VICTORIA UNIVERSITY OF WELLINGTON

GRADUATE SCHOOL OF BUSINESS
AND GOVERNMENT MANAGEMENT

WORKING PAPER SERIES 12/90

Gender Performance Differences in High School Economics and Accounting: Some Evidence from New Zealand

S.P. Keef *

* The Money and Finance Group,
Faculty of Commerce and Administration
Victoria University of Wellington

September 1990

ISSN 0114-7420
ISBN 0-475-11431-0

Gender Performance Differences in High School Economics and Accounting: Some Evidence from New Zealand

Abstract

This paper sought to determine whether performances in the University Bursaries examinations in economics and accounting exhibited any form of gender bias. A multivariate analysis of variance, with repeated measures, was used to determine gender performance differentials after controlling for the effects of academic ability, concurrent study of mathematics courses and inter-year variability. Although some minor gender biases were observed, the results were generally in accord with the thesis that significant gender biases were absent. In contrast, the popularity of this course combination exhibited a strong gender bias. There were twice as many males as females.

Keywords GENDER, ECONOMICS, ACCOUNTING, EXAMINATION PERFORMANCE, PRE-UNIVERSITY.

Gender Performance Differences in High School Economics and Accounting: Some Evidence from New Zealand

Introduction

It was widely held in the US that males generally do better than females in high school examinations in economics. Siegfried (1979, p.1), in an extensive literature survey, criticised most prior research on the basis that any reported gender performance differentials were a "... by-product of their primary objective.". He concluded that two-thirds of the relevant studies, and generally those that were more reliable, found that males outperformed females. A review of this literature shows that gender performance differentials were not present up to the age of 15 years (Davidson & Kilgore, 1971; MacDowell et. al., 1977). By the end of high school, however, gender performance differences appear to emerge. Although the majority of the US studies[1] found that males significantly outperformed females (Thornton & Vredeveld, 1971; Highsmith, 1974), other studies have reported that females were slightly, but not significantly, superior (Becker et. al., 1975). Hahn (1982) reported the absence of a gender differential.

Notwithstanding, the question of whether similar results would pertain in New Zealand has not been addressed in the literature. Indeed, empirical evidence on gender performance differentials in high school (Sixth and Seventh Forms) economics is, to the author's knowledge, non-existent in New Zealand. The situation is the same for the subject of accounting in the western world.

This paper seeks to address this deficiency in the literature. It addresses the simple question of whether there was a gender performance differential in the University Bursaries examinations in the subjects of accounting and economics. There are a number of reasons why this study provides a rigorous test of gender performance differentials.

(i) It links performances in both economics and accounting courses using a multivariate approach. This approach has

[1] It would appear that the UK was no different from the US (Attiyeh & Lumsden, 1971, pp. 85 and 1972, pp. 430-431).

not been previously adopted.

(ii) The analysis was based on a large sample size (n=6,499) covering a period of 5 years.

(iii) The performances were based on national examinations which were undertaken throughout New Zealand[2].

(iv) It was possible to control for concurrent proxies of academic ability and the concurrent study of two mathematics courses. Indeed, in the absence of such control, any observed gender performance differentials could be attributed to sampling[3] bias.

(v) It was not a by-product of another study.

Hypotheses

Given the conflicting empirical evidence, especially for economics examinations, it was decided to address the research question concerning gender performance differentials with the general hypothesis:

H_0: Gender performance differentials were absent.

Thus the hypothesis can be rejected if a gender characteristic[4] exceeded a specified significance level. Given the large sample size, a critical significance level of 0.01 was deemed appropriate. That is, there was a one in a hundred chance of a Type I error if H_0 was true. However, large samples can generate statistically significant results when the practical differences are almost minuscule. Thus significant results were investigated to determine their practical importance. The critical values were

[2] The curriculum for the economics examination was unchanged over the period. However, there were changes to the curriculum of the accounting examination. Since 1985, the curriculum has been financial accounting (60 per cent), management accounting (20 per cent) and accounting information systems (20 per cent). Prior to 1985, the curriculum was essentially financial accounting and bookkeeping.

[3] For, example, females could be found to outperform males since only the more intelligent females undertake the course (see Moyer & Paden, 1968, p. 875 for a similar argument as an explanation of their observed male superiority).

[4] Sex factor or any of its interactions with other factors.

the sum of squares explained (>1%) and differences in cell means (>1 residual mark)[5].

Methodology

The University Bursaries (UB) examinations' data tapes for academic years 1983 to 1987 were accessed to obtain data on each student that sat both the accounting and economics examinations in the same year. Cases were deleted that did not sit five examinations in the same year or one of the mathematics examinations. This resulted in a sample size of 6,499. There were approximately twice as many males as females in the sample[6].

A multivariate analysis of variance (MANOVA) with repeated measures was chosen to test the null hypotheses. The two dependent variables were the residual performances in the accounting (AC) and economics (EC) examinations. As repeated measures they were given the code name 'Mark'.

In essence a MANOVA with two repeated measures on the same subject is akin to two simultaneous analysis of variances (ANOVA) on: (i) the average of the two variables (AC+EC); and (ii) the difference between them (AC-EC). Thus the F_{Factor} and $F_{Interactions}$ ratios test for a mean effect. The $F_{Factor\ and\ Mark}$ ratios test for a difference effect. Thus a non-significant $F_{Factor\ and\ Mark}$ ratio would indicate that both variables had the same effect over the levels of the factor.

These residuals were obtained by eliminating the systematic effects of an academic ability covariate[7] on the reported raw mark. The covariate was the aggregate performance in the three other UB examinations taken in the same period. A separate model was used for each dependent course since there was evidence that

[5] However, sufficient detail is provided to allow the reader to arrive at independent conclusions.

[6] This ratio was relatively stable over the five years (chi-square(calculated) = 2.0 on 4 df, p>0.70).

[7] The rationale for such a procedure to control for potentially different samples of academic ability can be traced back to Spearman (1904). He argued that general academic ability, which he termed 'g', was the major determinant of academic performance.

the slopes were statistically different[8]. Reasons for the differences in the slopes are worthy of further research.

A three-way, mixed effects model was used. Sex (female/male) and Math (the three levels were the combinations of mathematics courses concurrently undertaken, that is, 1=Mathematics with Statistics, 2=Mathematics with Calculus and 3=Statistics+Calculus) were treated as fixed effects. Year (1983 to 1987) was treated as a random effect.

Because of the possibility that the slope[9] could differ between the sexes, it was necessary to test for homogeneity of slopes. The standard test (Hull & Nie, 1981, p. 16) indicated statistically significant denials of this assumption for each subject (p's<0.001). However, since the sums of squares explained by the interactions of the ability covariate and factors (Sex, Year and Math) were relatively small (accounting = 0.3 per cent and economics = 1.3 per cent) the use of a model of unequal slopes was not deemed appropriate.

There are three assumptions that should be met by the data for the correct use of a MANOVA test. The data should be: (a) independent[10]; (b) sampled from normal populations and (c) of homogeneous variance (compound symmetry). The first assumption is the most critical (Lindman, 1974). Given the the exceptionally low possibility that a student who has sat five UB examinations in one year will return in a later year to sit another five examinations, there was confidence that the independence assumption was not denied[11].

---

[8]The regression coefficients (and their standard errors) were:

	Intercept	Slope	r^2
Accounting	12.97 (0.60)	0.258 (0.004)	0.46
Economics	5.77 (0.69)	0.302 (0.004)	0.46

[9] Between performance in the examination and the proxy for academic ability.

[10] It is not necessary that the repeated measures on the same subject are independent. Indeed it is generally the case that they are not.

[11] There are other causes of dependence, for example, common teachers.

Preliminary tests on the uncorrected accounting and economics marks denied the normality assumption. The populations by Year, showed unexpected peaks especially at just above the 50 per cent mark. This, it was believed, was either due to scaling or the reconsideration of marginal failures. However, visual inspection, contrary to the more precise statistical tests[12], gave the impression of a not unreasonable approximation to normality. Notwithstanding, ANOVA's are generally accepted as being robust to mild departures from normality.

There was strong evidence that the compound symmetry assumption, the MANOVA repeated measures equivalent of homogeneity of variances for ANOVA, was denied for the total data (Box's M = 444, $p<0.001$). Such denials, however, may be an inherent problem with research into gender performance differentials (see Appendix 1).

The denial of compound symmetry combined with a large range in sample sizes[13] can result in unpredictable bias in the estimates of the F ratio (Winer, 1971, p. 205). There were two alternatives available. The first was to eliminate the heterogeneity of variances by an appropriate transformation. The second was to eliminate the non-orthogonality. Given the relatively large sample size, the latter was deemed to be the more appropriate approach.

Each of the 6,499 cases was assigned a pseudo random normal deviate. The cases in each cell were ranked by this deviate. The first 49 cases in each cell were sampled[14]. The resultant sample size was 1,470. To alleviate problems associated with sampling errors, the procedure was replicated ten times. The F ratios presented in this study are based on the average of these ten tests. Since there was not a 'considerable' difference in the cell sample sizes, the denial of the compound symmetry assumption[15] would result in a small positive bias against the null hypothesis (Winer, 1971, p. 205). That is, the observed F ratio will be overstated. In contrast, the cell means reported in the tables

[12] On average, the residuals exhibited a lesser deviation from normality than did the raw marks.

[13] The cell frequencies ranged from 49 to 654.

[14] This was the minimum cell sample size.

[15] As was to be expected, Box's M (mean = 231 with a range of 188 to 269) showed that the compound symmetry assumption was still denied ($p's<0.001$).

were based on the total sample (6,499). Nevertheless, the reader is cautioned to interpret the findings carefully in light of these departures.

Results

The results of the MANOVA with repeated measures are presented in Table 1. Statistically significant inter-year differences (Year, Year and Mark, p's<0.001) were to be expected due to difficulties associated with achieving perfect scaling over time[16]. There was one reason, however, to believe that the inter-year differences were of limited practical importance. The sums of squares explained were small[17]. The maximum observed yearly deviation from the grand mean was 1.5 residual marks (Table 2). Nevertheless, the statistical design would have controlled for these effects if they were indeed important.

In the discussion that follows it is important to be aware that a causal effect has not been proven. That is, the learning experience provided by the chosen mathematics course(s) directly influenced residual performance in the economics and accounting examinations. Whereas the degree association is characterised by the statistical tests, the underlying causal effect is not beyond doubt. However, it is unlikely that the results are a reflection of, say, intelligent students choosing a particular combination of mathematics courses. The use of a proxy of general academic ability to eliminate the systematic effects in the observed, raw performances was specifically incorporated in the methodology to control for such possibilities.

The choice of the mathematics course(s) undertaken at UB had a significant association with aggregate performance in the two courses and the difference between them (Math, Math and Mark; p's<0.001). The results (Table 3) are apparently complex. At the aggregate level, statistics was associated with an important disadvantage compared to either calculus or calculus+statistics (Math Contrast A, $p<0.001$; Δresidual mark>4.0).

[16] The examination authority carried out two types of scaling. The first was to eliminate between marker effects. The second sought to ensure that the aggregate Bursary mark was independent of the combination of subjects undertaken.

[17] 1.2 per cent for Year and 1.2 per cent for Year and Mark.

TABLE 1 MANOVA RESULTS WITH REPEATED MEASURES (AC, EC)

Source of variation	Sum of squares(a)	Degrees of freedom	Mean square	F ratio(b)	Significance of F ratio
Sex (S)	648	1	648	3.34	>0.10
Year (Y)	1,973	4	493	4.79	<0.001
Math (M) (c)	16,086	2	8,043	85.56	<0.001
Math Contrast A	15,939	1	15,939	169.56	<0.001
Math Contrast B	147	1	147	1.56	>0.10
S by Y	775	4	194	1.88	>0.10
S by M	376	2	188	2.54	>0.10
Y by M	750	8	94	0.91	>0.50
S by Y by M	593	8	74	0.72	>0.50
Residual	148,320	1,440	103		
Total	169,521	1,469	115		
Mark	147	1	147	2.29	>0.10
Sex and Mark	92	1	92	0.62	>0.25
Year and Mark	1,344	4	336	5.25	<0.001
Math and Mark (c)	14,054	2	7,027	92.46	<0.001
Math Contrast X and Mark	13,986	1	13,986	184.02	<0.001
Math Contrast Y and Mark	68	1	68	0.89	>0.25
S by Y and Mark	594	4	148	2.32	~0.05
S by M and Mark	655	2	328	2.73	>0.10
Y by M and Mark	609	8	76	1.19	>0.25
S by Y by M and Mark	957	8	120	1.87	>0.05
Residual	92,160	1,440	64		
Total	110,612	1,469	75		

Notes:

(a) The sums of squares were reconstituted from the average F ratios and the mean square of the residual.

(b) The F ratios, using a mixed-effects model, were calculated as follows (Winer, 1971, p. 346):

$F_{Factor} = MS_{Factor}/MS_{Factor*Year}$,

$F_{Year} = MS_{Year}/MS_{Residual}$,

$F_{Factor\ by\ Factor} = MS_{Factor\ by\ Factor}/MS_{Factor\ by\ Factor\ by\ Year}$,

$F_{Factor\ by\ Year} = MS_{Factor\ by\ Year}/MS_{Residual}$,

with the corresponding degrees of freedom.

(c) The Math contrasts were:

Level of Math factor	Contrast A	B	X	Y
1=Statistics	-2	0	-1	1
2=Calculus	1	-1	1	1
3=Statistics+Calculus	1	1	0	-2

TABLE 2 CELL MEANS: YEAR FACTOR[a]

Year	AC	EC	AC + EC	AC - EC
1983	+0.3	+1.1	+0.7	-0.8
1984	+0.1	+0.7	+0.4	-0.6
1985	+0.2	+1.5	+0.9	-1.3
1986	-0.5	-1.5	-1.0	+1.1
1987	-0.1	-1.3	-0.7	+1.2

Note:

(a) Deviations from the grand mean. Based on separate ANOVA's by Year after controlling for the effects of the Sex and Math factors.

TABLE 3 CELL MEANS[a]: MATH FACTOR

Level of Math factor	AC	EC	AC + EC	AC - EC
1 = Statistics	-0.7	-6.1	-3.4	+5.4
2 = Calculus	-1.8	+2.9	+0.6	-4.7
3 = Stat + Calc	+1.2	+0.6	+0.9	+0.6

Note:

(a) Adjusted for the effects of the Sex and Year factors.

However, this effect was not consistent between the two subjects (Math and Mark, $p<0.001$). Strictly interpreted, the results indicated that statistics and calculus, as the sole mathematics subject, had a different association with the two independent subjects (Math Contrast X and Mark, $p<0.001$). That is: (a) the study of statistics provided an important comparative advantage to accounting (Δresidual mark=5.4); and (b) the study of calculus provided an important comparative advantage to economics (Δresidual mark=4.7). These two comparative advantages tended to cancel each other. The study of statistics+calculus had the same effect, on average, as the study of the two subjects individually (Math Contrast Y and Mark, $p>0.25$).

If these results can be attributed to the benefits provided by the concurrent study of these mathematics courses and the goal is to maximise overall Bursary mark, then students, who attempt both the accounting and economics courses, are advised to avoid statistics as the sole mathematics subject. With the addition of a further assumption, that is, sample bias was not present[18] the following advice can be tentatively offered. Where the student intends to sit only one of the mathematics examinations and only one of accounting and economics, then accounting with statistics and economics with calculus are the permutations that will enhance the grades. Accounting and calculus would carry a small penalty; economics and statistics would be disastrous (Δresidual mark=6.1).

On average and after controlling for the effects of (i) academic ability; (ii) the concurrent study of mathematics; and (iii) inter-year variability, females outperformed their male peers in the combined accounting and economics courses (AC+EC)[19]. However, this result could be considered of minor importance in view of the following observations (Table 4).

(i) The statistical evidence was weak (Sex, $p>0.10$), that is,

[18] The sample is biassed in that it examines only those students who undertook both the accounting and economics courses. There is a possibility that these results may not extend to those who did not take one or other of these examinations. This is an area worthy of further investigation.

[19] The significance of the Sex by Year interaction ($p\approx0.05$, but explaining less than one per cent of the sum of squares) can be explained by the fact that the female superiority was beyond doubt in 1984 ($p<0.001$) and very close to the chosen confidence level in 1983 ($p=0.02$) yet was not statistically significant in the other three years ($p's\geq0.08$).

TABLE 4 CELL MEANS: SEX BY YEAR[a,b]

	AC		EC		AC + EC		AC - EC	
Year	Δ	prob	Δ	prob	Δ	prob	Δ	prob
1983	+1.4	0.02	+1.0	0.12	+1.2	0.02	+0.4	0.60
1984	+0.3	0.54	+3.8	<0.001	+2.0	<0.001	-3.4	<0.001
1985	-0.7	0.14	+1.0	0.08	+0.2	0.70	-1.7	0.01
1986	+0.6	0.26	+0.9	0.12	+0.7	0.08	+0.3	0.69
1987	-0.1	0.83	0.0	0.98	0.0	0.90	-0.1	0.85
Total	+0.2	0.28	+1.3	<0.001	+0.7	>0.10	-1.0	>0.25

Notes:

(a) Based on separate ANOVA's by Year.

(b) Δ represents female superiority.

there was at least one chance in ten of a difference greater than that observed when sampling errors alone are responsible.

(ii) The sum of squares of the sex factor (<0.4 per cent) was very low.

(iii) The difference in aggregate marks (Δresidual mark=0.7) was generally less than the observed inter-year variability.

On the other hand, there was some evidence of female superiority. In the economics course females consistently achieved higher marks than males ($p<0.001$, Δresidual mark=1.3). It was only in the accounting course that the sexes appeared evenly balanced ($p=0.28$, Δresidual mark=0.2). However, the evidence would suggest that these observed gender differences between the two course were essentially illusory (Sex and Mark interaction, $p>0.25$).

Concluding remarks

This study sought to characterise the degree of gender performance differentials in the University Bursaries accounting and economics examinations. Bearing in mind sample bias, the problems encountered with the statistical tests and the methodology adopted to overcome them, the following conclusions can be drawn. Firstly, there was strong support for the thesis that concurrent course work in mathematics was associated with performance in the two examinations. Although it seems sensible (to the author) that the probability of a causal effect exists, it is important to acknowledge that doubt must remain. Thus there remains the challenge to researchers to characterise the causal direction in the observed association. To put the association in context, in terms of the sums of squares explained, it was eight times greater than that observed for the Year effect. The relative unimportance of the Sex by Math factor ($p>0.05$) was consistent with the view that the possibility exists that previously reported gender performance differences in the US and UK can be attributed, in part, to differences in prior or concurrent course work exposure (see Pallas & Alexander, 1983 for a similar argument with the subject of mathematics).

Secondly, considering the accounting and economics courses as a combined unit, there was evidence that, on average, females achieved higher residual marks than males. However, the differences in marks were relatively small and the statistical significance was not convincing. To all intents and purposes, the accounting examination was devoid of gender bias. However, the economics examination exhibited a subtle gender bias with females consistently doing better than males. This latter result is contrary to the body of prior research in the UK and US which shows that males were superior to females.

The possibility exists, for some reason or other, that either New Zealand high school students or the structure of the courses are what can be termed the 'immature state of development' with respect to their counterparts in the UK and US. That is, the subjects analysed were at a stage where gender differences were absent (Davidson & Kilgore, 1971; MacDowell et. al., 1977). If this were the case, then it is necessary to search for gender performance differentials at the next level of education, that is, the first year at university. There is evidence that a significant gender performance differential does not exist in the first-level university course in accounting (Keef, 1989). However, to the author's knowledge, a similar study has not been carried out with university economics in New Zealand.

For those that subscribe to the view that the presence or absence of a magic chromosome has no influence whatsoever on academic performance, these results will come as no surprise. They would argue that environmental factors were the sole cause of any gender performance differentials that have been reported earlier. There was strong support for this view with the accounting examination. With the economics examination there was a modest degree of uncertainty. Although further research into gender performance differentials, particularly in economics would be worthwhile, there would appear to be a more pressing gender bias to be addressed. That is, why were there twice as many males as females in the accounting/economics course combination? Why was this ratio relatively stable over the five years studied?

Appendix 1

The data base permits exploratory analysis of the systematic effects of Sex, Subject, Math, and Year on the observed cell variances of the abnormal marks. The cell sample size was used as a covariate. Table A.1 presents the results of an orthogonal ANOVA to test the implicit null hypotheses.

The observation that the metric covariate of cell sample size was essentially independent of the cell variances ($p=0.53$) would point to the fact that any heterogeneity was not caused by unequal sample sizes.

The observation that the Year effect and its interactions (p's>0.22) were not important provides some solace for researchers analysing UB marks. Although the results obtain to the abnormal marks, it can be inferred that the scaling carried out by the examining authority would appear to have either:

(a) removed any heterogeneity by year, if it originally existed; or

(b) not created heterogeneity if it was initially absent.

The factors of Math and Course (the levels were: 1=economics and 2=accounting), together with their interaction, illustrated important differences in the variances (p's<0.001). Whilst these two main factors were moderately important (each explained 15 per cent of the sum of squares), their interaction was of paramount importance (explaining a quarter of the sums of squares). These results can be attributed, in the main, to a single effect. The variance of the abnormal marks of those studying both economics and statistics, Table A.1, footnote (c), were twice as large as any other combination. This result was consistent over the five years.

These visual interpretations were confirmed by the statistical analysis. Math Contrast #1 explained all of the factor's sum of squares. Thus the sole differences were between the study of statistics (only) and the other two levels of the Math factor. The Course by Math Contrast #1, which explained 90 per cent of the interaction's sums of squares, indicated that the major difference existed between the two courses (AC and EC).

TABLE A.1 ANOVA RESULTS, CELL VARIANCES AS THE DEPENDENT VARIABLE

Source of variation	Sum of squares	Degrees of freedom	Mean square	F ratio	Significance of F ratio
Sample size	94	1	94	0.41	0.53
Sex (S)	1,866	1	1,866	8.20	0.008
Math (M)	8,276	2	4,138	18.17	<0.001
Math Contrast #1	8,275	1	8,275	36.30	<0.001
Math Contrast #2	1	1	1	0.00	>0.95
Year (Y)	1,397	4	349	1.53	0.22
Course (C)	8,351	1	8,351	36.67	<0.001
S by M	104	2	52	0.23	0.80
S by Y	277	4	69	0.30	0.87
S by C	1	1	1	0.00	>0.95
M by Y	4,920	8	615	2.70	0.02
Y by C	1,144	4	286	1.26	0.31
C by M	14,165	2	7,082	31.10	<0.001
C by M Contrast #1	12,768	1	12,768	56.06	<0.001
C by M Contrast #2	1,397	1	1,397	6.10	0.02
Explained	49,398	30	1,646	7.23	<0.001
Residual	6,604	29	228		
Total	56,002	59	949		

Notes:

(a) Since the object was to investigate the effects of the factors within this specific data set, the Year factor has been treated as a fixed effect. If the Year factor were treated as a random effect, then the probabilities of the F ratios for the other factors would be reduced to: Sex (0.003), Math (0.025) and Course (0.003).

(b) Although the design was orthogonal, it was not possible to determine the sums of squares for the three-way, and higher, interactions due to a singular matrix being present. Given this problem, the sums of squares were estimated using an hierarchical method.

(c) The Math contrasts and the variances of the cells were:

	Contrast		Variances	
Level of Math factor	#1	#2	AC	EC
1=Statistics	-2	0	73	137
2=Calculus	1	-1	77	68
3=Statistics+Calculus	1	1	63	78

Explanations for this perplexing observation are not readily forthcoming. Not only did the concurrent study of economics and statistics (only) result in a greater variance in abnormal marks, but it also resulted in negative abnormal performance. It is tentatively suggested that further research is necessary before these results can be fully understood.

Having adjusted for the effects of the other factors, females were found to be 21 per cent less variable than males[20]. Whilst both the level of statistical support (p=0.008) and the differences in the variability appeared to be of some importance, the sums of squares explained was only modest at 3.3 per cent. Nevertheless, this gender effect was consistent across the other factors (p's>0.80). Thus, an important area to be addressed by future research is why males exhibit more variability than females in the University Bursaries examinations in accounting and economics.

[20] Forbes (1988) reports a much larger gender difference for the Universities Scholarships examination in Mathematics with Statistics in 1986.

References

Attiyeh, R. and K. G. Lumsden, (1971), "University students' initial understanding of economics: the contribution of the A level economics course and of other factors", Economica, Vol. 38, No. 149, February 1971, pp. 81-97.

Attiyeh, R. and K. G. Lumsden, (1972), "Some modern myths in teaching Economics: the U.K. experience" American Economic Review, Vol. 62, May 1972, pp. 429-433.

Becker, W. E., J. D. Helmberger and J. L. Thompson, (1975), "An evaluation project given limited data", Journal of Economic Education, Vol. 6, Spring 1975, pp. 120-125.

Davidson, D. G. and J. H. Kilgore, (1971), "A model for evaluating the effectiveness of economic education in primary grades", Journal of Economic Education, Vol. 3, Fall 1971, pp.17-25.

Forbes, S. D., (1988), "Gender differences in performance in the 1986 Mathematics with Statistics Scholarship examination", The New Zealand Statistician, Vol. 23, No. 1, 1988, pp. 17-25.

Hahn, C. L., (1982), "Are there gender differences in high school Economics in the Eighties?", Journal of Economic Education, Vol. 13, Summer 1982, pp. 57-65.

Highsmith, R., (1974), "A study to measure the impact of in-service institutes on the students of teachers who have participated", Journal of Economic Education, Vol. 5, Spring 1974, pp. 77-81.

Hull, C. H. and N. H. Nie, (1981), SPSS Update 7-9, McGraw-Hill, 1981.

Keef, S. P., (1989), "Does gender affect performance in accounting courses?", Pacific Accounting Review, Vol. 2, No. 1, December 1989, pp. 77-87.

Lindman, H. R., (1974), Analysis of Variance in Complex Experimental Designs, San Francisco: W. H. Freeman and Company, 1974.

MacDowell, M. A., P. R. Senn and J. C. Soper, (1977), "Does sex really matter?", Journal of Economic Education, Vol. 9, Fall 1977), pp. 28-33.

Moyer, M. E. and D. W. Padden, (1968), "On the efficiency of the high school economics course", American Economic Review, Vol. 58, September 1968, pp. 870-877.

Pallas, A. M. and K. L. Alexander, (1983), "Sex differences in quantitative SAT performance: new evidence on the differential coursework hypothesis", American Educational Research Journal, Vol. 20, No. 2, Summer 1983, pp. 165-182.

Siegfried, J. J., (1979), "Male-female differences in economic education: a survey", Journal of Economic Education, Vol. 10, No. 2, Spring 1979, pp. 1-11.

Spearman, C., (1904), "General intelligence", American Journal of Psychology, Vol. 15, No. 2, 1904, p. 201-293.

Thornton, D. L. and G. M. Vredeveld, (1977), "In-service education and its effect on secondary students: a new approach", Journal of Economic Education, Vol. 8, Spring 1977, pp. 93-99.

Winer, B. J., (1971), Statistical Principles in Experimental Design, second edition, McGraw-Hill, 1971.