Lancaster University Management School
Working Paper
2009/018

Estimating the Production Function of University Students

Kwok Tong Soo

The Department of Economics
Lancaster University Management School
Lancaster LA1 4YX
UK

The LUMS Working Papers series can be accessed at http://www.lums.lancs.ac.uk/publications/
LUMS home page: http://www.lums.lancs.ac.uk/

Estimating the production function of university students*

KWOK TONG SOO

Department of Economics,
Lancaster University Management School,
Lancaster LA1 4YX, U.K.
(e-mail: k.soo@lancaster.ac.uk)

Abstract

This paper estimates the production function for university students in English universities. Taking as the output the quality of a university degree and the dropout rate, we use as inputs teaching quality and quantity, entry qualifications, and the effort level. Our results uncover new findings regarding the importance of each of these elements in university performance. In particular, we find that the quality of teaching and entry qualifications affect degree performance, but not the number of hours of teaching or private study. Controlling for unobserved ability through a 2SLS/GMM estimator suggests that entry scores have no additional impact on degree performance beyond its role as a measure of student ability.

JEL Classification: I21, D24.
Keywords: Production function estimation; Higher education; Instrumental variables.

* Thanks to Bahram Bekhradnia for providing the data from HEPI on the student experience in English universities, and to Caroline Elliott, Geraint Johnes, Jill Johnes and Jim Taylor for helpful suggestions. The author is responsible for any errors and omissions.

I. Introduction

This paper seeks to discover how important is teaching to the performance of university students. The experience of university students has become a key policy issue in the UK, as universities, under pressure from the Research Asssessment Exercise, may have an incentive to de-emphasise teaching in favour of research. The student experience has been the subject of several important pieces of research in the UK; the National Student Survey (NSS), which has run annually since 2005, seeks to document the degree of student satisfaction with teaching provision at their university. Surridge (2006, 2007) summarises the findings of the 2005 and 2006 surveys. In 2006 and 2007 the Higher Education Policy Institute (HEPI) has conducted surveys of the student experience in higher education (Bekhradnia *et al,* 2006, Sastry and Bekhradnia, 2007).

In this paper we use data from the NSS and HEPI surveys of 2006 and 2007 to investigate the determinants of the performance of university students. This combined dataset allows us to explore the relative importance of teaching, private study, or prior education in determining student performance. By student performance, we mean the degree class obtained, and the percentage of dropouts. Our results may be expected to have implications for both policymakers and university administrators. For policymakers, information on the relative importance of each of these elements may help to direct funds and other resources towards the most beneficial way of improving university performance. For university administrators, it allows them to decide whether a concerted effort to improve teaching quality and quantity is the best strategy for improving students' performance.

Our main finding is that prior education and the quality of teaching are the most significant determinants of university degree performance measured as degree classification; the percentage of dropouts is influenced only by prior education. Teaching hours and hours of private study play statistically insignificant roles on degree performance. Controlling for unobserved student ability using a 2SLS/GMM approach indicates that it is student ability that drives the significance of prior education; once student ability has been controlled for, prior education no longer has any significant impact on degree outcomes. These results are robust to the inclusion of additional control variables. Dividing the sample into pre-92 and post-92 universities indicates that it is in pre-92 universities that teaching quality is important.

As discussed in greater detail in the conclusion, this suggests that policymakers may wish to improve secondary school provision especially to those from lower socio-economic groups. It also suggests that universities may be best off if they seek to improve the overall reputation of their institution, as this may improve the quality of applications, as well as to find ways of improving the quality (but not necessarily the quantity) of teaching provided.

There has been much prior research on the determinants of university student performance, measured both in terms of degree performance and dropouts. Much of this work is summarised in Naylor and Smith (2004). The majority of these studies make use of student-level data, enabling the researchers to address issues relating to personal characteristics and prior education, as well as differences across universities and subjects. However, most of these studies do not consider what happens once students arrive at university; a recent exception is Arulampalam *et al* (2007), who investigate the impact of absence from class on student performance among Economics students at a UK university.

Other work investigating the impact of class attendance on performance include Stanca (2006) on students taking a microeconomics module at an Italian university, and Martins and Walker (2005) on Economics students at a UK university. This work relates to the literature on educational production such as Lazear (2001) and Todd and Wolpin (2003). The present paper differs from previous literature by employing data across subjects and institutions. Whilst this prevents us from exploring the individual student characteristics that may influence performance, it allows us to draw more general conclusions based on analysis across different subjects and institutions, thus introducing another dimension to the empirical literature.

In terms of methods, this paper follows the method used in the economic growth literature by Mankiw *et al* (1992) and Islam (1995) by estimating a Cobb-Douglas production function to identify the relative contributions of the different inputs into the production of university degrees. This is distinct from the alternative levels accounting approach of Hall and Jones (1999), which imposes the assumption of the relative contributions of the different inputs. Since this approach has not to the best of our knowledge been used in the study of the contributors of student performance, one of the main objectives of this paper is precisely to uncover the magnitude of these different contributions, hence our choice of the Mankiw *et al* (1992) approach.

The next section describes the methods used in this paper. This is followed in Section 3 by a description of the data. Section 4 details the results, and the final section concludes.

II. Methods

Consider a Cobb-Douglas production function of the degree performance of students in a subject *i* in university *j* in year *t* as follows:

$$D_{ijt} = A_{ijt} E_{ijt}^{\alpha} S_{ijt}^{\beta} C_{ijt}^{\gamma} Q_{ijt}^{\delta} \tag{1}$$

Where D is the degree performance, E is the entry score, S is the hours of study, C is the number of classroom hours, Q is the quality of teaching, and A is a measure of the productivity of the students. The parameters α, β γ and δ are the contributions of each of the inputs into the degree result. Taking logs of equation (1) gives an expression which is linear in parameters:

$$\ln D_{ijt} = \ln A_{ijt} + \alpha \ln E_{ijt} + \beta \ln S_{ijt} + \gamma \ln C_{ijt} + \delta \ln Q_{ijt} \tag{2}$$

Suppose that student productivity A depends on some other factors so that $\ln A_{ijt} = \lambda_i + \eta_j + \varphi_t + \nu_{ijt}$, where λ_i is some subject-specific component, η_j is some university-specific component, φ_t is some year-specific component, and υ_{ijt} is some random component. Then equation (2) becomes:

$$\text{Ln}\, D_{ijt} = \lambda_i + \eta_j + \varphi_t + \alpha \ln E_{ijt} + \beta \ln S_{ijt} + \gamma \ln C_{ijt} + \delta \ln Q_{ijt} + \nu_{ijt} \tag{3}$$

Equation (3) is our basic empirical specification. If υ_{ijt} is uncorrelated with the other components on the RHS of equation (3), then it can be estimated using conventional fixed effects methods, treating the subject, university and year effects as fixed. The subject fixed effects control for the fact that subjects may differ in their characteristics; some subjects may be more difficult to study for, or may require more contact hours. The university fixed effects control for differences across universities; a more prestigious university may have higher entry requirements across the range of subjects. The year fixed effects control for any shock that impacts uniformly across universities and subjects, for example changes in government rules concerning university fees. The subject, university and year fixed effects mean that the coefficients of interest are therefore identified by changes within each subject-university pair.

The entry score captures the student's previous training, which reflects the student's past effort level and ability. It also partly captures the student's family background – there is evidence to suggest that students from higher social classes achieve better educational performance, which may be due to private schooling or other forms of advantages that greater financial resources may bring (see Chowdry *et al*, 2008).

A good measure of academic contact should capture the effectiveness of classroom hours. Therefore, our measure should not only include the number of classroom hours, but also some measure of how effective these hours are. As we detail below, our dataset provides data for both the quantity and quality of academic contact.

A similar argument can be made about the effectiveness of private study, but our dataset does not contain this information. The effectiveness of private study may differ across subject areas and/or universities, in which case it would be controlled for by the subject and university fixed effects. Otherwise, it would be captured in the error term[1]. The error term therefore serves a similar function to the Solow residual – it captures all other factors which influence degree performance but which we do not control for in our model. However, we are confident that by controlling for the students' past performance, their effort levels, the academic support that they receive, and the subject, university and year effects, we have controlled for the major factors that influence degree performance.

Unmeasured student ability may pose a serious omitted variable problem, since student ability is likely to be correlated with both entry scores and degree outcomes. If ability has no impact on degree performance beyond its impact on entry scores, then the coefficient on entry scores would be unbiased; however, if this is not the case, then conventional estimation methods lead to biased results. One solution to this problem is to use an instrumental-variables or 2SLS approach (see Angrist and Krueger, 2001 for an overview). Here, we use this approach in the context of an efficient feasible two-step GMM estimator following Baum *et al* (2007).

We use as instruments for entry scores in a department in a university, the entry scores for other departments in the university, and the entry scores for the same department in other universities. A good instrument is both highly correlated with the instrumented variable, and

[1] To the extent that the effectiveness of private study is correlated with the hours of private study (or with any of the other variables), our specification would suffer from omitted variable bias. The results should therefore be read with this in mind.

not directly correlated with the dependent variable. The instruments we use should satisfy both conditions, since a university with a strong reputation may be expected to attract good students across the range of subjects, and some subjects may attract better students than other subjects. At the same time, the ability of students in other subjects should have no impact on the degree performance in a particular subject, unless there are very strong peer-group effects. We also report tests for the validity of our chosen instruments.

III. Data

The data used in this paper comes from two primary sources: the surveys conducted by HEPI, and the National Student Surveys.

The HEPI surveys

For data on the student experience in English universities, we use data from the two HEPI reports by Bekhradnia *et al* (2006) and Sastry and Bekhradnia (2007). In these studies, the authors conducted surveys of first and second year students in English universities. In each year there were over 14,000 respondents from a sample of over 23,000. These respondents are distributed across 169 universities and all subject areas; see Bekhradnia *et al* (2006) and Sastry and Bekhradnia (2007) for details of the sample.

The surveys ask students questions regarding the workload that they experience, including the number of teaching hours, private study, outside employment, use of specialist equipment, and the level of satisfaction. Most questions (and all of those used in the present paper) were repeated in both surveys. Due to the changes in some questions and to a different weighting system for constructing descriptive statistics, Sastry and Bekhradnia (2007) caution against making comparisons between the results of the two surveys. They do however report that the results of the two surveys reveal only very small differences which they attribute to random variation or changes in the approach.

In the present study we use the results of both surveys, exploiting the similarity of results between them. We aggregate the individual responses to 15 broad subjects, which correspond closely to the JACS (Joint Academic Coding System) 19-subject level used by HESA

(Higher Education Statistics Agency) and other government bodies. The difference between our subjects and the JACS 19-subject level is that we omit four categories: first, the category of combined studies since all students are classified under one of the other categories; second and third, we omit students in veterinary science and agriculture since the survey combines these two categories, and there is no way to disentangle the them; and fourth, we omit students in medicine and dentistry since these degrees almost always do not follow the classification system used in other subjects. In practice this means deleting 828 students in 2006 and 689 students in 2007. To prevent outliers from unduly influencing the results, we also delete all university-subject observations for which the number of respondents was less than 5. This resulted in the deletion of a further 955 students in 2006 and 747 students in 2007.

As a result of these processes, our sample contains 1644 university-subject-year observations for the results of the surveys conducted by HEPI. After allowing for missing observations from our other data sources, our final sample contains 1312 observations: 630 in 2006 and 682 in 2007, from a total of 108 universities. From these HEPI reports we obtain the following variables. First, we obtain the number of teaching hours, given by the average number of hours attended[2], the variable C above. Second, we obtain the average number of hours of private study, the variable S above.

The NSS

Our second main data source is the National Student Survey (NSS), conducted by HEFCE (Higher Education Funding Council for England) in collaboration with the NUS (National Union of Students). This annual survey, conducted since 2005, asks students a set of questions relating to teaching, assessment and feedback, academic support, organisation and management, learning resources, and personal development. The question we use from the NSS is Question 22: “Overall, I am satisfied with the quality of my course”. The response is on a five-point scale, with higher values representing better quality. We use the average value as a summary measure of the quality of teaching provided by the institution in that subject, the variable Q above.

[2] The data also includes the number of timetabled teaching hours; we use the number of hours attended as this is a better reflection of how much academic contact students have.

The NSS data on the HEFCE website includes information on the median entry scores of students by university and subject. English students entering English universities almost always take the GCE A-level exam, taking three or four subjects. These exams are graded from A to E, with an A being worth 120 points, and each lower grade being worth 20 points less than the grade above, so that the lowest grade E is worth 40 points[3]. We use this median entry score as variable E, the measure of student entry grades.

The NSS data also includes information on the final degree outcomes by university and subject. In England, almost all degrees are classified as first class honours, upper second class honours, lower second class honours, third class honours, or ordinary or unclassified degrees. The NSS data gives information on the percentage of students that achieve each classification level. From this data we construct two measures of student achievement, variable D above. First, we calculate the percentage of students who achieve a "good degree", by which is meant upper second class honours or first class honours. As an alternative measure of student achievement, we use the average degree classification attained. This is calculated using a value of 5 for a first class degree, 4 for upper second class, 3 for lower second class, 2 for third class, and 1 for ordinary or unclassified degrees. The natural log of each of these measures is used as the dependent variable in our regressions, D.

As an alternative measure of student performance, we also obtain from the NSS data the percentage of students who leave their institution of higher education without an award, and the percentage of dormant students (a dormant student is defined as one who has ceased studying but has not formally de-registered)[4]. The natural log of the sum of these two percentages is used as an alternative dependent variable, D. This output may be thought of as a "bad", so that larger amounts of inputs may be expected to have a negative effect on this output.

Universities in the UK are prevented from giving good degrees to students who do not deserve them, by a system of external examiners who moderate university degrees. As a result of this moderation, degree results are not marked according to any distributional requirements. Also, whilst there is concern over grade inflation over time (see the survey in

[3] English pre-university education lasts two years, with the A-level exams occurring at the end of the second year. At the end of the first year, students take the AS-level exams, which can count towards their total score, with each AS-level grade counting for half of the equivalent A-level grade.
[4] Data is not separately available for dormant students and students who leave without an award.

Johnes, 2004), it may be argued that this is much less of a problem for the two year sample used here than for one with a longer time dimension.

Descriptive Statistics

Table 1 presents the descriptive statistics of the dataset, divided between 2006 and 2007. The variables have very similar means and standard deviations across both years, reinforcing the conclusion in Sastry and Bekhradnia (2007) that the results of the two HEPI surveys are very similar. The average number of timetabled teaching hours is between 13 and 14 hours per week, of which on average one hour is missed each week. Students then average between 12 and 13 hours of private study each week. The average of the median A-level score is between 300 and 320, while over half of all students obtain a second-class (upper or lower) degree. The average dropout rate is between 6 and 8 percent. Overall student satisfaction appears to be quite high, at approximately 4.0 on a 5-point scale.

Table 2 reports the simple correlations between the variables in the dataset, again divided between 2006 and 2007. Whilst this is no substitute for our econometric model outlined above, it does give some suggestive information. In both years average degree performance and the percentage of good degrees are highly positively correlated with each other, and both are highly negatively correlated with the percentage of dropouts. In both years A-level scores and student satisfaction are highly positively correlated with the percentage of good degrees and the average degree result, and highly negatively correlated with the percentage of dropouts. On the other hand, the average number of teaching hours and private study appear to only be weakly correlated with degree performance and dropouts. In our econometric analysis we will seek to explore whether such patterns in the data persist controlling for other factors.

A final note on the construction of the dataset. Although the data on entry qualifications, survey results, and degree performance are collected for the same years, they do not relate to the same students. Nevertheless, Tables 1 and 2 indicate that there is a large amount of persistence in all of the variables used in the analysis. This is confirmed in Table 3, which reports the Spearman rank correlation between the variables in 2006 and 2007, together with a test for the independence of the two years. As can be seen, for each variable the correlation is very high, and data for the two years are never found to be independent of each other. This again reinforces the idea that the data are highly persistent, and therefore that our results are similar to what they would have been had the data been for the same students. In the robustness section below we also use data on different years to check how reasonable this assumption is.

IV. Results

The results of estimating equation (3) using both OLS and 2SLS/GMM are reported in Table 4. Standard errors are clustered by university to allow for possible heteroskedasticity and within-university correlations in the error term. Columns (1) and (2) report the results using the log of the average degree classification as the dependent variable, columns (3) and (4) report the results using the log of the percentage of good degrees achieved as the dependent variable, and columns (5) and (6) report the results using the log of the percentage of dropouts as the dependent variable. All regressions are run with university, subject and year fixed effects.

We find that the results are very similar for both OLS and 2SLS/GMM for all estimated models, with one major exception: the A-level entry score, which is always a highly significant predictor of degree performance in the OLS models, is never significant in the 2SLS/GMM models. This may be regarded as evidence that unobserved ability leads to omitted variable bias in the OLS results, and that it is this unobserved ability that drives both A-level and degree performance. Nevertheless, the Hausman test shows that there is no significant difference between the OLS and 2SLS/GMM results. The results of the Hansen C test suggest that A-level score is correlated with the error term in the average degree regression, but is not correlated with the error term in the good degree and dropouts regressions. The Hansen J test of overidentification suggests that our instruments are jointly valid for all specifications, while the instruments easily pass the underidentification and weak identification tests. These results suggest that our identification strategy in the 2SLS/GMM model is appropriate.

Of the other variables included in the regressions, student satisfaction is positively and significantly related to degree performance in both the OLS and 2SLS/GMM models, and is negatively but not significantly related to the dropout rate. This therefore provides evidence in support of the idea that better teaching quality has a positive impact on degree performance. On the other hand, the number of hours attended is never significantly related to degree performance in any specification. Similarly, the number of hours of private study is never significantly related to degree performance in the 2SLS/GMM models. There is evidence of better degree performance in 2007 than in 2006, which may indicate grade inflation, but there is also a larger percentage of dropouts in 2007.

Given the specification of our econometric model, we can test whether the production function of university students exhibits constant returns to scale or not. This is equivalent to testing whether the coefficients on entry grades, teaching hours, private study and satisfaction sum to one or not (negative one in the case of dropouts). Table 4 reports the results of this test. For degree performance measured using either average degrees or percentage of good degrees, constant returns to scale is rejected in favour of decreasing returns to scale, whereas we cannot reject constant returns to scale for dropouts.

We can explore how well our model fits the data by comparing the predicted values from the model with the actual values of the dependent variables. Figure 1 plots predicted versus actual values for the 2SLS/GMM regressions in Table 4. As can be seen, in all three cases most of the observations have fitted values which are close to the actual values. This suggests that our econometric model enables us to capture most of the variation in the dependent variables.

Finally, we make some brief comments on the (unreported) coefficients of the subject and university fixed effects, for the 2SLS/GMM models. Students in Creative Arts and Design, and Mass Communication, perform better than students in other subjects, whereas students in Architecture, Business Studies and Law perform worse, when measured using either good degrees or average degrees. Amongst dropouts, Architecture students have higher dropout rates than other students, whereas students in Education have lower dropout rates than students in other subjects. Many of the university fixed effects are significantly different from each other. There is some evidence to suggest that universities in the Russell Group and the 1994 Group (two groups of research-led universities) have a higher percentage of good degrees and better average degrees, and lower dropout rates, than other universities in the sample.

Robustness

We perform four robustness checks on our results. First, we use different measures of teaching hours and student satisfaction, and see whether or not this changes the results. Second, we include additional variables in the regression to control for possible omitted variable bias. Third, we divide the sample into institutions that attained university status as a result of the Further and Higher Education Act of 1992 (known as post-92 universities; these are primarily former polytechnics) and those that had university status before this Act (pre-92 universities). A final robustness check partially addresses the timing of the data as noted in Section 3 by using data for different years for different variables.[5]

Although we have argued above that our use of the average number of classroom hours attended is the more appropriate measure of the amount of academic contact students get, to check the robustness of this, we use instead the average number of timetabled hours. In addition, instead of using the measure of student satisfaction from the NSS, we use a measure of student satisfaction obtained from the HEPI surveys; this is the response to the question "I am satisfied with the amount of timetabled sessions I have had this year". The response to this question is on a five-point scale. We use the logs of both variables[6].

The results of the regressions using these two alternative measures are reported in Table 5. The alternative definitions of the variables for student satisfaction and contact hours results in student satisfaction being an insignificant predictor of degree performance; this is true for both OLS and 2SLS/GMM. One possible reason for this could be that the NSS measure of student satisfaction better captures teaching quality than the measure from HEPI, which captures student satisfaction with teaching quantity. The other results remain largely unchanged, and the instruments for the 2SLS/GMM models remain valid with the alternative variables used in Table 5.

[5] An additional robustness check that was performed was to estimate the model using the sum of private study and classroom hours as one explanatory variable, on the assumption that it is the total hours of study that matters for degree performance. This variable was found to be statistically insignificant in all specifications, and did not change the other results of the model.

[6] A further alternative measure of student satisfaction is the response to another HEPI survey question: "To what extent do you feel you have received value for money on your present course". Responses to this question are on a five-point scale. Using the log of this variable as our measure of student satisfaction yields very similar results to those using the measure used in the text.

There remains the possibility of omitted variable bias in our results. Our next check therefore includes two additional variables in order to address this concern. Our two additional variables are obtained from the HEPI surveys. The first variable is the number of assignments per term; this variable may capture additional aspects of students' effort levels. The second variable is the number of hours per week students spend on paid employment unrelated to their course; this may capture the other responsibilities that students have which may detract from the effort they put into their studies. We use the logs of both variables.

Table 6 reports the results of regressions controlling for these additional variables. Neither of the additional variables has any significant impact on degree performance or dropout rates. As a result, the impact of the other variables on the dependent variables is very similar to Table 4. Once again students' A-level scores are significant in the OLS models but not in the 2SLS/GMM models. Student satisfaction remains a significant predictor of degree performance but not of dropouts.

Our regression results show the average effects of the explanatory variables on the dependent variables. It is possible that the explanatory variables may have different impacts on different groups of universities. To explore this possibility, we divide the sample into pre-92 and post-92 universities. There are a total of 45 pre-92 and 63 post-92 universities in our sample. The results of dividing the sample in this way are reported in Table 7, where Panel A refers to pre-92 universities, and Panel B refers to post-92 universities.

The results are quite striking. For post-92 universities, entry scores are a highly significant predictor of degree performance, but only when OLS is used; using 2SLS/GMM, none of the variables used have any significant explanatory power. On the other hand, for pre-92 universities, entry scores and student satisfaction are highly significant predictors of degree performance using any of the three dependent variables as measures of performance (entry scores lose significance when endogeneity is controlled for using 2SLS/GMM). Better entry scores and greater satisfaction with teaching increases degree performance and reduces dropouts. Also, for pre-92 universities, the number of classroom hours attended is a statistically significant explanatory variable for both the percentage of good degrees and the dropout rate. However, the sign of the coefficient on this variable is unexpected: more hours attended appears to reduce the percentage of good degrees and increases the percentage of dropouts. This suggests that more teaching hours may not be the optimal strategy for universities to pursue if they seek to improve student performance.

To investigate the impact of the timing of our data on our results, we use data on entry scores of entrants from the previous year (2005 entry scores for the 2006 survey, 2006 entry scores for the 2007 survey) and degree performance scores from the next year (2007 degree performance for the 2006 survey, 2008 degree performance for the 2007 survey). Whilst this still does not use data for the same students since most 2005 entrants would only graduate in 2008, it is the best available data, and at least narrows the time difference between cohorts and may be used to check the validity of the main results discussed above.

The results for this robustness check are reported in Table 8, where there are fewer observations than in the previous samples because some observations are lost when matching across years. Overall the results are broadly similar to the main results. For both good degrees and average degrees, entry scores are the only significant influence on degree performance, and this is true even after controlling for unobserved ability using 2SLS/GMM estimation. This suggests that the better time matching of the sample enables us to identify the (significant) impact of entry scores on degree performance independently of student ability. The results for dropouts is somewhat different, with entry scores playing no significant impact once endogeneity is controlled for, but the number of attended hours playing a positive and significant role on dropout rates. This latter result is similar to those for the subsample of pre-92 universities in Table 7.

V. Conclusions

This paper seeks to uncover the main determinants of university degree performance across English universities. In order to do so, a production function for university degrees is estimated, using as inputs entry qualifications, teaching quality and quantity, and private study. The approach follows that used in the economic growth literature to uncover the determinants of income levels. This is applied to data on 108 universities across 15 subject areas for 2006 and 2007. The estimation methods used are OLS and 2SLS/GMM. Using OLS, we find that entry qualifications have the largest and by far the most significant impact on degree results and dropout rates. However, once we control for unobserved ability using

2SLS/GMM, entry qualifications fail to have any impact on degree results and dropout rates. Once unobserved ability is controlled for, student satisfaction is the only significant predictor of degree performance. Dividing the sample into pre-92 and post-92 universities suggests that most of the results are driven by pre-92 universities, where student performance is positively related to student satisfaction, but negatively related to the number of classroom hours. These effects do not exist for post-92 universities.

Our results suggest that a student's ability as captured by his/her past performance is the best predictor of his/her future performance. Recent evidence by Chowdry *et al* (2008) suggests that students from lower socio-economic groups are less likely to enter higher education in the UK than those from higher socio-economic groups. Students from lower socio-economic groups are also less likely to enter better institutions. However, this difference is because those from lower socio-economic groups do not perform as well in secondary school. There is also evidence of low if not decreasing intergenerational social mobility as documented by Blanden and Machin (2007) and the continuing high return to university education as documented by Machin and McNally (2007). It therefore appears that the fact that students from low socio-economic groups have weaker secondary school performance, has a negative impact on the universities they attend, their degree performance, and hence their performance in the labour market, leading to lower social mobility. Calls for government policies to improve the secondary school performance of students from lower socio-economic groups would appear to be well-placed.

For university administrators, our results suggest that better teaching quality has a positive effect on degree results and so may be a good path for universities to take. Increasing the number of teaching hours does not have a positive effect on student performance and may even decrease it in pre-92 universities. Evidence from Elliott and Soo (2008) suggests that students are attracted to universities with better reputations and higher rankings in league tables. Efforts to improve a university's reputation so that better students are attracted to the university may therefore prove to be a viable strategy for improving student performance at one's own institution; this need not represent a zero-sum-game for the UK higher education system as a whole if it increases the attractiveness of UK universities to students from abroad.

References

Angrist, J. D. and Krueger, A. B. (2001). 'Instrumental variables and the search for identification: From supply and demand to natural experiments', *Journal of Economic Perspectives,* Vol. 15, pp. 69-85.

Arulampalam, W., Naylor, R. A. and Smith, J. (2007). 'Am I missing something? The effects of absence from class on student performance', Economic Research Paper No. 820, University of Warwick.

Baum, C. F., Schaffer, M. E. and Stillman, S. (2007). 'Enhanced routines for instrumental variables/GMM estimation and testing', *Stata Journal,* Vol. 7, pp. 465-506.

Bekhradnia, B., Whitnall. C. and Sastry, T. (2006). 'The academic experience of students in English universities', Higher Education Policy Institute Report.

Blanden, J. and Machin, S. (2007). 'Recent changes in intergenerational mobility in Britain', report for the Sutton Trust.

Chowdry, H., Crawford, C., Dearden, L., Goodman, A. and Vignoles, A. (2008). 'Understanding the determinants of participation in higher education and the quality of institute attended: Analysis using administrative data', mimeo, Institute for Fiscal Studies.

Elliott, C. and Soo, K. T. (2008). 'Does price matter? Overseas students in UK higher education', Working Paper No. 2008/017, Lancaster University Management School.

Hall, R. E. and Jones, C. I. (1999). 'Why do some countries produce so much more output per worker than others?' *Quarterly Journal of Economics,* Vol. 114, pp. 83-116.

Islam, N. (1995). 'Growth empirics: A panel data approach', *Quarterly Journal of Economics,* Vol. 110, pp. 1127-1170.

Johnes, G. (2004). 'Standards and grade inflation', in G. Johnes and J. Johnes (eds.) *International Handbook on the Economics of Education*, Edward Elgar, Cheltenham, UK.

Lazear, E. P. (2001). 'Educational production', *Quarterly Journal of Economics,* Vol. 116, pp. 777-803.

Machin, S. and McNally S. (2007). 'Tertiary education systems and labour markets', report for the OECD.

Mankiw, N. G., Romer, D. and Weil, D. N. (1992). 'A contribution to the empirics of economic growth', *Quarterly Journal of Economics,* Vol. 107, pp. 407-437.

Martins, P. and Walker, I. (2005). 'Student achievement and university classes: Effects of Attendance, Size, Peers and Teachers', Discussion Paper No. 2490, IZA.

Naylor, R. A. and Smith, J. (2004). 'Determinants of educational success in higher education', in G. Johnes and J. Johnes (eds.) *International Handbook on the Economics of Education*, Edward Elgar, Cheltenham, UK.

Sastry, T. and Bekhradnia, B. (2007). 'The academic experience of students in English universities', Higher Education Policy Institute Report.

Stanca, L. (2006). 'The effects of attendance on academic performance: Panel data evidence for Introductory Microeconomics', *Journal of Economic Education,* Vol. 37, pp.251-266.

Surridge, P. (2006). 'The National Student Survey 2005: Findings', HEFCE Report.

Surridge, P. (2007). 'The National Student Survey 2006: Full Report', HEFCE Report.

Todd, P. E. and Wolpin, K. I. (2003). 'Towards a unified approach for modelling the production function for cognitive achievement', *Economic Journal,* Vol. 113, pp. 3-33.

TABLE 1
Descriptive statistics.

Variable	*2006 (N = 631)* Mean	Std. Dev.	*2007 (N = 682)* Mean	Std. Dev.
Timetabled hours	13.25	3.66	13.71	3.69
Attended hours	12.23	3.48	12.47	3.51
Private study	12.70	3.49	12.18	3.66
First class (%)	10.99	6.97	12.42	7.69
Upper second class (%)	47.70	11.61	48.17	12.56
Lower second class (%)	33.14	10.26	31.59	11.09
Third class (%)	6.80	5.63	6.55	5.36
Ordinary degree (%)	0.99	3.43	0.88	2.57
Unclassified degree (%)	0.38	1.94	0.32	1.86
Dropouts (%)	6.13	3.17	7.85	3.84
Median A-level	306.33	76.22	316.21	83.77
Overall satisfaction (1-5)	4.00	0.26	4.03	0.25

Notes: In 2006 there were only 630 observations for the median A-level scores.

TABLE 2
Correlations between variables.

2006 (N = 630)

	Average degree	*Good degree*	*Dropout*	*Time-tabled*	*Attended hours*	*Private study*	*Median A-level*	*Satis-faction*
Average degree	1.0000							
Good degree	0.8999	1.0000						
Dropouts	-0.3488	-0.4057	1.0000					
Timetabled	0.0616	-0.0327	-0.0514	1.0000				
Attended hours	0.0649	-0.0214	-0.0515	0.9734	1.0000			
Private study	0.0304	0.0862	-0.0233	-0.1327	-0.1059	1.0000		
Median A-level	0.6125	0.6381	-0.5122	0.0861	0.0717	0.1007	1.0000	
Satisfaction	0.2493	0.2468	-0.2186	0.0321	0.0579	0.0330	0.3495	1.0000

2007 (N = 682)

	Average degree	*Good degree*	*Dropout*	*Time-tabled*	*Attended hours*	*Private study*	*Median A-level*	*Satis-faction*
Average degree	1.0000							
Good degree	0.9070	1.0000						
Dropouts	-0.4990	-0.5258	1.0000					
Timetabled	0.0691	-0.0717	0.0043	1.0000				
Attended hours	0.0562	-0.0696	0.0049	0.9720	1.0000			
Private study	0.1126	0.1682	-0.0850	-0.0442	-0.0231	1.0000		
Median A-level	0.6841	0.7022	-0.6279	0.0846	0.0720	0.2236	1.0000	
Satisfaction	0.3000	0.2985	-0.3785	0.0519	0.0585	0.1434	0.4333	1.0000

Notes: Average degree is calculated as the average degree classification. Good degree is the percentage of students with first class or upper second class honours.

TABLE 3
Spearman rank correlations between 2006 and 2007 for each variable.

Variable	*Average degree*	*Good degree*	*Dropouts*	*Median A-level*	*Attended hours*	*Satisfaction*	*Private study*
Corr(06,07)	0.8681	0.8625	0.8164	0.9514	0.8487	0.7422	0.4510
Independence p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: N = 509 for all correlations reported.

TABLE 4
Regression results.

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
	OLS	*2SLS/GMM*	*OLS*	*2SLS/GMM*	*OLS*	*2SLS/GMM*
Dependent variable	*Log average degree*		*Log good degree*		*Log dropouts*	
Log private study	-0.009	-0.008	0.004	0.009	0.021	0.019
	(1.82)*	(1.37)	(0.24)	(0.50)	(0.41)	(0.36)
Log attended hours	-0.004	0.006	-0.040	-0.010	0.104	0.080
	(0.35)	(0.49)	(1.22)	(0.25)	(1.26)	(0.98)
Log satisfaction	0.067	0.091	0.175	0.248	-0.363	-0.487
	(2.36)**	(2.51)**	(1.61)	(1.94)*	(1.14)	(1.60)
Log A-level	0.088	-0.069	0.392	-0.055	-0.854	-0.354
	(3.72)***	(0.96)	(5.05)***	(0.22)	(4.09)***	(0.60)
Year=2007	0.004	0.006	0.015	0.021	0.282	0.275
	(1.71)*	(2.10)**	(1.67)*	(2.14)**	(8.37)***	(8.80)***
Observations	1312	1312	1312	1312	1304	1304
R-squared	0.64	0.13	0.67	0.27	0.70	0.26
Subject fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
F-Test constraint	485.43	234.89	12.05	12.15	0.05	0.13
p-value	0.00	0.00	0.00	0.00	0.83	0.72
Hausman Test	5.76		3.93		1.38	
Hausman Test p-value	1.00		1.00		1.00	
Underidentification Test		32.71		32.71		31.66
Underid Test p-value		0.00		0.00		0.00
Weak id test		29.29		29.29		28.45
Hansen J test		0.40		0.23		0.85
J test p-value		0.52		0.63		0.36
Hansen C test		4.64		3.23		1.31
C test p-value		0.03		0.07		0.25

Notes: T-statistics clustered by university in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Estimation method is by OLS or efficient, feasible two-step 2SLS/GMM, with A-level scores assumed to be correlated with the error term and instrumented using the average A-level scores across departments within the same university, and across universities in the same subject area. The dependent variable is the average degree classification obtained by students (average degree), the percentage of students who get a first class or upper second class degree (good degree), or the percentage of students who drop out of university (dropouts). The F-Test of the constraint is the F-statistic of the test that the coefficients on private study, attended hours, satisfaction and A-level scores sum to one; p-value is the p-value of this test. The Hausman test is the Chi-squared of the test for whether the results of the OLS and 2SLS regressions are the same or not. The underidentification test is the Chi-squared of the Kleibergen-Paap rk LM test of underidentification. The Weak id test is the Kleibergen-Paap rk Wald F statistic for weak identification. The Hansen J test is the Hansen test of overidentification. The Hansen C test is the test for whether the instrumented variable (A-level score) is endogenous. See Baum, Schaffer and Stillman (2007) for further details.

TABLE 5

Regressions using alternative measures of satisfaction and teaching hours.

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	*2SLS/GMM*	*OLS*	*2SLS/GMM*	*OLS*	*2SLS/GMM*
Dependent variable	*Log average degree*		*Log good degree*		*Log dropouts*	
Log private study	-0.009	-0.007	0.005	0.011	0.021	0.019
	(1.76)*	(1.21)	(0.26)	(0.58)	(0.41)	(0.37)
Log timetabled hours	-0.003	0.007	-0.041	-0.010	0.076	0.054
	(0.25)	(0.60)	(1.26)	(0.27)	(0.86)	(0.64)
Log satisfaction	-0.003	-0.008	0.016	-0.002	-0.051	-0.031
	(0.28)	(0.72)	(0.45)	(0.06)	(0.59)	(0.34)
Log A-level	0.091	-0.069	0.402	-0.040	-0.869	-0.373
	(3.75)***	(0.91)	(5.11)***	(0.15)	(4.08)***	(0.61)
Year=2007	0.004	0.004	0.024	0.023	0.256	0.257
	(0.82)	(0.79)	(1.43)	(1.30)	(4.45)***	(4.62)***
Observations	1312	1312	1312	1312	1304	1304
R-squared	0.64	0.12	0.67	0.27	0.70	0.26
Subject fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
F-Test constraint	1355.18	181.32	49.26	16.58	0.44	1.09
p-value	0.00	0.00	0.00	0.00	0.51	0.30
Hausman Test	5.88		3.81		1.23	
Hausman Test p-value	1.00		1.00		1.00	
Underidentification Test		32.00		32.00		30.98
Underid Test p-value		0.00		0.00		0.00
Weak id test		28.39		28.39		27.62
Hansen J test		0.26		0.33		0.98
J test p-value		0.61		0.57		0.32
Hansen C test		4.49		3.05		1.22
C test p-value		0.03		0.08		0.27

Notes: T-statistics clustered by university in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Estimation method is by OLS or efficient, feasible two-step 2SLS/GMM, with A-level scores assumed to be correlated with the error term and instrumented using the average A-level scores across departments within the same university, and across universities in the same subject area. The dependent variable is the average degree classification obtained by students (average degree), the percentage of students who get a first class or upper second class degree (good degree), or the percentage of students who drop out of university (dropouts). The F-Test of the constraint is the F-statistic of the test that the coefficients on private study, attended hours, satisfaction and A-level scores sum to one; p-value is the p-value of this test. The Hausman test is the Chi-squared of the test for whether the results of the OLS and 2SLS regressions are the same or not. The underidentification test is the Chi-squared of the Kleibergen-Paap rk LM test of underidentification. The Weak id test is the Kleibergen-Paap rk Wald F statistic for weak identification. The Hansen J test is the Hansen test of overidentification. The Hansen C test is the test for whether the instrumented variable (A-level score) is endogenous. See Baum, Schaffer and Stillman (2007) for further details.

TABLE 6
Regressions including additional control variables

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	*2SLS/GMM*	*OLS*	*2SLS/GMM*	*OLS*	*2SLS/GMM*
Dependent variable	*Log average degree*		*Log good degree*		*Log dropouts*	
Log private study	-0.010	-0.007	0.005	0.011	0.022	0.016
	(1.87)*	(1.31)	(0.31)	(0.58)	(0.42)	(0.31)
Log attended hours	-0.006	0.003	-0.034	-0.007	0.098	0.069
	(0.47)	(0.26)	(0.93)	(0.18)	(1.21)	(0.85)
Log satisfaction	0.066	0.087	0.193	0.256	-0.387	-0.484
	(2.43)**	(2.53)**	(1.81)*	(2.07)**	(1.20)	(1.58)
Log A-level	0.089	-0.066	0.388	-0.040	-0.828	-0.282
	(3.66)***	(0.94)	(4.82)***	(0.15)	(3.93)***	(0.46)
Log number of	0.005	0.003	-0.012	-0.017	0.018	0.026
assignments	(0.84)	(0.47)	(0.59)	(0.76)	(0.41)	(0.60)
Log employment hours	0.001	-0.003	-0.001	-0.013	0.021	0.036
	(0.38)	(0.96)	(0.09)	(1.05)	(0.91)	(1.46)
Year=2007	0.006	0.010	0.009	0.022	0.277	0.263
	(1.15)	(1.68)*	(0.50)	(1.17)	(7.45)***	(7.49)***
Observations	1301	1301	1301	1301	1293	1293
R-squared	0.64	0.13	0.67	0.28	0.70	0.26
Subject fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Hausman Test	5.19		3.47		1.24	
Hausman Test p-value	1.00		1.00		1.00	
Underidentification Test		30.74		30.74		29.72
Underid Test p-value		0.00		0.00		0.00
Weak id test		26.32		26.32		25.64
Hansen J test		0.21		0.41		0.31
J test p-value		0.64		0.52		0.58
Hansen C test		4.47		2.78		1.35
C test p-value		0.03		0.10		0.24

Notes: T-statistics clustered by university in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Estimation method is by OLS or efficient, feasible two-step 2SLS/GMM, with A-level scores assumed to be correlated with the error term and instrumented using the average A-level scores across departments within the same university, and across universities in the same subject area. The dependent variable is the average degree classification obtained by students (average degree), the percentage of students who get a first class or upper second class degree (good degree), or the percentage of students who drop out of university (dropouts). The F-Test of the constraint is the F-statistic of the test that the coefficients on private study, attended hours, satisfaction and A-level scores sum to one; p-value is the p-value of this test. The Hausman test is the Chi-squared of the test for whether the results of the OLS and 2SLS regressions are the same or not. The underidentification test is the Chi-squared of the Kleibergen-Paap rk LM test of underidentification. The Weak id test is the Kleibergen-Paap rk Wald F statistic for weak identification. The Hansen J test is the Hansen test of overidentification. The Hansen C test is the test for whether the instrumented variable (A-level score) is endogenous. See Baum, Schaffer and Stillman (2007) for further details.

TABLE 7
Dividing the sample into pre-92 and post-92 universities.

Panel A: Pre-92 universities

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	2SLS/GMM	OLS	2SLS/GMM	OLS	2SLS/GMM
Dependent variable	*Log average degree*		*Log good degree*		*Log dropouts*	
Log private study	-0.014	-0.014	-0.011	-0.015	0.017	0.020
	(1.75)*	(1.65)	(0.43)	(0.53)	(0.28)	(0.31)
Log attended hours	-0.033	-0.033	-0.115	-0.103	0.310	0.296
	(1.38)	(1.49)	(2.29)**	(2.04)**	(2.36)**	(2.44)**
Log satisfaction	0.159	0.197	0.605	0.695	-0.840	-0.940
	(3.74)***	(3.29)***	(4.51)***	(4.59)***	(1.93)*	(2.39)**
Log A-level	0.144	0.042	0.546	0.235	-1.186	-0.924
	(3.93)***	(0.51)	(4.49)***	(0.95)	(3.09)***	(1.20)
Year=2007	0.004	0.005	0.015	0.017	0.238	0.234
	(1.06)	(1.38)	(1.20)	(1.28)	(5.67)***	(6.18)***
Observations	687	687	687	687	679	679
R-squared	0.59	0.30	0.65	0.44	0.74	0.31
Subject fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Panel B: Post-92 universities

	(7)	(8)	(9)	(10)	(11)	(12)
	OLS	2SLS/GMM	OLS	2SLS/GMM	OLS	2SLS/GMM
Dependent variable	*Log average degree*		*Log good degree*		*Log dropouts*	
Log private study	-0.003	-0.003	0.008	0.006	0.029	0.035
	(0.51)	(0.51)	(0.31)	(0.22)	(0.51)	(0.63)
Log attended hours	0.009	0.007	0.051	0.043	-0.193	-0.188
	(0.94)	(0.67)	(1.04)	(0.78)	(1.75)*	(1.47)
Log satisfaction	-0.027	-0.027	-0.146	-0.160	0.039	0.029
	(0.79)	(0.81)	(0.89)	(0.99)	(0.10)	(0.08)
Log A-level	0.032	0.060	0.223	0.340	-0.410	-0.327
	(1.43)	(0.67)	(2.62)**	(0.89)	(1.97)*	(0.34)
Year=2007	0.004	0.004	0.007	0.005	0.326	0.327
	(1.76)*	(1.34)	(0.57)	(0.34)	(6.40)***	(6.70)***
Observations	625	625	625	625	625	625
R-squared	0.49	0.01	0.53	0.02	0.55	0.23
Subject fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

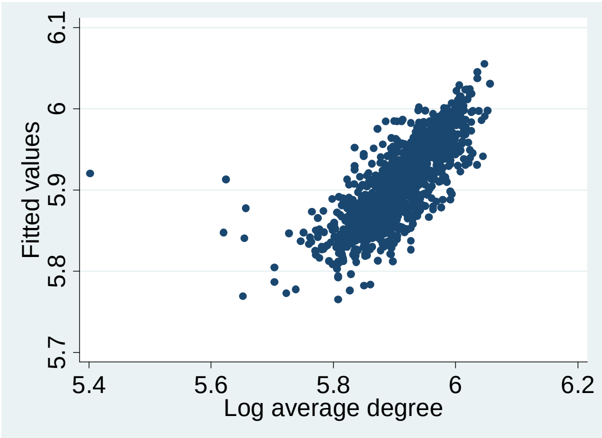
Notes: T-statistics clustered by university in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Estimation method is by OLS or efficient, feasible two-step 2SLS/GMM, with A-level scores assumed to be correlated with the error term and instrumented using the average A-level scores across departments within the same university, and across universities in the same subject area. The dependent variable is the average degree classification obtained by students (average degree), the percentage of students who get a first class or upper second class degree (good degree), or the percentage of students who drop out of university (dropouts).

TABLE 8
Performing the regressions with different time periods for the variables

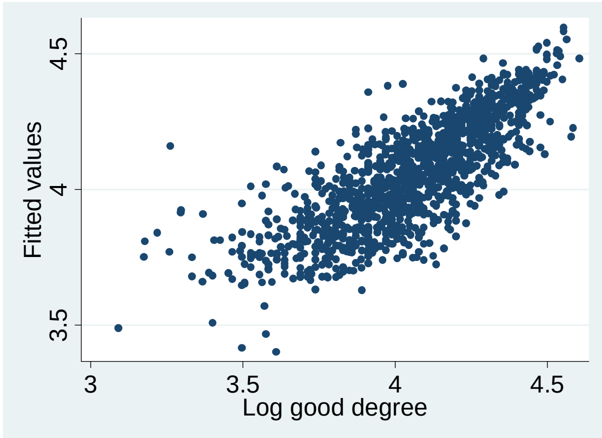
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
	OLS	*2SLS/GMM*	*OLS*	*2SLS/GMM*	*OLS*	*2SLS/GMM*
Dependent variable	*Log average degree*		*Log good degree*		*Log dropouts*	
Log private study	-0.008	-0.006	-0.019	-0.016	-0.052	-0.047
	(1.30)	(1.11)	(0.81)	(0.69)	(0.96)	(0.87)
Log attended hours	-0.003	-0.009	-0.050	-0.066	0.183	0.167
	(0.21)	(0.70)	(1.33)	(1.54)	(2.54)**	(2.25)**
Log satisfaction	0.044	0.020	0.119	0.065	-0.332	-0.392
	(1.34)	(0.57)	(0.89)	(0.43)	(1.22)	(1.25)
Log A-level	0.083	0.225	0.409	0.744	-0.845	-0.322
	(3.90)***	(3.24)***	(4.97)***	(2.65)***	(5.83)***	(0.41)
Year=2007	0.005	0.004	0.016	0.013	0.429	0.419
	(2.11)**	(1.14)	(1.53)	(0.99)	(17.84)***	(17.62)***
Observations	1065	1065	1065	1065	1056	1056
R-squared	0.66	0.08	0.66	0.26	0.74	0.40
Subject fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
University fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
F-Test constraint	405.23	142.71	11.16	1.36	0.02	0.35
p-value	0.00	0.00	0.00	0.25	0.89	0.56
Hausman Test	1.30		-32.72		-0.10	
Hausman Test p-value	1.00		1.00		1.00	
Underidentification Test		19.03		19.03		18.77
Underid Test p-value		0.00		0.00		0.00
Weak id test		12.77		12.77		12.51
Hansen J test		0.00		0.10		1.35
J test p-value		0.95		0.75		0.25
Hansen C test		6.08		1.81		0.52
C test p-value		0.01		0.18		0.47

Notes: T-statistics clustered by university in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Estimation method is by OLS or efficient, feasible two-step 2SLS/GMM, with A-level scores assumed to be correlated with the error term and instrumented using the average A-level scores across departments within the same university, and across universities in the same subject area. The dependent variable is the average degree classification obtained by students (average degree), the percentage of students who get a first class or upper second class degree (good degree), or the percentage of students who drop out of university (dropouts). The F-Test of the constraint is the F-statistic of the test that the coefficients on private study, attended hours, satisfaction and A-level scores sum to one; p-value is the p-value of this test. The Hausman test is the Chi-squared of the test for whether the results of the OLS and 2SLS regressions are the same or not. The underidentification test is the Chi-squared of the Kleibergen-Paap rk LM test of underidentification. The Weak id test is the Kleibergen-Paap rk Wald F statistic for weak identification. The Hansen J test is the Hansen test of overidentification. The Hansen C test is the test for whether the instrumented variable (A-level score) is endogenous. See Baum, Schaffer and Stillman (2007) for further details.

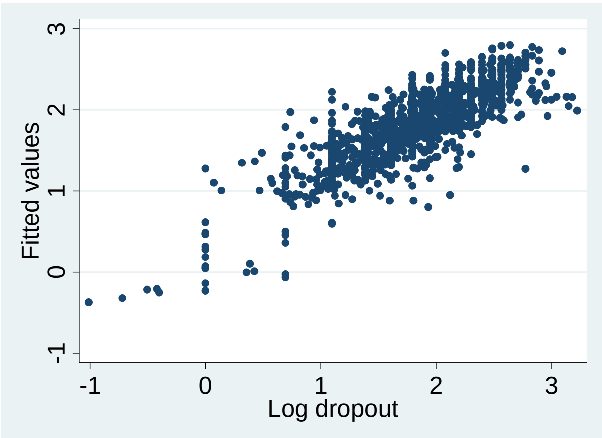
Figure 1. Actual versus fitted values, average degree performance, percentage of good degrees and percentage of dropouts: Full sample



Average degree performance.
N = 1312. Corr = 0.7727.



Good degrees.
N = 1312. Corr = 0.8039.



Dropouts.
N = 1304. Corr = 0.8336.