DOCUMENT RESUME

ED 377 239 TM 022 501

AUTHOR	Hou, Jack W.
TITLE	Class Size and Determinants of Learning Effectiveness.
PUB DATE	Apr 94
NOTE	32p.
PUB TYPE	Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Academic Achievement; *Class Size; *College Students; Comparative Analysis; *Economics; Females; Higher Education; High Schools; *Learning; Males; Questionnaires; Regression (Statistics); *Sex Differences

ABSTRACT

Based on the comparison of a regular-sized (n=25) and an oversized (n=54) class of college students taught by the same instructor, significant superior performance was found in favor of the large class. This disparity persisted even when a switching regression model (with endogenous switching) was performed. Only 39% of the learning disparity can be explained by characteristics of the student, including a 29.5% performance gap that is attributed to the potential "overachiever" trait of small class students. In addition, it was found that overall grade point average and performances in previous college economics courses did not have any significant impact on the students' current scores. Contrary to most of the findings in recent literature, high school economic education did not have any positive impact on the students' performance. In fact, it had a slightly negative effect. No evidence of gender difference was found for the large class, but a fairly significant lower performance was found for females in the small class. The student questionnaire is attached. Three tables are included. (Contains 28 references.) (Author/SLD)

Class Size and Determinants of Learning Effectiveness

Jack W. Hou

Department of Economics
California State University-Long Beach
1250 Bellflower Blvd.
Long Beach, CA 90840
TEL: (310)985-4710
FAX: (310)985-4979

April 1994

Working paper. Comments welcome.

Abstract

Based on the comparison of a regular sized class and an oversized one, significant superior performance was found in favor of the large class. This disparity persisted even when a switching regressions model (with endogenous switching) is performed. Only 39% of the learning disparity can be explained by characteristics of the student, including a 29.5% performance gap that is attributed to the potential "overachiever" trait of small class students.

In addition, it is found that overall GPA, performances in previous college economics courses do not have any significant impact on the students' current score. What is more interesting is that, contrary to most of the findings in recent literature, high school economic education does rot have any positive impact of the students' performance. In fact, it has a slightly negative effect. No evidence of gender difference is found for the large class, but a fairly significant lower performance is found for females in the small class.

Class Size and Determinants of Learning Effectiveness

By

Jack W. Hou*

A gradual increase in class size has been a national trend since the mid 1980s. The reason is obvious. Universities are facing budget cuts, while departments are trying to retain or increase enrollments to fight for operational funds and the ability to offset courses with lower enrollments. These, among others, certainly signals that oversized sections will be a fixture of academic life for at least the foreseeable future.

Economics aside, the concern we should have as educators is how this will affect the students' learning? This study is an attempt to evaluate this issue. It is surprising that direct measures of potential student learning differences between heterogeneous class sizes are conspicuously lacking, at least regarding college economics. Though this study is a comparison between two upper division economics classes, identical except for the class size, the experiment can easily be applied to any economics courses, and to other disciplines for that matter.

The objective of this study is to use a technique originally designed for discrimination analysis to decompose the measured learning difference (in terms of test scores) into a portion that is accountable by the differences in student characteristics and the residual, which can be interpreted as due to the difference in class size. For a more detailed discussion of this methodology, refer to Hou (1991, 1993ab).

---

* Associate Professor, Department of Economics, California State University, Long Beach. Partial support was provided by the Institute for Teaching and Learning, California State University System. I am grateful to Wayne Phillips for his help in providing some of the students background information. Sharon Yamada provided excellent assistant in terms of data entry. The usual disclaimer applies.

The size and direction of this residual will have significant impact on the direction of the classroom structure. If there is no difference in learnirg (or even in favor of larger classes), simple economics should dictate the gradual conversion towards larger class sizes, as the only cost is perhaps building bigger rooms while the benefit would be obvious. If larger class size does hamper the student's learning, but only minutely, one may still be able to justify the larger class size in terms of its economies of scale. However, if the negative effect on the students' learning is significant, a tough choice will have to be made in this time of budget difficulties.

The remainder of the paper is organized as followed. I will open with a brief description of the data. Section II will outline the methodology and the model. This is followed by the empirical estimates and discussion. A brief summary will conclude the paper.

I. Data Source

Ideally, we would like to measure the effect of class size not just in the contemporary disparity in learning, but also the difference in the half-life of the knowledge between students from the two classes.[1] However, clear and direct measures are hard to get, if they exist at all, especially regarding the half-life of learning.[2]

For realistic reasons, this study is based on the data collected from two classes (Managerial Economics or ECON333) in which I was the instructor: one was an "oversized" section, while the other was a normal section. This offers an excellent setting for examining the issues raised above. Since I am teaching both classes (back-to-back in the same semester), the usual problem of instructional heterogeneity is negligible. To further minimized any differences, I consciously attempted to make the material and presentation as homogeneous as possible.

A survey questionnaire (Appendix) was developed and implemented in both classes.[3] This generated detailed data regarding various relevant characteristics (both personal and family background) of the student. With the students' authorization, and the aid of the University Registrar, various information, such as current and high school GPA, SAT scores, etc., were extracted from the student files. Combined, these information formed the data base for the empirical study.

For the measurement of the students' learning, several quizzes were made identical between the two classes. In addition, one section of each midterm and of the final exam were also made identical between the two classes. These, combined with the fact that the lectures were maintained as homogeneous as possible between the two classes, allows for a meaningful comparison of the students' learning.

There are a total of six common quizzes and six common exam questions. A two-sample t test was calculated based on the following test statistics

$$t = \frac{\bar{Y}_L - \bar{Y}_S}{[\delta \times (1/n_L + 1/n_S)]}$$

where

$\bar{Y}_L$, $\bar{Y}_S$ = observed average grade of the large (L) and small (S) class

$$\delta = \frac{(n_L - 1)\sigma_L^2 + (n_S - 1)\sigma_S^2}{n_L + n_S - 2}$$

n_L, n_S = number of students in the large and small class

This statistic has a student t distribution with $n_L + n_S - 2$ degrees of freedom. For the twelve pair of comparison, the large class out-performed their small class counterpart in only two, and neither was statistically significant. In the ten pairs where the small class scored better, six of them were statistically significant at the (2-tailed) 5% level. Out of these, three were significant at the 1% level. The "observed" superior performance of the smaller class is evident. However, as will be explained in the next section, this may be quite misleading.

The common quizzes and exam questions were next standardized separately,[4] added up according to the weights indicated in the class syllabus, and the final score is standardized again.. Based on this procedure, for the combined sample, the mean or average is zero with a standard deviation of one. However, for the two classes separately, the difference is apparent. For the small/regular class, the mean is 0.40, as compared to the -0.18 for the over-sized class. The comparison for the standard deviation is 0.85 versus 1.02. Thus confirming intuition that students in the small class perform better and are more consistent. Based on the two-sample t test described above, the difference in the mean is significant at the 1% level.

A summary of several other key variables are presented in Table 1. As can be clearly seen, many characteristic or "endowment" differences exist between the two classes. In general, students in the small class are more likely to be males, Asians, those with betters grades (in terms of GPA, and performance in Principles of Microeconomics and Calculus), and transfers from either Junior College or other universities. This perhaps foresee the problem of selection bias that may hamper OLS estimates,[5] which will be discussed below.

An interesting point is observed if one contrasts the student self-reported GPA versus the GPA provided by the Registrar. The former is consistently higher than the latter. The two-sample t test between all three pairs of GPA(R) and GPA(A) indicate that the two measures are statistically different (at the 1% level) for all pairs. In other words, the student's "error" in reporting their GPA cannot be explained as "rounding error".

What is surprising is the high percentage (close to 80%) of students holding at least part-time jobs. Of these students, they work an average of 25 hours/week. This perhaps reflects both the nature of the institution and the difficult economic times we are in. Another startling statistics is the low percentage of students that started their undergraduate education at the University. Close to three-quarters are transfers either from JC (Junior College) or other four year universities. From a Departmental point of view, this certainly deals a heavy blow to the enrollment of lower division classes, and eventually affects FTE (Full-Time-Equivalent) and funding.

A word of caution. Due to the sample size, the empirical estimates and implications should be taken with a grain of salt as their robustness may be questionable. The main objective here is to initiate interest in addressing this question of the effects of class size on student learning, and to demonstrate the importance of using the correct methodology.

II. Methodology

The most intuitive, and often used measure, is the direct comparison of the observed average grade between two classes based on some uniform exam. This is crude and can be misleading as it ignores the differences in student composition (what shall be termed as characteristic "endowments") between the classes. Thus, the traditional method has been to estimate the following regression

(1) $Y_i = \beta X_i + \gamma D_i + \epsilon_i$

where Y_i = standardized test score of student i

X_i = a set of explanatory variables including the students personal/ family characteristics, GPA/SAT scores, etc.

D_i = binary class size variable (1 for small class, 0 otherwise)

e_i = random error term

The "β" coefficients reveal what determines the students learning potential, while the "γ" coefficient will indicate the effect of class size.

Though an improvement over the direct comparison of average performance, this measure of class size effect is crude/biased, and potentially erroneous. The reason is tri-folded. First, the exams/quizzes are different between classes. Second, The effects of "X_i" attributes may work differently in a small class relative to a large one. And finally, students in the small class may come from a different population that those in the large class (i.e. certain "intangibles" or unobservable attributes).

To bring the last point sharper into focus, let us assume that those more eager to learn (or, "better" and more self-motivated students) "self-select" into the small classes, while those that are passive and desire anonymity choose the large classes. Further, suppose the small class is observed to out-perform the large class. This may erroneous lead one to the conclusion that a larger class diminishes the students' learning potential. While in reality, it is possible that the seemingly superior performance of the smaller class is simply due to the fact that it is made up of more diligent students, and class size really have no significant effect at all. This will be dealt with later.

To compensate for the first two of the aforementioned problems with equation (1), separate regressions will be estimated for each class, limiting to only portions of the grades that come from identical quizzes and questions in exams. This leads to the following two regressions

(2) $Y_{Si} = \beta_S X_{Si} + \epsilon_{Si}$

(3) $Y_{Li} = \beta_L X_{Li} + \epsilon_{Li}$

where the subscript S and L stand for the small (i.e. normal or regular) class and the large or oversized class respectively.

Using a method originally designed for "discrimination" analysis in labor economics, one can then estimate the learning differential between the two classes. In a nutshell, the method estimates what the small class student's measured learning would be if he or she was in the large class. Deducting this from the observed learning difference is what I shall term as "class size" effect on student's learning.

More specifically, the "gross" learning differential is defined as $G = \overline{Y}_S - \overline{Y}_L$, where $\overline{Y}_S$ and $\overline{Y}_L$ are the average test scores of the small and the large class, respectively. By definition, these average test scores can be rewritten into

$$\overline{Y}_j = \hat{\beta}_j \overline{X}_j \qquad j = S,\ L\ (\hat{\beta}_j \text{ is the least squares estimate of } \beta_j)$$

Further define

$$\hat{Y}_{SL} = \hat{\beta}_L \overline{X}_S$$

This is the "expected" test score of the average student in the small class if he/she chose to attend the large class instead. The "gross" learning gap can be decomposed into two parts,

$$\text{(4)} \qquad G = \overline{Y}_S - \overline{Y}_L = (\overline{Y}_S - \hat{Y}_{SL}) + (\hat{Y}_{SL} - \overline{Y}_L)$$
$$= (\hat{\beta}_S - \hat{\beta}_L)\overline{X}_S + \hat{\beta}_L(\overline{X}_S - \overline{X}_L) = R + E$$

where "E" represents the learning differential which is due to, or can be "explained" by, differences in characteristics (i.e. "endowments"), while "R" is the residual difference which can be viewed as the class size effect.

However, due to the endogenous nature of the student's choice between attending the small versus the large section, the OLS estimates from the above regressions (equations (2) and (3)) will be biased.[6] To account for this endogenous class (size) selection decision, a probit model will be used to generate a selection-bias-correction variable in the Heckman-Lee tradition.[7]

To motivate the discussion, the students are assumed to be utility maximizers. The behavioral assumption is that he/she will choose the section that generates higher utility. Student i is assumed to have an indirect utility function of

$$V_i = V(Y_i,\ D_iB_i)$$

where

Y_i = learning (in terms of test scores)
D_i = binary class size dummy
B_i = net non-grade benefits associated with small classes

$V_{Si} = V(Y_{Si},\ B_i)$ and $V_{Li} = V(Y_{Li},\ 0)$ denote the student's utility level if he or she chose to be in the small and large class respectively. I shall further assume that the student's preference is characterized by a linear, additive indirect utility function of the form

$$V_i = Y_i + D_iB_i$$

Since the maintained behavioral assumption is that students will choose the class which generates higher utility, an indicator variable can be defined as

(5) $$I_i = V_{Si} - V_{Li} = Y_{Si} + B_i - Y_{Li}$$

where

I_i = comparative index ($I_i \geq 0$ implies that the small class is more attractive to the student and $I_i < 0$ implies the opposite)

Each student has a potential learning function for both the small and the large class, as shown in equations (2) and (3) above, and the net non-grade benefits can be written as

(6) $$B_i = \beta_b X_{bi} + \epsilon_{bi}$$

Substituting (2), (3), and (6) into equation (5) yields

(7) $$I_i = Y_{Si} - Y_{Li} + B_i = \beta_S X_{Si} - b_L X_{Li} + b_b X_{bi} - \eta_i = bZ_i - \eta_i$$

where α_j, X_{ji} = as defined in equations (2) and (3)

b = vector of coefficients

Z_i = all variables in X_{Si}, X_{Li} and X_{bi}

$\eta_i = -(\epsilon_{Si} - \epsilon_{Li} + \epsilon_{bi})$

If $I_i \geq 0$, the student derives higher utility from the small class, while if $I_i < 0$ then the large class is preferred. The binary class dummy is thus

$$D_i = \begin{cases} 1 \text{ if } I_i \geq 0 \\ 0 \text{ if } I_i < 0 \end{cases}$$

Assuming that η_i is a standard normal random variable, the probability that student i will choose the small class is given as

$$Pr(D_i = 1) = Pr(I_i \geq 0) = Pr(\theta_i \leq \beta Z_i) = \Phi(\beta Z_i)$$

where

$\Phi(\cdot)$ = cumulative density function (CDF) of the standard normal distribution

The probability that the student prefers the large class is

$$Pr(D_i = 0) = Pr(I_i < 0) = Pr(\theta_i > \beta Z_i) = 1 - \Phi(\beta Z_i)$$

The log-likelihood function then can be written as

$$(8) \qquad \mathcal{L} = \sum_i D_i[\ell n\Phi(\beta Z_i)] + \sum_i (1 - D_i)\ell n[1 - \Phi(\beta Z_i)]$$

The probit model will be estimated by maximum likelihood methods. Based on the results, the potential selectivity bias created by the endogeneity of the individual's sectoral choice can be corrected.

The Z variables that characterize the class selection are conceivably the same as those that determine their learning effectiveness (i.e. equations (2) and (3)). To identify equation (8) from the two learning regressions, the Z variables includes a class size "preference" variable.[8]

Based on the estimates of the probit function as shown in (8), the following variables can be defined to correct for the selectivity bias

$$(9) \qquad \lambda_{Si} = -\frac{\phi(\beta Z_i)}{\Phi(\beta Z_i)} \quad \text{and} \quad \lambda_{Li} = \frac{\phi(\beta Z_i)}{1 - \Phi(\beta Z_i)}$$

where

λ_{Si}, λ_{Li} = Heckman-Lee type selection bias correction terms (i.e. the inverse Mills ratio).

$\phi(\cdot)$, $\Phi(\cdot)$ = the PDF and CDF of the standard normal distribution.

β, Z_i = refer to equation (7) above.

Equations (2) and (3) can now be rewritten into

(10) $W_{Si} = \beta_S X_{Si} + \sigma_S \lambda_{Si} + \xi_{Si}$ iff $I_i \geq 0$

(11) $W_{Li} = \beta_L X_{Li} + \sigma_L \lambda_{Li} + \xi_{Li}$ iff $I_i < 0$

where

$\xi_{ji} = \epsilon_{ji} - \sigma_{j\eta} \lambda_{ji}$ $j = S, L$

$\sigma_j = \mathrm{Cov}(\epsilon_{ji}, \eta_i)$ $j = S, L$

OLS will generate consistent estimates of β_S, β_L, σ_S and σ_L since the error structures ξ_{Si} and ξ_{Li} have zero conditional means.[9] However, the standard errors will be biased, since they ignore the fact that the probit coefficients $\hat{\beta}$ (ML estimates of β) were estimated in the first stage. The standard errors will be correct following Lee, Maddala, and Trost (1980).

This will allow the explicit account for the student's class (size) selection choice (often termed in the literature as a "switching regressions model with endogenous switching"). The decomposition of the observed learning difference shown in equation (4) will result in the "corrected" measure of the effect of class size. This will effect will be purged of the intangibles that lead to the students selection between the class sizes.

III. Empirical Findings and Implications

As discussed in Section I, the grades for the total sample (i.e. both classes combined) was constructed such that it has a zero mean with a standard deviation of one. However, for the two classes separately, the difference is apparent. For the small/regular class, the mean is 0.40, as compared to the -

0.18 for the over-sized class. The comparison for the standard deviation is 0.85 versus 1.02. Based on the two-sample t test described above, the difference in the mean is significant at the 1% level.

This is a crude measure at best and is potentially misleading as, among other things, it ignores the differences in the composition of the two classes. If this was the only problem, a standard regression analysis (equation (1)), with a dummy variable (taking on the value of 1 for students in the small class) to capture the effect of the difference in class size, can handle it. Column (1) of Table 2 represents such a regression. As can be seen, the effect of class size (CS) is insignificant. This seems to suggest that after controlling for student characteristics, there is no statistically significant learning difference between the two classes.

This, however, may lead to an erroneous conclusion, as it assumes that the only difference in learning between the two classes is a "shift" disparity. In other words, characteristics are assumed equally "productive" in generating learning in both class settings. This is certainly not what one would commonly expect. In statistical terminology, this regression constrained $\beta_S = \beta_L$ (equations (2) and (3)) with the exception of the intercept.

To allow for "shadow" price differences, equations (2) and (3) are estimated separately and presented as columns (2) and (3) in Table 2. Clearly, the coefficient estimates are quite different between the two samples. Statistically, most of the difference is due to the effects of Load (number of units taken), Hispanic ethiniticity (Hisp.), transfer from another four year college/university (University), and number of siblings (SIBN).

As can be seen, the intercept (the "shift" variable) is not statistically different between the two samples. This clearly demonstrates the inadequacy of the simple model (shown in column (1)) where one tries to capture the class size effect with a mere binary dummy variable.

Several interesting points arise from the coefficient estimates of Columns (2) and (3). Somewhat surprisingly, the overall GPA had no significant predictive power on the student's learning. This is in sharp contrast to many studies on student performance (e.g. Raimondo, Esposito, and Gershenberg 1990). However, the student's grade in (Business) Calculus positively affected their learning, and was statistically significant for both samples. Though Principles of Microeconomics (P. Micro) showed no effect for the small class, it was an important determinant for the larger class. These are not surprising as Managerial Economics is a calculus-based and micro-oriented course.

The Load variable showed a sharp contrast between the two classes. It had no effect in the larger section, but was positive and very significant for the small class. This implies, for the students in the small class, the more units they are taking the better their performance. Since this is lacking in the large class, one may conjecture that the "overachievers" tend to self-select themselves into the small class. If this is the case, it indicates a potential biases in the OLS estimation (see Section II above).[10]

Though marginally significant at best, high-school economics (HSECON) exhibited a negative impact on the students' performance. The effect of high school economics course on the student's' performance in college has long been of interest to economists, and the evidence has been mixed. Palmer, Carliner, and Romer (1979) found that students with high-school economics background tend to do slightly (and statistically significantly) better in the "pre-test",[11] while Reid (1983) found the opposite that they achieve significantly lower grades in college level economics.

More recently, studies by Walstad and Soper (1988), Myatt, and Waddell

(1990), Becker, Greene, and Rosen (1990), Brasfield, Harrison, and McCoy (1993) all found that high school economics does have a positive effect on the students economic knowledge and performance, with Marlin (1991) as perhaps the sole exception. The negative effect of high school economics education found in this study does seem to be consistent with the earlier findings of Palmer, Carliner, and Romer.

Two (related) reasons may lie behind this. First of all, most of the aforementioned studies were regarding the effect on performance in principles courses or entry level tests on economic literacy, while the study here is on the effect of high school economics courses on the performance of students in an upper division economics course. Second, as Becker, Greene, and Rosen (1990) pointed out, the positive impact of high school economics may not have a lasting effect. The negative contribution of high school economics found here may well be due to these reasons.

Another marginally significant variable is whether the student went to a private high school (HSPRI). It also tends to negatively affect the student's performance. This is somewhat puzzling. The cost of private high schools are substantially higher than public schools. Granted, there may be prestige or safety associated with private schools, the expectation of a superior education (at least on average) is also undeniable. Then, why the observed negative effect on the student's performance? Without more rigorous testing, any statement is mere conjecture. With this in mind, one possibility could be that the upper end of the distribution of private high schools tend to enter the UC (University of California) system, while the mid to lower end of the distribution are sorted into the CSU (California State University) system.

Another interesting variable is the effect of transfers from other universities or four year colleges. It showed a negative effect (though not statistically significant) for the students in the large class, however, for the small class, the effect was positive and very significant. This can again be viewed as a sign of self-selection.

Before I turn to the selection bias stemming from this potential self-selection, the observed learning or performance differential between the two classes will be decomposed according to equation (4)

$$(4) \qquad G = \overline{Y}_S - \overline{Y}_L = (\overline{Y}_S - \hat{Y}_{SL}) + (\hat{Y}_{SL} - \overline{Y}_L) = (\hat{\beta}_S - \hat{\beta}_L)\overline{X}_S + \hat{\beta}_L(\overline{X}_S - \overline{X}_L)$$

$$= 0.3973 - (-0.1839) = (0.3973 - [-0.1270]) + (-0.1270 - [-0.1839])$$

Thus, the decomposition become

$$0.5812 = G = R + E = 0.5243 + 0.0569$$

As can be seen, the explainable portion of the difference is very small indeed (less than 10%) relative to the residual (more than 90%). This latter difference can categorically be attributed to the effect of the class size. Compared to the binary dummy approach embodied in column (1) of Table 2, the effect of class size is consistent (i.e. smaller class size lead to higher learning).

However, both the magnitude and the significance level of the learning difference is markedly higher once the difference in learning effects of the characteristics are allowed to differ between the two classes. In terms of the size of the learning difference, the dummy variable approach lead to an estimated gap of 0.2851 (or 49% of the observed difference), while the "switching" regressions model led to a much higher 0.5243 (or 90%) learning differential content due to class size. Statistically, the dummy variable estimate had a mere 1.27 t statistic, implying that the difference is not statistically different, even at the 20% significance level, from zero (as in no learning difference due to class size). While in contrast, the difference implied in the switching regressions model complied a 13.10 (two sample t test) test statistic, which is significant at the 1 % level.

This result is hardly surprising. Recall, the binary dummy approach essentially assumes that the shadow price (or coefficients) for the characteristics are the same between the two classes, with the only difference in the "shift" or intercept. Casual examination of columns (2) and (3) in Table 2, clearly refutes this. The bulk of the difference appears to be in the shadow price or slope coefficients, while the intercept is similar (statistically insignificantly different).

However, due to the possibility of the student's endogenous self-selection in terms of what type of class to enroll in, the above results may be biased. To account for this self-selection, the probit function as shown in equation (8) is estimated via ML methods. The Heckman-Lee selection bias correction term (equation (9)) were calculated and incorporated into equations (10) and (11). The estimated results are shown in Table 3.

Column (1) presents the estimates of the probit function (equation (8)). As expected (from the examination of Table 2), students with larger loads tend to select the small class (perhaps signifying "overachievers"), and transfers from other universities also tend to prefer the small class. Other variables that show significant explanatory power in terms of this selection process includes the ethnic variable of Hispanics (Hisp.) and the binary dummy for Junior College (J.C.) transfers. The former was not all surprising as the OLS estimates in Table 2 foresaw this. The latter is somewhat unexpected as the percentage of students being J.C. transfers was about the same (Table 1), and there was no statistically significant difference between their OLS estimates (Table 2), though they were of opposite signs.

The sole other variable that had marginal predictive power is the order the student ranks among his siblings, i.e. first born, second born, etc. The positive coefficient implies that the lower the order in the rank (i.e. third born, as compared to say, first born) the more likely the student will select the small class. Though one can certainly offer a number of speculation as to why we observe this, but they are just that, speculations.

The identification variable chosen is the student's class size preference. It is assumed that this affects their choice between the two classes, but have no effect on their learning effectiveness. The log likelihood of the probit estimation (column (1) of Table 3) is contrasted with that of a constrained version (where all coefficients are forced to zero with the exception of the intercept) to test for the significance of the overall model suggested by equation (8). This likelihood ratio test resulted in a rejection of the null hypothesis (that both versions had statistically similar explanatory power) at the 1% significance level.

The selection bias correction variables (equation (9)) are calculated based on the above probit estimates and entered as an additional variable (SBCV) in the two learning regressions (equations (10) and (11)). The estimated results are shown as columns (2) and (3) in Table 3. Several things stand out. First of all, the explanatory power of the regression for the small class improved significantly. This can be seen in the higher adjusted R-squared, and the increase in the significance level of coefficient estimates. The OLS regression had six regressor statistically significant at the 10% or better level. These same six regressors are still significant after correcting for the selection bias, and at a higher significance level (e.g. Calculus had a significant effect at the 5% level in the OLS regression, while the significance level is raised to the 1% level here).

In addition, two previously unimportant variables have emerged with strong predictive powers: Gender and J.C. transfers. The former is of special interest, as a sizable literature has accumulated regarding gender difference in performance, especially in high school or below. Earlier studies tend to show a superior learning of economics by males. Moyer and Paden (1968) and Highsmith (1974) found that high school males outperformed their female counterparts. Siegfried (1979) surveyed previous studies and concluded gender differences in learning and understanding of economics were absent at the elementary school level, but by high school a distinct gap appears and persists into college.

However, subsequent studies painted a somewhat different picture. MacDowell, Senn, and Soper (1978) found no evidence of gender difference among junior high school economic education. Jackstad and Grootaen (1980) and Hahn (1982) found that the gender variable had no predictive power regarding the learning of economics among high school students. Both Ferber, Birnbaum, and Green (1983) and Lumsden and Scott (1987) found evidence that males outperform females in multiple choice exams, while indifferent or inferior to females in essay exams. However, Williams, Waldauer, and Duggal (1992) found no significant gender difference in college economics performances.

Much of the above discussion mainly pertain to the regression of the small class, a similar pattern is categorically absent for the large class. This is consistent with the coefficient of the SBCV, which was much more significant for the small class than it was for the large class, indicating a significant positive self-selection[12] among students in the former but lacking in the latter class. This finding is not unexpected either, since the large classes are almost like defaults, while a student will have to "seek out" a small class. In a typical semester, the department offers ten sections of ECON333 (Managerial Economics). Out of these ten, seven will be oversized sections, while only three would be small or regular sized classes. This combined with the fact that oversized sections usually have the capacity of at least twice as much as a small class, students may indeed have to "seek out" these scarce seats in the small sections.

Based on the selection-bias-corrected estimates (columns (2) and (3) in Table 3), the observed learning difference was once more decomposed. After accounting for the potential "overachiever" (or at least higher motivated) nature of students in the small class, the following was found:

$$0.5812 = G = R + E = 0.3530 + 0.2282$$

The "class size" effect (i.e. the term "R") dropped significantly from 0.5243 (or more than 90% of the observed difference) to 0.3530 (or approximately 61%). The t statistics also decreased to 5.35 (relative to the 13.10 under the OLS estimate), but still statistically significant at the 1% level.

In other words, though the potential higher self-motivation of the students in the small class (which implies that these same students will still perform adequately even if they were in a large class setting) does explain significantly why the small section out-performed the larger section, however, a significant portion of the difference remain unexplainable and are attributed to the effect of class size.

Conceptually, if one compares the OLS estimated "class size" effect of 0.5243 with its counterpart of 0.3530 after the selection bias was corrected (i.e. accounting for the potential "overachiever" nature of the small class), the difference may be loosely interpreted as the effect of this higher motivation of the small section students. In other words, the superior motivation of the students account for 0.1713 or 29.5% of the apparent learning difference.

Summary and Conclusion

Due to budgetary problems and a host of other factors, a national trend of moving towards larger classes seem seductive and inevitable. Oversized classes has been a rather common for lower division courses for quite some time. The new tendency is to extend this practice into intermediary or even upper division class. This paper attends to use rigorous econometric methods to analyze the effect of different class size settings on the students' learning for a calculus based intermediate level economics class.

As expected, the small class students performed significantly better than their large class counterpart in a set of standardized quizzes/exams. This observed difference was then decomposed into an explainable portion and a residual. This latter portion can be interpreted as the learning disparity caused by the difference in the class setting. It was found that only 10% of the difference can be explained by observable heterogeneity between the students of the two classes, leaving 90% of the difference as due to class size. Even after accounting for the fact that those students who self-select themselves into the small class might be potential overachievers (or at least more self-motivated and disciplined), 61% of the learning disparity remain unexplained.

In conclusion, a disparity in learning is evident in favor of the smaller class. Though the size of this disparity is smaller (roughly 61%) than the observed difference, it is sizable and statistically significant (at a 1% significance level). Serious thoughts need to be devoted in the direction resource management. On the one hand, larger class sizes have its obvious economic efficiency in terms of its cost effectiveness, but on the other hand, it is our duty as educators to provide the best learning environment possible for our students.

Whether this "cost" of lower learning capacity is enough to outweigh the "benefit" of savings in operation budget is not the objective of this study, nor is it within the ability of this author. This study is merely designed to provide the best possible measure of the effect of class size on student learning, given the various constrains the data poses.

On this note, it is prudent to remind the readers once again that due to the size of the sample, the robustness of some of the findings are suspect. However, the objective of this study is to demonstrate the methodology and model than can and should be used in such studies. It is meant to motivate further research in hope of serving as a demonstration. In light of this, sufficient empirical evidence should have been presented.

In addition to expanding the size of the study to test the robustness of the findings here, further research can move in several directions. For example, it would be of great interest to compare the magnitude of the class size effect between different levels of economics courses. In other words, have a minimum of one large and one small classes for two distinct level of classes (e.g. a pair of principles, versus a set of intermediate). It might well be that the class size effect is much smaller, or even negligible for lower division principles courses. Another extension is to extend such studies into different disciplines. There is no reason to expect just because their is a significant class size effect in the economic disciple there is a comparable difference in, say, political science.

Reference

Becker, W., W. Greene, and S. Rosen (1990). "Research on High School Economic Education," _Journal of Economic Education_, 21(3):231-245.

Borg, Mary O., Paul M. Mason, and Stephen L. Shapiro (1989). "The Case of Effort Variables in Student Performance," _Journal of Economic Education_, 20(3):309-313.

Brasfield, David W., Dannie E. Harrison, and James P. McCoy (1993). "The Impact of High School Economics on the College Principles of Economics Course," _Journal of Economic Education_, 24(2):99-111.

Craig, Eleanor D., James B. O'Neill, and Douglas W. Elfner (1979). "Large-Class Retention: The Effects of Method in Macroeconomics," _Journal of Economic Education_, 10(3):12-21.

Ferber, Marianne A., Bonnie G. Birnbaum, and Carole A. Green (1983). "Gender Differences in Economic Knowledge: A Reevaluation of the Evidence," _Journal of Economic Education_, 14(2):24-37.

Griffin, Peter, Lisa Grobar, and Jack W. Hou (1991). "Issues in Teaching Over-sized Economics Principles Classes", presented at the 66th Annual Conference of the Western Economic Association, June 29 - July 3, Seattle, Washington, (Department of Economics, California State University-Long Beach).

Hahn, Carole (1982). "Are There Gender Differences in High School Economics in the Eighties?" _Journal of Economic Education_, 13(2):57-65.

Heckman, James J. (1976). "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models", _Annuals of Economic and Social Measurements_, 5:475-492.

Highsmith, Robert J. (1974). "A Study to Measure the Impact of In-Service Institutes on the Students of Teachers Who Have Participated," _Journal of_

Economic Education, 6(2):77-81.

Hou, Jack W. (1991). "Wage Comparison by Gender and the Effect of Job Segregation: The Case For Taiwan," *China Economic Review*, 2(2):195-214, 1991.

——— (1993a). "Public-Private Pay Comparison: The Case of Taiwan", *Journal of Asian Economics*, 4(2):347-362.

——— (1993b). "Wage Comparison and Job Distributional Differences Between the Public and Private Sectors," Taiwan Economic Review, 21(3):249-287.

Jackstad, Stephen L., and Christian Grootaen (1980). "Gender, Gender Stereotyping, and Socioeconomic Background of Economic Knowledge and Learning," *Journal of Economic Education*, 12(1):34-40.

Lee, Lung Fei (1978). "Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables," *International Economic Review*, 19:415-33.

———, G.S. Maddala, and Robert Trost (1980). "Asymptotic Covariance Matrices of Two-State Probit and Two-Stage Tobit Methods for Simultaneous Equations Models with Selectivity." *Econometrica*, 48(2):491-503.

Lumsden, Keith G., and Alex Scott (1987). "The Economics Student Reexamined: Male-Female Differences in Comprehension," *Journal of Economic Education*, 18(4):365-375.

MacDowell, Michael A., Peter R. Senn, and John C. Soper (1978). "Does Sex Really Matter?" *Journal of Economic Education*, 9(1):28-33.

Marlin, James (1991). "State-Mandated Economics Education, Teacher Attitudes, and Student Learning," *Journal of Economic Education*, 22(1):5-14.

Montmarquette, Claude, and Rachel Houle (1986). "An Empirical Note on Selection Bias in Educational Production Functions," *Journal of Economic Education*, 17(2):99-105.

Moyer, M.E., and Donald W. Paden (1968). "On the Efficiency of the High School Economics Course," *American Economic Review*, 58(2):870-877.

Myatt, A., and C. Waddell (1990). "An Approach to Testing the Effectiveness of the Teaching and Learning of Economics at High School," *Journal of Economic Education*, 21(3):355-363.

Palmer, John, Geoffrey Carliner, and Thomas Romer (1979). "Does High School Economics Help?" *Journal of Economic Education*, 10(2):58-61.

Peterson, Norris A. (1992). "The High School Economics Course and Its Impact on Economic Knowledge," *Journal of Economic Education*, 23(1):5-16.

Raimondo, Henry J., Louis Esposito, and Irving Gershenberg (1990). "Introductory Class Size and Student Performance in Intermediate Theory Courses," *Journal of Economic Education*, 21(4):369-381.

Reid, R. (1983). "A Note on the Enviroment as a Factor Affecting Student Performance in Principles of Economics," *Journal of Economic Education*, 14(4):18-22

Siegfried, John J. (1979). "Male-Female Differences in Economic Education: a Survey," *Journal of Economic Education*, 10(2):1-11.

Walstad, William B., and John C. Soper (1988). "A Report Card on the Economic Literacy of the U.S. High School Students," *American Economic Review*, 78(2):251-256.

Williams, Mary L., Charles Waldauer, and Vijaya G. Duggal (1992). "Gender Differences in Economic Knowledge: An Extension of the Analysis," *Journal of Economic Education*, 23(3):219-231.

Footnotes

1 Craig, O'Neill, and Elfner (1979) found that class size did not affect student retention of knowledge.

2 There are a handful of surveys (e.g. TUCE and the revised TUCE) that attempt to generate these measures but are generally plagued by many problems.

3 I benefited greatly from the study conducted by Jane Lopus and Nan Maxwell

at the California State University, Hayward.

4 The two classes are pooled together. The mean and standard deviation are calculated. Based on these, the grades are converted into a standard normal distribution.

5 Refer to Heckman (1976) and Lee (1978). For a more recent discussion in terms of "discrimination" analysis, see Hou (1991, 1993ab).

6 This selectivity bias is different from the one raised by Montmarquette and Houle (1986). Their concern was the bias rising from the "sample" selection, while the issue here is the bias stemming from the student's endogenous self-selection of class size.

7 Peterson (1992) provides a condensed and informative description of the source of the selection bias. For a full discussion, refer to Heckman (1976) or Lee (1978).

8 In the survey questionnaire, the students were asked whether they preferred large or small classes.

9 Though unbiased, these OLS estimates are less efficient due to heteroscedasticity between n_{Si} and n_{Li}. To gain full efficiency, weighted least squares should be used in place of OLS. Conventionally, the literature ignores this heteroscedasticity and I will follow suit.

10 This could, alternatively, be due to explicit differences in "effort" between the students of the two classes (Borg, Mason, and Shapiro 1989) rather than (or in addition to) unobservable traits of "overachievers".

11 But slightly (but not statistically significant) worse in the "post-test".

12 The "selectivity effects" are simply $\sigma_S\lambda_S$ and $\sigma_L\lambda_L$ (refer to equations (9), (10), and (11)). A "positive" self-selection occurs when the selectivity effect is positive. This implies that the students are choosing the classes based on their relative comparative advantage, i.e. those students that would do better in a small class setting choose to enroll in the small class.

Table 1 Summary Statistics of Selected Variables

	Total Sample Mean (S.D.)	Small Class Mean (S.D.)		Large Class Mean (S.D.)
Score	0.00 (1.00)	0.40 (0.85)	***	-0.18 (1.02)
Age	23.49 (4.24)	23.98 (4.33)		23.27 (4.22)
Gender	0.44 (0.50)	0.36 (0.49)	***	0.48 (0.50)
Race:				
White	0.51 (0.50)	0.52 (0.51)	*	0.50 (0.50)
Asian	0.32 (0.47)	0.40 (0.50)	***	0.28 (0.46)
Hispanic	0.09 (0.29)	0.08 (0.28)	*	0.09 (0.30)
Black	0.06 (0.25)	0.00 (0.00)	***	0.09 (0.30)
Other	0.03 (0.16)	0.00 ((0.00)	***	0.04 (0.19)
Married	0.06 (0.25)	0.04 (0.20)	***	0.07 (0.26)
Grades:				
GPA(R)[a]	2.86 (0.45)	2.93 (0.46)	***	2.83 (0.44)
GPA(A)	2.79 (0.44)	2.87 (0.42)	***	2.75 (0.45)
P. Micro[b]	2.72 (0.77)	2.84 (0.75)	***	2.67 (0.78)
Calculus	2.87 (0.87)	2.92 (0.94)	*	2.84 (0.85)
Transfers	0.72 (0.45)	0.92 (0.28)	***	0.63 (0.49)
Junior College	0.79 (0.41)	0.78 (0.42)		0.79 (0.41)
University	0.25 (0.43)	0.30 (0.47)	***	0.21 (0.41)
Load[c]	13.08 (3.14)	13.96 (3.49)	**	12.67 (2.90)
Work	0.78 (0.41)	0.80 (0.41)	*	0.78 (0.42)
Hours[d]	24.77 (8.56)	22.25 (8.21)		25.84 (8.50)
Private HS	0.16 (0.37)	0.12 (0.33)	***	0.19 (0.39)
# of Siblings	3.19 (1.82)	3.92 (2.47)	***	2.85 (1.32)
Rank[e]	2.51 (1.68)	3.30 (2.25)	***	2.15 (1.19)
Sample Size	79	25		54

Standard deviation in parentheses.

*, **, *** denote statistically significant difference between the two classes at the 10%, 5%, and 1% level, respectively.

Notes: a) GPA(R) and GPA(A) refer to the student's self-reported and actual (obtained from the Registrar) GPAs.
b) Principles of Microeconomics
c) # of units taken in the Fall semester of 1992.
d) # of hours of work per week (conditional mean).
e) Order of birth, 1st born, 2nd born, etc.

Table 2 Regession Estimates

	(1) Total	(2) Small	(3) Large
Intercept	-4.0580 ***	-5.4136 ***	-3.5157 ***
	(0.632)	(1.306)	(0.780)
GPA	0.5132	0.1254	0.5236
	(0.322)	(0.428)	(0.443)
Calculus	0.4522 ***	0.6602 **	0.4281 **
	(0.126)	(0.241)	(0.172)
P. Micro	0.1773	-0.0140	0.3254
	(0.163)	(0.219)	(0.219)
Load	0.0618 *	0.1767 ***	+ 0.0331
	(0.034)	(0.055)	(0.043)
Hisp.	0.4937	2.3175 **	++ 0.0134
	(0.380)	(0.815)	(0.437)
Asian	0.4720 **	0.4963	0.3678
	(0.204)	(0.400)	(0.260)
Gender	0.0542	-0.4455	0.0815
	(0.191)	(0.360)	(0.234)
HSECON	-0.2067	-0.0746	-0.4026
	(0.211)	(0.427)	(0.260)
HSPRI	-0.2777	-1.2371 *	-0.2185
	(0.211)	(0.599)	(0.298)
Uni.	0.1204	1.4103 ***	++ -0.4070
	(0.271)	(0.448)	(0.359)
J.C.	-0.0302	0.5718	-0.0916
	(0.214)	(0.390)	(0.254)
SIBN	-0.1121	0.2573	+ -0.1233
	(0.109)	(0.250)	(0.133)
Rank	0.1078	-0.2261	0.0059
	(0.122)	(0.275)	(0.156)
CS	0.2851		
	(0.216)		
Adj. R^2	0.4336	0.5061	0.4558

Standard errors in parentheses.
*, **, *** denote that the coefficient is statistically significant (2-tailed t test) at the 10%, 5%, and 1% level, respectively.

Table 3 Probit and Corrected Regession Estimates

	(1) Probit	(2) Small	(3) Large
Intercept	-1.5732	-8.0795 ***	-3.3286 ***
	(1.754)	(1.765)	(0.821)
GPA	-0.7982	0.0201	-0.6048
	(0.654)	(0.383)	(0.457)
Calculus	0.3908	0.8594 ***	0.3753 *
	(0.268)	(0.236)	(0.186)
P. Micro	-0.1292	-0.0067	0.3345
	(0.325)	(0.194)	(0.220)
Load	0.1893 **	0.2108 ***	-0.0016
	(0.080)	(0.052)	(0.063)
Hisp.	1.9543 **	3.4404 ***	-0.3909
	(0.785)	(0.915)	(0.686)
Asian	0.7434	0.4944	0.2591
	(0.450)	(0.354)	(0.297)
Gender	-0.2002	-0.7802 *	0.1203
	(0.405)	(0.360)	(0.241)
HSECON	-0.5554	-0.4551	-0.3197
	(0.427)	(0.424)	(0.283)
HSPRI	-0.7958	-1.6720 **	-0.0989
	(0.588)	(0.574)	(0.337)
University	1.6514 ***	2.4776 ***	-0.6096
	(0.614)	(0.665)	(0.447)
J.C.	1.3147 **	1.6693 **	-0.2678
	(0.499)	(0.648)	(0.343)
SIBN	-0.0882	0.3821	-0.1018
	(0.225)	(0.231)	(0.137)
Rank	0.4905 *	-0.2178	-0.0590
	(0.264)	(0.243)	(0.178)
Preference	0.5342		
	(0.376)		
SBCV		-0.3833 *	0.4216
		(0.192)	(0.549)
L. Likelihood	-34.2431		
Adj. R^2		0.6121	0.4502

Standard errors in parentheses.
*, **, *** denote that the coefficient is statistically significant (2-tailed t test) at the 10%, 5%, and 1% level, respectively.

Appendix: Student Questionnaire

You are being asked to participate in an important pilot experiment to evaluate the determinants of teaching and learning, aimed at developing methods to improve the teaching of Managerial Economics. All your responses will be kept strictly confidential. The actual use of this survey will not be until the start of the next semester (after the course grades are given) at the earliest. In order to obtain some crucial information, we also need your authorization to access your student records at the University Registration. This will also, of course, be implemented next semester. Once the records are matched, all traces of your personal identity will be destroyed. Your complete anonymity is assured. Based on the methodology, it should be clear that your grades in this course will not be affected in any fashion. Your cooperation is greatly appreciated.

Yes, I authorize permission to access my student records for analytical purposes only.

Signature: ______________________

1. Name (print): ______________________ (Last) ______________________ (First) ______________________ Middle
2. Student I.D. # ______________
3. Gender (circle one): Male Female
4. Ethnic Origin (circle 1): White, Black, Hispanic, Asian, Other (specify: ____________)
5. Date of Birth __________.
6. Were you born in the U.S.? Yes No If no, how old were you when you (and/or your family) moved to the U.S.? ___/19___. Is English your native language? Yes No
7. Are you married? Yes No If yes, number of children younger than 6: ___
8. Have you taken a high school economics course? Yes No.
   If yes, how long was it? 1 year, 1 semester, shorter.
   Was it mainly Macro, Micro, Both?
9. What type of high school did you attend: Public, Private.
   If private, was is Catholic/Christian? Yes No.
10. Between graduation from high school and now, how long have you been out of school? __________
11. Have you taken the Scholastic Aptitude Test (SAT)? Yes ____ No ____
    If yes, what was your score? _____

12. Have you taken the American College Test (ACT)? Yes ____ No ____
    If yes, what was your score? ______

13. Are you a (circle one): Sophomore, Junior, Senior,
    Other (specify: ______________)

14. Are you a transfer? Yes, No. If yes, from a Junior College, another 4-year college/university (______________________________)

15. What is your declared major (Accounting, Economics, etc.): ______________.
    If undeclared, what is your planned major: ______________

16. What is your current Grade Point Average (GPA): _____

17. When did you take ECON201: __________ of 19____ (with a grade of ____)
    When did you take MATH115b: __________ of 19____ (with a grade of ____)

18. Have you taken ECON333 before? Yes, No. If yes, when: ________ of 19____

19. How many economics courses (other than 201, 202, 333) have you taken? ____

20. How many calculus courses have you taken: ____

21. How many units of college course work are you taking this semester? ____

22. Are you working at a job(s) in addition to taking college courses?
    Yes, No. If yes, how many hours a week do you work? ____. Is your work related to your major? Yes, No.

23. Everything else equal, would you prefer a large or small class? Large, Small, Indifferent. Why: ______________

24. What is your main consideration in choosing which ECON333 section to take?
    Day (___), Time(___), Class Size(___), Instructor (___)

25. How often do you read the Wall Street Journal, or the financial pages of major newspapers? Daily, Frequently, Occasionally, Seldom/Never.

    How often do you read Forbes/Fortune/Businessweek (etc.), or other trade journals? Frequently, Occasionally, Seldom/Never.

    How often do you read Times/Newsweek, etc.? Frequently, Occasionally, Seldom/Never.

26. Typically where do you sit in the classroom? Front Middle Back

27. Family background:

    Father's education ________________, occupation ________________

    Mother's education ________________, occupation ________________

    Number of children in your family _____, you are # ____ (1 for oldest, 2 for 2nd oldest, etc.)