voter*

Let us try to interpret the results under the assumption that there is an unchanging median voter whose preferences for the gender of candidates are unchanged over the entire sample period. To the extent that women have productive characteristics that appeal to that voter independent of scholarly productivity, so that male and female candidates are not perfect substitutes, one would expect that voter to have a decreasing willingness to choose a female candidate as the number of females on the ballot increases. To examine this possibility we estimate a probit describing a woman's chances of winning an election as a function of her scholarly quality (citation share) and the total number of women on the ballot. The sample consists only of the 46 female candidates; in years when at least one woman is on the ballot the total number of female candidates has a minimum of 1 (and in this sample) a maximum of 3 from among the 8 candidates.

The mean probability that a female candidate wins if she is the only female on the ballot is 0.81; it falls to 0.75 if there are two women on the ballot, and to 0.56 if there are three women.

Estimated probit derivatives describing the effects of the number of female candidates on the ballot on a woman's chance of victory are shown in the first column of Table 3. They suggest that the median voter's willingness to vote for a female candidate diminishes as the number of women on the ballot increases, with the estimate being nearly statistically significant at conventional levels despite the small sample size. Moreover, the decline in women's electoral success as more women are added to the ballot is not the definitional result of the presence of more women competing for the same two positions: When we substitute the number of women competing for the other office in place of the total number of women, the results, shown in the second column of the Table, are essentially unchanged. Voters appear to pay attention to the number of women on the ballot, although the impact is not highly significant.

This apparent eventual satiation with female candidates is unique to that group. In the remaining columns of Table 3 we present estimates of similar probits over samples of Honorable candidates, academics at Top 5 Schools, nonacademics and theorists/econometricians.<sup>7</sup> The number of other candidates on the ballot who have the same characteristic as a particular candidate matters to the voter, if at all, only in the case of female candidates.<sup>8</sup>

Let us model the apparent attention that voters pay to the number of females on the ballot. Consider the median voter, whose CES preferences are defined (as is implicit in the results in Table 3) as a function of the candidates' gender and their scholarly productivity:

$$(5) \quad U = \{ \alpha [M \cdot G(C_M)]^\rho + [1 - \alpha] [F \cdot G(C_F)]^\rho \}^{1/\rho}$$

where  $M$  and  $F$  are the number of male and female candidates,  $\alpha$  is the preference parameter for otherwise identical male and female candidates,  $\rho = 1 - 1/\sigma$ , where  $\sigma$  is the median voter's substitution elasticity

---

<sup>7</sup>We do not show results for African-Americans, since in no year was more than one African-American on the ballot.

<sup>8</sup>Only those observations are included in which fewer than three of the four candidates had the particular characteristic, so that each observation in these samples could conceivably have won the election. This restriction is binding only for the category Top 5 School.

between male and female candidates, and  $G$  is some increasing function of the candidate's share of citations. Letting  $M \equiv 8 - F$  and  $G(C_M) = G(1 - C_F)$ , the median voter's marginal rate of substitution between male and female candidates is:

$$(6) \quad MRS_{MF} = [\alpha/[1 - \alpha]] \cdot \{[F/M] G^*(C_F)\}^{1-\rho}$$

where  $G^*$  is an increasing function of  $C_F$ . The willingness of the median voter to substitute female for male candidates is a decreasing function of the median voter's inherent preference for males and an increasing function of the female candidates' average share of citations.

The supply of candidates to the voters is not competitive. Instead, the AEA Nominating Committee (a rotating group designated by the current President of the Association) chooses candidates. Assuming rationality on the part of the Committee, we model its choices of candidates as if it makes forecasts of their probabilities of victory, so that it implicitly sets those probabilities,  $p_F$  and  $p_M$ , for female and male candidates in each election. As the estimates of the probits in the first columns of Table 3 show, these probabilities depend on the number of female candidates, so we can write them as:

$$p_F = p_F(F) \text{ and } p_M = p_M(8-F) .$$

The median voter sets his/her marginal rate of substitution between the male and female candidates equal to this implicit price ratio, yielding:

$$(7) \quad p_F(F)/p_M(8-F) = [\alpha/[1 - \alpha]] \cdot \{F G^*(C_F)/[8-F]\}^{1-\rho} .$$

Setting the relative quality (measured by  $C$ ) of female and male candidates equal (implicit in the estimates in Table 3), so that  $G^*(C_F) \equiv 1$ , and taking logarithms yields:

$$(8) \quad \ln[p_F(F)/p_M(8-F)] = \ln(\alpha/[1 - \alpha]) - [1/\sigma] \ln([F/8-F]) .$$

We can treat this as an estimating equation, albeit one on which there are useful observations on  $F$  ranging only from 1 through 3. Thus merely as an exercise (with this tiny sample) we can estimate equation (8) over these three values using the implied  $p_F$  and  $p_M$  calculated from the estimates in the first column of Table 3. This exercise yields an adjusted  $R^2 = 0.55$  ( $p=.32$ ) and estimates of  $\alpha = 0.57$  and  $\sigma = 4.27$ . The estimates are sensible economically, but not at all statistically meaningful.

These results suggest that, if there were a median voter whose views on the gender of candidates were unchanging over the entire period, that voter would view males and females as highly substitutable and would prefer to have a nearly 50-50 split of winning candidates. Under the assumption of an unchanging median voter they also indicate that it is unsurprising that women's electoral chances are so high: That hypothetical voter's welfare would be maximized if he/she could elect candidates, roughly half of whom were women. Taking this view one might infer that Nominating Committees have supplied too few female candidates for these offices to satisfy the voters' preferences.

*C. Is There an Unchanging Median Voter Along the Dimension of Gender?*

To examine whether our assumption of an unchanging median voter makes sense, consider first whether female candidates' electoral chances have changed over time. Figure 1 decomposes the sample period from Table 1 into sub-periods of eight to ten years and also includes two earlier sub-periods, 1935-1948 and 1949-1958, for comparison purposes. The graph shows the fraction of females among all candidates and among winners. The coincidence of the fractions of female nominees and winners before the mid-1970s shows that a female candidate's chances of winning an election in the Association were remarkably close to those of a male candidate. Since the mid-1970s female candidates have had a decided electoral advantage, the results of which are apparent in the overall means for 1959-2004 in Table 1.

To examine further the possibility of a structural break we re-estimated the multinomial multiple-response model shown in Column (3) of Table 2 over each of a large number of pairs of sub-periods, beginning with the pair 1959-1966, 1967-2004, and ending with the pair 1959-1996, 1997-2004. Likelihood-ratio tests make it absolutely clear that a structural change took place at some point in the early to mid-1970s, with the highest (and statistically significant) likelihood ratio being for a structural break between 1974 and 1975. Before the mid-1970s women's chances of being elected, given their other objective characteristics, did not differ from those of men; thereafter women had a huge electoral advantage. The median voter's views on gender do not appear to have been constant over this entire period. Before the mid-1970s the median voter seemed indifferent to gender; thereafter he/she preferred women at the margin (of the supply of candidates).

What might have caused this change? One possibility is that the identity of the median voter changed toward someone who is more likely to favor female candidates, perhaps a female voter. Since we cannot observe individual ballots, we cannot be certain about the gender of voters in this Association; but we can use published information from the Association's *Directories* or *Handbooks* to infer the share of women in the potential electorate, the AEA membership. Taking all of the available issues beginning in 1936, we sampled members' names randomly and in each case tried to infer from their first names whether they were men or women. While it is not always possible to make these inferences correctly, so that measurement error is added to sampling error, there is no reason to believe that the estimates are biased down.

The first column of Table 4 presents our best estimates of the representation of women in the Association's membership. Even today, women account for no more than one-sixth of AEA members.<sup>9</sup> The growth in female representation since the 1960s (which occurred exclusively in the 1970s and 1980s) may mean that the gender identity of the median voter, and thus perhaps his/her preferences, changed over this period.<sup>10</sup>

The fractions female before and after the structural break did not differ that greatly—rising from perhaps 8 percent to perhaps 13 percent. Could this small increase have made such a huge difference? Voter turnout in these elections is not large, as the data in Column (4) of Table 4 show—by the 1970s turnout was 30 percent. Even if nobody's preferences changed, the small increase in female representation in the membership coupled with the low turnout could have combined to alter substantially the preferences of the median voter because his/her identity changed.

---

<sup>9</sup>We conducted similar tabulations for earlier decades. Counts show that the share of women in the Association in 1894 was 4.55 percent, and in 1906 was 3.55 percent, while samples indicate that in 1910 the share was 4.05 percent, in 1919, 3.33 percent, and in 1928, 4.51 percent. Since in each of the last three cases the standard error was about 1 percent, we may infer that the association's membership was roughly 4 percent female from its inception through the first third of the 20<sup>th</sup> century.

<sup>10</sup>The data provide an interesting perspective on how World War II changed the gender composition of the profession. While the membership of the AEA grew steadily, World War II saw a sharp rise in the fraction female (with "Edna the Economist" perhaps an analog to "Rosie the Riveter"). Absent a *Directory* or *Handbook* between 1948 and 1957, we cannot tell whether, as Goldin (1991) showed generally, women left the profession disproportionately after the War or whether the old patterns of inflow reasserted themselves.

It is also possible that the gender of the median voter did not change, but that his preferences did. Male voters (who surely constituted the majority of the voters through the late 1980s) may in the 1970s either have begun to believe that female candidates possess unobservable productive characteristics, or they may perhaps have suddenly become willing to indulge in reverse discrimination. Like the possibility that the median voter's identity changed, this one too is consistent with the data.

Regardless of the identity of the median voter, for the past thirty years the evidence shows that at the margin he/she has a preference for electing otherwise identical female candidates. Since 1974 women have accounted for only 16.2 percent of all candidates. Indeed, even during the last ten years of the sample women comprised only 21.5 percent of the candidates. Although these percentages exceed women's representation in the Association's membership, the suppliers of candidates—the Nominating Committees—might be viewed as having supplied too few female candidates to satisfy voters' revealed preferences to vote for women. Whether or not the median voter's preferences on gender have changed during this period, from this viewpoint one could interpret the evidence presented here not as reverse discrimination in favor of women by the electorate, but rather as discrimination against women by the suppliers of candidates for office in the Association—the Nominating Committees and the Association Presidents who select them.<sup>11</sup>

This apparent change in preferences is not a result of the changing gender mix of the suppliers of candidates—the Nominating Committees. Taking the sub-periods shown in Figure 1, in 1935-1948, 1949-1958 and 1959-1966 there was one woman (out of between five and seven people) on the Committee in most years of the latter two sub-periods. In every year during the sub-periods 1967-1974 and 1975-1984 there was exactly one woman. Since then the number of women on the Committee has ranged from zero to four (out of seven or eight members). Women have been well represented among the suppliers of candidates, given their representation in the Association; but their presence among the suppliers only

---

<sup>11</sup>One reader argued that lumpiness and the thinness of the distribution of qualified female candidates might have induced this apparent undersupply of female candidates. This explanation is logically possible; but even restricting the choice set to female full professors in the top 17 public university economics departments, the Ivy League, Stanford, Chicago and MIT, in 2002 there were 31 women (calculated from Hasselback, 2002). Moreover, there is no trend in the average citations share of female candidates.

affected electoral outcomes beginning in the mid-1970s, and only then did the representation of women on the ballot increase rapidly.

We cannot determine the ultimate cause of the apparent electoral advantage of women in this association. All the facts together, however, militate toward an interpretation that the median voter's (probably a male's) attitude toward gender in these elections changed in the early 1970s and yielded the apparent reverse discrimination that we now observe.

## **VI. Conclusions—Implications for Studying Discrimination**

We have examined the determinants of victory in elections to office in the American Economic Association. To measure these properly we have developed an estimator that we believe is useful in a large variety of other contexts. The estimates show that, while standard measures of scholarly productivity affect outcomes, so does the gender of the candidate, an effect that became apparent only beginning in the mid-1970s. The preference for women, given their representation on the ballot, may stem from unchanging behavior by an unchanging median voter; the structural shift in the mid-1970s suggests, however, that it arose either from the small increase in female representation in the electorate or more likely from changing preferences among male voters.

This examination should demonstrate that the notion of discrimination is a very slippery empirical issue. While it appears superficially that these elections suggest the existence of reverse discrimination in favor of women, our inability to control for many characteristics that might be argued are indicators of productivity in the particular “job” means that reverse gender discrimination may not exist in this sample. Indeed, assuming that the electorate is rational and can assess the productivity of the candidates presented to it, the apparent demand for more female candidates than have generally been provided may mean that the Association has discriminated against women by failing to nominate them in numbers sufficient to satisfy the electorate.

This conclusion may satisfy the priors of many observers of this Association and of labor markets generally. What if, however, we had shown that women's (or some other group's) electoral chances were significantly below 50 percent and that, like female candidates in this Association, they were at least

proportionately represented among the nominees? Would the analogous inference, that the suppliers of candidates had failed to accommodate voters' preferences and had been nominating too many women, be as appealing? Put in the context of labor markets more broadly, if we measured market discrimination against a minority group, an argument analogous to the one made here might reasonably point out that this outcome simply satisfies the tastes of the median consumer. In sum, the results should hardly reassure anybody who has thought about issues of discrimination in this profession, in the electoral process or in labor markets more generally.

The inability to control for productivity differences in empirical studies of possibly discriminatory outcomes is hardly unique to this study, to studies of electoral outcomes generally, or even to studies of economic outcomes. In the end our notions of what constitutes discrimination in specific instances must be cognizant of the possibility that we have been unable to account for differences that may be correlated with the characteristic of the group that we are studying and are productive in the jobs under study. The only alternatives are to concentrate on those few cases where we can be sure that we have properly adjusted for inter-group differences in productivity or to demonstrate in other cases that those differences are unimportant. The former approach is severely limiting, while the latter approach seems highly uncertain.

## REFERENCES

- Gary Becker, *The Economics of Discrimination*. Chicago: University of Chicago Press, 1957.
- Francine Blau and Lawrence Kahn, "Gender Differences in Pay," *Journal of Economic Perspectives*, 14 (Fall 2000): 75-99.
- David E. Bloom and Christopher L. Cavanagh, "An Analysis of the Selection of Arbitrators," *American Economic Review*, 76 (June 1986): 408-22.
- James Butcher and Paolo Pancheri, *A Handbook of Cross-National MMPI Research*. Minneapolis: University of Minnesota Press, 1976.
- 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.
- Elisabeth Gerber, Rebecca Morton and Thomas Rietz, "Minority Representation in Multimember Districts," *American Political Science Review*, 92 (March 1998): 127-44.
- Claudia Goldin, "The Role of World War II in the Rise of Women's Employment," *American Economic Review*, 81 (Sept. 1991): 741-56.
- Joan Guilford, Wayne Zimmerman and J.P. Guilford, *The Guilford-Zimmerman Temperament Survey Handbook*. San Diego: EdITS Publishers, 1976.
- Daniel Hamermesh and Peter Schmidt, "The Determinants of Econometric Society Fellows Elections," *Econometrica*, 71 (January 2003): 399-407.
- James Hasselback, *The 2002-2003 Economics Faculty Directory*. Upper Saddle River, NJ: Prentice-Hall, 2002.
- Lawrence Kahn, "Discrimination in Professional Sports: A Survey of the Literature," *Industrial and Labor Relations Review*, 44 (April 1991): 395-418.
- Richard Niemi, Jeffrey Hill and Bernard Grofman, "The Impact of Multimember Districts on Party Representation in U.S. State Legislatures," *Legislative Studies Quarterly*, 10 (November 1985): 441-55.
- Jeffrey Wooldridge, *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press, 2002.

## APPENDIX. Specific Functional Forms for the Multinomial Multiple-Response Estimator

In the case where  $\varepsilon_{ji} \sim N(0,1)$  we can use the symmetry of the normal distribution to specify equation (4) as:

$$(A1) \quad E\left(\left(1-\Phi(y_{wi}^* - x_{1i}\beta)\right)\left(1 - \Phi(y_{wi}^* - x_{2i}\beta)\right) \mid x_i\right) = E\left(\Phi(x_{1i}\beta - y_{wi}^*)\Phi(x_{2i}\beta - y_{wi}^*) \mid x_i\right).$$

Letting candidates 1 and 2 in each election be the winners, we can write

$$\begin{aligned} P(z_i\{1,2\} = 1 \mid x) &= \int \Phi(x_{1i}\beta - y_{wi}^*)\Phi(x_{2i}\beta - y_{wi}^*)\Phi(y_{wi}^* - x_{3i}\beta)\varphi(y_{wi}^* - x_{3i}\beta)dy \\ &\quad + \int \Phi(x_{1i}\beta - y_{wi}^*)\Phi(x_{2i}\beta - y_{wi}^*)\Phi(y_{wi}^* - x_{4i}\beta)\varphi(y_{wi}^* - x_{4i}\beta)dy. \end{aligned}$$

The log-likelihood is then:

$$(A2) \quad \log L(\beta) = \sum_{i=1}^n z_i\{1,2\} \log(P\{z_i\{1,2\} = 1 \mid x_i\}).$$

In the case of the extreme value distribution the error term is distributed:

$$\varepsilon \sim g(\varepsilon) = \exp(-\varepsilon)G(\varepsilon),$$

$$G(\varepsilon) = \exp(-\exp(-\varepsilon)).$$

A typical expression is then:

$$\exp(x_{3i}\beta) / [\exp(x_{1i}\beta) + \exp(x_{3i}\beta) + \exp(x_{4i}\beta)],$$

and the left-hand side of (A1) reduces to:

$$\begin{aligned} E(G(y_{wi}^* - x_{1i}\beta)G(y_{wi}^* - x_{2i}\beta)) &= \\ &= [\exp(x_{3i}\beta) + \exp(x_{4i}\beta)] / [\exp(x_{1i}\beta) + \exp(x_{2i}\beta) + \exp(x_{3i}\beta) + \exp(x_{4i}\beta)]. \end{aligned}$$

We then have  $P(z_i\{1,2\} = 1 \mid x) =$

$$\begin{aligned} &1 - [\exp(x_{3i}\beta)\exp(x_{4i}\beta)] / [\exp(x_{1i}\beta) + \exp(x_{3i}\beta) + \exp(x_{4i}\beta)] \\ &- [\exp(x_{3i}\beta)\exp(x_{4i}\beta)] / [\exp(x_{2i}\beta) + \exp(x_{3i}\beta) + \exp(x_{4i}\beta)] \\ &+ [\exp(x_{3i}\beta) + \exp(x_{4i}\beta)] / [\exp(x_{1i}\beta) + \exp(x_{2i}\beta) + \exp(x_{3i}\beta) + \exp(x_{4i}\beta)]. \end{aligned}$$

The log-likelihood function is calculated using these expressions in (A2).

**Table 1. Fractions of Candidates by Type, and Their Winning Chances and Shares of Citations, Contested AEA Elections 1959-2004 (N = 368)**

<b>Characteristic</b>	<b>Share of Candidates</b>	<b>Win Probability</b>	<b>Share of Citations</b>
Female	.125	.739 (.065)	.102 (.011)
Honorable	.125	.696 (.069)	.196 (.023)
Top 5 School	.370	.574 (.043)	.320 (.016)
Nonacademic	.092	.471 (.087)	.151 (.025)
African-American	.046	.412 (.123)	.087 (.029)
Theory/Econometrics	.209	.416 (.057)	.305 (.020)
Future Nobelist	.103	.605 (.080)	.414 (.029)
Share of Citations:			
Top Quartile		.620 (.051)	.495 (.012)
2 <sup>nd</sup> Quartile		.533 (.052)	.283 (.004)
3 <sup>rd</sup> Quartile		.456 (.052)	.166 (.003)
Bottom Quartile		.391 (.051)	.057 (.003)

\*Standard errors in parentheses below the means.

**Table 2. Estimates of the Determinants of Electoral Victory,  
92 Contested AEA Elections 1959-2004 (N = 368)**

Characteristic	Ordinary probit		Multinomial Multiple Response	
Share of Citations	2.634 (.444)	2.629 (.493)	2.771 (.455)	2.680 (.546)
Female	1.146 (.232)	1.156 (.234)	1.359 (.306)	1.349 (.322)
Honorable	0.691 (.218)	0.660 (.228)	0.843 (.280)	0.782 (.306)
Top 5 School	-----	0.170 (.152)	-----	0.225 (.205)
Nonacademic	-----	-0.101 (.259)	-----	-0.040 (.307)
African-American	-----	0.358 (.343)	-----	0.139 (.379)
Theory/Econometrics	-----	-0.315 (.185)	-----	-0.332 (.272)
Future Nobelist	-----	0.162 (.250)	-----	0.118 (.292)
Constant	-0.879 (.145)	-0.896 (.165)		
Log L	-226.41	-223.74	-140.64	-138.72

\*Standard errors in parentheses below the coefficient estimates.

**Table 3. Probit Estimates of Candidates' Winning Chances as Functions of "Crowding," 1959-2004\***

	<b>Female</b>		<b>Honorable</b>	<b>Top 5</b>	<b>Nonacademic</b>	<b>Theory/ Econometrics</b>
Share of citations	3.07 (1.13)	2.84 (1.13)	.484 (.481)	.826 (.275)	.381 (.604)	1.175 (.384)
Number on ballot	-.127 (.080)		-.008 (.102)	-.067 (.050)	-.031 (.117)	.044 (.062)
Number in other election	-.128 (.097)					
Number of Observations	46		46	106	34	71
Pseudo-R <sup>2</sup>	.177	.163	.019	.079	.009	.105

\*The estimates are of the impact of a one-unit increase in the independent variable on the probability of winning. Their standard errors are in parentheses.

**Table 4. Estimates of AEA Female Membership, and Voter Turnout in AEA Elections, 1936-2002<sup>a</sup>**

Year of Directory	Percent Female	Std. Error	Sample Size	Percent Voter Turnout
1936	5.56	0.97	558	46.52
1938	6.36	1.07	519	53.51
1940	6.45	1.07	527	48.82
1942	7.93	1.03	694	21.44
1946	9.90	1.04	818	41.16
1948	9.38	0.99	864	42.82
1957	8.08	1.05	668	34.86
1964	8.16	1.08	637	36.48
1966	7.92	0.97	770	35.41
1969	6.33	1.02	569	29.37
1974	9.38	0.97	895	30.28
1978	10.53	0.99	959	31.09 <sup>b</sup>
1981	13.05	1.06	1019	29.22
1985	12.79	1.04	1040	27.96
1989	15.13	1.13	1011	23.88
1993	15.15	1.08	1096	22.50
1997	15.52	1.06	1173	23.23
2002	16.08	0.60	3794	14.87

<sup>a</sup>The estimates of the percent female in the membership are based on tabulations of the first names of random samples of members from AEA *Handbooks* and *Directories*. The voter turnouts are calculated from reports contained in various issues of *American Economic Association, Papers and Proceedings*.

<sup>b</sup>1977.

**Figure 1. Women as Fractions of Candidates and Winners, AEA 1935-2004**

