The Journal of Economic Education

ISSN: 0022-0485 (Print) 2152-4068 (Online) Journal homepage: http://www.tandfonline.com/loi/vece20

The Effects of Attendance on Academic Performance: Panel Data Evidence for Introductory Microeconomics

Luca Stanca

To cite this article: Luca Stanca (2006) The Effects of Attendance on Academic Performance: Panel Data Evidence for Introductory Microeconomics, The Journal of Economic Education, 37:3, 251-266, DOI: 10.3200/JECE.37.3.251-266

To link to this article: https://doi.org/10.3200/JECE.37.3.251-266

Research in Economic Education

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

The Effects of Attendance on Academic Performance: Panel Data Evidence for Introductory Microeconomics

Luca Stanca

Abstract: The author presents new evidence on the effects of attendance on academic performance. He used a large panel data set for introductory microeconomics students to explicitly take into account the effect of unobservable factors correlated with attendance, such as ability, effort, and motivation. He found that neither proxy variables nor instrumental variables provide a solution to the omitted variable bias. Panel estimators indicate that attendance has a smaller but significant impact on performance. Lecture and classes have a similar effect on performance individually, although their impact cannot be identified separately. Overall, the results indicate that, after controlling for unobservable student characteristics, attendance has a statistically significant and quantitatively relevant effect on student learning.

Key words: attendance, performance, teaching, undergraduate
JEL codes: A22, I21

It is commonly assumed that university students benefit from attending lectures. This assumption, however, needs to be tested because developments in information technology call for a reassessment of the traditional approach to

Luca Stanca *is an associate professor of economics at the University of Milan Bicocca, (e-mail: luca.stanca@unimib.it).* *The author is grateful for the comments of Peter Kennedy, Piergiovanna Natale, and two anonymous referees.*

university education, which mostly uses physical attendance of lectures and classes. A number of alternative educational models, using distance learning, are being introduced. Nevertheless, as pointed out by Romer (1993), until recently there was relatively little evidence about attendance and its effects on student learning.[1]

In the past decade, a number of researchers have examined the relationship between students' attendance (or absenteeism) and academic performance, generally finding that attendance matters for academic achievement (Durden and Ellis 1995; Devadoss and Foltz 1996; Marburger 2001; Rodgers 2001; Dolton, Marcenaro and Navarro 2003; Kirby and McElroy 2003). This kind of evidence has led some authors to call for measures to increase student attendance and even to consider the possibility of making attendance mandatory in some undergraduate courses.[2]

The main problem in assessing the effects of attendance on academic performance is that attendance levels are not exogenous, given that students *choose* whether to attend lectures and classes and this choice is affected by unobservable individual characteristics, such as ability, effort, and motivation, that are also likely to determine performance. Better students, who are more able, work harder, or are more motivated, tend to have higher attendance levels, other things being equal. This implies that estimates of the impact of attendance on academic performance are likely to be subject to omitted variable bias.

Most authors of existing studies either brush aside this problem or attempt to disentangle the impact of attendance on performance from unobservable ability and motivational factors by including in the set of regressors proxies of capability (students' grade-point-averages (GPA), scores on college entry exams, etc.), effort (homework assignment completion), and motivation (students' self-reported interest in the course). However, such indicators are generally an imperfect measure of ability and motivation. As a consequence, ordinary least squares (OLS) estimates of the returns to attendance obtained from specifications that include appropriate control variables are still likely to be biased and inconsistent, to the extent that they incorrectly attribute to attendance the effect of the components of ability, effort, and motivation not captured by the controls.

One possible solution would be to find appropriate instruments for attendance. However, it is generally quite difficult to find variables correlated with attendance but uncorrelated with unobservable ability, effort, and motivation. An alternative route, that I followed in this study, was to exploit the variability of attendance and performance in the time dimension, if a panel data set is available. This allowed me to take into account time-invariant unobservable factors that affected both attendance and performance, and therefore I could deal with the omitted variable bias that characterizes estimates of the effect of attendance on performance based on cross-sectional data.

For the analysis presented in this article, I collected observations on the performance of 766 introductory microeconomics students on several tests and their attendance levels at lectures and classes covering the material examined on those tests. I also had information on proxies for ability (high school grade, GPA, exam speed, and proficiency in calculus), effort (number of study hours), motivation (subject and teacher evaluation), candidate instruments for attendance, and a number of other individual characteristics. It was therefore possible to compare the results obtained with three approaches: OLS controlling for unobservable factors with proxy variables, instrumental variables for attendance, 2SLS (two-stage least squares), and panel estimators (random effects and fixed effects).

LITERATURE REVIEW

In a widely cited study, Romer (1993) reported evidence on absenteeism in undergraduate economics courses at three major U.S. universities, citing an average attendance rate of about 67 percent. He also presented regression results, on the basis of a sample of 195 intermediate macroeconomics students, indicating a positive and significant relationship between student attendance and exam performance. He found this result to be qualitatively robust for inclusion among the explanatory variables of students' GPA and the fraction of problem sets completed.[3] On the basis of these findings, Romer suggested that measures aimed at increasing attendance, including making attendance mandatory, could be considered.[4]

Prior to Romer, Schmidt (1983) investigated student time allocation in a sample of 216 macroeconomic principles students and found that time spent in lectures and discussion sections had a positive and significant effect on exam performance, even after controlling for hours of study. Park and Kerr (1990) found an inverse relationship between students' attendance and their course grades in a money and banking course over a 4-year period, even after controlling for the effect of unobservable motivation by means of students' self-reported hours of study and their perceived value of the course.

Following the controversial conclusions of Romer (1993), in the past decade a number of empirical studies in the economic education literature have examined the relationship between student attendance and academic performance. Durden and Ellis (1995) investigated the link between overall course grade and self-reported attendance levels in a sample of 346 principles of economics students over three semesters. Their results, on the basis of OLS controlling for ability and motivational factors (GPA, college-entrance exam scores, having had a course in calculus), indicated that attendance mattered for academic performance. In particular, whereas low levels of absenteeism had little effect on the eventual outcome, excessive absenteeism had a large and significant effect.

Devadoss and Foltz (1996) examined attendance in a sample of about 400 agricultural economics students at four large U.S. universities. They found that, even after controlling for prior GPA and the degree of motivation, on average, students who attended all classes achieved a full letter grade higher than students who attended no more than 50 percent of the same classes. A positive and significant relationship between attendance and academic performance was also found by Chan, Shum, and Wright (1997) in a sample of 71 principles of finance students.

More recently, Marburger (2001) investigated the relationship between absenteeism and exam performance in a sample of 60 students of a principles of microeconomics course. He matched information on student attendance at each class during the semester with records of the class meeting when the material corresponding to each question was covered. The results indicate that students who missed class on a given date were significantly more likely to respond incorrectly to questions relating to material covered that day than were students who were present. Rodgers (2001) found that attendance had a small but statistically significant effect on performance in a sample of 167 introductory statistics students. Kirby and McElroy (2003) studied the determinants of levels of attendance at lectures and classes and the relationship with exam performance in a sample of 368 first-year economics students and found that hours worked and travel time were the main determinants of class attendance and that the latter, in turn, had a positive and diminishing marginal effect on grade.

Among studies that have reached less-robust conclusions about the positive effect of attendance on performance, Bratti and Staffolani (2002) argued that estimates of student performance regressions that omit study hours might be biased, given that hours of study are a significant determinant of lecture attendance. Using a sample of 371 first-year economics students, they found that the positive and significant effect of lecture attendance on performance was not robust to the inclusion of the number of hours of study. Dolton, Marcenaro, and Navarro (2003), applying stochastic frontier techniques to a large sample of Spanish students, found that both formal study and self-study were significant determinants of exam scores but that the former might be up to four times more important than the latter. However, they also found that self-study time might be insignificant if corrected for ability bias.

All of these studies, with the exception of Marburger (2001) and Rodgers (2001), were based on cross-sectional data sets. As a consequence, as observed by Romer (1993), the possibility that the estimated relationship between attendance and exam performance reflected the impact of omitted factors rather than a true effect cannot be ruled out. In this article, I report results obtained using panel data on introductory microeconomics students to estimate the *net* effect of attendance on academic performance.

DATA

I collected the data for this study by conducting a survey among students attending the introductory microeconomics course at the University of Milan between 2001 and 2004. This course, taught over 12 weeks in the spring semester, is taken by all students in the first year of study for the economics degree. The course is also repeated by students in their second (or higher) year who have not passed the exam in their first year (22 percent of the sample). There were seven parallel sections of this course offered to about 1,500 first year students. I analyzed about 200 students in one section, surveyed in each academic year between 2001 and 2004. The resulting sample was 766 individuals.

In each academic year the sample students' exam was based on four mid-term tests, which were administered every three weeks, covering equal fractions of the course and carrying the same weight for the final grade. I distributed questionnaires to the students with each of the four test papers, and they were completed before the students started the tests. This produced four independently pooled panels (one for every year), each with a cross-section of about 200 students observed over four tests, resulting in a potential balanced panel of 3,064 observations ($N = 766$ times $T = 4$). The number of observations for the actual (unbalanced) panel was 2,913 ($\bar{T} = 3.8$) because of incomplete questionnaires and students dropping out before the end of the course.[5]

Table 1 shows summary statistics for the main variables. Academic performance was measured by the students' test scores (SCO). The actual test score ranged potentially from -36 to $+36$, as it resulted from summing the outcome of 24 independent true/false questions, with 1.5 marks for correct responses, -1.5 marks for wrong responses, and 0 for no response. Note that the expected. score from random guessing implied that differences in attitudes toward risk did not affect exam performance. This ruled out the possibility of a spurious link between attendance and performance reflecting students' risk behaviors. Test scores were rescaled to the range of $-100, +100$ to make them more easily interpretable and comparable with the results reported in the literature. In the sample, the average rescaled test score was 58.7, with a range between -41.7 and 100 (Table 1).[6]

TABLE 1. Descriptive Statistics

Variable	Mean	SD	Min.	Max.
Test score (%)	58.71	24.45	–41.67	100.00
Lectures attended (%)	70.82	27.78	0.00	100.00
Classes attended (%)	67.36	35.42	0.00	100.00
High school grade (%)	77.24	12.34	60.00	100.00
Grade point average (%)	76.86	8.72	60.00	100.00
Exams per annum	2.05	1.20	0.00	6.00
Hours of study (per week)	10.85	5.69	1.00	35.00
Subject evaluation	73.73	11.36	10.00	100.00
Teacher evaluation	80.89	12.20	10.00	100.00
Travel time (min.)	46.91	26.88	1.00	150.00
Age	20.43	0.83	19.00	27.00
Years since registration	1.31	0.62	1.00	4.00
High school type	3.06	2.12	1.00	10.00
Father education	2.76	0.91	1.00	6.00
Mother education	2.63	0.89	1.00	6.00
Father occupation	4.03	2.57	1.00	11.00
Mother occupation	4.31	2.37	1.00	11.00
Province	1.53	1.23	1.00	7.00
Female	0.47	0.50	0.00	1.00
Foreign language	0.06	0.23	0.00	1.00
Calculus	0.36	0.48	0.00	1.00
Work	0.43	0.49	0.00	1.00
Web at home	0.74	0.44	0.00	1.00
Live away from home	0.10	0.30	0.00	1.00

Notes: Number of observations: 3,064 ($N = 766$, $T = 4$).

The figures on lecture (LEC) and class (CLA) attendance are estimated percentages, as reported by the students, out of a variable number of lectures and classes for each of the three-week course units (generally eight lectures and three classes for each course unit). Lectures covered theory, and classes focused on exercises and questions similar to those on which the tests were based. All lectures and classes lasted two hours and were taught by the same lecturer and class teacher, respectively. On average, the students in the sample attended about two-thirds of the classes (67.4 percent) and a slightly higher percentage of the lectures (70.8 percent).[7]

Ability was proxied by four main indicators on the basis of the student's past performance in both high school and university: *high school grade* (HSG), ranging between 60 and 100, was the leaving certificate score; *grade point average* (GPA) was the average mark on exams passed before taking introductory microeconomics (the actual GPA, defined on a 18–30 scale, was rescaled to a 60–100 scale to ease comparability); *exams per annum* (EPA) was a measure of speed in completing course work, defined as the number of exams passed divided by the years of registration;[8] *calculus* was a dummy variable equal to 1 if the student had passed the first-year (first semester) calculus exam. The means of both GPA and EPA appeared to be quite low (76.9 and 2.1, respectively). Effort was measured by the average number of *hours of study* (SSH) per week (in addition to attendance) for introductory microeconomics, ranging between 1 and 35 around a mean value of 10.9. Motivation was measured by two indicators: *subject evaluation* (SEV) and *teacher evaluation* (TEV), self-reported assessments defined on a 0 to 100 scale (average values of 74.7 percent and 80.9 percent, respectively).

Additional quantitative variables included *travel time* to reach the university (in minutes), *age*, and *years since registration* (number of years since student started university). A number of dummy variables provided information on student characteristics, such as gender (1 = female), *foreign language* (1 = non-native speaker), *work* (1 = worked while taking the course), *Web* (1 = internet available at home), and *live away from home*. Further information on the background of students was provided by categorical variables referring to *high school type*, *education,* and *occupation* for both father and mother and *province of residence*.[9]

METHOD

I was interested in estimating the parameters characterizing the relationship between teaching and learning. I assumed that learning is the output of an educational production function that reflects the match between two types of factors: academic input and student input.[10] Academic input broadly refers to teaching (lectures, classes, seminars, tutorials, office hours, etc.). I assumed student input to reflect a number of individual factors, among which the three main ones were ability, effort, and motivation. Assuming linearity, the relationship can be described as

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i, \tag{1}$$

where y_i is learning for individual i, with $i = 1,\ldots,N;$ x_{1i} is academic input; x_{2i} is student input; and ε_i is an error term reflecting all other factors that affect learning.

Learning was measured by academic performance (test score) and teaching by lecture and class attendance. It was more difficult to find an appropriate measure for student input, given that factors such as ability, effort, and motivation are not directly observable. This would not be a problem for the estimation of β_1 if student input and attendance were uncorrelated. However, ability, effort, and motivation are all likely to be positively correlated with attendance: Students who are more able, work harder, or are more motivated tend to have higher attendance levels. As a consequence, the omission of x_2 from equation (1) would make the OLS estimator of β_1 biased and inconsistent, as the latter would attribute to attendance an effect that is actually a result of unobservable student characteristics. In short, this is a classic example of omitted variable bias.

One possible solution is to find appropriate *proxy variables* for student input. This implies estimating

$$y_i = \beta_1 x_{1i} + \beta_2 x^*_{2i} + \varepsilon_i, \tag{2}$$

and assuming that $x_{2i} = \gamma_0 + \gamma_1 x^*_{2i} + v_i$ describes the relationship between the unobservable factors and the proxy variables, where v_i is an error term reflecting all other factors that affect student input. To obtain a consistent estimator for β_i, x_{1i} and v_i must be uncorrelated: The proxy variables must capture all of the correlation between the unobserved factors (student input) and the regressor of interest (attendance). I used *high school grade*, *GPA*, *exams per annum,* and *calculus* as proxies for ability, *hours of study* as a proxy for effort, and *subject and teacher evaluation* as proxies for motivation.[11]

If no proxy variables are available, or the ones available are not suitable because they do not capture all the correlation between the regressor of interest and the omitted factors, an alternative solution is to find appropriate *instrumental variables* (IV) for attendance. The instruments net out the correlation of student input with attendance, so that $\hat{\beta}_1^{IV}$ would measure the net effect of attendance on academic performance. However, the consistency of the IV estimator relies on the assumption of instrument validity, which is often difficult to maintain in practice. In addition, even if the assumption of instrument validity is satisfied, the instruments can be weakly related to the endogenous variables, resulting in imprecise estimates. In the following discussion, I consider estimates of equation (2), with and without the inclusion of the proxy variables, obtained by 2SLS, using *travel time*, *work,* and *Web* as instruments for attendance. The choice of the instruments was based on the assumption that longer travel time, being a working student, and having internet at home should be negatively related to attendance but not have a direct impact on performance. In particular, in the case of work, this assumption can be maintained, given that (in the complete specification) I am controlling for the number of hours of study.

An alternative possibility is to exploit the time dimension of the data set, assuming that the omitted variables do not change over time, to eliminate the effect of unobservable factors, using a *panel estimator* in the following specification:

$$y_{it} = \beta_1 x_{1it} + \beta_2 x_{2i} + a_i + \eta_{it}, \tag{3}$$

where x_{1it} and x_{2i} are time varying and time invariant regressors, respectively; η_{it}~$(0, \sigma^2_\eta)$ is the idiosyncratic error component, uncorrelated with $(x_{1it}, x_{2i}, \alpha_i)$; and α_i denotes unobservable fixed effects potentially correlated with the regressors.

The fixed-effect (FE) estimator is based on the assumption that α_i represents fixed constants and is obtained as OLS on the data transformed in deviations from individual means (the time invariant terms drop out).

$$y_{it} - \bar{y}_i = \beta_1(x_{1it} - \bar{x}_{1i}) + (\eta_{it} - \bar{\eta}_i). \tag{4}$$

This estimator is consistent even in the presence of unobservable effects correlated with the regressors, provided η_{it} and x_{1it} are uncorrelated at all leads and lags. However, the FE estimator wipes out time invariant regressors and is not efficient. I therefore also considered the random-effects estimator (RE), on the basis of the assumption that α_i~$(0, \sigma^2_a)$, obtained as OLS on the data transformed in quasi-deviations from individual means.

$$y_{it} - \theta\bar{y}_i = \beta_1(x_{1it} - \theta\bar{x}_{1i}) + \beta_2(x_{2i} - \theta\bar{x}_{2i}) + (\alpha_i - \theta\alpha_i) + (\eta_{it} - \theta\bar{\eta}_i), \tag{5}$$

where $\theta = 1-(\sigma^2_\eta/\sigma^2_\eta + T\sigma^2_a)^{1/2}$ is a measure of the weight of the between component in the total variability of the error term. This estimator is inconsistent in the presence of unobservable effects correlated with the regressors. It is, however, efficient (it is the generalized least squares [GLS] estimator), and it allows estimation of the parameters of time-invariant regressors.[12]

All the specifications estimated below include time-fixed effects, allowing for heterogeneity among the four cross-sections (one for each academic year between 2001 and 2004) and among the different tests (four for each cross-section). The time-fixed effects λ_t are modeled by means of 16 year test-specific dummies:

$$\lambda_t = \sum_{j=1}^{4}\sum_{k=1}^{4}\delta_{jk}D_{jk},$$

where D_{jk} is a time dummy for year j and test k, and δ_{jk} is the corresponding parameter (in practice, the dummy for test 4 in 2004 was omitted). I also controlled for individual characteristics such as years since registration, gender, foreign language, and live away from home, and I included sets of dummy variables for high school type, parental education and occupation, and province of origin.

RESULTS

OLS estimates of alternative specifications of the relationship between academic performance and attendance are provided in Table 2. All specifications produced a coefficient estimate for attendance that was positive and statistically significant at the one percent significance level. In the basic univariate specification (column 1), the point estimate indicated that one additional percentage point

TABLE 2. Determinants of Academic Performance: Ordinary Least Squares (OLS) Estimates

Variable	(1)	(2)	(3)	(4)	(5)
Lectures attended					
OLS	0.090	0.090	0.076	0.087	0.073
t	5.757	5.729	4.858	5.547	4.687
High school grade					
OLS			0.182		0.197
t			4.799		5.146
Grade point average					
OLS			0.231		0.223
t			4.240		4.100
Exams per annum					
OLS			1.217		1.185
t			2.973		2.875
Calculus					
OLS			3.838		3.421
t			3.804		3.391
Hours of study					
OLS				0.169	0.139
t				2.298	1.913
Subject evaluation					
OLS				0.084	0.066
t				2.371	1.887
Teacher evaluation					
OLS				0.144	0.155
t				3.817	4.149
Year of registration					
OLS		−3.243	−2.671	−3.242	−2.618
t		−4.735	−3.849	−4.790	−3.824
Female					
OLS		0.641	−0.869	0.161	−1.334
t		0.773	−1.043	0.194	−1.590
Foreign language					
OLS		−3.424	−4.144	−4.235	−5.026
t		−1.510	−1.826	−1.886	−2.216
Away from home					
OLS		−4.085	−3.718	−4.795	−4.450
t		−2.499	−2.237	−2.925	−2.667
Adjusted R^2	0.253	0.276	0.302	0.283	0.308

Notes: Dependent variable: test score. Number of observations: 2913. All specifications include time-fixed effects. Models (2) to (5) also include dummy variables for high school type, parental education and occupation, and province of residence.

of lecture attendance corresponded to a 0.09 percent improvement in performance. As reported in column 2, the addition of a set of controls for individual characteristics did not affect the estimated coefficient for attendance. In this specification, *years since registration* and *live away from home* were negatively and significantly associated with performance.

Next, I considered how controlling for unobservable factors, such as ability, effort, and motivation, affected the estimated coefficient for attendance. Adding either the set of ability proxies (column 3) or the set of effort and motivation indicators (column 4), the estimated coefficient for lecture attendance fell to 0.076 and 0.087, respectively.[13] Adding both sets of indicators (column 5) left the coefficient virtually unchanged (0.073) relative to the specification that included only ability controls. These results, consistent with the findings in Romer (1993), suggest that ability is positively related to both attendance and performance, so that in estimating the effect of attendance on performance it is crucial to take into account the effect of unobserved ability. Controlling for effort and motivation, instead, does not seem to have a major impact on the estimated coefficient for attendance.[14]

Focusing on the complete specification (column 5), all the ability indicators had a positive and significant coefficient. One additional percentage point of HSG or GPA corresponded to 0.20 and 0.22 percent improvements in test score, respectively. The point estimates for exam speed and calculus were also quantitatively large: one additional exam per annum was associated with a 1.19 percent higher test score, and students who had passed calculus had a test score 3.42 percentage points higher than the others had.[15] The indicators of motivation and effort also had the expected sign: One additional hour of study per week produced a 0.14 percentage point increase in performance, although the coefficient was only marginally significant. Subject evaluation had a positive coefficient (0.07), significant at the 10 percent level, whereas teacher evaluation had a significant and larger coefficient: One additional percentage point in teacher evaluation corresponded to a 0.16 percentage point increase in test score.

Looking at the other controls in column 5, an additional year since registration had a significant negative impact on test score of 2.6 percentage points. This result can be interpreted as an indication that slow completion of exams reflects lower ability or effort. Speaking a foreign language and living away from home both had very large negative and statistically significant effects on test score (–5.03 and –4.45, respectively). The coefficient on the female dummy, on the other hand, was negative but not statistically significant, indicating that gender did not have a significant effect on performance, consistent with the results in Williams, Waldauer, and Duggal (1992) and Durden and Ellis (1995).[16]

The results in Table 2 indicate that controlling for ability, effort, and motivation by means of proxy variables lowered the estimated coefficient for attendance from 0.090 to 0.073. This result could be interpreted positively, as in Romer (1993, p. 173), as a sign that "an important part of the relationship reflects a genuine effect of attendance." An alternative, more plausible interpretation, is that, despite the introduction of a set of control variables, the relationship still reflects the effect of omitted factors correlated with regressors: To the extent that, despite the control factors, there are still unobservable fixed effects correlated with attendance, $\hat{\beta}^{OLS}$ remains biased and inconsistent (likely to be overestimated). I thus considered the results obtained by IVs (2SLS), using *travel time*, *work,* and *Web* as instruments for attendance. The results indicate that the estimated coefficient for attendance was very sensitive to the set of controls included in the specification,

TABLE 3. Attendance and Academic Performance: Instrumental Variables (IV) Estimates

Variable	(1)	(2)	(3)	(4)	(5)
Lectures attended					
IV	0.148	0.125	0.038	0.148	0.065
t	2.095	1.719	0.515	2.049	0.873
Sargan test (χ^2)					
IV	18.109	12.116	7.956	111.131	5.838
t	0.000	0.002	0.019	0.000	0.054

Notes: Dependent variable: test score. Number of observations: 2,913. All specifications include time-fixed effects and the same set of regressors as in the corresponding columns of Table 2. Instruments for *lectures attended: travel time, work, web*.

TABLE 4. Attendance and Academic Performance: Panel Estimates

Variable	Fixed effects	Random effects
Lectures attended		
Panel	0.039	0.070
t	2.158	4.446
Adjusted R^2	0.292	0.322

Notes: Dependent variable: test score. Number of observations: 2,913. Both specifications include time-fixed effects and the same set of additional regressors as in Table 2, column 5. Hausman test: 29.28 (p = .01). Breusch-Pagan LM test: 85.9 (p = .00).

suggesting that the instruments could be invalid (Table 3). This was confirmed by the Sargan tests of overidentifying restrictions, that rejected the null hypothesis of instrument validity for all models (except, marginally, for the full specification in column 5).

Given that IV estimation does not provide a solution to the omitted variable bias, I then turned to estimates obtained by exploiting the panel structure of the data set. Estimates of the FE and RE models are presented in Table 4.[17] The RE estimates were similar to the OLS estimates, indicating that the weight of the between component in the error term was small relative to that of the within component. The RE model indicated that attending an extra 1 percent of lectures increased test score by 0.07 points. The FE model produced an estimated coefficient for attendance that was positive and statistically significant at the 5 percent level. The point estimate (0.039) was about half the size of the OLS and RE estimates, suggesting that there was indeed positive correlation between unobserved effects and time varying regressors, even after controlling for ability, effort, motivation, and other individual characteristics. This was confirmed by the Hausman test statistic (29.28), that strongly rejected the null hypothesis of unobservable characteristics uncorrelated with attendance (p = .01). This result was quite important, as it indicated that the inclusion of proxy variables was not sufficient to capture all the correlation between the regressor of interest and unobservable ability, effort, and motivation.

Beside statistical significance, is the estimated effect of attendance on performance quantitatively relevant? Given that each two-hour lecture was equivalent to 12.5 percent of total attendance, the 0.04 estimate in the FE model implied that missing one lecture was associated with about a half percentage point drop in test score.[18] This also implies that an average student who had not missed any lectures obtained a test score 1.2 percentage points higher than a student who had the average attendance level (70.8 percent).

To provide a complete description of the relationship between attendance and performance, I turned to the analysis of class attendance. In particular, I first examined whether class attendance had an impact on performance comparable to that of lectures and then whether the respective roles of lectures and classes could be identified separately. Table 5 contains estimates obtained by replacing lecture attendance with class attendance (columns 1–3) and by including classes and lectures jointly (columns 4–6), comparing in both cases the results for the OLS, RE, and FE models (all results refer to the complete specification that included the full set of controls).

The coefficient for class attendance was positive and statistically significant in all models (Table 5, columns 1–3). The point estimate was about 0.05 for both the OLS and RE estimators and only slightly lower (0.037) for the FE estimator, remarkably close to the 0.039 FE estimated for lecture attendance reported in Table 4. It is interesting that, in this case, the Hausman test did not reject the RE model against the FE model. This result indicated that, contrary to lecture attendance, class attendance was not significantly correlated with unobservable factors. One possible explanation for this result is that the decision to attend classes is less related to ability, given that it is commonly believed by students that class

TABLE 5. Lectures, Classes, and Academic Performance: Panel Estimates

Variable	OLS (1)	FE (1)	RE (1)	OLS (2)	FE (2)	RE (2)
Lectures						
Panel				0.053	0.026	0.051
t				2.763	1.198	2.777
Classes						
Panel	0.052	0.037	0.050	0.029	0.029	0.030
t	4.522	2.828	4.248	2.054	1.804	2.148
Adjusted R^2	0.308	0.295	0.321	0.310	0.039	0.324

Notes: Dependent variable: test score. Number of observations: 2,896. All specifications include time-fixed effects and a set of additional regressors as in Table 2, column 5. Hausman test (model 1): 2.16, p = .99, (model 2) 30.23, p = .01. Breusch-Pagan LM test: (model 1) 89.23 (p = .00), (model 2) 86.84 (p = .00). FE = Fixed effects; RE = Random effects attendance has a higher return for exam performance than lecture attendance. Overall, the results suggested that the effect of class attendance on performance was significant and quantitatively similar to that of lecture attendance.

Next, I considered the estimates obtained by inserting lectures and classes *jointly* in the full specification, to assess whether the respective roles of lectures and classes could be identified independently. As in the previous case, the OLS and RE estimates were quite similar (about 0.05 and 0.03 for lecture and class attendance, respectively). These results would seem to indicate that lecture and class attendance had independent effects on performance and that lectures had a larger impact than did classes.[19] However, the Hausman test rejects the consistency of the RE model. The FEs model provided estimates of about 0.03 for both lectures and classes, but the coefficients are no longer statistically significant. This result indicated that, after controlling for omitted variable bias, it is not possible to identify separately the effects of lecture and class attendance.

DISCUSSION AND CONCLUSIONS

The results of the empirical analysis suggest two main conclusions. First, neither proxy variables nor IVs provide a viable solution to the omitted variable bias in estimating the effect of attendance on academic performance. A proposed alternative solution is to exploit the panel structure of the data set to explicitly take into account the effect of unobservable factors correlated with attendance, such as student ability, effort, and motivation. Second, after controlling for unobservable factors, attendance at either lectures or classes was found to have a smaller but significant impact on test scores in an introductory microeconomics course. On the basis of this evidence, can one conclude that teaching has a positive effect on student learning?

One possible objection could be that test scores are not a good measure of learning: Attendance could affect exam performance because students learn how to do well on the exam, without any actual effect on the quality of learning (Deere 1994). This can be true if, for example, lectures only improve exam-taking skills or provide information on the topics and types of questions that will be in the exam or, more generally, lectures present examinable material that is not covered in the textbook.[20] This critique, however, does not apply to the data set I investigated: All students had access to detailed lecture notes and past exam papers on the course Web site, so that attendance did not reveal any private information. In addition, lectures and classes followed very closely the textbooks, so that all exam questions could be answered correctly by students not attending lectures or classes, who instead relied exclusively on the texts to prepare for the exams. It should also be observed that the marking scheme was fully objective, so that test scores could not be used to reward students for attendance.

A second possible argument is that, although the coefficient for attendance was significant, it was quantitatively small. My results indicate that the estimated effect of attendance can be considered quantitatively relevant: Missing one lecture was associated with about a half percentage point drop in test score. The opportunity cost of missing lectures is relevant not only in absolute terms but also in relative terms: The return to each hour of self-study is substantially lower than that to each hour spent attending lectures or classes. In assessing the size of the estimated coefficient for attendance, one should also consider that measurement error, because of the self-reported nature of attendance, is likely to produce a downward bias in the estimate of its effect on performance. In addition, to the extent that regressors, such as GPA and exam speed, reflect the effect of attendance in other courses (and that attendance is positively correlated across courses), the inclusion of these regressors could lead to underestimating the effect of attendance on performance in introductory microeconomics.

Can we conclude then that we, as academics, are doing something useful for student learning? According to the results of this study, the answer is "Yes." Alternative educational schemes, such as e-learning, would imply a positive and significant cost in terms of the quality of student learning. When considering the introduction of alternative educational models, the benefits of distance learning in terms of cost reduction for suppliers and time saving for students should be carefully weighed against the loss for student learning.

Should, then, anything be done to increase attendance? The answer may be "Yes." The costs of absenteeism are significant and quantitatively relevant. In addition, consider that absenteeism implies not only a direct negative effect on learning, as reported in this study but also significant negative externalities, such as the nuisance to the rest of the class and the high costs to the lecturer outside class (Brauer 1994).

Does this mean that attendance should be made compulsory? The answer is "Definitely not." A compulsory attendance policy would distort the opportunity cost of absenteeism and impose a welfare loss on students.[21] In addition, beside the fact that a captive audience is not a good learning environment, compulsory attendance would take away an important signal for lecturers on the quality of their teaching.[22] The solution to the problem of high levels of academic absenteeism is not to make attendance compulsory, nor to design exams so as to make attendance necessary but to improve the quality of instructors' teaching, in terms of both content and format, to provide students with the right incentives and let them vote with their feet.

NOTES

1. "Even though teaching is a very large part of what we do, we know very little about many aspects of instruction and learning" (Romer 1993, 214).
2. See Romer (1993) and the following discussion in Brauer (1994).
3. To control for the effects of motivation, Romer also examined the results obtained by restricting his sample to students who had completed all the problem sets assigned during the semester.
4. "I believe that the results here both about the extent of absenteeism and its relation to performance are suggestive enough to warrant experimenting with making class attendance mandatory in some undergraduate lecture courses" (Romer 1993, 173).
5. Note that only the three main variables of interest (test score, lecture attendance, and class attendance) are time varying, whereas all other variables are time invariant.
6. Test scores of 100 represented less than 2 percent of the observations in the sample. The pass mark for the rescaled score was 50, given that in the Italian university system the exam pass score is 18.
7. The figures for lectures were very similar to those reported by Romer (1993) and Rodgers (2001) and substantially higher than those reported by Kirby and McElroy (2003). The higher attendance

rate for classes relative to lectures is the opposite of what was found by Rodgers (2002) and Kirby and McElroy (2003).

8. The number of *exams per annum* was used as a measure of ability, despite the fact that the majority of the students were in their first year, because introductory microeconomics is taken in the second semester of the first year. By that time, first-year students could (and should) have passed up to four exams. Among the first-year students in the sample, the sample average of *exams per annum* was 2.08, and only about 6 percent had not passed any exams.
9. The legend for the categorical variables is as follows (sample frequencies in brackets): high school types: technical (0.41), vocational (0.10), general (0.46), other (0.03). Parental education (father, mother): primary (0.09, 0.10), lower secondary (0.28, 0.34), upper secondary (0.43, 0.44), university (0.18, 0.11). Parental occupation (father, mother): manual worker (0.13, 0.08), clerk (0.24, 0.31), executive (0.14, 0.02), self-employed (0.12, 0.09), retired (0.21, 0.06), unemployed (0.11, 0.05), other (0.04, 0.39).
10. See, for example, Lazear (2001), Todd and Wolpin (2003), and Coates (2003).
11. Subject and teacher assessment provide information about the match between academic and student inputs. They are, therefore, a measure of the suitability of the student for the subject, which is what I refer to by the term *motivation*.
12. I report Hausman tests of the null hypothesis that the individual-specific component of the error term (α_i) is uncorrelated with the regressors, on the basis of the comparison of the estimates obtained for the FE and RE models. I also report Breusch-Pagan Lagrange Multiplier tests of the null hypothesis of zero variance of the RE ($\sigma^2_\alpha = 0$).
13. To the extent that GPA reflects the effect of attendance in other courses, and that attendance is positively correlated across courses, the inclusion of this variable could lead to an underestimate of the effect of attendance on performance in introductory microeconomics.
14. This could be interpreted as indicating either that effort and motivation are not correlated with attendance or that student and teacher evaluation and hours of study are not good proxies for motivation and effort.
15. This finding is consistent with the results in Brasfield, McCoy, and Milkman (1992) and Durden and Ellis (1995).
16. Other studies, however, report significant gender-related differences in performance, for example, Sigfried (1979), and Lumsden and Scott (1987).
17. Both models refer to the complete specification that includes the whole set of controls, as in Table 2, column 2.
18. It is interesting to observe that the return to each hour spent attending lectures (0.04*6.25) = 0.25 was substantially higher than the return to each hour of self-study (0.14 and 0.17 in the OLS and RE models, respectively).
19. The difference between the 0.051 and 0.030 estimates for lectures and classes in the RE model was not statistically significant.
20. It is also possible that grades are used, either explicitly or implicitly, to reward for attendance.
21. As observed by Deere (1994, 210), a compulsory attendance policy would contradict many of the principles typically taught in introductory economics courses: "While students may not always make the wisest use of their time, it seems rather arrogant to suggest that we faculty know better the value of our subject just because we know our subject."
22. See the comments in Brauer (1994) for a comprehensive discussion of the arguments against enforcing attendance.

REFERENCES

Brasfield, D., J. McCoy, and M. Milkman. 1992. The effect of university math on student performance in principles of economics. *Journal of Research and Development in Education* 25 (Fall): 240–47.

Brauer, J. 1994. Correspondence. *Journal of Economic Perspectives* 8 (3): 205–15.

Bratti, M., and S. Staffolani. 2002. Student time allocation and educational production functions. Working Paper. 170, Economics Department, University of Ancona, Italy.

Chan, K. C., C. Shum, and D. J. Wright. 1997. Class attendance and student performance in principles of finance. *Financial Practice and Educations* 7 (2): 58–65.

Coates, D. 2003. Education production functions using instructional time as an input. *Education Economics* 11 (3): 273–92.

Deere, D. 1994. In Brauer, J. Correspondence. *Journal of Economic Perspectives* 8 (3): 205–15.

Devadoss, S., and J. Foltz. 1996. Evaluation of factors influencing student class attendance and performance. *American Journal of Agriculture Economics* 78 (3): 499–507.

Dolton, P., O. D. Marcenaro, and L. Navarro. 2003. The effective use of student time: A stochastic frontier production function case study. *Economics of Education Review* 22 (6): 547–60.
Durden, G. C., and L. V. Ellis. 1995. The effects of attendance on student learning in principles of economics. *American Economic Review* 85 (2): 343–46.
Kirby, A., and B. McElroy. 2003. The effect of attendance on grade for first year economics students in University College Cork. *The Economic and Social Review* 34 (3): 311–26.
Lazear, E. P. 2001. Educational production. *Quarterly Journal of Economics* 116 (3): 777–803.
Lumsden, K. G., and A. Scott. 1987. The economics student reexamined: Male-female differences in comprehension. *Journal of Economic Education* 18 (Fall): 365–75.
Marburger, D. R. 2001. Absenteeism and undergraduate exam performance. *Journal of Economic Education* 32 (Spring): 99–110.
Park, K. H., and P. M. Kerr. 1990. Determinants of academic performance: A multinomial logit approach. *Journal of Economic Education* 21 (Spring): 101–11.
Rodgers, J. R. 2001. A panel-data study of the effect of student attendance on academic performance. Unpublished manuscript.
Rodgers, J. R. 2002. Encouraging tutorial attendance at university did not improve performance. *Australian Economic Papers* 41 (3): 255–66.
Romer, D. 1993. Do students go to class? Should they? *Journal of Economic Perspectives* 7: 167–74.
Schmidt, R. 1983. Who maximizes what? A study in student time allocation. *American Economic Review* Papers and Proceedings 73 (2): 23–28.
Siegfried, J. 1979. Male-female differences in economic education: A survey. *Journal of Economic Education* 10 (Spring): 1–11.
Todd, P., and K. Wolpin. 2003. On the specification and estimation of the production function for cognitive achievement. *Economic Journal* 113: F3–F33.
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.