Short-Run Production of Academic Achievement: Does Lecture Attendance Matter for Grades?

Liam Delaney[a], Colm Harmon[b] and Martin Ryan[c†]

[a] Division of Economics, Stirling University, Scotland. Geary Institute, University College Dublin, Ireland.

[b] School of Economics and Geary Institute, University College Dublin, Ireland. School of Economics, University of Sydney, Australia. IZA, Bonn, Germany.

[c] School of Economics and Geary Institute, University College Dublin, Ireland.

18/01/'12

Abstract: This is the first longitudinal study to examine the relationship between lecture attendance and grades across multiple subject areas. There are a number of unobserved characteristics that may affect the decision to attend, as well as affecting exam-performance. Therefore, the econometric specification benefits from repeated measures of attendance and achievement. Besides the inclusion of individual fixed effects (which account for the stable traits of students that cannot be directly measured), the specification benefits from the direct measurement of noncognitive traits. This means that traditionally unobserved time-varying traits can be included in the specification: students' attitude to risk and their future-orientation. In addition, an approximate measure of time-varying class-rooms is included in the specification (this is an interaction between subject area, university affiliation and year of enrolment). This accounts for time-varying factors such as quality of teaching, class-size, or assessment procedure. Results from fixed effects regression show that lecture attendance is not associated with higher grade-scores. Academic achievement appears to be mainly driven by unobserved individual differences, at least in the short-run.

JEL Classification: I21, J2, D90

Keywords: student achievement, lecture attendance, longitudinal analysis, fixed effects, time-varying noncognitive traits, future-orientation, attitude to risk, higher education production functions

[†] **Corresponding author:** Martin Ryan, Desk 7.1, 2nd Floor, Geary Institute, University College Dublin, Dublin 4, Ireland. Tel: 00-353-1-716-4615. Fax: 00-353-1-716-1108. Email: martin.ryan@ucd.ie. The corresponding author is a Ph.D. student at the UCD School of Economics; and a Ph.D. affiliate at the UCD Geary Institute. He was supported for three years of his Ph.D. by the Irish Research Council for the Humanities and Social Sciences (IRCHSS).

Acknowledgements: Thanks to seminar participants at the UCD School of Economics and the Geary Institute for providing comments; and to participants at the annual conference of the Irish Society of New Economists (Trinity College Dublin; September 2010); to participants at the Irish Economics and Psychology Conference (UCD; November 2010); to participants at the First Lisbon Workshop on the Economics of Education (ISEG; January 2011); to participants at the 6th Ph.D. Meeting of the Royal Economic Society (City University London; January 2011); to participants at the *Critical Issues in Irish Society* Ph.D. Seminar Series (UCD; March 2011); and to participants at the annual conference of the Irish Economics Association (Limerick; April 2011). Thanks to the Irish Universities Association for sponsoring the field-work that produced the data for this paper.

1. Introduction

This paper extends the literature on lecture attendance and grades beyond cross-sectional correlations and individual subject areas. There is a large literature on the relationship between attendance and achievement in higher education; and it is commonly reported that attendance is associated with higher grades. Longitudinal analysis features rarely in the literature; however, a small number of recent studies have used fixed effects regression models (Gendron & Pieper, 2005; Cohn & Johnson, 2006; Martins & Walker, 2006; Stanca, 2006; Arulampalam, Naylor & Smith, 2008). Other studies have examined student achievement in higher education using fixed effects estimation; but without a focus on lecture attendance (Arcidiacono & Nicholson, 2005; Arcidiacono, Foster, Goodpaster & Kinsler, 2009; Bandiera, Larcinese & Rasul, 2010; Foster & Kinsler, 2011). Notably, this is the first longitudinal study to examine the relationship between attendance and achievement across multiple subject areas. Furthermore, this is the only study to estimate a higher education production function across more than one subject area; with the exceptions of Betts and Morell (1999), Dolton, Marcenaro and Navarro (2003), Arcidiacono, Foster, Goodpaster and Kinsler (2009) and Bandiera, Larcinese and Rasul (2010).

There are a number of unobserved characteristics that may affect the decision to attend, as well as affecting exam-performance. A similar situation arises in the literature on the returns to education, where there is a concern that higher-ability individuals are more likely to attain higher levels of education as well as higher levels of earnings. In this paper there is a concern that students with higher levels of motivation are more likely to attend their lectures as well as achieve higher grades.[1] In fact, there is some evidence that unobserved heterogeneity explains more about student achievement than observable inputs such as attendance (Martins & Walker, 2006). Also, it is generally accepted that more able and motivated students are more likely to both attend and to score highly in their courses (Arulampalam, Naylor & Smith, 2008). Finally, Bandiera, Larcinese and Rasul (2010) state that underlying student characteristics (such as ability or motivation) are the most important determinants of academic achievement. However, unlike the situation in many studies on the returns to education, it is difficult in this case to find instruments that explain variation

[1] Even when students are in attendance, motivation has a further role to play. As Crede, Roch and Kieszczynka (2010) note, "Physical presence in a classroom… encapsulates a very wide range of possible student behaviours, ranging from students who listen to the professor, take notes, and attempt to understand the material and integrate it with their existing knowledge structure to students who may be physically present but engage in few of the behaviours or cognitive processes that are likely to result in learning."

in lecture attendance, and that are also unrelated to students' performance except through their effect on attendance. Nevertheless, the econometric specification in this paper benefits from repeated measures of attendance and achievement. This accounts for the stable traits of students that cannot be directly measured; and which may influence both attendance and achievement.[2]

Besides individual fixed effects for students, the specification also benefits from the direct measurement of noncognitive traits. This means that traditionally unobserved time-varying traits can be included in the specification: students' attitude to risk and their future-orientation. These traits are not expected to vary a great deal over the short time-period in question. Cunha, Heckman, Lochner and Masterov (2006) have demonstrated the relative malleability of noncognitive traits during the early years of human development. Furthermore, Rivkin, Hanushek and Kain (2005) make the assumption that noncognitive traits (such as motivation and personality factors) do not change during the (later) school years considered for their model of academic achievement. However, while attitude to risk and future-orientation may be relatively stable over a short time-period, it is difficult to rule out any changes occurring in these variables. This is especially the case for the students examined in this paper, given that their traits are interacting with incentives that discourage the smoothing of academic engagement over the entire duration of higher education; there is an end-loading of overall assessment towards the final year of university in the institutional setting of this study. Traditional strategies to estimate education production functions do not allow for time-varying noncognitive traits; however, it has been demonstrated that it is important to account for these time-varying characteristics (Almlund, Duckworth, Heckman & Kautz, 2011). Cobb-Clark and Schurer (2011) provide a detailed discussion on the issue. Finally, an approximate measure of time-varying class-rooms is also included in the specification (this is an interaction between subject area, university affiliation and year of enrolment). This accounts for time-varying factors such as quality of teaching, class-size, or assessment procedure.[3] [4]

---

[2] An important point is that the fixed effects estimation controls for any factors that are stable during the student's time in higher education. Those factors include gender, fathers' education, prior achievement and area of study.

[3] Poor lecture quality is a reason often reported by students for non-attendance (Romer, 1993; Friedman, Rodriguez & McComb, 2001; Dolnicar, 2005; Kottasz, 2005; Clay & Breslow, 2006; Massingham & Herrington, 2006; Arulampalam, Naylor & Smith, 2008; Lang Joyce, Conaty & Kelly 2008).

[4] Larger class-size is associated with non-attendance (Friedman, Rodriguez & McComb, 2001). Grise and Kennedy (2003) show that students perceive smaller theatres to allow for greater interaction between lecturers and students. Students may be less attentive in larger classes, or may compensate for larger classes by exerting more effort outside of lecture times (Bolander, 1973; Feldman, 1984; McConnell & Sosin, 1984). Instructors may be better able to identify the ability and interests of the median student in smaller classes, or be more able to answer students' questions directly (Bandiera, Larcinese & Rasul, 2010).

The use of a web-survey means that the results in this paper are not affected by the selection bias which arises from data-collection in the class-room.[5] This selection bias would have been a major problem for class-room studies that collected data on students' motivation in the past. In fact, it might seem difficult to conceive of a study on the relationship between attendance and grades (across multiple subject areas and institutions) that would not use self-reported data.[6] The results from fixed effects regression in this paper show that lecture attendance is not associated with higher grade-scores. Until longer panels of more objective data are used to investigate this relationship, this study (the first longitudinal analysis across multiple subject areas: measuring noncognitive traits, and avoiding class-room selection bias) shows that attendance is not associated with higher grade-scores. The next section of the paper reviews the relevant literature on the relationship between attendance and achievement. The third section discusses the measurement of time-varying noncognitive traits and the collection of the survey data. The fourth section presents the methodology and results. The fifth section concludes with a discussion.

2. Literature Review

There is a large literature on the relationship between attendance and achievement in higher education; and it is commonly reported that attendance is associated with higher grades. In the economics of education, a production model is prevalent in research on student achievement. Production functions in economics represent the process by which an institution -in this case a school or college- transform inputs into outputs (Hopkins, 1990). Some common inputs are school resources, teacher quality, and family attributes, and the outcome is student achievement. Much of the work using this model of production has concentrated on the educational attainment of pupils in compulsory schooling, with less attention paid to higher education (Arulampalam, Naylor & Smith; 2008). However, there is a precedent for higher education

---

[5] A valid concern is whether survey respondents selected into the survey. However, monetary incentives were used to encourage participation in the survey, which alleviates this concern to a considerable extent.

[6] While self-reported variables provide particular measurement challenges (discussed later in the paper), the use of self-reported data provides the advantage of collecting information across multiple institutions relatively easily. However, there are recent technological advances which substantially ease the burden of collecting more objective data; these are discussed later in the paper.

production functions (Hopkins, 1990; Douglas & Sulock, 1995). There is also a wide empirical literature -often not making reference to a production function- in which economists give attention to the student inputs in the production of achievement: lecture attendance and additional hours of study.

Mandatory attendance policies are rare, in the UK and Ireland at least (Allen & Webber, 2010). Indeed most UK tutors are not in a position to implement a mandatory attendance policy on their own modules as such a strategy would be against the ethos of their university (Allen & Webber, 2010). According to a meta-analysis by Crede, Roch and Kieszczynka (2010), mandatory attendance policies have a small positive impact on grades. However, there is much debate on what incentives or penalties are appropriate in this regard, as penalising students for not showing up can be seen as *double jeopardy*: that is, students would be likely to get lower grades as well as being affected by an attendance-penalty. Stephenson and Deere (1994) suggest that lecture attendance should not be mandatory by making the following arguments: students are missing the least productive classes, a captive audience is not an ideal learning environment, students should be allowed to maximise utility, attendance policies are difficult to implement.

The contemporary literature on attendance and grades begins with a paper examining student time allocation in a *Macroeconomics Principles* course (n=216) (Schmidt, 1983). Schmidt (1983) reports that hours spent attending lectures (and class-discussions) positively affects course grades, even after controlling for additional hours of study (that is, personal study). Park and Kerr (1990) use a multinomial logit model in order to identify the determinants of academic achievement in a *Money and Banking* course (n=97). Their results show that more attendance is associated with higher achievement, although students' GPA and college entrance exam scores are more important factors overall. Romer (1993) is a widely cited study; he surveys attendance at all undergraduate economics classes during one week at a large public institution, a medium-sized private university, and a small liberal arts college. Romer (1993) shows that the effect of class attendance is positive and significant; however, its magnitude is greatly reduced by the inclusion of proxies for motivation.

Following on from the initial few papers by Schmidt (1983), Park and Kerr (1990) and Romer (1993); more cross-sectional studies address the issue of whether lecture attendance matters for grades. Durden and Ellis (1995) use students' self-reported number of absences to explore the relationship between absenteeism and achievement in a *Principles of Economics* course (n= 346). Controlling for student differences in background, ability and motivation, Durden and Ellis (1995) find a nonlinear effect of attendance: while a few absences do not lead to worse grades, excessive absenteeism does. Using data on a sample of 400 *Agricultural Economics* students at four large U.S. universities, Devadoss and Foltz (1996) find that attendance has a substantial effect upon grade-scores, after taking into account motivational and aptitude differences across students. Chan, Shum and Wright (1997) examines the relationship between attendance and student achievement in a *Principles of Finance* course (n=71). After correcting for selectivity bias (due to student withdrawals) by using Tobit and Heckman two-stage models, the results show a positive relationship between attendance and achievement. Maloney and Lally (1998) find that both lecture attendance and previous results are positively and significantly related to examination results for second and third year economics students (n=121) at the National University of Ireland (Galway).

More recently, Marburger (2001) examines the effect of absenteeism on achievement in a *Principles of Microeconomics* course (n=60). Student's attendance records over the semester are matched with records of the class meetings when the material corresponding to each question of three multiple-choice exams was covered. Results show that missing class on a specific day significantly increases the likelihood to respond incorrectly to a multiple-choice question based on the material covered that day (compared to students who were present). Rodgers (2001) finds a small but statistically significant impact of attendance on student achievement in a sample of students enrolled in an *Introductory Statistics* course (n=167). Dolton, Marcenaro and Navarro (2001) use data from a survey conducted at the University of Malaga, on first and final year students. Their sample includes 3722 observations across 40 subject areas. They find that lectures are four times more productive than additional study-hours (personal study). Using a sample of first-year *Economics* students from Italy (n= 71), Bratti (2002) finds that the positive and significant effect of attendance on achievement is not robust to the inclusion of additional study-hours.

Existing panel studies on the relationship between lecture attendance and grades include Gendron and Pieper (2005), Cohn and Johnson (2006), Martins and Walker (2006), Stanca (2006) and Arulampalam, Naylor and Smith (2008). All of these studies use fixed effects regression models. Gendron and Pieper (2005) estimate a model using data from an *Introductory Microeconomics* course in Canada (n=429). Their results show a strong impact of attendance on final grade. Cohn and Johnson (2006) examine the relationship between attendance and achievement in a sample of 347 economics students. Their findings indicate that there is a strong positive correlation between attendance and achievement. Martins and Walker (2006) use records of student attendance at class meetings of all 1st and 2nd year undergraduate modules offered in *Economics* at a UK university (n=1700).[7] They find that class attendance is statistically insignificant after controlling for student fixed effects. Stanca (2006) uses a panel data-set collected from

[7] The literature for the US typically measures attendance rates aggregated over all forms of meetings, and labels these as 'classes'. There is, however, a potentially important distinction to be drawn (in the UK and Ireland) between attendance at lectures – typically large group meetings – and at classes, which are typically small group meetings (Arulampalam, Naylor & Smith, 2008).

an *Introductory Microeconomics* course at an Italian university (n=766); combining administrative and survey sources. Attendance at classes and tutorials is self-reported by students in this data-set. Applying instrumental variables and fixed effects to address the endogeneity of attendance, Stanca (2006) finds that attendance has an important independent effect on student achievement. Arulampalam, Naylor and Smith (2008) use an administrative panel dataset for cohorts of *Economics* students at a UK university (n=444). They find that there is a significant effect of class absence on students' performance after controlling for unobserved individual effects.

3. Data Description

The Irish Universities Study (sponsored by the Irish Universities Association) is a large scale web-survey that the authors, and other researchers at the UCD Geary Institute, designed to elicit feedback from students attending the seven universities in Ireland. The data used in this paper are taken from the longitudinal component of the web-survey: the first wave was conducted during spring 2009; the second wave (the follow-up) was conducted during spring 2010. The original wave of the longitudinal component (spring 2009) contains 4,770 observations;[8] the follow-up (spring 2010) contains 1,622 observations. This paper proceeds with the use of a balanced panel (n=1,622). Attrition between the survey-periods is substantial: 66 percent of the original sample did not respond adequately to the follow-up.[9] However, of those students who qualify for the balanced panel, every one of these answers questions about their age and gender. Furthermore, most students in the panel answer the survey comprehensively. Nonetheless, in addition to the attrition (unit non-response), there are missing values (item non-response) in each survey-period.[10] Overall however, no more than 10 percent of the data is missing for any variable: in the sample used for this paper. In fact, the extent of missing values is generally much less than 10 percent.

When the characteristics of the attritors are different to those in the retention sample, the representativeness of a sample deteriorates. To allay concerns about a loss of representativeness,

[8] Delaney, Harmon and Ryan (2010) use an analytical sample of 2,867 full-time undergraduates, based on the first wave of the longitudinal component. The first wave collected information on the 'Big Five' personality factors; Delaney, Harmon and Ryan (2010) use this information for their investigation into the determinants of study-behaviours.

[9] Attrition is defined here as the probability of an individual leaving the sample as the number of periods increases.

[10] Some panel studies observe the joint decline of item and unit response rates over time (Van de Ven & Van Praag, 2002). This finding may be explained by self-selection of respondents: over time only the motivated respondents stay in the group of panel participants and they have low item non-response propensities (Van de Ven & Van Praag, 2002).

robustness checks between the data from the balanced panel and the data available for the population of Irish university students are presented in Appendix A. Overall, across gender, institution and area of study, the panel-sample is broadly representative of its underlying population.[11] Analysis is restricted to full-time undergraduates because part-time students and postgraduates are characteristically different groups.[12] All of the students in the sample are studying for honours bachelor degrees (n = 803; 1,606 observations); Table 1 shows summary statistics. The average percentage grade-score is 65-58, the average percentage of lectures attended is 83-84, students' average age is 21-22, and the average year in university is 2-2.5.[13] About 36 percent of students are male, approximately 50 percent of students' fathers have some higher education (this is a binary indicator), and average family-income is in the range of €60,000-80,000. Family-income is measured in brackets of €20,000 and top-coded at €140,000+. Students' average points-score in the Leaving Certificate (an exam taken at the end of secondary-school in Ireland) is 475; the maximum score on this measure of achievement is 600 points. The average amount of time that students spend studying per week is 15 hours. Additional study-hours are measured in a grid comprised of hours per week, categorised as follows: 0, 1-5, 6-10, 11-15, 16-20, 21-30, 31-40, 41-50, 51-60, 60+. Future-orientation is measured on a scale ranging from 5-20; willingness to take risks is measured on a scale ranging from 0-10.[14] In the second survey-period students are (on average) achieving higher grades, spending more time studying, and are slightly less willing to take risks. This change in the pattern of students' behaviour is consistent with underlying (deteriorating) macro-economic conditions.[15]

[11] Males are slightly under-represented. One university is somewhat under-represented. Another university is somewhat over-represented. Re-weighting the sample to be fully representative of its underlying population makes no difference to the overall pattern of results.

[12] These students are likely to be older, more career-focused, and to be paying tuition fees.

[13] Most courses are 3-4 years in duration; a small number (such as Medicine) last 6 years.

[14] Section 1.2 of the thesis discusses the measurement of noncognitive traits.

[15] Family-income is also down slightly in the second survey-period.

Table 1: Summary Statistics

Variable	W1: Mean	W1: S.D.	W1: N	W2: Mean	W2: S.D.	W2: N	W2-W1
Average grade	65.1	9.87	604	68.5	8.34	681	+77
Lecture attendance	83.4	15.7	741	84.3	18.2	683	-56
Year of enrolment	2.00	0.97	801	2.64	0.99	799	-2
Age of student	21.2	5.07	803	21.9	5.08	803	0
Student is male	0.35	0.47	803	0.35	0.48	803	0
Student's father has HE	0.50	0.47	752	0.50	0.47	756	+4
Family-income bracket	4.21	2.51	773	4.09	2.35	755	-18
Future-orientation	13.9	3.54	751	14.0	3.51	732	-19
Willing to take risks	6.33	1.66	755	5.96	1.88	745	-10
Prior achievement (LC)	474	76.0	719	476	73.9	706	-13
Study-time interval	2.99	1.30	733	3.44	1.52	769	+36

Notes: W1 = Wave 1. W2 = Wave 2.
W2-W1 is the difference in the number of observations between survey-periods.
HE = higher education. LC = Leaving Certificate

The column headed "W2-W1"in Table 1 is the difference in the number of observations between survey-periods. This is the temporal aspect of item non-response. Most of the item non-response is greater in the second survey-period. However, there are more observations for grade-scores and additional study-hours in Wave 2. This is the case because the question relating to additional study-hours occurs later in Wave 1 than it does in Wave 2. In relation to the grade-score variable, information on average percentage grade-score for Wave 1 is reported during Wave 2 (i.e. it is recalled during Wave 2). Hence, there are more observations for grade-score in Wave 2. There is an ordered measure of average grade-score reported during Wave 1, which could have been used instead. However, the continuous grade-score measure is preferred because it contains more information, and does not require the use of discrete choice modelling.[16] In addition, the continuous measure for Wave 1 (recalled in Wave 2) is extremely consistent with the ordered measure of grade-score reported during Wave 1. Table 2 illustrates this by comparing continuous and ordered measures of grade-scores. (An ordered measure is also reported in Wave 2). The continuous measure is based on the following question: "On a scale of 0-100, what is your average grade-score at university?" The ordered measure takes the form "1h1, 2h1, 2h2, Pass" or "A+, A, A-, B+, B, B-, C+, C, C-, D+, D, D-"; depending on the grading-scheme the student is actually

[16] This will be an important point when applying fixed effects regression to the panel.

marked on. The ordered measure of grade-score comes from a question at a different position in the survey. Therefore, Table 2 shows that students are answering questions about their grade-scores in a consistent manner. This also indicates that students generally treated the survey with care.

Table 2: Ordered Grade-Scores vs. Continuous Grade-Scores

	Ordered Categories	**W1: Continuous Measure**	**W2: Continuous Measure**
Grade Option 1:	1st	73	74
	2:1	64	66
	2:2	59	60
	Pass	51	54
	Fail	50	-
Grade Option 2:	A+	81	80
	A	78	80
	A-	75	77
	B+	70	73
	B	68	69
	B-	65	66
	C+	60	63
	C	59	58
	C-	56	50
	D+	56	52
	D	52	44

Notes: W1 = Wave 1. W2 = Wave 2.

However, it is possible that students may be over-stating their grades (albeit consistently so).[17] Maxwell and Lopus (1994) discuss the *Lake Wobegon Effect* in student self-reported data; demonstrating that below-average students tend to inflate their academic achievements. Cassady (2001) finds that low-performing students over-report their grade-scores, more than high-performing students. Dobbins, Farh and Werbel (1993) suggest that students tend to inflate their past performance scores to a level that they consider socially acceptable (or desirable). Social desirability bias is a term used to describe the tendency of respondents to reply in a manner that will be viewed favourably by others; see Bound, Brown and

[17] Despite assurances to students (before starting the web-survey) that their data would be anonymised.

Mathiowetz (2001) for a discussion. Haley, Johnson and McGee (2010) examine whether using student-survey data in place of official records data biases regression estimates; these authors find that it is not necessary to correct for bias from the *Lake Wobegon* effect.

Nonetheless, it is possible that grade “mis-reporting” may vary from sample to sample; and that grades are being over-stated in this sample. A related issue is students’ perception of their academic ability. Chevalier, Gibbons, Thorpe, Snell and Hoskins (2009) show that students are not good at predicting their own performance. Also, students generally over-estimate their own ability (Falchikov & Boud, 1989); and tend to be overconfident about their future academic performance (Zafar, 2011). However, in the data used here, students are not required to anticipate their future performance; they simply report their average percentage grade-score (we assume that students use information from the immediately preceding semester in the provision of this self-report). Finally, lecture attendance (the independent variable of particular interest) may also be overstated. Similar to the theory surrounding self-reported grades, attendance may be overstated due to the presence of *social desirability bias*.[18]

Table 3 shows the between- and within-variation of the two-period panel (as well as the overall means and standard deviations). The top half of Table 3 shows variables where within-variation is expected over time (these variables can be considered “time-varying”). The bottom half of the table shows variables where within-variation is unexpected over time (these variables can be considered “time-invariant” – that is, they take values of zero in the “within” column). Two of the “time-invariant” variables are students’ gender and students’ prior academic achievement. This is an indication that the panel-data are reliable; that is, that students are not providing random answers to the survey. One of the “time-invariant” variables, whether the students’ father has some higher education, exhibits some variation over time in its raw form. It is possible that fathers could be entering into higher education in the second survey-period; however, any observations for father’s education that do not match up between survey-periods are replaced with missing values.[19]

[18] However, students may also be attending more of their lectures in the recession than they used to beforehand. University students in the UK studied for two hours and 12 minutes more (per week) in 2009 than they did in 2007, according to the Higher Education Policy Institute (Bekhradnia, 2009).

[19] Also, any observations for gender and prior achievement - that do not match up between survey-periods - are replaced with missing values. These adjustments are necessary to ensure that time-invariant variables drop out of fixed effects regression. The adjustments apply to both time-periods.

Table 3: Between and Within Variation of the Panel

	Mean	Overall S.D.	Between S.D.	Within S.D.
Time-Varying:				
Average grade	66.9	9.25	8.18	4.65
Lecture attendance	83.8	17.0	15.9	6.74
Age of student[20]	21.6	5.09	5.10	0.52
Year of enrolment	2.33	1.03	0.84	0.61
Family-income bracket	4.15	2.44	2.31	0.74
Study-time interval	3.23	1.44	1.30	0.67
Future-orientation	14.0	3.53	3.22	1.51
Willing to take risks	6.15	1.78	1.60	0.80
Time-Invariant:				
Student is male	0.36	0.48	0.48	0
Student's father has HE	0.51	0.50	0.50	0
Prior achievement (LC)	484	71.7	71.7	0

4. Method and Results

4.1 Methodology

The empirical analysis is concerned with the role of student inputs in a higher education production function. The main econometric specification is a fixed effects (within-student) regression. A naïve OLS regression is also used for comparison purposes. OLS is applied to the pooled data first, as follows:

$$Y_{ijt} = \alpha_i + \beta_1 att_{ijt} + \beta_2 family_{ijt} + \beta_3 college_{ijt} + \beta_4 noncog_{ijt} + \pi_1 year + \mu_{ijt} \quad (1)$$

[20] Date of birth is time-invariant. Age-cohorts are time-invariant in some studies. However, age varies across time in this case.

where Y_{ijt} is a measure of educational achievement (average percentage grade-score) for student *i* in university *j* in year *t*; att_{ijt} is students' lecture attendance; specifically, average percentage of lectures attended; $family_{ijt}$ is a matrix of family background variables (and student demographics): age, gender, year of enrolment, subject area, father's education, family-income, study-hours, prior (second-level) achievement; $college_{ijt}$ is a matrix of institutional and class-room effects; and $noncog_{ijt}$ is a matrix of variables related to students' noncognitive traits: attitude to risk and future-orientation. Class-room effects are approximately derived from a three-way interaction between subject area, university affiliation and year of enrolment. This accounts for factors such as quality of teaching, class-size, or assessment procedure. There is also a dummy-indicator for the survey-period, $\pi_1 year$, included as a control for unobservable factors that vary across time but are constant throughout each survey-period (Martins and Walker, 2006).[21] The error term, μ_{ijt}, represents all unobserved factors; this is assumed to be independent across individuals (and survey-periods).

The main econometric specification, a fixed effects (within-student) regression, is applied to the panel data as follows:

$$Y_{ijt} = \alpha_i + \beta_1 att_{ijt} + \beta_2 family_{ijt} + \beta_3 college_{ijt} + \beta_4 noncog_{ijt} + \pi_1 year + \pi_2 ind + \mu_{ijt} \qquad (2)$$

In Eq. 2, the only differences are the inclusion of an individual fixed effect, $\pi_2 ind$; and the fact that the coefficients of the regressors on the time-invariant variables are not identified: because they have no within-variation. The first difference, the inclusion of a student fixed effect, allows for the possible correlation between some of the explanatory variables and any stable characteristics of students that have not been directly measured. The second difference (lack of within-variation) primarily affects the "family" matrix; so that students' gender, subject area, fathers' education, and students' prior achievement are not included in the fixed effects specification. In addition, with reference to the "college" matrix, students' university-affiliation is not included in the fixed effects specification. An important point to emphasise is that the fixed effects estimation controls for any factors relating to students that are stable during the student's time in higher education. Those factors include gender, fathers' education, prior (second-level) achievement, area of study and university-affiliation. Of note, the three-way

[21] This is an important control for the deteriorating macroeconomic conditions between survey-periods. Students may be attending more of their lectures in the recession than they used to beforehand. University students in the UK studied for two hours and 12 minutes more (per week) in 2009 than they did in 2007, according to the Higher Education Policy Institute (Bekhradnia, 2009).

interaction between university, subject area and year of enrolment is included in the fixed effects estimation; as students' class-room changes from one academic year to the next. Students' attitude to risk and future orientation are also included as time-varying factors. The error term, μ_{ijt}, represents all unobserved factors; this is assumed to be independent across individuals and serially uncorrelated. The error in the fixed effects estimation is likely to be clustered over time for a given individual, so cluster-robust standard errors are used. Finally, *t* is a time index that spans two time periods; as a result, the fixed effects regression shown above (Eq. 2) is analogous to a first-differenced approach. With two-period panel data, differencing results in one cross-sectional equation, where time-varying factors represent the change over the two periods.[22]

Some consideration must be given to the initial conditions problem which arises in longitudinal analysis; this occurs when the start of the observation period does not coincide with the start of the stochastic process generating individuals' outcomes (Heckman, 1981). In this case the outcome is academic achievement; and any individual who is observed with a certain level of academic achievement at the start of the observation period -- may be observed as such because of an earlier history of achievement. Furthermore, an individual who has a low level of academic achievement in an early time-period may struggle to engage with the process of learning in subsequent time-periods; especially if the curriculum in later time-periods requires knowledge of content from curricula in earlier time-periods. This is often the case. In addition, an individual who has a low level of academic achievement in an early time-period may feel discouraged from attending lectures in subsequent time-periods.[23] Most critically, an individual who has a low level of academic achievement in an early time-period may be more likely to drop out of education altogether. Unfortunately, longer panel duration is necessary to adequately address the issue of persistence (or state dependence) in students' academic achievement. Also, estimation over two time-periods (as in this case), prevents one from observing a complete cohort trajectory; that is, the movement of individuals from the start of higher education right through to the end.

To address the role of previous conditions in the empirical strategy, prior academic achievement (at the end of second-level schooling) is included in the OLS specification; this represents a cumulative

[22] We do not estimate a between effects model because we do not think there could be many individual-level factors (that we do not measure) which change from one period to the next, for such a short duration. By extension, we do not estimate a random effects model (which is a combination of fixed effects and between effects).

[23] Conversely, an individual who has a low level of academic achievement in an early time-period may strengthen their determination to perform better in the next time-period. However, the effect of this resolve on achievement may be dampened by learning difficulties associated with curricular prerequisites.

function of prior family, community, and school experiences (Rivkin, Hanushek & Kain, 2005).[24] However, the coefficient on prior academic achievement cannot be identified in the fixed effects specification, because the measure of prior achievement (at the end of second-level) has no within-variation; that is, it does not change over time. Nonetheless, any stable qualities relating to prior achievement are controlled for in the fixed effects specification. Fixed effects estimation controls for all time-invariant differences between individuals; therefore the estimated coefficients from such a model cannot be biased due to omitted time-invariant characteristics. With respect to prior achievement *during* higher education; Achen (2001) discusses the drawback of using a lagged dependent variable in longitudinal analysis, particularly in a short panel that runs over two or three time-periods.[25] When serial correlation is high and the exogenous variables are heavily trended, as will happen frequently in panel data, the lagged variable will falsely dominate the regression and suppress the legitimate effects of the other variables (Achen, 2001). According to and Bedi (2011), estimates from a specification (for a student achievement model) that includes a lagged dependent variable are likely to be inconsistent as the lag is correlated with unobserved ability. Todd and Wolpin (2003) also raise concerns about lagged measures of academic performance.

Finally, the use of a web-survey means that the results in this paper are not affected by the selection bias which arises from data-collection in the class-room.[26] This selection bias would have been a major problem for class-room studies that collected data on students' motivation in the past. Of course, the use of a web-survey also means that measurement error is more likely (compared to the use of administrative records). Measurement error bias is considered to be more exaggerated in fixed effects estimates because random misclassification in two periods will produce a larger number of misclassified variables (Griliches & Hausman, 1986; Swaffield, 2001). Freeman (1984) shows that the effect of unions on wages is biased downward in panel studies due to misclassification of union status. Hamermesh (1989) shows that imprecision in fixed effects estimates can arise from measurement error in dependent variables. However, descriptive statistics from the previous section show that students answered survey-questions consistently across time-periods (Table 3, with respect to time-invariant variables), and also within the same time-period (Table 2, with respect to ordered and continuous measures of grade-score).

---

[24] There is evidence that students overestimate their performance in secondary education. In England, 96 percent of secondary school pupils believe that they are "Average" or above when asked how good they are at their school work (Gibbons & Silva 2007). Therefore, the measure of prior academic achievement used in this paper may suffer from upward-biased measurement error. Nonetheless, it is still a useful barometer of prior conditions.

[25] The inclusion of a lagged dependent variable would essentially mean the estimation of a "value-added" model. Todd and Wolpin (2003) discuss the undesirability of such an approach when unobserved characteristics can be accounted for.

[26] A valid concern is whether survey respondents selected into the survey. However, monetary incentives were used to encourage participation in the survey, which alleviates this concern to a considerable extent.

4.2 Analysis

Table 4 shows several models explaining students' academic achievement, using data on the same individuals over two periods in time: spring 2009 and spring 2010. Results from the OLS regression (Column 1) show that grades are predicted by students' lecture attendance, their prior achievement, their additional hours of study, and their future-orientation. These are findings which are strong in the literature; see Martins and Walker (2006), Stinebrickner and Stinebrickner (2008), Arulampalam, Naylor and Smith (2008) and Grave (2011) - on the role of additional study-hours in academic achievement; and Joireman (1999) and Peters, Joireman and Ridgway (2005) - on the role of future-orientation in academic achievement. The coefficient on the lecture attendance variable in the OLS regression suggests that a one percentage increase in lecture attendance predicts an extra 0.11 percent of a student's grade score. This means that students could gain an extra percentage point in their grade-scores by attending 10 percent more of their lectures; this is not a very substantial effect-size. The effect of being future-orientated is much more substantial. The coefficient on future-orientation in the OLS regression suggests that a one point increase in future-orientation predicts an extra 2.4 percent of a student's grade score. This means that students could move into a higher award category by being four points higher on the scale for future-orientation (which has sixteen points). The effect of being in a higher category of study-hours improves a student's grade-score by 0.4 percent. The effect of an extra ten points in the examination-score at the end of second-level improves a students' grade-score by 0.2 percent. It would take very unlikely changes to study-hours or prior achievement to make a substantially different improvement to students' grade-score.

Results from fixed effects regression (Column 2) show that lecture attendance is not associated with higher grade-scores, once unobserved stable traits are taken into account by using fixed effects estimation.[27] This finding is in opposed to the majority of existing research on the relationship between lecture

[27] A Hausman test is performed, and this indicates that the fixed effects model is more efficient than the random effects model. Therefore, it is statistically appropriate to focus on the fixed effects regression.

Table 4: Longitudinal Regressions Explaining Students' Average Grade at University: (Irish Universities Study: Spring 2009 and Spring 2010)

	(1)	(2)	(3)	(4)
	OLS on Pooled Data	Fixed Effects Regression	Fixed Effects: STEM	Fixed Effects: Non-STEM
Lecture attendance	0.109***	0.018	0.039	0.011
	(0.017)	(0.029)	(0.052)	(0.035)
Year of enrolment	26.967**	-4.788	-9.452	-5.697
	(8.705)	(0.000)	(0.000)	(4.592)
Age of student	-0.013	-0.150	-0.751	0.396
	(0.055)	(0.951)	(1.684)	(1.067)
Student is male	0.549			
	(0.318)			
Student's father has HE[a]	0.951			
	(1.021)			
Family-income bracket[b]	0.058	-0.236	-0.133	-0.128
	(0.095)	(0.361)	(0.687)	(0.319)
Prior achievement (LC)[c]	0.025***			
	(0.006)			
Study-time interval[d]	0.425***	-0.305	-0.474	-0.064
	(0.118)	(0.282)	(0.461)	(0.398)
Willing to take risks	0.155	-0.313	-1.193	0.431
	(0.403)	(0.598)	(1.118)	(0.606)
Future-orientation	2.435***	0.841	1.903*	-0.002
	(0.564)	(0.584)	(1.006)	(0.728)
STEM subject[e]	0.994			
	(1.977)			
Period 2	3.235***	9.204***	10.108***	5.139
	(0.305)	(3.101)	(3.835)	(3.875)
Constant	13.053	98.474***	80.653**	51.843**
	(11.818)	(20.785)	(36.963)	(23.950)
Observations	1,285	1,285	602	683
R-squared	0.277	0.220	0.219	0.304

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: In the first column, academic achievement is modelled using OLS regression; robust standard errors are clustered by students' university affiliation. The second column shows results from fixed effects regression. A class-room effect based on a three-way interaction between university-affiliation, subject area and year of enrolment is not shown. In the third column, the same fixed effects regression is applied to students who are enrolled in STEM subjects only. In the fourth column, the same fixed effects regression is applied to students who are enrolled in non-STEM subjects only. All standard errors in the fixed effects regression are robust. Missing values are replaced by a value of zero: to keep as many observations in the sample as possible. Dummy variables are included to take account of where this is done; this method is known as *dummy variable adjustment*. Where they apply to incomplete cases, control variables for missing value adjustment are not shown above. Outliers and missing values are adjusted for independent variables only. Attitude to risk and future-orientation are standardized using z-scores.

[a]HE is higher education [b]Income-brackets are in categories of €20,000 [c]LC = Leaving Certificate. Leaving Cert. points are a continuous measure of academic achievement at the end of second-level education in Ireland [d]Additional study is extra hours of personal study [e]STEM subjects are science, technology, engineering and maths

attendance and grades. However, the use of fixed effects regression for models of student achievement is a relatively recent development. Existing panel studies on the relationship between lecture attendance and grades are rare. Gendron and Pieper (2005), Cohn and Johnson (2006), Stanca (2006) and Arulampalam, Naylor and Smith (2008) all find a strong relationship between attendance and achievement, using fixed effects regression. However, Martins and Walker (2006) find that class attendance is statistically insignificant after controlling for student fixed effects. There is now a clear division in the small literature that uses repeated measures of attendance and achievement to investigate this topic.

Furthermore, the results from Column 2 in Table 4 show that academic achievement appears to be mainly driven by unobserved individual differences. This is in keeping with a recent finding by Bandiera, Larcinese and Rasul (2010). In fact, there are no statistically significant explanatory variables in the main fixed effects regression in Table 4; besides the dummy-indicator for survey-period. While this dummy-indicator is a useful control variable, it should be interpreted with caution. Although students in the second survey-period are more likely to achieve higher grades, those students are the "non-attritors"; that is, they stayed in the sample after Wave 1 (and may have more desirable attributes).

Finally, given the fact that this is the first longitudinal study across multiple subject areas, additional results are provided in Table 4 in sub-samples defined by STEM (science, technology, engineering and maths) and non-STEM enrolment (Columns 3 and 4, respectively).[28] This distinction is important because STEM students are required to attend more lectures than non-STEM students.[29] The results show that STEM-enrolled students with higher levels of future-orientation are somewhat more likely to achieve higher grades. As was the case in Column 2, the coefficient on the attendance variable is insignificant in both of the sub-samples that are defined by subject choice.

Appendix B shows a number of additional fixed effects estimations; based on the inclusion of interaction terms between lecture attendance and the statistically significant independent variables from the OLS regression in Table 4. The statistically significant independent variables from the OLS regression are: year of study, prior (second-level) achievement, additional study-hours, and future-orientation. As mentioned at the outset, students' non-cognitive traits are interacting with incentives that discourage the smoothing of their academic engagement over the entire duration of their higher education. This is because

[28] Betts and Morell (1999) find that grades are lowest in science and engineering and highest in the arts and humanities. There is a common belief that the sciences and maths grade harder than the social sciences, which in turn grade harder than the humanities (Achen & Courant, 2009). In addition, students are required to attend more lectures if they enrol in a STEM subject. By extension, STEM students have less time for additional study.

[29] There is also a common belief that the sciences and maths grade harder than the social sciences, which in turn grade harder than the humanities (Achen & Courant, 2009).

there is an end-loading of overall assessment towards the final year of university in the institutional setting of this study. Therefore, there is a good theoretical reason for interacting year of study with lecture attendance. There is also a good theoretical reason for interacting prior achievement with attendance. This is because there may be some students who are more intelligent or hard-working; and these students can still perform well despite lower levels of attendance. Furthermore, these students may even be bored in class; and therefore not as inclined to attend. Additional study-hours are interacted with attendance as students who revise material or who prepare in advance for lectures, may get more benefit from attending. Finally, future-orientation is interacted with attendance as an approximate indication for more motivation during lectures. It is possible that lectures matter more for students who engage more with lectures. So it is the students who engage less with lectures who might put a downward bias on the effect of attendance. Appendix B shows that none of the interaction terms are statistically significant; nor does the coefficient on lecture attendance become statistically significant in any of the additional estimations containing the interaction terms.

Appendix C returns to the issue that students who are more intelligent can still perform well despite lower levels of attendance. The use of prior academic achievement as a multiplier for lecture attendance (in an interaction term) went some way towards tackling this issue (in Appendix B). However, the particular concern here is that the effect of lecture attendance may differ throughout the distribution of academic achievement. To account for this, Appendix C breaks down academic achievement into three sub-samples of award status: second class honours (division one), first class honours, and students who report their average percentage grade-score as being higher than 75%.[30] The effect of attendance on grades is statistically insignificant in each of these sub-samples. However, it is noteworthy that the sample of students used in this paper is characterised by high-achievement; a comparison to official grade data from Irish universities confirms this.[31] The high achievement levels that are reported in the sample may be due to students over-stating the level of their academic performance; students selecting into the sample based on higher levels of

[30] Second-class honours (division two) is defined as an average percentage grade-score between 50% and 60%. Second-class honours (division one) is defined as an average percentage grade-score between 60% and 70%. First-class honours is defined as an average percentage grade-score higher than 70%. Unfortunately the number of observations in the sub-sample for second-class honours (division two) is too small to use for estimation.

[31] Information about official grade data relating to the population of university students in Ireland is taken from this page on the website of the Irish Higher Education Authority (HEA): http://www.hea.ie/en/node/289. The official grade distribution is as follows: first-class honours: 14%; second-class honours (division one): 43%; second-class honours (division two): 23%; pass or other: 20%. The sample grade distribution is as follows: first-class honours: 40%; second-class honours (division one): 46%; second-class honours (division two) or lower: 14%.

achievement; or a combination of both these factors.[32] This is an issue that would benefit from a comparison between self-reported and administrative data for the same individuals. Finally, Appendix D considers non-linearity in students' lecture attendance; it transforms lecture attendance into the following dummy-indicators (none of which are statistically significant): (i) 0 = attend less than 50% of lectures/1 = attend more than 50% of lectures (ii) 0 = attend less than 80% of lectures/1 = attend more than 80% of lectures (iii) 0 = attend less than 90% of lectures/1 = attend more than 90% of lectures.

5. Conclusion

The major contribution of this paper to examine the relationship between lecture attendance and grades: longitudinally and across multiple subject areas; it is the first study to do this. In addition, the specification in this paper includes repeated measures of students' attitude to risk and their future-orientation. Traditional strategies to estimate education production functions do not allow for such (time-varying) noncognitive traits. Also, an approximate class-room effect (derived from a three-way interaction between subject area, university affiliation and year of enrolment) accounts for time-varying class-room factors such as quality of teaching, class-size, or assessment procedure. Finally, the use of a web-survey means that the results in this paper are not affected by the selection bias which arises from data-collection in the class-room. This would have been a major problem for class-room studies in the past. The results show that lecture attendance is not associated with higher grade-scores. In fact, academic achievement appears to be mainly driven by unobserved individual differences.

While the results of this paper do not demonstrate any benefit from attending lectures over a two-year time-period, no inference can be drawn about the effect of attendance over the entire duration of higher education. Also, lectures may have other benefits besides their potential effect on grade-scores. Attending lectures could be beneficial for adjustment into college, the formation of peer-networks, the accumulation of social capital, and student well-being more generally. Nonetheless, if academic achievement in higher education is mainly driven by unobserved individual differences (as shown in this paper), then that would under-score the importance of fostering non-cognitive skills earlier in the life-cycle of the student.

[32] A similar explanation can be offered for the seemingly high levels of lecture attendance in the sample-data. This may be due to students over-stating the level of their lecture attendance; students selecting into the sample based on higher levels of attendance; or a combination of both these factors. It doesn't add value to re-weight the data by grade-score, because unlike age, gender or institution, there is good reason to suspect mis-reporting (over-stating) of grade-scores. Therefore, one would not necessarily be re-weighting to correct a selection problem.

Information on grades and attendance was acquired from students' self-reports; this provides challenges for measurement; but also provides the advantage of collecting information across multiple institutions relatively easily. In fact, it might seem difficult to conceive of a study on the relationship between lecture attendance and grades (across multiple subject areas and institutions) that would not use self-reported data. It would not be easy to count attendance in every class-room at a given university. However, there are recent technological advances which substantially ease the burden of collecting more objective attendance data. Smart-card technology is available specifically for the purpose of measuring student attendance.[33] If this electronically recorded attendance data was linked to administrative grade information, then measurement error would be absent. However, there would still be the issues of unobserved time-invariant characteristics (which necessitates the use of panel data); and traditionally unobserved time-varying characteristics (which necessitates the measurement of self-reported traits). While it would be interesting to see the results from a regression of electronically recorded attendance data on administrative grade information, such an approach would have to be longitudinal in order to account for potentially unobserved characteristics of a stable nature. It would also be necessary to link such an 'objective panel' to self-reported noncognitive traits; in order to account for (unmeasured) time-varying characteristics.

Future research should replicate the collection of self-reported data on attendance and grades (as demonstrated in this paper); with an emphasis on: (i) the measurement issues associated with the potential over-statement of these variables, and (ii) the need for comparisons with other data-sources. If it could be shown that self-reports of grades do not differ systematically from administrative data on student performance, then a reliable measure of achievement could be included in many of the surveys conducted by higher education institutions every year (without researchers having to link to administrative data).[34] However, if there was still a misrepresentation of achievement in sample-data; then that would indicate the existence of a self-selection process. Such a process could explain why there is no effect of attendance upon grades in the fixed effects regression results of this paper. That is, one would imagine that lectures might matter more for lower-achieving students; but if a sample does not contain a representative amount of lower-achieving students; then an effect of attendance on grades might not be observed. Of course, over-statement in the self-reporting of grades is another potential explanation for why an effect of attendance upon achievement is not observed in the fixed effects regression results of this paper.

Finally, it is worth noting the short duration of the two-period panel used in this paper. Future

[33] There are new electronic systems which are being used to detect the ID cards students are carrying as they enter classrooms at Arizona University, and at one Irish institution of higher education.

[34] Of course, self-reporting of grades could vary across different institutions and countries. This would also have to be considered.

research should endeavour to use panels of longer duration. This is necessary to adequately address the issue of persistence (or state dependence) in students' academic achievement. A student who has a low level of academic achievement in an early time-period may struggle to engage with the process of learning in subsequent time-periods; especially if the curriculum in later time-periods requires knowledge of content from curricula in earlier time-periods.[35] Also, estimation over two time-periods (as done in this paper) does not allow one to observe a complete cohort trajectory; that is, the movement of individuals from the start of their higher education through to the end.[36] Until longer panels of more objective data are used to investigate the relationship between lecture attendance and grades, this study (the first longitudinal analysis across multiple subject areas: measuring noncognitive traits, and avoiding class-room selection bias) shows that attendance is not associated with higher grade-scores. Academic achievement appears to be mainly driven by unobserved individual differences, at least in the short-run.

References

Achen, A.C. & Courant, P.N. (2009). What Are Grades Made Of? *Journal of Economic Perspectives, 23* (3), 77-92.

Achen, C.H. (2001). Why Lagged Dependent Variables Can Suppress the Explanatory Power of Other Independent Variables. *University of Michigan mimeo*. Presented at the Annual Meeting of the American Political Science Association (UCLA, 2000).

Allen, D.O. & Webber, D.J. (2010). Attendance and exam performance at university: a case study. *Research in Post-Compulsory Education, 15* (1), 33–47.

Almlund, M., Duckworth, A., Heckman, J.J., & Kautz, T. (2011). Personality Psychology and Economics. *Handbook of the Economics of Education*. Edited By Erik A. Hanushek, Stephen J. Machin & Ludger Woessmann, forthcoming.

Arcidiacono, P. & Nicholson, S. (2005). Peer Effects in Medical School. *Journal of Public Economics, 89*, 327-350.

---

[35] In addition, an individual who has a low level of academic achievement in an early time-period may feel discouraged from attending lectures in subsequent time-periods, or may simply feel unmotivated even if he or she attends lectures and spends additional time on study. Most critically, an individual who has a low level of academic achievement in an early time-period may be more likely to drop out of education altogether.

[36] The observation of an entire cohort trajectory would enable researchers to learn more about the time-varying nature of noncognitive traits during higher education. One hypothesis is that these traits might be stronger in later years of study, due to the end-loading of overall assessment towards the final and penultimate years of higher education. The more valuable a course is to a student's final degree-score, the more future-orientation and risk-aversion may characterise their behaviour.

Arcidiacono, P., Foster, G. Goodpaster, N. & Kinsler, J. (2009). Estimating Spillovers in the Classroom with Panel Data. *Duke University,* mimeo.

Arulampalam, W., Naylor, R. & Smith, J. (2008). Am I Missing Something? The Effects of Absence from Class on Student Performance. *IZA Discussion Paper* 3749, Institute for the Study of Labor (IZA).

Bandiera, O., Larcinese, V. and Rasul, I. (2010). Heterogeneous class size effects: new evidence from a panel of university students. *Economic Journal, 120* (549), 1365-1398.

Bekhradnia, B. (2009). The Academic Experience of Students in English Universities. *The Higher Education Policy Institute*. United Kingdom.

Betts, J.R. & Morell, D. (1999). The Determinants of Undergraduate Grade Point Average: The Relative Importance of Family Background, High School Resources, and Peer Group Effects. *Journal of Human Resources, 34 (2)*: 268-293.

Bolander, S.F. (1973). Class size and levels of student motivation. *Journal of Experimental Design, 42*, 12–8.

Bound, J. Brown, C. & Mathiowetz, N. (2001). Measurement Error in Survey Data. in J. Heckman and E. Leamer (eds.) *Handbook of Econometrics*, Volume 5. Amsterdam: North Holland.

Bratti, M. (2002). Does the choice of university matter? A study of the difference across UK universities in life science students' degree performance. *Economics of Education Review, 21*, 431–443.

Cassady, J.C. (2001). Self-Reported GPA and SAT: A Methodological Note. *Practical Assessment, Research & Evaluation, 7* (12).

Chan, K.C., Shum, C. & Wright, D. J. (1997). Class Attendance and Student Performance in Principles of Finance. *Financial Practice and Education, 7*, 58-65.

Chevalier, A., Gibbons, S., Thorpe, A. Snell, M. & Hoskins, S. (2009). Students' academic self-perception. *Economics of Education Review, 28* (6), 716-727.

Clay, T. & Breslow, L. (2006). Why Students Don't Attend Class. *MIT Faculty Newsletter*, XVIII (4).

Cohn, E. & Johnson, E. (2006). Class attendance and performance in principles of economics. *Education Economics*, *14* (2): 211–233.

Cobb-Clark, D. & Schurer, S. (2011). Two Economists' Musings on the Stability of Locus of Control. *IZA DP* No. 5630.

Crede, M., Roch, S., & Kieszczynka, U. (2010). Class Attendance in College: A Meta-Analytic Review of The Relationship of Class Attendance With Grades and Student Characteristics. *Review of Educational Research, 80* (2), 272-295.

Cunha, F., Heckman, J.J., Lochner, L.J. & Masterov, D.V. (2006). Interpreting the evidence on life cycle skill formation. In E. A. Hanushek & F. Welch (Eds.), *Handbook of the Economics of Education, Chapter 12*, pp. 697–812. Amsterdam: North-Holland.

Deere, D. (1994). Correspondence: Should Class Attendance be Mandatory? *Journal of Economic Perspectives 8* (3): 210-211.

Devadoss, S. & Foltz, J. (1996). Evaluation of factors influencing student class attendance and performance. *American Journal of Agricultural Economics, 78*, 499-507.

Dobbins, G.H., Farh, J. & James D.W. (1993). The Influence of Self-Monitoring and Inflation of Grade-Point Averages for Research and Selection Purposes. *Journal of Applied Social Psychology*, 321-334.

Dolnicar, S. (2005). Should we still lecture or just post examination questions on the web? The nature of the shift towards pragmatism in undergraduate lecture attendance. *Quality in Higher Education, 11* (2), 103-115.

Dolton, P., Marcenaro, O. & Navarro, L. (2003). The Effective Use of Student Time: A Stochastic Frontier Production Function Case Study. *Economics of Education Review, 22* (6), 547-60.

Douglas, S. & Sulock, J. (1995). Estimating educational production functions with correction for drops. *Journal of Economic Education, 26*, 101-112.

Durden, G. C. & Ellis, L. V. (1995). The effect of attendance on student learning in Principles of Economics. *American Economic Review, 85*, 343-346.

Falchikov N. & Boud, D. (1989). Student Self-Assessment in Higher Education: a Meta Analysis. *Review of Educational Research, 59*, 395-430

Feldman, K.A. (1984). Class size and college students' evaluations of teachers and courses: a closer look. *Research in Higher Education, 21*, 45–91.

Freeman, Richard B. (1984). Longitudinal Analyses of the Effects of Trade Unions. *Journal of Labor Economics, 2* (1), 1-26.

Friedman, P., Rodriguez, F., & McComb, J. (2001). Why students do and do not attend classes: Myths and realities. *College Teaching, 49* (4), 124-134.

Foster, G. & Kinsler, J. (2011). Estimating Peer Spillovers in Australian Undergraduate Classrooms. *University of New South Wales*, mimeo.

Gendron, P., & Pieper, P. (2005). Does attendance matter? Evidence from an Ontario ITAL. *Discussion Paper, The Business School, Humber institute of Technology and Advanced Learning*. Toronto, Canada.

Gibbons, S. & Silva, O. (2007). Enjoyment, choice and achievement. *London School of Economics, Centre for Economic Performance*, mimeo.

Goulart, P. & Bedi, A.S. (2011). The Impact of Interest in School on Educational Success in Portugal. *IZA DP* No. 5462.

Griliches, Z., & Hausman, J.A. (1986). Errors in Variables in Panel Data. *Journal of Econometrics 31* (1), 93-118.

Grise, D. J., & Kennedy, A. M. (2003). Nonmajors' performance in biology: Effects of student based initiatives and class size. *Journal of College Science Teaching, 33* (2), 18-21.

Haley, M.R., Johnson, M.F. & McGee, M.K. (2010) A Framework for Reconsidering the Lake Wobegon Effect. *The Journal of Economic Education, 41* (2), 95-109

Hamermesh, D.S. (1989). Why Do Individual-Effects Models Perform So Poorly? The Case of Academic Salaries. *Southern Economic Journal, 56* (July), 39-45.

Heckman, J. J. (1981). Heterogeneity and state dependence, in Studies in Labour Markets, ed. S. Rosen, Chicago, Chicago Press.

Hopkins, D. S. P. (1990). The higher education production function: Theoretical foundations and empirical findings. In S. A. Hoenack & E. L. Collins (Eds.), The Economics of American Universities: Management, Operations, and Fiscal Environment (pp. 11-32). Albany: State University of New York Press.

Kottasz, R. (2005). Reasons for Student Non-Attendance at Lectures and Tutorials: an analysis. *Investigations in university teaching and learning, 2*, 5-16.

Lang, M., Joyce, A., Conaty, F. & Kelly, B. (2008). An Analysis of Factors Influencing the Attendance of First Year University Students. In Pieterick, J. et al. (eds), *Proceedings of European First Year Experience Conference*, Wolverhampton, UK, May 7-9, 141-147.

Marburger DR (2001). Absenteeism and Economic Exam Performance. *The Journal of Economic Education 32*, 99-109.

Martins, P. & Walker, I. (2006). Student Achievement and University Classes: Effects of Attendance, Size, Peers, and Teachers. *IZA Discussion Papers* 2490, Institute for the Study of Labor (IZA).

Massingham, P., & Herrington, T. (2006). Does attendance matter? An examination of student attitudes, participation, performance and attendance? *Journal of University Teaching and Learning Practice, 3* (2), 82-103.

Maxwell, Nan L. & Jane S. Lopus (1994). The Lake Wobegon Effect in Student Self-Reported Data. *American Economic Review Papers and Proceedings*, *84*, 201-205.

McConnell, C.R. & Sosin, K. (1984). Some determinants of student attitudes towards large classes. *Journal of Economic Education, 15*, 181–90.

Park, K.H. & Kerr, P. (1990). Determinants of Academic Performance: A Multinomial Logit Approach. *Journal of Economic Education, 21*, 101-111.

Rivkin, S.G., Hanushek, E.A. & Kain, J.F. (2005). Teachers, Schools and Academic Achievement. *Econometrica, 73* (2), 417-458.

Rodgers, J.R. (2002). Encouraging tutorial attendance at university did not improve performance. *Australian Economic Papers, September*, 255–266.

Romer, D. (1993). Do Students Go to Class? Should They? *Journal of Economic Perspectives, 7* (3), 167-74.

Romer, D. (1994). Response in 'Correspondence'. *Journal of Economic Perspectives, 8*, (3), 214-215.

Schmidt R.M. (1983). Who maximises what? A study of student time allocation. *The American Economic Review, 73*, 23-28.

Stanca, L. (2006). The effects of Attendance on Academic Performance: Panel Data Evidence for Introductory Microeconomics. *The Journal of Economic Education, 37* (3), 251-266.

Stephenson, K. & Deere, B. (1994). Correspondence. *Journal of Economic Perspectives, 8*, 207–208.

Stinebrickner, R. & Stinebrickner, T.R. (2008). The Causal Effect of Studying on Academic Performance. *The B.E. Journal of Economic Analysis & Policy, 8* (1), 1-53.

Swaffield, J.K. (2001). Does measurement error bias fixed-effects estimates of the union wage effect? *Oxford Bulletin of Economics and Statistics, 63* (4), 437-456.

Todd, P. & Wolpin, K.I. (2003). On the Specification and Estimation of the Production Function for Cognitive Achievement. *The Economic Journal, 113* (485), 13-33.

Zafar, B. (2011). How Do College Students Form Expectations? *Journal of Labor Economics, 29* (2), 301-348.

Appendix A: Representativeness of the Survey Data

	IUS: Wave 1 (Spring 2009)	**HEA (2008/09)** (Official Data)	**IUS: Wave 2** (Spring 2010)	**HEA (2009/10)** (Official Data)
Gender				
Male	35%	42%	35%	43%
Female	65%	58%	65%	57%
University				
DCU	6%	9%	6%	9%
NUIG	16%	16%	16%	16%
NUIM	10%	7%	10%	8%
TCD	28%	15%	26%	15%
UCC	14%	18%	15%	18%
UCD	21%	23%	22%	23%
UL	5%	11%	5%	12%
Subject				
Education	4%	5%	3%	4%
Humanities & Arts	22%	25%	22%	25%
Social Science	13%	6%	13%	7%
Business	10%	13%	8%	13%
Law	5%	7%	5%	6%
Science	16%	11%	20%	12%
Maths	2%	1%	2%	1%
Computing	3%	3%	3%	3%
Engineering	6%	8%	6%	8%
Agriculture	2%	2%	2%	2%
Health	13%	18%	14%	18%
Sport	0%	0%	0%	0%
Other	4%	2%	2%	2%

Appendix B: Fixed Effects Regressions Explaining Students' Average Grade at University

(Irish Universities Study: Spring 2009 and Spring 2010)

	(1) Interaction with year of study	(2) Interaction with prior achievement	(3) Interaction with additional study-hours	(4) Interaction with future-orientation
Lecture attendance	0.045	0.024	0.041	0.024
	(0.046)	(0.038)	(0.041)	(0.030)
Year of enrolment	-4.032	-1.199	0.000	0.000
	(4.405)	(0.000)	(4.238)	(4.200)
Age of student	-0.157	-0.149	-0.143	-0.109
	(0.948)	(0.953)	(0.952)	(0.949)
Family-income bracket[a]	-0.226	-0.231	-0.217	-0.248
	(0.361)	(0.364)	(0.362)	(0.360)
Study-time interval[b]	-0.286	-0.309	0.345	-0.309
	(0.283)	(0.284)	(0.846)	(0.282)
Willing to take risks	-0.296	-0.310	-0.309	-0.295
	(0.596)	(0.601)	(0.595)	(0.597)
Future-orientation	0.816	0.851	0.836	2.992
	(0.583)	(0.587)	(0.583)	(1.999)
Interaction term[c]	-0.009	-0.000	-0.008	-0.027
	(0.012)	(0.000)	(0.009)	(0.024)
Period 2	9.271***	9.187***	9.032***	9.196***
	(3.121)	(3.103)	(3.148)	(3.117)
Constant	67.541**	69.224***	67.460**	69.123***
	(26.495)	(26.319)	(26.544)	(26.222)
Observations	1,285	1,285	1,285	1,285
R-squared	0.221	0.220	0.221	0.222

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Results from fixed effects regression; including and a class-room effect based on a three-way interaction between university-affiliation, subject area and year of enrolment (not shown). All standard errors in the fixed effects regression are robust. Where they apply, control variables for missing value are not shown above. Outliers and missing values are adjusted for independent variables. Attitude to risk and future-orientation are standardized using z-scores.

[a]Income-brackets are in categories of €20,000 [b]Additional study is extra hours of personal study [c]Interaction terms are lecture attendance multiplied by the following: year of study, prior achievement, additional study-hours, future-orientation.

Appendix C: Fixed Effects Regressions Explaining Students' Average Grade at University

(Irish Universities Study: Spring 2009 and Spring 2010)

	(1)	(2)	(3)
	2:1 Award Status	1:1 Award Status	Grade > 75%
Lecture attendance	-0.021	-0.071	0.060
	(0.032)	(0.091)	(0.103)
Year of enrolment	3.642	5.620	0.000
	(9.821)	(0.000)	(0.000)
Age of student	0.496	-1.625	8.383**
	(0.997)	(2.100)	(3.695)
Family-income bracket[a]	-0.021	-0.029	-1.107
	(0.287)	(0.496)	(1.027)
Study-time interval[b]	0.266	-0.142	-0.818
	(0.309)	(0.720)	(1.237)
Willing to take risks	-0.514	-1.086	1.255
	(0.768)	(1.587)	(2.833)
Future-orientation	0.107	-0.493	-0.610
	(0.684)	(1.377)	(2.771)
Period 2	1.826	1.613	-2.408
	(6.809)	(2.824)	(4.491)
Constant	45.113**	76.434	-104.116
	(20.567)	(46.532)	(113.744)
Observations	591	516	386
R-squared	0.310	0.302	0.668

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Results from fixed effects regression; including and a class-room effect based on a three-way interaction between university-affiliation, subject area and year of enrolment (not shown). All standard errors in the fixed effects regression are robust. Where they apply, control variables for missing value are not shown above. Outliers and missing values are adjusted for independent variables. Attitude to risk and future-orientation are standardized using z-scores.

[a]Income-brackets are in categories of €20,000 [b]Additional study is extra hours of personal study

Appendix D: Fixed Effects Regressions Explaining Students' Average Grade at University

(Irish Universities Study: Spring 2009 and Spring 2010)

	(1) >50% Attendance	(2) >80% Attendance	(3) > 90% Attendance
Lecture attendance*	0.050	0.545	-0.657
	(1.537)	(0.955)	(0.650)
Year of enrolment	-4.748	0.000	-4.932
	(4.280)	(0.000)	(4.332)
Age of student	-0.151	-0.154	-0.162
	(0.953)	(0.954)	(0.953)
Family-income bracket[a]	-0.228	-0.225	-0.232
	(0.362)	(0.356)	(0.357)
Study-time interval[b]	-0.290	-0.297	-0.287
	(0.280)	(0.282)	(0.280)
Willing to take risks	-0.333	-0.329	-0.307
	(0.599)	(0.600)	(0.600)
Future-orientation	0.864	0.833	0.896
	(0.579)	(0.598)	(0.578)
Period 2	9.256***	9.209***	9.339***
	(3.108)	(3.090)	(3.128)
Constant	70.259***	70.653***	71.049***
	(26.327)	(26.592)	(26.421)
Observations	1,285	1,285	1,285
R-squared	0.220	0.220	0.221

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

*Lecture attendance categorised as follows:
(i) 0 = attend less than 50% of lectures/1 = attend more than 50% of lectures
(ii) 0 = attend less than 80% of lectures/1 = attend more than 80% of lectures
(iii) 0 = attend less than 90% of lectures/1 = attend more than 90% of lectures

Note: Results from fixed effects regression; including and a class-room effect based on a three-way interaction between university-affiliation, subject area and year of enrolment (not shown). All standard errors in the fixed effects regression are robust. Where they apply, control variables for missing value are not shown above. Outliers and missing values are adjusted for independent variables. Attitude to risk and future-orientation are standardized using z-scores.

[a]Income-brackets are in categories of €20,000 [b]Additional study is extra hours of personal study