Research Repository UCD

Title	Does it pay to attend a prestigious university?
Authors(s)	Chevalier, Arnaud, Conlon, Gavan
Publication date	2003-07
Publication information	Chevalier, Arnaud, and Gavan Conlon. “Does It Pay to Attend a Prestigious University?” (July, 2003).
Series	ISSC Discussion Paper Series, WP 2003/13
Publisher	University College Dublin. Institute for the Study of Social Change (Geary Institute)
Item record/more information	http://hdl.handle.net/10197/1857

Downloaded 2024-03-28T04:02:09Z

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa)

© Some rights reserved. For more information

Institute for the Study of Social Change

ISSC DISCUSSION PAPER SERIES

DOES IT PAY TO ATTEND A PRESTIGIOUS UNIVERSITY?

Dr. Arnaud Chevalier
Dr Gavan Conlon

This paper is produced as part of the Policy Evaluation Programme at ISSC; however the views expressed here do not necessarily reflect those of ISSC. All errors and omissions remain those of the author.

ISSC WP 2003/13

Does it pay to attend a prestigious university?

ARNAUD CHEVALIER[♣]

AND

GAVAN CONLON[♦]

July 2003

Abstract:
This paper provides evidence of heterogeneity in the returns to higher education in the UK. Attending the most prestigious universities leads to a wage premium of up to 6% for males. The rise in participation in higher education also led to a greater sorting of students and an increase in the returns to quality. These results somehow justify the recent introduction of top-up fees.
Additionally, identification strategy matters and OLS estimates may be severely biased. However, our estimates, based on propensity score matching, are imprecise due to the thinness of the common support.
Key Words: Higher Education Quality, Tuition fees
JEL: I20, I22

Acknowledgments:
The authors are indebted to Peter Elias for facilitating the access to the 1995 graduate survey and Vincent O'Sullivan for excellent research assistance. We also thank Clive Belfield, Mark Berger, Augustin de Coulon, Deborah Garniss, Gauthier Lanot, Steve Machin, Olive Sweetman, Tarja Viitanen and participants at presentations at the Centre for the Economics of Education (LSE) and European Network of Micro-econometrics (Dublin) for their valuable comments on previous versions of this paper. All remaining mistakes are solely ours.
This project was partially funded by the Department for Education and Skills and the Council for Industry and Higher Education.

♣ Arnaud Chevalier is a research fellow at the Institute for the Study of Social Change, University College Dublin. He is also a research associate at the Centre for the Economics of Education, London School of Economics and the Institute for the Study of Labour, IZA, Bonn. Corresponding address: ISSC, UCD, Dublin 4, Ireland. arnaud.chevalier@ucd.ie

♦ Gavan Conlon is a research officer at the Centre for the Economics of Education, London School of Economics. Corresponding address: Centre for Economic Performance, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom E-mail: g.conlon@lse.ac.uk

1 Introduction

The literature on the returns to education has mostly estimated private returns for an average individual. These estimates have been widely used to encourage individuals to seek tertiary education. Recently, a group of prestigious universities has argued that their graduates were benefiting from higher financial returns since they provided higher quality education. Contrary to the US, the claim that university quality has positive effect on the labour market has not been analysed in the UK, with the exception of Belfield and Fielding (2001) or Naylor et al. (2000) who find no effect. The difficulty in estimating the returns to university quality originates from the heterogeneity and sorting of students. As students tend to attend a university that matches their ability (Hoxby, 1997), a simple comparison of the graduates' earnings between institutions is uninformative.

First, as in the bulk of the US literature, we assume linear selection on observables; i.e. we control for characteristics, mostly academic ability, affecting the choice of institution. However, Black and Smith (2003) have recently shown the drawbacks of such estimates; in particular, two primary weaknesses are identified. First, even if the selection to an institution is based on observable variables, the estimates may be biased if the quality effect on earnings is non linear. Second, let us assume that selection is solely based on academic ability and that there are only two types of institutions (prestigious vs. standard). Assuming no supply constraint at prestigious institutions, all good students would be observed in a prestigious institution and all other students would be in a standard institution. With perfect sorting, the institutional quality effect is perfectly correlated with ability and is identified by the imposed functional form only. This is the lack of common support problem.

An alternative identification strategy, relies on propensity score matching (Rosenbaum and Rubin, 1983) and identifies institution effects by pairing each individual in a prestigious

institution with a "similar" individual in a less prestigious university, thus not imposing a functional form. This estimation also highlights the difficulties associated with the absence of common support between the two populations of students as it can only be implemented if common support is indeed found. OLS and matching are biased if selection is on the unobservables.

Our analysis is based on three cohorts of UK graduates (1985, 1990 and 1995), who were surveyed 11, 6 and 3 years after leaving university respectively. Returns to higher education vary by the type of institution attended even after accounting for students' characteristics. Male graduates from Russell group institutions (thereafter RG), the most prestigious universities, earn 4% to 12% more than those from Modern universities. Matching estimates range from 1% to 6%, but lack precision due to the thinness of the common support. The difference between the two sets of estimates originates both from the sorting of students and the imposed functional form. We drop females as the analysis was biased by selection into the labour force effects (see Chevalier and Conlon, 2003).

The return to quality is higher for the younger cohort, which is consistent with an increased segregation of students by ability as the system expanded (Hoxby, 1997). We also tentatively support that the quality effect originates from an increase in human capital rather than a signalling or network effect.

By currently imposing a unique price for higher education, an implicit subsidy is provided to graduates attending 'prestigious' institutions. Assuming a lifetime premium of 2% to 5% on earnings for attending a prestigious institution, the value of this subsidy ranges from £2,950 to £7,100 per annum for the three years of the typical degree. This is less than the top-up fees currently proposed by the government[3].

[3] The White Paper on higher education (HMO 2003) states that from 2004, universities will be free to charge top up fees of up to £3,000 provided they meet specific requirements on widening participation for individuals from 'non-traditional' university backgrounds. This reform aims to reduce the shortage of funding in higher education. For example, between 1989 and 1997 per student funding decreased by 36% (HMO, 2003). The interested reader is referred to Greenaway and Haynes (2003) for details about trends in participation and funding in HE in the last two decades.

2 The United Kingdom Higher Education system

University education is almost universally provided in publicly funded institutions. A constant policy of most UK governments over the past four decades has been to promote participation to tertiary education. This led to the creation of several universities in the Sixties. The most recent change was the 1992 Higher Education Act, which granted university status and degree awarding power to all higher education institutions. Universities with a post-1992 charter, used to be mostly vocational and we refer to them as Modern Universities. Within the Old universities, with a charter prior to 1992, a self-selected group encompassing the oldest and most prestigious higher education institutions in the country has been formed. This group of institutions is referred to as the Russell Group[4].

Currently, Old universities account for almost half of the undergraduate students in the country, almost evenly split between RG and other old universities (From now one, by Old universities we will refer to non RG old universities only). This hierarchy of institutions is used as a comprehensive measure of quality. The three types of university differ by the degree subjects offered and the emphasis placed on research and as a result the sources and amount of public funding available. Current mean statistics on A-level scores of intake students, pupil-staff ratio, research assessment and destination of graduates are statistically different for the three types of institutions (see Table 1). For each measure, RG dominates Old universities, and Modern universities always have the lowest mean quality. We believe this measure of quality also captures some unobservable characteristics

[4] The members of the Russell Group are as follows: University of Birmingham, University of Bristol, University of Cambridge, Cardiff University, University of Edinburgh, University of Glasgow, University of Leeds, University of Liverpool, University of Manchester, University of Newcastle upon Tyne, University of Nottingham, University of Oxford, University of Sheffield, University of Southampton, University of Warwick, Imperial College, King's College London, London School of Economics and University College London.

such as reputation or network effect and therefore limit bias due to selection on unobservable characteristics[5].

3 'Teaching Quality' and returns to Higher Education

Heterogeneity in the returns to a degree by subject has long been recognised (see Walker and Zhu, 2002 for example) but evidence concerning institution effect is sparse and inconclusive. Naylor *et al*. (2000) use administrative data on the population of individuals graduating from RG and Old universities in 1994. Among this selective group of universities, the mean weekly earnings range from £370 to £430 with no significant institution effect. This study suffers from important caveats: first, the selected institutions are more homogenous than the universe of higher education institutions in the UK; second graduates earnings are imputed from the occupation occupied six months after graduation. Belfield and Fielding (2001) rely on a graduate survey and measure quality by the student-staff ratios and subject-adjusted resources. They find no effect of quality on graduates' financial outcomes, which may reflect the specificity of their measure of quality.

These UK results contrast with the US, where college quality effects are prevalent; Brewer *et al*. (1999) for example conclude that even correcting for selection into the type of university, prestigious private institutions provide significantly higher financial returns compared to low cost public institutions. These returns increase between cohorts concomitantly with the evolution of the fee differential. Whilst Daniel *et al*. (1997) confirm that fee differentials are in line with quality differentials, they criticise the parametric approach used in the rest of the literature. Using propensity score matching, Black and Smith (2003) estimate that in the long run attending a top quality university increases earnings by about 6% for men and 10% for women. However, their estimates are imprecise

[5] Heckman *et al*. (1996) suggest that estimates of the effect of quality of education are sensitive to the choice of quality measures and the level of aggregation of the data. Our measure of quality incorporates several

due to small sample sizes. Unobservable characteristics of students may also bias these estimates upwards. Berg-Dale and Krueger (2002), using information on all applications, control for selectivity on unobservables and find no financial return to attending a more selective institution.

4 Identifying Strategy

To extend Montmartquette *et al*. (2002), we assume that for an individual *i* the choice of an institution *j*, is based on the expected probability of graduation (g_{ij}) and expected lifetime earnings (w_{ij}). We define *A* and *X*, as the determinants of respectively g_{ij} and w_{ij}. The quality of the institution j affects the probability of graduation and earnings upon graduation[6]. Since tuition fees are equal in all institutions, individual *i* expected utility of graduation at an institution *j* is simply:

$$E(U_{ij}) = g_{ij}(A)\, w_{ij}(X) + (1 - g_{ij}(A))\, w_{i0}(X) \qquad (1)$$

where w_{i0} is the expected lifetime earning of i if dropping out of university. If the supply at each university were perfectly elastic, individual *i*, would choose to register at the university maximising the utility of graduation (U_{ij}). For high ability individuals, the choice will be based on expected lifetime earnings; whilst for less able, the probability of graduation may be a more important determinant.

In the UK, application to university is centralised and prospective students are limited to a maximum of six choices. The probability of acceptance at institution j (l_{ij}) depends on

dimensions that may or may not be correlated with observable characteristics of the institutions such as peer group or reputation effects.

[6] The difference in the probability of graduation between institutions reflects variations in the subjects offered and threshold to pass the test. Light and Strayer (2000) for example shows that the probability of graduating is higher when observed ability and quality are matched.

the individual characteristics but also on the characteristics of the other applicants (*Z*). Thus, for individual *i*, the utility of graduation at institution *j* becomes:

$$E(U_{ij}) = l_{ij}(A,Z)\, g_{ij}(A)\, w_{ij}(X) + (1 - l_{ij}(A,Z))\ \mathrm{w}_{i0}(X) + (1 - g_{ij}(A))\, l_{ij}(A,Z)\ \mathrm{w}_{i0}(X) \quad (2)$$

The second and third terms of (*2*) represent the earnings if not attending tertiary education and the earnings if dropping out of university respectively. Since, applications are limited, and individuals may have poor information on their relative characteristics, risk aversion and strategy affect the mix of institutions applied to. The application process creates some disparities between the academic ability of individuals and the quality of the institution attended. Hence the common support assumption required for propensity score matching is likely to be fulfilled.

The simplest model to estimate institution effects on wages is simply to rely on a log wage model and include variables for the type of institution attended (Specification 1)[7]. Throughout this discussion, returns are estimated relative to an individual who graduated from a Modern university.

$$\ln W_i = I_i \beta_I + X_i \beta_X + c + \varepsilon_i \quad (3)$$

where, *ln(W)* is the natural logarithm of gross wage, X_i is a vector of idiosyncratic characteristics affecting wages, *I* indicates the university type, *c* is a constant and ε is an error terms measuring the impact of individual non-observable characteristics on wages.

Specification 1 would provide unbiased estimates of institution effects if students where randomly allocated to universities. As stated above, both students and institutions

choose and the sorting is mostly based on academic ability. Therefore, in specification 2, we include the student's A-level score. To reflect the heterogeneity in the returns to a degree, we also control for degree class and subject[8] (specification 3). Finally, since strategic behaviour may be at play during the application process, we include family background and type of schools attended, as they affect risk aversion and information (specification 4). Thus, our final specification has the following form.

$$\ln W_i = I_i\beta_I + X_i\beta_X + A_i\beta_A + P_i\beta_P + S_i\beta_S + c + \varepsilon_i \qquad (4)$$

where, A is the A-level score achieved, P is a vector of parental characteristics measured by parental social class and S is the type of school attended prior to university.

The covariates included in X impact on the estimates. If quality affects wages directly but also through characteristics that are included in X, our estimates of institutional effects would be biased downwards. To capture the total effect of quality on wages, X is restricted to the following covariates: post-graduate qualifications, a quadratic function of labour market experience since graduation, employer size, type of contract (permanent / temporary), self-employment status and current region of residence[9].

As stated above, this identifying strategy may lead to bias results, thus we estimate institutional effects on wages by propensity score matching. In other words, we estimate the earning differential between graduating from a RG university (Y_1) and a Modern university (Y_0). D is a dichotomous variable taking the value one for an individual graduating from a prestigious institution, we then estimate the Average Treatment Effect on the Treated.

---

[7] Since we observed clusters of respondents by institution, the standard errors of ordinary least square have to be corrected (Moulton, 1987).

[8] Datcher-Loury and Garman (1995) shows that for white omitting major, degree grade and parental income bias upwards the positive effect of college selectivity on earnings by up to 50%.

$$ATET = E(Y_1 - Y_0 / X, D = 1) = E(Y_1 / X, D = 1) - E(Y_0 / X, D = 1) \qquad (5)$$

As individuals are observed in only one state of the treatment, $E(Y_0 / X, D = 1)$ is never observed. In order to identify this parameter, we assume, as with ordinary least squares, the Conditional Independence Assumption (CIA), which is equivalent to assuming that the selection is only based on observable characteristics.

$$(Y_1 - Y_0) \perp D / X \qquad \text{(CIA)}$$

Assuming CIA, the treated and non-treated population have on average the same outcome regarding the treatment effect. Ordinary linear least squares further assumes homogeneity of the effect of the treatment; conditioning on *X*, the effect of attending a prestigious institution is identical for all individuals. Matching assumes that the CIA holds for individuals who have "similar" *X*. As a large number of covariates is usually required to satisfy CIA, Rosenbaum and Rubin (1983) show that it is equivalent to condition on the estimated probability of being treated (Pr (D=1/X)=P(X)) or on all the dimensions of *X*. For each treated individual i, the counterfactual outcome of non-treatment is a weighted average of a selection of control observations.

$$\hat{E}(Y_0 / \hat{P}(X_i)) = \sum_{j=1}^{J} w(\hat{P}(X_i), \hat{P}(X_j)) Y_{oj} \qquad (6)$$

The weight being attached to a given control (*j*) is a function of the distance between the propensity score of individuals *i* and *j*. Individuals who are the most similar to *i* in

[9] Estimates of quality when X only includes experience and region where not statistically different from those

terms of observable characteristics (X) are given the largest weight[10]. In order to match a treated individual, at least one individual in the control population must have a "similar" propensity score: this is the common support assumption. As only matched treated individuals are used, it is crucial to check for common support, otherwise, the estimate of the treatment would be biased[11]. Whilst neither OLS nor matching account for possible selection on the unobservable characteristics, matching identifying power does not come from the functional form imposed and it highlights any breach of the common support.

5 Data

We use two nationally representative postal surveys of graduates: the 1985-1990 Graduate Cohort Studies and the 1995 Graduate Cohort Study. Alumni from participating institutions were surveyed in 1996 (cohort 1985 and 1990) and 1998 for the 1995 cohort. The raw sample sizes for the 3 cohorts of interest are: 5,835, 9,688 and 10,575. The two surveys are similar but not directly compatible[12]. In particular, the sampling time frame differences make it difficult to separate cohort and 'time since graduation' effects. The main discrepancies relate to the type of school attended which is not available for the 1995 cohort, and the A' Level score, parental occupation and subject of degree which are not identically defined across surveys. Whilst both surveys rely on categorical annual gross wage, it is possible to approximate hourly wage for the 1985/1990 cohorts but not for the 1995 cohort, as weekly hours worked are not recorded.

---

presented.

[10] In this paper, the weighting functions used are: nearest neighbour match with replacement and calliper, and Epanechnikov kernel. Further details are available in Chevalier and Conlon (2003).

[11] In case of a lack of common support, the estimate becomes the Average Treatment Effect on the Matched Treated. Heckman et al, (1997) decompose the bias of a propensity score estimate into its three basic components: B_1 is the biased that occurs due to lack of common support, B_2 arises from different distributions of X within the two populations on the common support, and B_3 is due to differences in outcomes that remain even after conditioning on observables and making comparisons on a region of common support (due to selection on the unobservables). The authors compare matching results with experimental data and conclude that the first two terms of the bias may be substantial.

[12] The 1996 survey and the 1998 survey differ in the institutions included in their target population, with the latter survey excluding Open University, a distance learning centre, but also some specialist colleges. To make the cohorts more compatible we exclude all these institutions from the 1996 survey.

The working sample is restricted to male who obtained an undergraduate degree before the age of 25, were in full time employment in the UK, provided information on earnings at the time of the survey and were not affected by health problems. These restrictions reduce the sample size to 1,247, 1,635 and 2,482 observations respectively[13].

We first examine the relationship between academic ability and our measure of quality. As expected from the aggregate data, the mean A-level scores are significantly higher in RG than in Modern universities (+40%). The dispersion in the average score is lower for the 1995 cohort than for previous cohorts, but this is likely due to measurement differences. Increasing competition for students has lead universities to admit individuals without the traditional entrance requirements. Theoretically, this strategy should be adopted by institution of lower quality (Hoxby, 1997) and we observe that Old and Modern graduates are respectively 2 and 3 times more likely not to possess A-levels than graduates from RG.

As well as evidence on the heterogeneity of the student population, Table 2 reports the first substantiation of common support. While the distribution of ability differs by institution type, the sorting is not perfect. Due to possible grade drift, it is unclear whether the segregation of students has increased. Tentatively, we note that the proportion of students from the top half of the A-level distribution rose from less than 60% to 85% at RG universities. Contrary to Black and Smith (2003) or Light and Stayer (2000), we do not find asymmetry of sorting with more able students in institutions of lower quality than low ability students in "prestigious" institutions. The common support may nevertheless be thin, with only 4% of the most able students in Modern universities.

Gross wages are inflated to 2002 prices using the Retail Price Index and are reported in Table 3. For the first two cohorts, we reproduce annual and hourly wage, whilst for the

[13] The important selection of the sample is due to limiting the population to first degree holders; in the 1996 survey, this restriction eliminates 40% of the male population as the sampled population also includes diploma holders and post-graduates.

1995 only annual wage is available. Wages are reported 3 to 11 years after graduation, so cross cohort comparisons are difficult to interpret. Graduates from RG universities earn more than those from Modern universities at each percentile of the income distribution. For 1985 graduates, the average wage following graduation from a RG institution is £16.93 (hourly) or £39,659 (annual), which is equivalent to a premium of 14% over graduates from a modern university[14]. This premium declines to 9% for the 1995 cohort.

Evidence of a pay premium for attending an Old university is tenuous. At most the difference between graduates from Old and Modern universities reaches 4% but is negative in 1985. The preliminary analysis confirms that attending a RG university is correlated with higher earnings. Since students are heterogeneous in academic ability, this gap over-estimates the effect of university quality on earnings. Importantly for our estimation strategy, the sorting of students in not perfect and students of all ability levels are found in each institution type. This common support is nevertheless likely to be thin.

6 Quality effect on wages

6.1 Linear selection

First we estimate the model presented in (4) and establish the importance of controlling for pre-university ability (specification 2), subject of graduation and grades (specification 3) and parental background (specification 4) compared to the base model (specification 1).

Quality of higher education matters, as reported in Table 4, in the base specification, this quality premium ranges from 9% to 12% for RG and 3% and 8% for other Old universities (not significant in 1985 and 1990). As expected, controlling for pre-university educational achievement (specification 2) reduces the quality premium by about 50% in the

[14] Since the wage premium for graduating from a prestigious institution is similar for hourly and annual wage, the number of hours worked is not dependent on the type of university attended, at least for this selected group of full-time workers.

1985 and 1990 cohorts. This is consistent with the sorting governing university admission being mostly based on academic ability. When controlling for A-level score, the quality premium becomes insignificant for these two cohorts. For the 1995 cohort, adding controls for A-level score reduces the premium to quality, which remains significant and reaches 10% for RG and 8% for Old universities.

Adding further controls for subject choice, class of degree achieved and parental background does not substantially affect the estimated quality effects. These results confirm that the selection to university is essentially based on the A-level score. Contrary to previous UK evidence, we find that for the most recent cohorts, quality of higher education had an impact on labour market outcomes. These returns to quality are in line with evidences based on linear model in the US.

In the final specification, returns to quality are higher for the younger cohorts. Two competitive explanations are put forward: either quality effects decrease with time on the labour market or the expansion of higher education led to an increase in the competition between universities and an increased sorting of students. We come back to this point in the next section.

6.2 Propensity score matching

The institutional characteristics of the UK application process, makes the claim of selection on observables plausible. As presented above, while the segregation of students by ability is high, the sorting is not perfect. We estimate the propensity score with a probit where the covariates are ethnicity, paternal socio-economic group, paternal education (1985 and 1990 cohort only), home ownership (1985 and 1990 cohort only), type of school attended (1985 and 1990 cohort only), and use of careers service (cohort 1995 only). These variables attempt to capture academic and financial constraints as well as information and motivation. Each covariate is interacted with the A-level score. We separately consider two

treatments, attending a RG or an Old university while the control group is composed of graduates from Modern universities.

The distributions of propensity scores are reported in Figure 1. In these histograms, each bin has a width of 0.05. Cohort 1995 excepted there is evidence of a thin common support throughout the distribution of propensity score but the small numbers of useful control observations constrains the choice of procedure to matching with replacement.

With the expansion of higher education, RG have become less elitist and poached the best students out of Modern universities. In 1985, 90% of RG graduates have a propensity score greater than 0.90, this proportion falls to 45% and 20% by 1990 and 1995 respectively[15]. Similarly, the distribution of propensity score at Modern universities was almost uniform with 20% having propensity score lower than 0.20. By 1990, ability sorting is striking with 70% of graduates with a propensity lower than 0.20. The growth in the Modern university sector was made possible by attracting students of predominantly lower ability. The effect of the increased competition for students on the returns to quality depends on the relative variations in admission standards. The evolution of the distribution of propensity score is similar at Old universities. Here again, common support is universal but thin.

Various estimates of the effect of quality on labour market outcome are reported in Table 5, as well as two indicators of the match quality. We report the proportion of matched treated observations (with nearest neighbour) and, as an indicator of the thinness of the common support, the number of control observations accounting for 50% of the matches. This indicates how sensitive our estimates are to the few observations guaranteeing common support. With a few controls being used several times, the precision

[15] It has to be noted that the distribution of the 1995 cohort is affected by the fact that A-level score was reported in a categorical format rather than in a continuous form. Thus, we observed clusters of propensity scores rather than a distribution.

of our estimates suffers (Abadie and Imbens, 2002)[16]. Nearest neighbour matches are reported with a calliper of 0.1 and 0.01[17]. Similarly, kernel estimates use a bandwidth of 0.1 and 0.01. Additionally, OLS results based on the sample of matched pairs with the tightest calliper are also included. Standard errors are obtained by bootstrap with 500 replications. Each bootstrap re-estimates the propensity scores within the new sample before matching.

Since we have few control observations, nearest neighbour matching is likely to produced biased estimates. Our favoured specification is therefore the Epanechnikov kernel match. Local Linear Matching usually performs the best when the distribution of the propensity scores is clustered around either 0 or 1. However, results were similar to those obtained with kernel estimates and are therefore not reported. The choice of the kernel bandwidth is also of interest, with a trade-off between bias and precision. Heckman et al. (1997) decompose the bias originating from propensity score matching into three components: B_1 is due to lack of common support, B_2 arises from different distributions of X within the two populations, and B_3 is due to selection on the unobservables.

With the larger bandwidth, there is no problem of common support and all treated observations are matched. With the tighter bandwidth, we match more than 90% of the treated observations; the dropped observations are at the top end of the propensity score distribution. Since, we do not observe that the effect of attending a prestigious university is heterogeneous in academic ability (see next section), dropping these variables is unlikely to bias our estimates significantly. Since B1 is likely to be small, kernel match with the tighter calliper is our favourite specification as it reduces the match bias (B2).

[16] Results presented do not include degree subject, which has been shown to be a determinant of earnings (Walker and Zhu, 2001). Adding subjects made the common support even thinner, so that for the 1985 cohort, 1 observation represented more than 40% of the matches. For the 1995 cohort, where the support was thicker, estimates including subject of graduation were 50% higher than those presented.

[17] Propensity score matching with larger callipers (up to 0.20) lead to similar results than those with a calliper or bandwidth of 0.10.

Using our favoured estimate, the returns to quality are much smaller than when estimated by OLS. The imprecision of our estimates is a direct consequence of the thin support. To test the robustness of these results, we add a normal random error term to the propensity score. While changing the matched pairs, estimates based on the modified propensity score were similar to those presented. Hence our estimates appear reliable and the imprecision is solely due to the small number of control observations. The quality premium differs by cohort, from 1% in 1985, 4.5% in 1990 and 5.8% in 1995. While the returns to quality are larger for the younger cohort, it is impossible to conclude whether this trend relates to the increase heterogeneity of students between institution types.

For each cohort, the last line reports the OLS estimate based on the matched population rather than the full population. The estimates on the selected population differ from those obtained on the full sample, indicating that the original results were affected by the lack of common support. In a number of cases, the matched OLS estimates are out of line with propensity score matching estimates, which indicates that the assumption of linear selection on the observables should be rejected. Previous studies relying on OLS estimating methods may therefore be seriously biased.

Similarly, we estimate the effect of graduating from an Old university as opposed to a Modern university. The common support is rather thin and the standard errors large. For most estimates, returns to graduating from an Old university are lower than those for a RG, which supports our classification as a measure of quality. Returns to Old universities may even be negative for the older cohort. As with RG institutions, there is no quality premium for older workers and the trends are similar to those previously described. Focusing on kernel match with the tighter bandwidth, there is no clear evidence that graduating from an Old university rather than a Modern university leads to a financial premium. The largest estimates are obtained for the most recent cohort (+2% for men), which is consistent with an improvement in the relative quality of the population of graduates from Old universities.

Since, we do not find positive effects of attending an Old university, the remaining of the paper focuses on Russell Group institutions solely.

6.3 Who benefits from attending a "prestigious" institution?

The effect of university quality may be heterogeneous. First, RG universities have so far been treated as a homogenous group, while there are obvious variations in their observed quality. We test this assumption of heterogeneous returns by focusing on two large institutions. The second source of heterogeneity is between students. We test whether returns are correlated with ability and family background. If the wage premium from attending a prestigious university stems from a network effect then graduates from higher social class may reap higher benefits from the "old boy network".

First, we isolate two RG institutions for which we have more than 100 observations that differ in their observed quality[18]. Graduates from these two institutions are matched following the same procedure as the one presented above. We believe the institutions differ sufficiently to insure that selection is mostly based on observables, but we can not ruled out that unobservable characteristics of students or the institutions do not determine the application decision. Since we only have 234 observations, we use larger bandwidth (0.20 and 0.10). The common support is rather thick. Despite these two institutions being member of the Russell group, graduates from the higher quality university earn between 9% and 10% more than those from the control institution. While the rest of the discussion carries on focusing on differences between institution types, it is worth remembering that heterogeneity in quality within institution group (at least for the prestigious group) is also marked.

Secondly, we focus on two observable characteristics of graduates: ability and social background. If the quality effect on earnings is correlated with ability, this will be

informative on how the resources are allocated in different institutions. Similarly, the quality effect on earnings may just be due to a reputation or a network effect, which may benefit graduates according to their socio economic background. For each individual, we compute the effect of the treatment as the difference between the observed and the control wages using estimates obtained with the kernel matching and a bandwidth of 0.01. Quality premium is never correlated with either ability or family background. The quality effect is homogenous and therefore more likely to steam from improved teaching rather than a network effect. Prestigious universities allocate their extra resources evenly and do not solely focus on a few prominent students. While RG universities are rather unequal in access, with more able pupils from better socio-economic background being over-represented, the wage premium for attending them is independent of the student's characteristics. Prestigious universities level the playing field within their intake.

6.4 Wage growth effects

Several times in this discussion, differences in the returns to quality by cohorts have been noted. In this section, we test whether these differences originate from differences in labour market experience or cohort effect, by looking at wage growth. The financial return to university quality has potentially three origins. A pure quality effect, which improves graduates' human capital, hence the premium to quality should be constant over the lifetime. The other two hypotheses have conflicting predictions regarding relative wage growth of graduates from prestigious universities. In a signalling model, the prestige of an institution may be used by employers to differentiate between new graduates when hiring. With time, the true ability of individuals is revealed and the value of the signal tends to zero, thus the wage growth gap in quality should decrease with time. Alternatively if the returns to attending a prestigious institution stem from peer group effects, one may expect

[18] University A has a student staff ratio that is half of the one observed in university B. Its average A-level

the returns to the peer group to increase with time on the labour market. Thus, looking at wage growth allows us to conjecture on the origin of the pay gap between graduates from different types of institutions.

Graduates on the 1996 survey report annual earnings up to three points in time (1, 6 and 11 years after graduation. For individuals with positive earnings in two consecutive periods, we calculate wage growth[19]. The earning profile of UK graduates is steep, with a mean growth ranging from 60% to 100% for the period between the first and the sixth year following graduation[20]. Wage growth is concave in labour market experience as predicted by the human capital model and is reduced by about half for the second period (6 to 11 years after graduation). We follow the same identifying strategy has the one presented for the effect of quality on earnings.

Table 6, summarises the results by cohort and time period for different estimators. Due to the small number of observations and selection issues, we do not attach too much weight to individual estimate. Depending on the estimation method, the estimates vary substantially but are never statistically significant. Whilst for the 1985 cohort, they may have been some positive effect on initial wage growth for attending a college of higher quality, over the 1991-1996 period university type has no effect on wage growth. These results are consistent with the view that college quality increases the human capital of graduates, which has a constant positive effect on their wage, but does not alter wage growth.

While this conclusion is interesting in itself, it also gives us the opportunity to comment on the between cohorts differences in the effect of university quality on wage.

---

score of entrants differ by 2 points and its research assessment is also higher.

[19] The earnings variables are categorical, and mid-points are used to calculate earning growth which leads to mis-measurement of the true earning growth. Individuals reporting earnings in the same bin at two points are estimated to have experienced negative growth due to the deflation of the mid-point value over-time. However, only 1% of the individuals remain in the same earning category between 1986 and 1991, and 3% between 1991 and 1996.

[20] These estimates of wage growth are affected by selection, since only individuals reporting earnings in the two periods are selected.

Since, the quality gap is constant over the lifetime of graduates, the observed differences in the returns to quality are due to cohort specific effects. Returns to quality have increased from 1 to 6% as the number of student expanded. This is consistent with the increased sorting of students that we described in section 6.2. As predicted by theoretical models and experienced in the US (Hoxby, 1997) the increased competition for students has lead to an increased heterogeneity between institutions and thus higher returns to quality.

7 Fee differentials

The most recent cohort observed was the first affected by the increased in the numbers of universities. This is also the cohort for which we observe the highest returns to university quality. Since 1995, the number of students in Higher Education has been stable (Greenaway and Haynes, 2003) so this cohort provides valuable information for the current debate in higher education in the UK. As stated above, from 2004, universities will be able to charge top-up fees. We now estimate the tuition fee that a representative individual entering a RG university would be willing to pay over and above those charged at a Modern university in order to capture the quality premium.

We estimate the earnings profile of male graduates compared to those whose highest qualification is a GCE A levels using the Labour Force Survey. This is used as the profile for a Modern university graduate. We create a profile for RG attendance by adding a constant wage premium[21]. We also assume that the likelihood of employment is the same for all graduates, irrespective of the institution attended. Based on official real earnings growth rate (2.0%) and real discount rate (3.5%) and using the current tax allowance and rates, we calculate the difference in the net present value of the lifetime earnings between graduates from Modern and RG universities. Assuming that a degree takes three years to

[21] As the Labour Force Survey includes graduates from various type of institution, the mean graduate premium estimated is an upward value of the returns to graduating from a polytechnic, which would bias our estimates of the fee differential upwards.

complete, we convert the net present values of the earning premium to a yearly fee differential. Based on a premium of 2% to 5% over the lifetime, the annual fee differential between RG and Modern universities should range from £2,950 to £7,100.

How credible are these estimates? In the US, the market for higher education is competitive and it is usually acknowledged that the fee differential matches the quality differential (Daniel *et al*., 1997). The inter-quartile ranges in 2002 (in pounds) were £1,000 in the public sector and £5,000 for private institutions[22]. Two UK evidences can also be used. Fist, despite the unique price charged in higher education, RG institutions spend £4,000 more on their academic expenditure per student compared to Modern universities[23]. Assuming this cost differential reflects higher quality of undergraduate teaching, Russell Group universities are more efficient providers of higher education if a wage premium of at least 2.8% is obtained by their graduates. Second, universities are free to set their fees for the competitive market of non-European Union undergraduates. For this market, the mean fee differential between Russell Group and Modern universities ranges from £1,400 to £2,900 depending on faculty. So our estimates are globally in line with available evidences on fee differential.

The government's plan is to allow universities to charge fees up to £3,000 per annum. Not considering the general equilibrium that will result from the differentiation of fees between institutions, our estimates suggest that RG universities should set their fees at the new maximum, while less prestigious institution may be tempted by a reduction of their fees (currently set at £1,100) in order to attract more students. If the earning premium for attending a RG university is greater than 2% over the life-time, students attending them would still be more subsidised than their peers graduating from less prestigious

[22] Fees were obtained from the National Center for Education Statistics website. For public institution no break down by in or out of state residency status was available.

[23] These figures were obtained from the Resources of Higher Education, 1997. Differences in academic expenditures are a crude measure of the relative efficiency of quality since expenditures on academic department are only one component of the quality of the teaching at a given institution.

universities. Unfortunately, our estimates are not precise enough to put a definite value on the premium to quality in the UK.

8 Conclusions

This paper provides evidence of heterogeneity in the returns to higher education by type of institution in the UK. Our main findings based on three cohorts of UK graduates can be summarised as follow.

The matching of students to an institution's quality is largely put not perfectly due to academic ability. This imperfect sorting may be due to some institutional features and provides an alternative estimation strategy that had not been implemented to estimate the labour market impact of university quality in the UK previously.

Using propensity score matching we find some evidence of financial premium for attending more prestigious universities. In our preferred specification, the estimated premium ranges from 1% to 6% for males. The estimates are imprecise since the common support while not failing is rather thin. Linear estimates are more precisely estimated but the identification comes from the imposed functional form, and are therefore biased.

The financial benefit of attending a Russell Group university is neither dependent on previous academic achievement nor parental background. In some sense, prestigious universities level the playing field for their graduates. However, heterogeneity between prestigious institutions is large, with returns varying by up to 10%.

With the increase in the number of students, the higher education system has become more polarised and the students more homogenous within a university type. This segregation of students has increased the premium to quality.

There is no evidence that university quality affects wage growth. This constant effect of quality on earnings is consistent with the premium originating from an increase in human capital rather than a signalling or a network effect.

The net present value of the quality premium is equivalent to a tuition fee differential between RG and Modern universities ranging from £2,950 to £7,100. This differential is broadly consistent with available evidence. By implementing a unique price in higher education (the current practice), the government subsidises graduates attending more prestigious institutions more generously than others. Thus the claim that institutions should be allowed greater freedom in setting their tuition fees has some justifications.

However, introducing price competition will drastically affect higher education in the UK. As mentioned in the current White Paper (HMO, 2003), with the introduction of top-up fees, students will become more exigent customers. Thus while average quality is likely to increase, the heterogeneity between institutions may also augment with some institutions deciding to compete on price and others on quality. The choice of subject provided by universities may also be affected with less popular courses being dropped out in favour of high revenue courses. Hoxby (1997) demonstrates that these market mechanisms took place in the US and are responsible for higher average quality, a homogenisation of the students within institutions but also a greater variation in the quality of the degree provided between institutions and account for up to 40% of the explained growth in the dispersion of returns to higher education (Hoxby and Terry, 1999). A negative consequence of the increased competition between providers of tertiary education is to make the signal attached to a degree less precise to employers.

References

Belfield C. and A. Fielding, 2001, Measuring the relationship between resources and outcomes in higher education in the UK, *Economics of Education Review*, 20, 589-602

Berg Dale S. and A. Krueger, 2002, Estimating the pay off to attending a more selective college: An application of selection on observables and unobservables, *Quarterly Journal of Economics*, 1491-1527

Black D. and J. Smith, 2003, How robust is the evidence on the effects of college quality? Evidence from matching. *University of Maryland*, mimeo

Brewer D., E. Eide and R. Ehrenberg, 1999, Does it pay to attend an elite private college?, *Journal of Human Resources*, 33, 104-123

Chevalier A and G. Conlon, 2003, Does it pay to attend a prestigious university, Centre for the Economics of Education, London School of Economics, DP

Daniel K., D. Black and J. Smith, 1997, College quality and the wages of young men, University of Pennsylvania, mimeo

Datcher Loury L. and D. Garman, 1995, College selectivity and earnings, *Journal of Labor Economics*, 13, 289-308

Greenaway D. and M. Haynes, 2003, Funding higher education in the UK: the role of fees and loans, *The Economic journal*, 113, F121-F149

Heckman J., A. Layne-Farrar and P. Todd, 1996, Human capital pricing equations with an application to estimating the effect of schooling quality on earnings, *The Review of Economics and Statistics*, 78, 562-610

Heckman J., H. Ichimura, P. Todd, 1997, Matching as an econometric evaluation estimator: Evidence from evaluating a job training program, *Review of Economic Studies*, 64,605-654

HMO, 2003, *The future of Higher Education*, Department for Education and Skills.

Hoxby C., 1997, How the changing market structure of US Higher Education explains college tuition, *National Bureau of Economic Research*, WP 6323

Hoxby C. and B. Terry, 1999, Explaining rising income and wage inequality among the college educated, , *National Bureau of Economic Research*, WP 6873

Light A. and W. Strayer, 2000, Determinants of college completion: School quality or student ability?, *Journal of Human Resources*, 35, 299-332

Montmarquette C., K. Cannings and S. Mahseredjian, 2002, How do young people choose college majors? *Economics of Education Review*, 21, 543-556

Moulton B., 1987, Diagnostic for group effects in regression analysis, *Journal of Business and Economic Statistics*, 5, 275-282

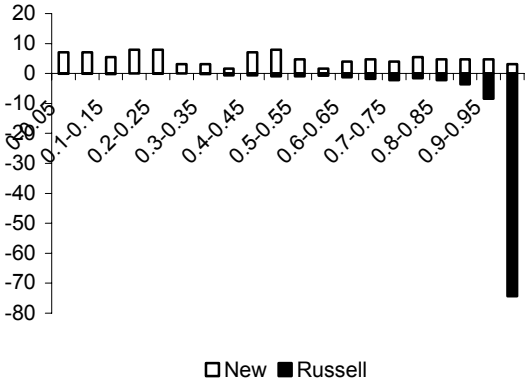
Naylor R., J. Smith and A. McKnight, 2000, Occupational earnings of graduates: evidence for the 1993 UK university population, *University of Warwick*, mimeo

Rosenbaum P. and D. Rubin, 1983, The central role of the propensity score in observational studies for causal effects, *Biometrika*, 70, 41-55

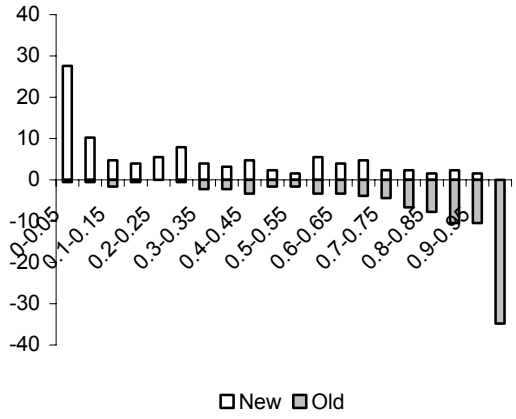
Walker I. And Y. Zhu, 2001, The returns to education: Evidence from the Labour Force Surveys, *University of Warwick*, mimeo

Figure 1: Distribution of propensity scores between Russell Group and Modern Universities and between Old and Modern universities

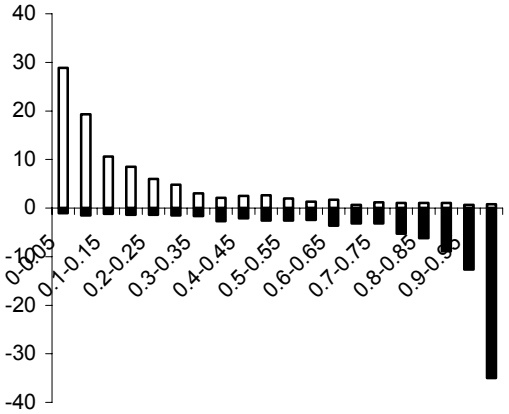
Male, 1985 Cohort



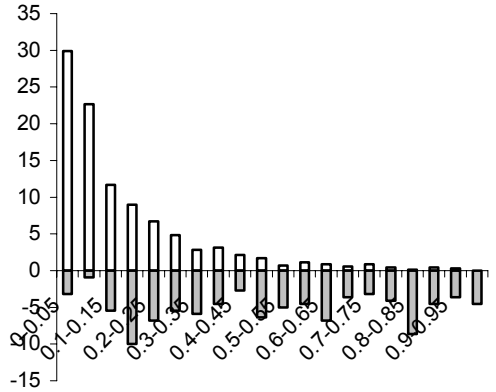
Male, 1985 Cohort



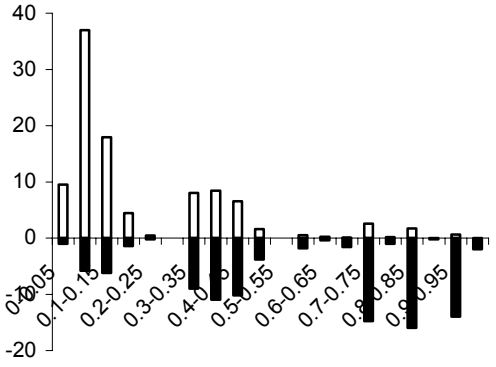
Male, 1990 Cohort



Male, 1990 Cohort



Male, 1995 Cohort



Male: 1995 Cohort

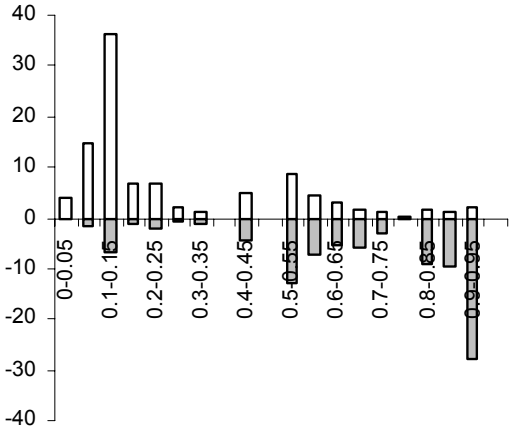


Table 1: Mean quality by institution type (Standard Deviation) in 2002

	Russell Group	Old university	Modern University
Number of Undergraduates	268,479	294,576	592,432
Number of Institutions	19	36	55
Mean A-level score of entrants	25.20 (2.19)	20.53 (2.84)	14.28 (1.96)
Student/ staff ratio	13.74 (2.30)	15.62 (2.59)	19.01 (2.85)
Research Assessment	5.65 (0.42)	4.92 (0.54)	2.42 (0.59)
Destination of graduates	78.22 (4.49)	75.54 (5.54)	70.89 (5.75)

Note: Data provided by the Times Higher Education Supplement, League Table, May 2003 (No 1588) at the Institution level. Aggregate means presented in this Table have been weighted by number of undergraduates. The results are similar if unweighted means are used. Destination of graduates is the proportion of students in a graduate job 6 months after leaving university.

Table 2: Distribution of Academic ability by institution type

	Russell Group Universities	Old Universities	Modern Universities
1985 Cohort			
No A-level	0.024	0.071	0.079
Mean A-level score (out of 30)	23.59 (5.14)	17.83 (6.22)	14.00 (6.34)
Quartile 1	12.91	45.99	74.02
Quartile 2	29.67	35.02	17.32
Quartile 3	27.63	10.97	4.72
Quartile 4	29.78	8.02	3.94
Observations	883	237	127
1990 Cohort			
No A-level	0.016	0.118	0.149
Mean A-level score (out of 30)	24.32 (4.80)	19.97 (5.84)	14.21 (6.10)
Quartile 1	3.02	18.64	43.75
Quartile 2	12.67	32.27	37.37
Quartile 3	34.99	30.00	14.76
Quartile 4	49.32	19.09	4.12
Observations	663	220	752
1995 Cohort			
No A-level	0.069	0.080	0.303
Mean A-level score (out of 30)	22.71 (4.29)	22.84 (4.04)	17.00 (3.84)
Quartile 1	6.92	8.02	30.31
Quartile 2	8.08	5.36	39.96
Quartile 3	41.94	44.68	25.36
Quartile 4	43.06	41.94	4.37
Observations	515	902	1143

Note: The mean A-level is based on individuals with a positive score. In 1995, A level results are reported in a categorical variable, which are used rather than quartile. To calculate the means, we use the band mid-point. Individuals with no A-levels are allocated to the 1st quartile.

Table 3: Distribution of hourly and annual pay (in 2002 price) by institution type

	Russell Group Universities		Old Universities		Modern Universities	
1985 Cohort	Hourly pay	Annual wage	Hourly pay	Annual wage	Hourly pay	Annual wage
10th percentile	9.86	24,760	8.12	19,000	9.34	21,881
50th percentile	16.03	35,125	12.99	28,215	13.56	31,670
90th percentile	26.58	69,098	22.15	51,823	22.15	69,098
Mean	16.93	39,659	14.20	32,668	14.60	35,011
	(6.38)	(15,443)	(5.87)	(13,716)	(5.26)	(15,035)
Observations	883		237		127	
1990 Cohort						
10th percentile	7.47	15,547	6.96	15,547	6.77	15,547
50th percentile	11.90	24,760	10.98	24,760	10.58	24,760
90th percentile	18.46	43,186	16.00	35,125	15.71	35,125
Mean	12.64	29,220	11.78	26,574	11.29	25,564
	(4.99)	(12,425)	(5.14)	(11,541)	(4.40)	(9,912)
Observations	663		220		752	
1995 Cohort	Annual wage		Annual wage		Annual wage	
10th percentile	14,531		14,531		11,839	
50th percentile	20,989		20,989		20,989	
90th percentile	37,134		33,905		30,676	
Mean	23,962		22,706		21,908	
	(9,679)		(9,329)		(8,588)	
Observations	502		870		1116	

Note: Gross wages are expressed in GBP (price 2002). For all cohorts, annual gross wage is reported in categorical form.

Table 4: Linear estimates of institutional effects on gross wage

	Specification (1)	Specification (2)	Specification (3)	Specification (4)
1985 Cohort				
RG University	0.111	0.038	0.052	0.047
	(0.026)	(0.023)	(0.028)	(0.029)
Old University	0.027	-0.010	0.003	0.004
	(0.028)	(0.031)	(0.032)	(0.032)
R^2	0.30	0.32	0.36	0.36
1990 Cohort				
RG University	0.091	0.046	0.049	0.044
	(0.024)	(0.032)	(0.035)	(0.035)
Old University	0.071	0.044	0.033	0.032
	(0.046)	(0.043)	(0.036)	(0.035)
R^2	0.27	0.27	0.32	0.33
1995 Cohort				
RG University	0.124	0.111	0.101	0.094
	(0.031)	(0.036)	(0.031)	(0.030)
Old University	0.084	0.071	0.070	0.064
	(0.028)	(0.031)	(0.030)	(0.028)
R^2	0.25	0.25	0.31	0.32

Note: Hourly wages are used for the 1985 and the 1990 cohort. Model (1) controls for a quadratic function in labour market experience, firm size, type of contract self-employment, race, region of residence and post-graduate qualifications. In Model (2) we add A-level scores. Model (3) is similar to (2) but also includes subject of graduation and degree grade. The full model (4) adds controls for type of school attended (cohort 1985 and 1990) or visit to information services (cohort 1995) and father's occupation. Standard errors are reported in parentheses.

Table 5: Matching estimates of university quality on gross wage

	Russell Group		Old	
	Calliper 0.1	Calliper 0.01	Calliper 0.1	Calliper 0.01
Cohort 1985				
Nearest Neighbour	-0.006	-0.006	0.003	-0.002
	(0.100)	(0.100)	(0.076)	(0.066)
Kernel	0.060	0.011	-0.072	0.061
	(0.061)	(0.087)	(0.058)	(0.061)
OLS	0.008	0.027	-0.018	-0.007
	(0.044)	(0.046)	(0.057)	(0.065)
Nbr Treated	883	846	237	208
% matched treated	100	96	100	88
Nbr of controls for 50% match	5	6	13	16
Cohort 1990				
Nearest Neighbour	0.088	0.056	0.003	0.009
	(0.039)	(0.036)	(0.049)	(0.047)
Kernel	0.040	0.043	0.020	0.014
	(0.028)	(0.031)	(0.035)	(0.036)
OLS	0.047	0.053	-0.003	0.007
	(0.034)	(0.034)	(0.043)	(0.043)
Nbr Treated	663	618	219	210
% matched treated	100	93	100	95
Nbr of controls for 50% match	14	17	36	37
Cohort 1995				
Nearest Neighbour	0.042	-0.008	0.068	-0.004
	(0.099)	(0.089)	(0.095)	(0.095)
Kernel	0.074	0.058	0.039	0.021
	(0.041)	(0.041)	(0.035)	(0.038)
OLS	0.102	0.091	0.096	0.014
	(0.066)	(0.069)	(0.066)	(0.061)
Nbr Treated	502	491	870	865
% matched treated	100	98	100	99
Nbr of controls for 50% match	9	9	8	8

Note: The comparison group is drowned out of graduates from New Universities. The heading Calliper refers to the size of the calliper in nearest neighbour match or the size of the bandwidth in the case of kernel match. Kernel estimates are obtained using Epanechnikov kernel. Standard errors are calculated by bootstrap (500 replications). OLS uses the full control specification on the population of matched treated and control observations.

Table 6: Matching estimates of graduating from a Russell Group institution on wage growth

	Men	
	Estimate	% matched[A] (50%)[B] Obs[C]
1985 Cohort		
Growth 1986-1991		
Nearest Neighbour Cal 0.1	0.037 (0.316)	100% (5)
Kernel Epan. Band 0.1	0.292 (0.117)	682
Growth 1991-1996		
Nearest Neighbour Cal 0.1	0.047 (0.117)	100% (5)
Kernel Epan. Band 0.1	0.001 (0.083)	834
1990 Cohort		
Growth 1991-1996		
Nearest Neighbour Cal 0.1	0.071 (0.180)	100% (15)
Kernel Epan. Band 0.1	0.031 (0.128)	462

Note: The comparison group is drowned out of graduates from Modern Universities. Standard error calculated by bootstrap (500 replications). Kernel estimates are obtained using Epanechnikov kernel.
[A] Percentage of treated observations matched to a control observation (nearest neighbour)
[B] Number of control observations responsible for 50% of the matches.
[C] Number of treated observations