In this section, the *Journal of Economic Education* publishes original theoretical and empirical studies of economic education dealing with the analysis and evaluation of teaching methods, learning, attitudes and interests, materials, or processes.

PETER KENNEDY, Section Editor

Exploring the Gender Gap on the GRE Subject Test in Economics

Mary Hirschfeld, Robert L. Moore, and Eleanor Brown

On average, women achieve lower scores on the Graduate Record Exam (GRE) Subject Test in Economics than men. From 1989 through 1992, 5,815 men and 2,164 women nationwide took the exam; the mean score for women was 603, in contrast to a mean score of 651 for men.[1] In this article, we look for explanations behind the gender gap in economics GRE test scores. Our data include GRE scores, overall college GPA as well as grades in all college economics and math courses, and Scholastic Aptitude Test (SAT) scores for economics majors at Occidental College, all of whom are required to take the GRE Subject Test. We also look at economics majors at Pomona College, who take an alternative standardized test, also produced by Educational Testing Service (ETS), known as the Major Field Achievement Test (MFAT) in Economics.

We find no easy answers. In particular, our results do not support the notion that men are simply better than women at economics; that the difference can be explained by such things as grades in economics courses, SAT scores, or even math backgrounds; or that women are substantially handicapped by a multiple-choice format alone.

We briefly review the GRE exam and the literature on gender and performance in economics, and preview our key empirical results. We then describe the Occi-

Mary Hirschfeld and Robert L. Moore are professors of economics at Occidental College. Eleanor Brown is a professor of economics at Pomona College. *The authors would like to thank Mike Tamada for providing some of the data used in this article; Lori Snyder and Jennifer Imazeki for excellent research assistance; and Rob Dorso and Dawn S. Robinson, who graciously performed some statistical analysis using ETS data. They would also like to thank associate editor Peter Kennedy and three referees for useful suggestions.*

dental data and estimate GRE score-generating functions for men and women economics majors. The Pomona College experience with the MFAT exam is contrasted with the Occidental GRE results. We then propose a possible explanation for our results.

THE GRE AND THE GENDER GAP

According to the descriptive booklet distributed to all economics GRE test takers by ETS, "scores on the test are intended to indicate students' abilities and their mastery of the subject matter emphasized in many undergraduate programs, with a focus on analytic methods for dealing with economic problems." The booklet goes on to describe the content of the exam:

> The committee of examiners has felt that the primary concern in most graduate school admissions is the student's aptitude for and competence in the basic skills of economic analysis. Of secondary importance is his or her knowledge of economic history, institutions, and terminology. Economic analysis is broadly defined. It includes interpreting and manipulating diagrams and simple mathematics; explaining and predicting economic behavior, given certain assumptions; prescribing appropriate action and policy; and drawing conclusions from specified economic information and data. (Educational Testing Service 1979, 5; and 1987, 6)

Given this orientation, some might conclude that the differential performance by gender must reflect the fact that men have better aptitude for, and competence in, the basic skills of economic analysis. The results we present using data from Occidental College indicate that this differential performance on the GRE persisted even when we controlled for undergraduate grades in economic courses and SAT scores.

Others might point to prior studies, most of which used results from principles of economics courses, that have found that women do worse on standardized multiple-choice exams. A good summary of these past studies on the effect of gender on scores on tests of economic understanding can be found in Williams, Waldauer, and Duggal (1992, 219–20). In general, such studies have used standardized tests to identify the determinants of such economic knowledge. The most comprehensive survey to date (Siegfried 1979) found that men performed better than women in a majority of cases. More recent studies (Anderson et al. 1994; Heath 1989; Soper and Walstad 1988; Walstad and Soper 1989; Ferber, Birnbaum, and Green 1983; Gohmann and Spector 1989; Lumsden and Scott 1987; and Watts and Lynch 1989) seem to confirm the previous findings. Evidently, these men-women differentials are reduced (Ferber, Birnbaum, and Green 1983) or even reversed (Lumsden and Scott 1987) when essay questions are used instead. The study by Williams, Waldauer, and Duggal (1992) extends the research to include intermediate theory and economic statistics courses, and includes essay questions. According to the authors, this provides better information on "the issue of whether observed differences in performance on economic exams are the result of greater male skills in quantitative and spatial relationships compared with greater female skills in verbal relationships" (p. 220). Their study finds "no evi-

dence to support the hypothesis that significant and consistent gender differences exist in college students' performances on economic exams" (p. 229).

Our results, using data from Pomona College, where almost all economics majors take the Major Field Achievement Test in Economics, indicate virtually no differential performance by gender. The MFAT is a shorter and easier version of the GRE Subject Test in Economics prepared by ETS and contains some of the same GRE questions, but it has the same multiple-choice format.[2] Thus, a more complex set of considerations may lie behind the differential performance by gender on the GRE Subject Test than simply the multiple-choice format.

GENDER DIFFERENCES ON THE GRE SUBJECT TEST IN ECONOMICS AT OCCIDENTAL

Since the early 1980s, every economics major at Occidental has taken the GRE Subject Test administered in February as one part of their comprehensive exam. In addition, each Occidental major takes principles of microeconomics, principles of macroeconomics, intermediate microeconomics, intermediate macroeconomics, a course each in statistics and econometrics, and at least three upper-division courses in economics. Two terms of calculus are also required as is the writing of a 50-page senior thesis.[3]

The mean GRE score by gender at Occidental over the past five years is similar to the national averages of the 8,000 individuals who took the exam nationally over the 1989–92 period. In particular, the mean for men at Occidental was 646 (with a standard deviation of 73) while that for women was 598 (with a standard deviation of 58). Such scores are unadjusted and thus are of only limited usefulness. However, they do show that the 48-point difference in the means of the men and women who took the exam nationally (651 versus 603) is almost exactly the same for Occidental students over the past five years. In this sense, the Occidental results are not an anomaly.

We were able to get detailed information about our economics graduates during the last five years, including their high school SAT scores, GPA in all economics courses taken at Occidental, and specific math and economics courses taken (and respective grades). The results of our basic equation are shown in Table 1. We used OLS to regress the GRE Subject Test in Economics score (GRE) on a student's GPA in all economics courses taken at Occidental (GPA), Math SAT (SATM), Verbal SAT (SATV), a gender dummy variable (MALE), and a constant term (column 1). Separate OLS equations for men and women are also displayed in Table 1. Presumably SAT scores allowed us to control for verbal and quantitative aptitude as of the senior year in high school as well as skills in taking standardized multiple-choice exams.[4]

The data in Table 1 demonstrate that the differential effect of gender remains quite strong, even after adjusting for economics GPA and SAT scores. Women with the same GPA in economics and SAT scores as men scored about 40 points lower, or about 94 percent of the score for men.[5] Further, this coefficient is significantly different from zero, with a t statistic of 3.6. The other variables were all significantly different from zero and of the expected sign.

TABLE 1
Results of Estimating Basic Equation for Occidental College Economics Majors
(GRE Score = Co + C1 MALE + C2 GPA + C3 SATM + C4 SATV + error term)

Variable	All	Men	Women
MALE	39.7	—	—
	(10.9)		
GPA	78.9	77.2	85.8
	(10.4)	(12.1)	(20.3)
SATM	.203	.212	.162
	(.071)	(.084)	(.122)
SATV	.110	.125	.067
	(.058)	(.072)	(.087)
Constant term	172.4	202.6	199.7
Observations	149	118	31
% male	79	100	0
R^2	.46	.40	.47
Mean GRE	636	646	598
SEE	53.4	56.4	42.1

Notes: Standard errors are in parentheses.
Definitions: MALE = dummy variable, male = 1. GPA = Grade point average of all economics courses (4 = A, 3 = B, etc.). SATM = score on mathematics portion of SAT. SATV = score on verbal portion of SAT.

Of course, we were implicitly assuming that this equation that translates relevant economics skills and knowledge (in the form of GPA), and math and verbal aptitudes plus standardized test-taking skills (in the form of SAT scores) into generating points on the GRE, was identical for men and women. The separate results by gender (second and third columns) allowed the coefficients on each of the independent variables to differ by gender. The Chow test resulted in a rejection of the null hypothesis that the regressions are the same.[6]

In general, following Oaxaca (1973), we could decompose the difference in the means for men and women in the GRE into two parts: (1) the portion that resulted from the difference in the returns of a student's investment in economics-relevant skills and knowledge in generating points on the GRE, such as GPA in economics and (2) the portion that resulted from the different skill levels. When we carried out the calculations, we found that only 13.8 percent of this difference was due to the different skill levels, while 86.2 percent was due to the different returns from these variables.

The approximately 40-point differential that remained after controlling for GPA and SAT scores turned out to be a remarkably robust result. Because prior studies showed that math courses taken and math course performance were sometimes significant variables explaining multiple-choice exam results (for example, Anderson et al. 1994), we added variables to the basic equation to capture the number of math courses taken beyond calculus as well as the grade in calculus. The addition of such variables altered the basic equation only slightly, with the coefficient on MALE falling from 39.7 to 37.8 while all the other coefficients were hardly changed. The coefficient on MALE was also changed only slightly when we limited the sample to those students with an economics GPA of 3.5 or below. Because grades are truncated (there is no way to acknowledge better than

A work), it has been suggested that such constraints could be masking differences among the very best students. Finally, the coefficient on MALE was little changed when we substituted overall GPA (for economics GPA) or substituted course grades in the three core courses of intermediate microeconomics, intermediate macroeconomics, and econometrics.[7] (Details of these results are available from the authors.)

Thus, the central question remains, Why do women do less well on the GRE, even after controlling for economics grades, SAT scores, the number of advanced math courses, and calculus grades? Of course, grades in economics courses could be measuring many other potential differences between men and women, such as writing skills or effort that the GRE test does not measure. We, therefore, turned to the results at Pomona College, where we had information on economics GPA and scores on the Major Field Test in Economics for a sample of Pomona economics majors.

THE POMONA COLLEGE RESULTS WITH THE MAJOR FIELD TEST IN ECONOMICS

Although we knew that the gender difference in unadjusted test score means for Occidental students was in line with differences observed nationally, we wondered whether our results, controlling for grades in economics courses, could be replicated at a similar liberal arts college. Our search led us to Pomona College and a rather different piece of evidence. All senior economics majors at Pomona are required to take a standardized test.[8] Unlike Occidental students, however, most take the Major Field Achievement Test (MFAT) in Economics, which like the GRE, is developed by ETS.[9] This test, according to ETS, is a shortened and easier version of the GRE. Although it includes some actual GRE questions directly, it is designed to be a better indicator of the knowledge that an undergraduate economics major has acquired.

The Pomona data had MFAT scores for 93 students from the graduating classes of 1990, 1991, and 1992. Of the sample, 26 percent were female, compared to 23 percent in the Occidental sample. The Pomona men and women had remarkably similar performance averages: the women had a slightly lower MFAT mean, 169.8 versus 171.8 for men, and a slightly higher grade point average, 9.65 versus 9.58 on a 12-point scale (A = 12, A− = 11, B+ = 10, and so on). Nationally, although the women/men ratio of MFAT test scores was a bit lower than at Pomona, it was still quite a bit above the ratio for the GRE Subject Test.[10]

Regressing MFAT scores on economics GPA and the gender dummy variable (MALE) yielded the results reported in Table 2. (The results of a similar equation using the Occidental sample are also shown in Table 2 for comparison purposes.) As in the raw data, the MFAT scores at Pomona did not seem to be much affected by gender. The regression of the full sample's test scores on GPA and MALE are reported in Table 2, column 1; the MALE variable was not statistically significant at the .05 level. In this data set, females scored only slightly lower than their male counterparts on the MFAT, with or without controlling for grades in economics courses.

TABLE 2
Results of Estimating Basic Equation for Major Field Test in Economics (MFAT) — Pomona College Economics Majors (MFAT Score = Ao + A1 MALE + A2 GPA + error term)

Variable	Pomona College—MFAT			Occidental—GRE		
	All	Men	Women	All	Men	Women
MALE	2.53	—	—	46.5	—	—
	(1.89)			(11.2)		
GPA	6.77	6.83	6.52	94.0	93.2	98.3
	(.66)	(.76)	(1.39)	(10.1)	(11.6)	(19)
Constant term	104.4	106.4	106.8	312.1	361.2	299.2
Observations	93	69	24	149	118	31
R^2	.53			.42	.35	.48
% male	74	100	0	79	100	0
Mean MFAT	171.3	171.8	169.8	636 (GRE)	646	598
Mean GPA	9.6	9.58	9.65	3.05	3.06	3.04
SEE				55.7	58.7	42.7

Notes: Standard errors are in parentheses.
Definitions: MALE = dummy variable, male = 1. MFAT = Major Field Achievement Test Score in Economics. The scale ranges from 120 to 200. GPA = Grade point average in all economics courses. Pomona uses a 12-point scale: A = 12; B = 9; C = 6, etc.

The regressions for Pomona men and women run separately are reported in columns 2 and 3 of Table 2. The two groups had quite similar constants and intercepts. Indeed, the Chow test for comparing two regressions indicated that we could not reject the null hypothesis that the Pomona equations are the same.[11] The regression represented by column 1 in Table 2 implies that women with the same GPA as men at Pomona scored almost as well on the MFAT.[12] The results made it clear that the Pomona-MFAT data and the Occidental-GRE data gave quite different pictures of the relationship between gender and test scores. It is possible, of course, that the difference arose from some sort of difference between the economics programs and economics students at Occidental and Pomona, but these schools have quite similar philosophies and attention to pedagogy, and in both samples women had slightly higher GPAs than men. In addition, the national averages on the GRE and MFAT tests mirror our results from Occidental and Pomona. As such, we feel that the interesting hypotheses link gender differences in performance to differences across the two tests.

A POSSIBLE EXPLANATION WITH TESTABLE IMPLICATIONS

In searching for an explanation of this persistent gender gap in test performance on the GRE, two hypotheses suggested themselves. First, it may be that on average men and women self-select economics with somewhat different interests, and the GRE is more nearly aligned with the subset of economics that appeals to men. If the GRE accurately reflects the skills and interests relevant to further productivity in economics, the gender gap is a signal that the areas of economics that attract women as undergraduates receive less emphasis in advanced study. Alternatively,

the content and or structure of the GRE exam may have inherent biases that lead it to underestimate, on average, the potential of women economists. If, for example, some economists excel at manipulating analytical tools and others at connecting analytical techniques to important problems, a test that focuses on the first set of skills will underrate the second set of economists.

Under the first hypothesis, the GRE is doing its job.[13] Under the second hypothesis, the GRE is systematically misevaluating candidates for graduate school, with potentially costly implications for women and the economics profession. Because its consequences are more serious, we pursue the second hypothesis here. We offer a set of hypotheses about the potential sources of bias in the GRE, outline some potential tests (both informal and formal) of the hypotheses, and report some very preliminary results.[14]

Men also outperform women on the SAT exams. For example, in 1980 the average verbal score for men was 8 points higher than that of women, and the gap between the means on the mathematical score was 48 points. Studies indicate that these results are biased in favor of men. An examination of the numerous tests of mathematical ability shows a slight advantage (which does not approach the difference measured by the SAT) for men that has been diminishing over time. There is a similarly slight advantage for women in tests of verbal ability. Thus, the SAT appears to overstate male abilities in both categories, and is considered by many psychologists to be a flawed instrument (Fausto-Sterling 1985, 262).[15] The GRE Subject Test in Economics is structurally similar to the SAT and may carry the same biases.

Our results have shown that even holding constant economics GPA and SAT scores (along with whatever bias inheres in those scores), men score on average 40 points higher than women. A straightforward response to our results might be to ask whether any advantages enjoyed by men on the SAT exam are magnified on the GRE Subject Test in Economics. There are two obvious differences between the exams. First, the GRE exam is taken at the end of the college career, rather than at the beginning. Second, the GRE exam is in a field that continues to be dominated by men; in 1988, only 32.8 percent of all bachelor's degrees in economics were received by women, up from 9.8 percent in 1966 (U.S. Department of Education 1990).

Both the SAT and the GRE Subject Test are multiple-choice exams administered within a rigid time constraint. Both exams present some problems that appear unfamiliar, but that can be analyzed using familiar tools. Both reward a willingness to guess when some choices have been ruled out. On exams like this, confidence and competitiveness are two attributes that might enable students to get maximum scores for their ability level. There are several reasons for this. First, questions that appear unfamiliar may scare off students lacking confidence. In contrast, students confident of their abilities may be energized by the challenge. Thus students with confidence would tend to answer more questions than students with less confidence. The confidence gap would be exacerbated by the time constraints of the exam. Students with less confidence may be more afraid of running out of time, and so might skip questions they are, in reality, capable of handling.[16] Students who thrive in competitive situations are likely to maintain a more confi-

dent attitude under the stress of the exam than those students who tend to avoid competition. The smaller gender gap on the MFAT, which allows more time per question and is a less intimidating exam, is consistent with this explanation.[17]

Differences in confidence might lead to differences in guessing strategies. On the SAT and GRE, students score a point for each correct answer and lose one-fourth point for each wrong answer. If one potential answer can be eliminated, guessing leads to a higher expected score. Students with confidence might be more willing to risk being wrong than students without confidence. In this light, it is interesting to note that the MFAT, which has a much smaller gender gap, has no penalty for wrong answers.

The gap in confidence and competitiveness is likely to be greater when students take the GRE economics exam than when they take the SAT. The GRE comes at the end of the college career. It appears that women systematically lose confidence over the course of their college years, while men maintain or strengthen their confidence. For example, Karen Arnold studied a group of high school valedictorians through their college careers. The women earned higher grades than the men on average. However the men had sharply higher estimates of their own intelligence as they moved through college. The percentages are as follows (Arnold 1992, 398):

Intelligence self-estimate	High school senior		College sophomore		College senior	
	Male	Female	Male	Female	Male	Female
Average/slightly above average	30	35	28	54	19	30
Above average	47	44	50	42	56	70
Far above average	23	21	22	4	25	0

Hall and Sandler (1983) have also argued, based on substantial empirical work, that the classroom environment in college (differential responses from professors, harassment from male students) tends to discourage women.[18]

Pursuing this hypothesis in a highly informal manner, the two authors at Occidental tried to predict which students would have large residuals based on their personalities. This exercise was quite successful, including the identification of two women who had the type of confidence we have discussed.

We also wondered about the gender gap on the GRE Subject Test in other disciplines. The hypothesis above would suggest that this gap should be widest in those disciplines where there are fewer women relative to men, ceteris paribus. The gap should also be wider in those disciplines in which test problems appear unfamiliar but which can be analyzed using familiar tools.[19] It turns out that the gender gap is largest in math and physics (100 and 70 points, respectively) where men outnumbered women test takers by a 2 to 1 and 4 to 1 margin, respectively. The gender gap is smallest for psychology, education, and sociology where women score 10 points less, 5 points less, and 5 points more than men, respectively. In each of these latter three disciplines, more women than men were test takers.[20]

More formal tests of the confidence gap hypothesis are theoretically possible. It would be useful, for example, to test the following predictions: (1) men attempt more questions and are more willing to make informed guesses, and therefore

leave fewer questions unanswered; (2) the gender gap would be smaller if the GRE Subject Test were administered without, or with less severe, time constraints; and (3) correlation between reported self-esteem (or similar measures from the psychology literature) and GRE scores would be positive, all else held constant.

None of these tests lies within the scope of our data. We asked ETS if they could help us with the first prediction and they graciously agreed to run a limited test using data from one recent administration of the GRE Subject Test in Economics.[21] Specifically, the results show that of the 130 individual test questions, men left 17 percent blank, while the women left 20 percent blank. Unfortunately, we were unable to obtain the exact sample sizes used by ETS in calculating these figures, but on average about 500 men and 180 women would be taking the GRE Subject Test in Economics at any one administration. Because sample sizes of only 10 correspond to a t value of 1.97, we believe it is safe to conclude that this difference is significant.[22] The statisticians also calculated for each question the percentage of men who left it blank and the percentage of women who left it blank. For 117 of the questions, the difference between these percentages was less than 10 percentage points, for example, the percentage of men who left it blank might have been 6 percent, while the percentage of women, 8 percent, for a difference of 2 percentage points. For the remaining 13 questions, however, the difference in these percentages was greater than 10 percentage points; 23 percent of the men left one of these questions blank while 39 percent of the women did, for a difference of 16 percentage points. In all these cases where the difference was greater than 10 percentage points, the men always had a lower percentage left blank.[23]

CONCLUSIONS AND IMPLICATIONS

In trying to account for the fact that women score about 50 points lower than men on the GRE Subject Test nationally, we have tried to show that some of the more obvious explanations do not explain the differential. Our analysis of the data from senior economics majors at Occidental College over the past five years shows that although economics GPA and SAT scores are significant variables in predicting the GRE Subject Test score, women with the same economics GPA and SAT scores, and even math background and ability, as their male counterparts still score about 40 points lower on the GRE Subject Test. Although earlier research has suggested that women do worse on standardized multiple-choice exams, our analysis of data from senior economics majors at Pomona College over the past three years indicates that this differential between men and women all but disappears on the Major Field Test in Economics, a shorter and easier version of the GRE Subject Test.

The explanation we offer suggests that the GRE Subject Test is a somewhat flawed instrument for measuring a student's aptitude for, and competence in, the basic skills of economic analysis. Because the GRE Subject Test is a multiple-choice exam administered with a rigid time constraint involving problems that appear initially unfamiliar but can be analyzed using familiar tools, confidence

and competitiveness are two attributes that might be more important to GRE performance than to performance as an economist. We conducted several very informal investigations of our hypothesis and had ETS perform a more formal test of one of our predictions. All the results were consistent with our confidence gap hypothesis. Naturally, further investigation of this issue would be useful.

We hope our empirical results will spur further investigation by others. It would be useful to know if our findings at Occidental and Pomona can be duplicated at other institutions that require their students to take the GRE Subject Test or the MFAT. Perhaps just as important, we hope that graduate schools that currently use the GRE Subject Test as part of their admissions process, and undergraduate programs that, like Occidental, have used the GRE Subject Test as part of their evaluation of senior economics majors, might keep our results in mind when considering the scores of women that appear out of line with the other measures they use.[24]

NOTES

1. This information comes from phone conversations with Dawn S. Robinson, Associate Program Director of the GRE Program Staff.
2. According to the ETS (1990, 2), "the Major Field Achievement Tests are based on the GRE Subject Tests, but they are shorter and less difficult, making them suitable for all seniors majoring in the field, not just those planning graduate study." The MFAT results from Pomona College also make it difficult to accept an amended version of the first explanation above. Such an explanation would start with the assumption that men have better aptitude for, and competence in, the basic skills of economic analysis, but add that no independent variables, such as economics grades and SATs, can adequately control for such aptitude and competence.
3. Several individuals, including some at ETS, wondered whether our results were due to the fact that there could be differences by gender in the economics courses taken. For example, it was suggested that a smaller proportion of women might take econometrics. Although such a conjecture might be important for a national sample, the breadth of the Occidental requirements, that require every major to take a course in statistics and another in econometrics, would seem to rule this out for our sample.
4. It should be noted that some psychologists contend that the SAT exam itself is biased in favor of men.
5. We obtained this percentage by subtracting the coefficient on MALE from the mean for men, and then dividing this difference by the mean for men.
6. The F statistic was 3.32 with 4, 141 degrees of freedom. The critical F-statistic value at the .05 level for 4, 120 *d.f.* is 2.45.
7. We also ran all equations in Table 1 with additional dummy variables for the graduation year of each student. This was to control for the possibility that different classes that have slightly different means on the GRE also might have different proportions of men. The coefficients of these dummy variables never had significant t statistics nor did their inclusion alter the coefficients of the independent variables in Table 1 in any significant way.
8. All Pomona College economics majors take principles of macroeconomics, principles of microeconomics, statistics, intermediate microeconomics, intermediate macroeconomics, one semester of calculus, and attend a senior colloquium. Students have the option of either taking five additional upper division courses in economics or pursuing an alternative set of requirements that includes substantially more mathematics (one course beyond linear algebra and differential equations) and one fewer economics courses. Most Pomona economics students planning to go to graduate school take the GRE, and some of these students (three men and two women from the graduating classes in our sample) exercised their option to substitute the GRE for the MFAT in fulfilling their testing requirement.
9. The MFAT has only 90 questions and lasts 120 minutes (1.33 minutes per question) while the GRE Subject Test has 135 questions and lasts 170 minutes (1.26 minutes per question), according to the materials provided to test takers (Educational Testing Service 1987, 6).

10. Nationally on the MFAT for the period 1989–1992, the mean for 1,175 men was 154, and the mean for women was 148. Thus, the women's mean was 95.8 percent of the men's. For the Pomona students, this ratio was a bit higher at .988. Both are well above the .92 value nationwide on the GRE Subject Test.
11. The Chow test *F* statistic to determine if the regression for men is different than the regression for women has a value of only .016 with 2, 89 degrees of freedom. This is far below the critical value of 1.40 for the .25 level. Thus, we cannot reject the null hypothesis that the Pomona equations are the same. For the Occidental results, however, the *F* statistic was 8.52 with 2, 145 degrees of freedom. Thus, we could reject the null hypothesis that the Occidental equations are the same.
12. If we subtract the coefficient on MALE from the mean MFAT score for men and then divide by the mean MFAT score for men, the resulting percentage is 98.5. This compares with a percentage of 92.6 using the GRE results in Table 2, column 4, at Occidental.
13. Schneider and Briel (1990, 9) in their most recent validity report of the GRE find that the GRE Subject Test in Economics has an average correlation with first year graduate school grades of .42, while the average correlation between first year graduate school grades and GRE Verbal, Quantitative, and Analytical test scores are .23, .26, and .27, respectively. Undergraduate GPA had a correlation coefficient of .34. This, by itself, does not mean that the GRE Subject Test is a good measure to use for admission because graduate schools might have the same biases in terms of confidence and competitiveness that we suggest the GRE may have. Further, it may be that graduate schools themselves, especially in the first year, may be emphasizing a particular subset of skills that are not important for the ultimate success of Ph.D. economists.
14. We should briefly note that the Educational Testing Service uses a statistical procedure on the GRE General Test known as Differential Item Functioning Analysis (DIF Analysis) to test for bias on any specific question. In a technical report, *Sex, Race, Ethnicity, and Performance on the GRE General Test,* DIF Analysis is described as based "on a comparison between groups of test takers of the same overall ability, as determined by their performance on the test as a whole. A DIF statistic is computed for each test question, indicating the extent to which members of the focal group perform differently from members of the reference group who have the same scores on the test as a whole. On the basis of this type of analysis, any questions that members of the focal group miss substantially more often than members of the reference group are deleted from the test scores used to match the two groups on ability" (Educational Testing Service 1989, 5). Upon closer questioning, it turns out that this DIF analysis is not performed on the GRE Subject Test in Economics (although it is on some other Subject Tests such as in psychology) and the number of females taking the GRE Subject Test in Economics on any given date is less than the 200 necessary for statistical validity. Dawn Robinson, of ETS, did have the statisticians at ETS perform this analysis for us for the December 1992 and December 1991 administrations of the Subject Test even though the number of women taking the test was too small for statistical validity. The results indicated that none of the questions would have been deleted based on the standard DIF analysis. A few were in the B or second category where there was some evidence of differential results but not serious enough to warrant deletion of the question (and some of these had women performing better than men). However, the failure of this DIF analysis to find evidence of a gender bias in any particular question does not necessarily imply that there are not inherent biases in the content and/or structure of the GRE exam taken in its entirety, as our discussion will demonstrate.
15. It is worth noting that women outperform men in the classroom as measured by GPA in both high school and college. This female advantage is observed in math courses (at all levels through graduate school) (Fausto-Sterling 1985, 266).
16. One of the authors teaches a mathematics course for economics majors who are intimidated by math. In this course, students complain particularly about the deleterious effects of time constraints on their performance. The author has found that exams administered to students without time constraints tend to give a better reading of the student's ability (as demonstrated on problem sets and by the student's responses in class) than do exams with time constraints.
17. The only careful analysis of the time-constraint issue that we discovered was a 1979 report by ETS entitled, *Effect of Increased Test-Taking Time on Test Scores by Ethnic Group, Age, and Sex* (Wild and Durso 1979). It investigated the effects of increasing the test time to reduce the time pressures of the verbal and quantitative experimental sections of the GRE Aptitude (but not Subject) Test. The results indicated that "although the 20-minute experimental tests (are) generally more speeded(sic) than the 30-minute tests, the 10 additional minutes resulted in a small score increase for all groups, and differential score gains were not found between the subgroups."
18. Of course, other more speculative reasons have been advanced for the type of confidence that enhances exam performance by men. For example, as far back as the turn of the century, Helen

Bradford Thompson devised some early versions of IQ tests. On these tests, she found that men and women performed equally well, except on puzzles. Her explanation for this disparity was that boys were "encouraged to independence and individuality, while girls were trained to 'obedience, dependence, and deference'—hardly the tools for effective problem-solving" (Russett 1989, 168). Others (Siegfried and Strand 1977, 247) cite as a standard hypothesis (for the better performance of men on economics exams) that "women students have grown up in a cultural environment in which girls are not supposed to like business and thus have a disadvantage in business or economics courses." Still others would cite the fact that women tend to be a smaller proportion of economics majors and thus find themselves in the uncomfortable position of being in the extreme minority in their economics classes. However, while these last two rationales could easily have negative consequences for the self-confidence of the women who do persevere in the major, it is also possible that only confident women stick it out in this environment.

19. We thank associate editor Peter Kennedy for directing our attention to this possible additional test of our hypothesis.
20. This information comes from a telephone conversation with Rob Dorso at ETS and is based on all students taking these exams during 1988–91. The data are not yet available, but are slated to be published in the *Faculty Guide to the GRE Subject Test*, forthcoming.
21. The information reported here was based on a phone conversation with Dawn S. Robinson. She indicated that prior to our inquiry, ETS had not investigated the number of questions left blank by gender.
22. Using a sample size of 500 men and 180 women, the t statistic for the difference between the means comes out to 10.075.
23. No further information was available in which to carry out a formal significance test. Further, we were not able to examine the 13 questions where the difference in percent left blank was greater than 10 percentage points to determine if there was any similarity in the type of question. We were informed that there was no discernable pattern as to when in the exam these 13 questions occurred. For example, it was not the case that the 13 questions came toward the end of the exam. Note that the confidence gap hypothesis does not require or even predict that women will have more unanswered questions at the end of the exam because of time pressure. All that matters is a perception that time constraints will become binding.
24. Readers whose undergraduate or graduate departments compile statistics relating test scores to performance are invited to report to us what data are available. Correspondence should be sent to Professor Robert L. Moore, Department of Economics, Occidental College, Los Angeles, CA 90041.

REFERENCES

Anderson, G., H. Benjamin, and M. A. Fuss. 1994. The determinants of success in university introductory economics courses. *Journal of Economic Education* 25 (Spring): 99–119.

Arnold, K. 1992. Academically talented women in the 1980's: The Illinois valedictorian project. In *Women's lives through time: Educated American women of the twentieth century*, ed. Kathleen Dan Hulbert and Diane Tickton Schuster. San Francisco: Jossey-Bass.

Educational Testing Service. 1979 and 1987. GRE: A description of the advanced economics test, 1980–1982, and 1988–90. Princeton, N.J.: Educational Testing Service.

———. 1989. Sex, race, ethnicity, and performance on the GRE general test: Technical report. Princeton, N.J.: Educational Testing Service.

———. 1990. Major field achievement tests: Assessment in the major. Princeton, N.J.: Educational Testing Service.

Fausto-Sterling, A. 1985. *Myths of gender: Biological theories about women and men.* New York: Basic Books.

Ferber, M. A., B. G. Birnbaum, and C. A. Green. 1983. Gender differences in economic knowledge: A reevaluation of the evidence. *Journal of Economic Education* 14 (Spring): 24–37.

Gohmann, S., and L. C. Spector. 1989. Test scrambling and student performance. *Journal of Economic Education* 20 (Summer): 235–38.

Hall, R. M., and B. R. Sandler. 1983. The classroom climate: A chilly one to women. Washington, D.C.: Project on the status and education of women. Association of American Colleges.

Heath, J. A. 1989. An econometric model of the role of gender in economic education. *American Economic Review* 79 (May): 226–30.

Lumsden, K. G., and A. Scott. 1987. The economics student reexamined: Male-female differences in comprehension. *Journal of Economic Education* 18 (Fall): 365–75.

Oaxaca, R. 1973. Male-female wage differentials in urban labor markets. *International Economic Review* 14 (3): 693–709.

Russett, C. E. 1989. *Sexual science: The Victorian construction of womanhood.* Cambridge: Harvard University Press.

Schneider, L. M., and J. B. Briel. 1990. *Validity of the GRE: 1988–89,* summary report. Princeton, N.J.: Educational Testing Service.

Siegfried, J. J. 1979. Male-female differences in education: A survey. *Journal of Economics Education* 10 (Spring): 1–11.

Siegfried, J. J., and S. H. Strand. 1977. Sex and the economics student. *Review of Economics and Statistics.*59 (May): 247–49.

Soper, J. C., and W. B. Walstad. 1988. What is high school economics? Posttest knowledge, attitudes, and course content. *Journal of Economic Education* 19 (Winter): 37–51.

U.S. Department of Education, National Center for Education Statistics, 1990. *Digest of Education Statistics.* Washington, D.C.: Government Printing Office.

Walstad, W. B., and J. C. Soper. 1989. What is high school economics? Factors contributing to student achievements and attitudes. *Journal of Economic Education* 20 (Winter): 23–28.

Watts, M., and G. J. Lynch. 1989. The principles course revisited. *American Economic Review.* 79(2): 236–41.

Wild, C., and R. Durso. 1989. Effect of increased test-taking time on test scores by ethnic group, age, and sex. *GRE Board Research Report No. 76–6R.* Princeton, N.J.: Educational Testing Service.

Williams, M. L., C. Waldauer, and V. G. Duggal. 1992. Gender differences in economic knowledge: An extension of the analysis. *Journal of Economic Education* 23 (Summer): 219–31.