Heterogeneity, Selection and Advantage in the Graduate and Non-Graduate Labour Market

Cathy Redmond

January 10, 2014

Abstract

Recent econometric developments have allowed the full distribution of causal treatment effects to be estimated under certain conditions. This paper uses the marginal treatment effect (MTE) approach to estimate the distribution of monetary benefits of higher education in the United Kingdom, providing evidence on the shape of the MTE curve for females for the first time and providing additional evidence on the shape of the MTE curve for males. Significant heterogeneity in the returns from higher education were found due to observables characteristics, with weaker evidence on heterogeneity from unobservables. The paper also analyses whether (non) graduates have an advantage in the (non) graduate labour market, providing evidence on whether a two-skill market exists. This paper measured the extent to which individuals select into higher education on the basis of their heterogeneous returns, with results suggesting that in fact negative selection exists, with individuals who stand to benefit the most from higher education not attending. This is largely explained through observed ability; highly able individuals earn more in the non-graduate labour market, but earn similar incomes compared to less able individuals in the graduate labour market. Since highly able individuals are more likely to attend college, this explains why you find negative selection estimates. Highly able individuals give up more (in the non-graduate labour market) to earn the same (in the graduate labour market). Possible explanations for these surprising findings are discussed.

This research was supported by the Centre for Microdata Methods and Practise (CEMMAP). I would like to thank my supervisors, Richard Blundell and Adam Rosen for discussions, and Jon Johnson at the Centre for Longitudinal Studies for generously giving his time in facilitating access to the NCDS data.

1 Introduction

The marginal treatment effect (MTE) approach to programme evaluation allows for heterogeneity in the response to treatment, and under weak assumptions identifies the distribution of heterogeneous treatment effects. This approach combines advantages of both the reduced form approach and the structural method approach, and knowledge of the distribution of marginal treatment effects allows for estimation of a range of treatment effect parameters of interest (Heckman and Vytlacil, 1999). This paper applies the MTE approach to estimate the distribution of monetary benefits from higher education in the United Kingdom.

While most research analysing the returns to education uses a reduced form approach, recently a series of papers (Heckman and Vytlacil, 1999, 2001b,c,a, 2005, 2007; Heckman et al., 2006; Carneiro et al., 2010, 2011a) have developed and applied a method of estimating treatment effects that falls somewhere in the middle of the two extremes of reduced form and structural methods. This approach allows for the estimation of a distribution of treatment effects which can be used to estimate various parameters of interest.[1] This middle ground approach to economic estimation falls into the 'sufficient statistic approach' overviewed in Chetty (2009). This paper aims to apply the marginal treatment effect (MTE) approach to estimate the distribution of the impact of college education on labour income in the United Kingdom.

Recent empirical applications of the MTE approach have found a positive relationship between the return to college and values of the unobservables which make it more likely an individual attends college, both in the United States (Carneiro et al., 2010) and in the United Kingdom (Moffitt, 2008). A positive relationships between the return to college and values of the unobservables which makes college attendance more likely may be interpreted as follows; individuals with values of the unobservables that make it more likely they attend college might be individuals with higher unobserved ability, and these high ability individuals may benefit more from participation in higher education than individuals with lower levels of unobserved ability. However, unlike previous empirical work, this paper finds little evidence of any positive relationship between the return to college and values of the unobservables which makes college attendance more likely in the United Kingdom. However, evidence of heterogeneous returns due to observable characteristics were found, surprisingly, with individuals who stand to gain the most from higher education choosing not to attend. Finally, when earnings levels were compared in the graduate/non-graduate labour market, there does not seem to be significant differences in the potential earnings of graduates compared to non-graduates in either sector.

This paper proceeds as follows: section 2 discusses the MTE approach, section 3 describes the data set used in this analysis, section 4 describes the estimation and results and section 5 concludes.

[1] Under certain conditions, outlined in section 2, MTE can be used to identify treatment parameters including the average treatment effect (ATE), the average treatment effect on the treated (ATT), the average treatment effect on the non-treated (ATNT), the local average treatment effect (LATE), the policy relevant treatment effect (PRTE) and the marginal policy relevant treatment effect (MPRTE)

2 Marginal Treatment Effects

Björklund and Moffitt (1987) first introduced the concept of a marginal treatment effect by identifying a treatment effect for marginal individuals in a normal selection model. Heckman and Vytlacil (1999, 2001b,c,a, 2005, 2007); Heckman et al. (2006); Carneiro et al. (2010, 2011a) develop this approach, allowing for identification of MTE under more general specifications. Furthermore, Heckman and Vytlacil (1999) show that commonly reported treatment effects (ATE, ATT, LATE) can be expressed as a weighted average of the MTE. When the full distribution of MTE cannot be estimated, often it is possible to estimate bounds of the treatment effects (Heckman and Vytlacil, 1999, 2001b). In addition, the MTE approach has facilitated the definition and estimation of a policy relevant treatment effect (PRTE), (Heckman and Vytlacil, 2001b). Rather than using commonly reported treatment effect parameters to try and answer policy relevant questions, this approach advocates estimating the treatment effect that would result due to implementation of the new policy. A marginal policy relevant treatment effect (MPRTE) has also been defined by Carneiro et al. (2010, 2011a), and aims to estimate the impact of a marginal increase/decrease in the currently implemented policy.

The MTE was introduced initially in the context of binary treatment models. More recent papers, Heckman et al. (2006) and Heckman and Vytlacil (2007), extend this binary treatment case to models with more than two treatments, however this paper makes use only of the binary treatment case. The MTE approach is founded on a potential outcome model combined with a latent variable model of treatment choice. Under the potential outcome model, each individual is associated with two potential outcomes Y_{1i}/Y_{0i} which is the outcome that would be observed had individual i received/not received treatment. The treatment we are considering in this analysis is whether or not an individual has obtained a college degree. Let $C_i = 1$ represent college education and let $C_i = 0$ represent less than college education, therefore the outcome Y_i can be modelled as follows:

$$Y_i = C_i Y_{1i} + (1 - C_i) Y_{0i}$$

Let the potential outcomes be modelled:

$$Y_{1i} = \mu_1(Xa_i, Xb_i, U_{1i})$$

$$Y_{0i} = \mu_0(Xa_i, Xb_i, U_{0i})$$

Where Xa and Xb are vectors of observed characteristics determined outside of the model. Xa is the set of observed characteristics that will enter both the college choice model and the potential outcome model, whereas the set Xb enter only the potential outcome model (either vector can be null). This treatment is somewhat non-standard, however since U_{1i} and U_{0i} are unobserved random variables. Since this is a model with non-separable errors, continuous, discrete and ordered outcome variables can all be modelled within this framework.

Much of the following analysis rests on the assumption of a latent variable/index model for selection into treatment. Assume the following model for selection into college education:

$$C_i^* = \mu_c(Xa_i, Z_i) - Uc_i$$
$$C_i = 1[C_i^* > 0]$$
$$\Rightarrow C_i = 1[\mu_c(Xa_i, Z_i) - Uc_i > 0]$$

Where Z_i is a non null vector of exogeneous instruments, of which at least one element is required to be continuous. It is useful to interpret $\mu_c(Xa_i, Z_i)$ as a measure of the ease of attending college, since those with high values of $\mu_c(Xa_i, Z_i)$ tend to go to college. For instance, in the empirical specification used in this paper, distance to nearest university at age 16 is one of the instruments used, and the identification strategy rests on the intuitive assumption that the nearer you live to university the easier it is to attend college, for example, due to reduced transport costs or living costs. Uc_i is an unobserved random variable that impacts upon the treatment choice decision. It might be useful to interpret Uc_i as an unobservable measure of inability, or an unobservable measure of the cost of effort, as those individuals with high values of Uc_i tend not to go to college.

Given the notation defined above, we can now define the marginal treatment effect. The marginal treatment effect at a point $(Xa = xa, Xb = xb, Uc = uc)$ is defined:

$$MTE(xa, xb, uc) = E[Y_1 - Y_0 | Xa = xa, Xb = xb, Uc = uc]$$

The marginal treatment effect can be interpreted in a number of ways (Heckman and Vytlacil, 2005). MTE(xa, xb, uc) can be interpreted as the average treatment effect for an individual who has observable characteristics $(Xa = xa, Xb = xb)$ and unobservable costs of attending college of $Uc = uc$. Alternatively, the MTE(xa, xb, uc) can be interpreted as the average treatment effect of an individual who has observable characteristics $(Xa = xa, Xb = xb)$ and who would be indifferent between attending college and not if they were randomly assigned instrument value $Z = z$, where $\mu_c(xa, z) = uc$.

In order to derive an estimator for the MTE, the following assumptions as stated in Heckman and Vytlacil (2005) are imposed:

A1: $\mu_c(Xa, Z)$ is a nondegenerate random variable conditional on Xa, Xb
A2: The random vectors (U_1, Uc) and (U_0, Uc) are independent of Z conditional on Xa, Xb
A3: The distribution of Uc is absolutely continuous with respect to Lebesgue measure
A4: Both potential outcomes Y_1 and Y_0 have finite first moments
A5: $1 > P(C = 1|Xa = xa, Xb = xb) > 0$ for all $(xa, xb) \in \Omega(Xa, Xb)$, where $\Omega(.)$ denotes support

The propensity score, or the probability that an individual with observed characteristics $(Z = z, Xa = xa, Xb = xb)$ attends college is given by:

$$\begin{aligned} P(C = 1|Xa = xa, Xb = xb, Z = z) &= P(Uc < \mu_c(Xa, Z)|Xa = xa, Xb = xb, Z = z) \\ &= F_{Uc|Xa,Xb,Z}(\mu_c(Xa, Z)) \end{aligned}$$

Also, notice we can rewrite the selection into college equation as follows:

$$C = 1[F_{Uc|Xa,Xb,Z}(\mu_c(Xa, Z)) - F_{Uc|Xa,Xb,Z}(Uc) > 0]$$

Following the notation in Carneiro and Lee (2009), define $V = F_{Uc|Xa,Xb,Z}(Uc)$ and $P = F_{Uc|Xa,Xb,Z}(\mu_c(Xa, Z))$, where P is the propensity score. Therefore we can write the selection into college equation as:

$$C = 1[V < P]$$

Therefore, an individual with $V < P$ attends college, and an individual with $V = P$ is just indifferent between attending college or not. The MTE distribution will now be defined over (Xa, Xb, V), rather than over (Xa, Xb, Uc). Later we will use the fact that the distribution of V conditional on $(Xa, Xb, Z) \sim Unif[0, 1]$. To see this notice that for any $a \in [0, 1]$

$$\begin{aligned} P[F_{Uc|Xa,Xb,Z}(Uc) < a|Xa, Xb, Z] &= P[Uc < F^{-1}_{Uc|Xa,Xb,Z}(a)|Xa, Xb, Z] \\ &= F_{Uc|Xa,Xb,Z}(F^{-1}_{Uc|Xa,Xb,Z}(a)) \\ &= a \end{aligned}$$

The next step in deriving an estimator of the MTE is to write the expectation of the outcome, conditional on the observables (Xa, Xb) and conditional on the propensity score P.

$$\begin{aligned} E[Y|Xa = xa, Xb = xb, P = p] &= E[Y_0|xa, xb, p] \\ &+ E[Y_1 - Y_0|xa, xb, p, C = 1] * Pr[C = 1|xa, xb, p] \end{aligned}$$

Or,

$$\begin{aligned} E[Y|Xa = xa, Xb = xb, P = p] &= E[\mu_0(xa, xb, uc)|xa, xb, p] \\ &+ E[\mu_1(xa, xb, uc) - \mu_0(xa, xb, uc)|xa, xb, p, C = 1] * Pr[C = 1|xa, xb, p] \end{aligned}$$

Since (U_1, Uc) and (U_0, Uc) are independent of Z conditional on Xa, Xb, it is also true that (U_1, Uc) and (U_0, Uc) are independent of any function of Z conditional on Xa, Xb. P is simply a function of

Z once you have already conditioned on (Xa, Xb). Additionally, plugging in the condition for selection into college we can rewrite the above as:

$$E[Y|Xa = xa, Xb = xb, P = p] = E[\mu_0(xa, xb, uc)|xa, xb] \\ + E[\mu_1(xa, xb, uc) - \mu_0(xa, xb, uc)|xa, xb, V < p] * Pr[V < p|xa, xb]$$

Notice this can also be written as:

$$E[Y|Xa = xa, Xb = xb, P = p] = E[\mu_0(xa, xb, uc)|xa, xb] \\ + \frac{\int_0^p E[\mu_1(xa, xb, uc) - \mu_0(xa, xb, uc)|xa, xb, V = p']f(V = p'|xa, xb)dp' * [\int_0^p f(V = p'|xa, xb)dp']}{\int_0^p (V = p'|xa, xb)dp'}$$

Cancelling above and below the line and substituting for $f(V = p'|xa, xb)$:

$$E[Y|Xa = xa, Xb = xb, P = p] = E[\mu_0(xa, xb, uc)|xa, xb] \\ + \int_0^p E[\mu_1(xa, xb, uc) - \mu_0(xa, xb, uc)|xa, xb, V = p'] \int_{\mathbb{Z}} f(V = p'|xa, xb, z)f(z|xa, xb)dzdp'$$

Since V is uniform conditional on (Xa, Xb, Z) this simplifies to:

$$E[\mu_0(xa, xb, uc)|xa, xb] + \int_0^p E[\mu_1(xa, xb, uc) - \mu_0(xa, xb, uc)|xa, xb, V = p']dp'$$

If we now take the derivative of this conditional expectation with respect to $P = p$ we are left with an estimator for the MTE at the point $(xa, xb, V = p)$.

$$\begin{aligned} \frac{\partial E[Y|Xa = xa, Xb = xb, P = p]}{\partial p} &= E[\mu_1(xa, xb, uc) - \mu_0(xa, xb, uc)|xa, xb, V = p] \\ &= E[Y_1 - Y_0|Xa = xa, Xb = xb, V = p] \\ &= MTE(Xa = xa, Xb = xb, V = p) \end{aligned}$$

For every (Xa, Xb) combination, V ranges from 0 to 1, but the range over which $MTE(xa, xb, V)$ can be estimated depends on the support of P conditional on (Xa, Xb), and can only be estimated at points of continuity of P. Heckman and Vytlacil (1999), Heckman and Vytlacil (2001b) and Heckman and Vytlacil (2005) show that various treatment parameters can be represented as weighted averages of the MTE(Xa,Xb,V) (e.g. the ATE(Xa,Xb), TT(Xa,Xb), LATE(Xa,Xb), TUT(Xa,Xb), PRTE(Xa,Xb), IV(Xa,Xb) and OLS(Xa,Xb)). The support of P determines which of these treatment parameters can be estimated from the MTE estimates, see Heckman and Vytlacil (1999) and Heckman and Vytlacil (2001b) for further discussion.

3 Estimation

Estimation of the MTE generally requires assumptions in addition to A1-A5. This is because estimation of the propensity score and the mean outcome conditional on (Xa, Xb, Z) suffers from the curse of dimensionality in finite samples, particularly when the conditioning set contains continuous variables. In addition, non-parametric estimation of the MTE imposes heavy requirements on the support of the propensity score conditional on (Xa, Xb). Figure 1 shows the support of the propensity score conditional on a linear index of (Xa, Xb), and it can be seen that over some ranges of the index the support of the propensity score is quite small. A number of different approaches have been followed in the literature to avoid these issues. Typically, the propensity score is estimated parametrically in the first step, using a logit or probit model (e.g. Carneiro et al. (2010), Carneiro et al. (2011b), Moffitt (2008)), however Carneiro and Lee (2009) discuss semiparametric estimation of the propensity score using a partially linear additive model. In addition, it is often assumed that the potential outcomes are linear in parameters and have separable errors, and the independence assumption (A2) is often strengthened to full independence; (U_1, Uc) and (U_0, Uc) are independent of (Xa, Xb, Z).

Given these assumptions, a number of different estimation approaches have been followed in the literature. In the following, 3 approaches are discussed. When normality is not assumed the MTE distribution can be estimated via Robinson's partially linear model or sieve estimation. Alternatively, if joint normality of the errors are assumed the MTE distribution can be estimated using the normal selection model.[2]

3.1 Robinson's Partially Linear Model

In the first stage of all three approaches the propensity score is estimated typically using a logit or probit model.

$$Y_0 = (Xa\ Xb)\beta_0 + U_0$$
$$Y_1 = (Xa\ Xb)\beta_1 + U_1$$

Therefore the expectation of Y conditional on (Xa, Xb, P) can be written as:

$$\begin{aligned} E[Y|Xa = xa, Xb = xb, P = p] = &(xa\ xb)\beta_0 + E[U_0|xa, xb, p] \\ &+ (xa\ xb)(\beta_1 - \beta_0)P(C = 1|xa, xb, p) \\ &+ E[U_1 - U_0|xa, xb, p, C = 1]P(C = 1|xa, xb, p) \end{aligned}$$

Which, due to the stronger independence assumption simplifies to:

$$\begin{aligned} E[Y|Xa = xa, Xb = xb, P = p] = &(xa\ xb)\beta_0 + E[U_0] \\ &+ (xa\ xb)(\beta_1 - \beta_0)P(V < p) + E[U_1 - U_0|V < p]P(V < p) \end{aligned}$$

Also, since $V = F_{Uc|Xa,Xb,Z}(Uc)$ was uniformly distributed on [0,1] conditional on (Xa, Xb, Z), and since U_c is now independent of (Xa, Xb, Z), this implies that $V = F_{Uc}(Uc)$, which is now unconditionally uniform. Also, denoting $E[U_1 - U_0|V < p]P(V < p)$ as K(p) the above can be rewritten as:

$$E[Y|Xa = xa, Xb = xb, P = p] = (xa\ xb)\beta_0 + E[U_0] + (xa\ xb)(\beta_1 - \beta_0)p + K(p)$$

Splitting out the constant terms (since Xa,Xb contains a constant) rewrite the above as

$$E[Y|Xa = xa, Xb = xb, P = p] = \alpha_0 + (\tilde{Xa}\ \tilde{Xb})\tilde{\beta_0} + E[U_0] + (\alpha_1 - \alpha_0)P + (\tilde{Xa}\ \tilde{Xb})(\tilde{\beta_1} - \tilde{\beta_0})P + K(P)$$

[2] see Brinch et al. (2012) for estimation of MTE with finite support of the propensity score, and Carneiro and Lee (2009) for estimation of distributions of potential outcomes under the above assumptions

Where α_0, α_1 represent the constant coefficient in the potential outcome models, $\tilde{\beta_0}, \tilde{\beta_1}$ represent the remaining coefficients in the potential outcome models and where $\tilde{Xa}\,\tilde{Xb}$ represents the variables excluding the constant.

Taking the derivative with respect to p we get:

$$\begin{aligned}\frac{\partial E[Y|Xa = xa, Xb = xb, P = p]}{\partial p} &= (\alpha_1 - \alpha_0) + (\tilde{xa}\,\tilde{xb})(\tilde{\beta_1} - \tilde{\beta_0}) + K'(p)\\ &= MTE(xa, xb, V = p)\end{aligned}$$

We have that

$$\begin{aligned}&E[Y|Xa = xa, Xb = xb, P] = \alpha_0 + (\tilde{Xa}\,\tilde{Xb})\tilde{\beta_0} + E[U_0] + (\alpha_1 - \alpha_0)P + (\tilde{Xa}\,\tilde{Xb})(\tilde{\beta_1} - \tilde{\beta_0})P + K(P)\\ &\Rightarrow Y = \alpha_0 + (\tilde{Xa}\,\tilde{Xb})\tilde{\beta_0} + E[U_0] + (\alpha_1 - \alpha_0)P + (\tilde{Xa}\,\tilde{Xb})(\tilde{\beta_1} - \tilde{\beta_0})P + K(P) + \epsilon\end{aligned}$$

Where $E[\epsilon|Xa, Xb, P] = 0$

Taking expectations conditional on the propensity score:

$$\begin{aligned}&E[Y|P] = \alpha_0 + E[\tilde{Xa}\,\tilde{Xb}|P]\tilde{\beta_0} + E[U_0] + (\alpha_1 - \alpha_0)P + E[\tilde{Xa}\,\tilde{Xb}|P](\tilde{\beta_1} - \tilde{\beta_0})P + K(P)\\ &\Rightarrow Y - E[Y|P] = ((\tilde{Xa}\,\tilde{Xb}) - E[\tilde{Xa}\,\tilde{Xb}|P])\tilde{\beta_0} + ((\tilde{Xa}\,\tilde{Xb}) - E[\tilde{Xa}\,\tilde{Xb}|P])(\tilde{\beta_1} - \tilde{\beta_0})P + \epsilon\end{aligned}$$

In order to estimate $\tilde{\beta_0}$ and $\tilde{\beta_1} - \tilde{\beta_0}$, $E[W|P = \hat{p}]$ is calculated nonparametrically for $W = (Y, \tilde{Xa}, \tilde{Xb})$, (for instance using kernel estimation with cross validated bandwidth), using the fitted p̂ calculated in the first step. This is then used to calculate $W - E[W|P = \hat{p}]$. Then $Y - E[Y|P = \hat{p}]$ is regressed on $(\tilde{Xa}\,\tilde{Xb}) - E[\tilde{Xa}\,\tilde{Xb}|P = \hat{p}]$ and $((\tilde{Xa}\,\tilde{Xb}) - E[\tilde{Xa}\,\tilde{Xb}|P = \hat{p}]) * \hat{p}$, dropping a small fraction of the data for which there is low estimated density of p̂.

The K(P) is estimated in an extension of Robinson's partially linear model developed in Heckman et al. (1998). This approach is as follows:

$$Y - (\tilde{Xa}\,\tilde{Xb})\tilde{\beta_0} - (\tilde{Xa}\,\tilde{Xb})(\tilde{\beta_1} - \tilde{\beta_0})P = E[U_0] + (\alpha_1 - \alpha_0)P + K(P) + \epsilon$$

Using P̂ estimated in the first step, and $\hat{\tilde{\beta_0}}, (\widehat{\tilde{\beta_1} - \tilde{\beta_0}})$ estimated in the second step, the residual $Y - (\tilde{Xa}\,\tilde{Xb})\hat{\tilde{\beta_0}} - (\tilde{Xa}\,\tilde{Xb})(\widehat{\tilde{\beta_1} - \tilde{\beta_0}})\hat{P}$ is computed. Call this residual $(E[U_0] + (\alpha_1 - \alpha_0)P + K(P) + \epsilon)$ $L(\hat{P})$. The $L(\hat{P})$ function is then estimated nonparametrically by analysing the relationship between this residual and P̂, using, for example, locally quadratic regression with cross validated bandwidth. L'(P̂) is then estimated either using analytical or numerical first order differentiation of the estimated L(P) function. This derivative estimates $(\alpha_1 - \alpha_0) + K'(\hat{P})$.

Recall $MTE(xa, xb, V = p) = (xa\;xb)(\beta_1 - \beta_0) + K'(p)$. Therefore, combining the above we can estimate

$$\begin{aligned}\hat{MTE}(Xa = xa, Xb = xb, V = \hat{p}) &= (\tilde{xa}\,\tilde{xb})(\widehat{\tilde{\beta_1} - \tilde{\beta_0}}) + \hat{L}'(\hat{p})\\ &= (\tilde{xa}\,\tilde{xb})(\widehat{\tilde{\beta_1} - \tilde{\beta_0}}) + (\alpha_1 - \widehat{\alpha_0 + K'}(\hat{p}))\end{aligned}$$

Typically, what is reported in empirical work is MTE(V=p), which is the average marginal treatment effect for those individuals with V=p. Since Uc is independent of (Xa,Xb,Z), and $V = F_{Uc|Xa,Xb,Z}(Uc)$, this implies V is independent of (Xa,Xb,Z), and since $L'(P)$ is only a function of P, the average MTE can be derived as follows:

$$MTE(V=\hat{p}) = E_{Xa,Xb|V}[E[Y_1 - Y_0|Xa, Xb, V=\hat{p}]]$$
$$\Rightarrow MTE(V=\hat{p}) = E[(\tilde{X}a\ \tilde{X}b)|V=\hat{p}](\tilde{\beta}_1 - \tilde{\beta}_0) + L'(\hat{p})$$
$$\Rightarrow M\hat{T}E(V=\hat{p}) = (\hat{\bar{\tilde{X}}}a\ \hat{\bar{\tilde{X}}}b)(\widehat{\tilde{\beta}_1 - \tilde{\beta}_0}) + (\alpha_1 - \widehat{\alpha_0 + K'}(\hat{p}))$$

3.2 Sieve Method:

As in the previous approach the propensity score is estimated in the first stage. As derived above, we have:

$$E[Y|Xa=xa, Xb=xb, P=p] = (xa\ xb)\beta_0 + E[U_0] + (xa\ xb)(\beta_1 - \beta_0)p + K(p)$$

Again, splitting out the constants

$$E[Y|Xa=xa, Xb=xb, P=p] = \alpha_0 + (\tilde{X}a\ \tilde{X}b)\tilde{\beta}_0 + E[U_0] + (\alpha_1 - \alpha_0)P + (\tilde{X}a\ \tilde{X}b)(\tilde{\beta}_1 - \tilde{\beta}_0)P + K(P)$$

$$\Rightarrow Y = \alpha_0 + (\tilde{X}a\ \tilde{X}b)\tilde{\beta}_0 + E[U_0] + (\alpha_1 - \alpha_0)P + (\tilde{X}a\ \tilde{X}b)(\tilde{\beta}_1 - \tilde{\beta}_0)P + K(P) + \epsilon$$

Where $E[\epsilon|Xa, Xb, P] = 0$.

The sieve approach to estimating the MTE estimates the $K(P)$ using a flexible functional form such as series/spline specifications. For example, Moffitt (2008), estimating the MTE of college education in the UK estimates using a number of different $K(P)$ specifications; a linear function of the propensity score, a quadratic function, a cubic function, a quadratic function with a median spline break and a quadratic function with quartile spline breaks. Given the specific functional form assumed the model is estimated by regressing Y on $(Xa\ Xb), (Xa\ Xb)\hat{P}, K(\hat{P})$ where $\hat{P}$ is the estimated propensity score from the first step. Suppose for instance that a quadratic function is assumed for the $K(P)$ function[3]:

$$K(P) = \pi_0 + \pi_1 P + \pi_2 P^2$$

Therefore we have the following model

$$E[Y|Xa=xa, Xb=xb, P=p] = \alpha_0 + (\tilde{X}a\tilde{X}b)\tilde{\beta}_0 + E[U_0] + (\alpha_1 - \alpha_0)P + (\tilde{X}a\tilde{X}b)(\tilde{\beta}_1 - \tilde{\beta}_0)P + \pi_0 + \pi_1 P + \pi_2 P^2$$

And taking the derivative with respect to p we get:

$$\frac{\partial E[Y|Xa=xa, Xb=xb, P=p]}{\partial p} = (\alpha_1 - \alpha_0) + (\tilde{xa}\ \tilde{xb})(\tilde{\beta}_1 - \tilde{\beta}_0) + \pi_1 + 2\pi_2 p$$
$$= MTE(xa, xb, V=p)$$

Estimate the following by OLS:

$$Y = \theta_0 + (\tilde{X}a\ \tilde{X}b)\beta_0 + \theta_1 P + (\tilde{X}a\ \tilde{X}b)(\tilde{\beta}_1 - \tilde{\beta}_0)P + \pi_2 P^2 + \epsilon$$

Where $\theta_0 = (\alpha_0 + E[U_0] + \pi_0)$, and $\theta_1 = (\alpha_1 - \alpha_0 + \pi_1)$. The MTE is then estimated using the estimated OLS coefficients from the above model as:

$$\frac{\partial E[Y|Xa=xa, Xb=xb, P=\hat{p}]}{\partial \hat{p}} = M\hat{T}E(Xa=xa, Xb=xb, V=\hat{p})$$
$$= \hat{\theta}_1 + (\tilde{X}a\ \tilde{X}b)(\widehat{\tilde{\beta}_1 - \tilde{\beta}_0}) + 2\hat{\pi}_2\hat{P}$$

[3]The K(P) specification can be chosen by least squares cross-validation as suggested in Belloni et al. (2011)

As before, the average MTE(V=p) can be estimated:

$$
\begin{aligned}
& MTE(V=\hat{p}) = E_{Xa,Xb|V}[E[Y_1 - Y_0|Xa, Xb, V=\hat{p}]] \\
& \Rightarrow MTE(V=\hat{p}) = \theta_1 + E[(\tilde{Xa}\;\tilde{Xb})|V=\hat{p}](\tilde{\beta}_1 - \tilde{\beta}_0) + 2\pi_2\hat{p} \\
& \Rightarrow M\hat{T}E(V=\hat{p}) = \hat{\theta}_1 + (\hat{\hat{\tilde{X}}}a\;\hat{\hat{\tilde{X}}}b)(\widehat{\tilde{\beta}_1 - \tilde{\beta}_0}) + 2\hat{\pi}_2\hat{p}
\end{aligned}
$$

3.3 Normal Selection Model

As in the previous two approaches the propensity score is estimated in the first stage. Using results from Heckman (1979) the MTE distribution can be estimated under the additional assumption of joint normality of the errors using normal selection models:

$$
\begin{aligned}
& Y_0 = (Xa\;Xb)\beta_0 + U_0 \\
& Y_1 = (Xa\;Xb)\beta_1 + U_1 \\
& C = 1[Uc_i < \mu_c(Xa_i, Z_i)] \\
& \begin{pmatrix} U_1 \\ U_0 \\ U_c \end{pmatrix} \Bigg| (Xa, Xb, Z) \sim N\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{10} & \sigma_{1c} \\ \sigma_{01} & \sigma_0^2 & \sigma_{0c} \\ \sigma_{c1} & \sigma_{c0} & 1 \end{pmatrix}\right]
\end{aligned}
$$

$$
\begin{aligned}
MTE(Xa, Xb, Uc) &= E[Y_1 - Y_0|Xa, Xb, Uc] \\
&= E[Y_1|Xa, Xb, Uc] - E[Y_0|Xa, Xb, Uc] \\
&= (Xa\;Xb)(\beta_1 - \beta_0) + E[U_1|Uc] - E[U_0|Uc]
\end{aligned}
$$

Where

$$
\begin{aligned}
& E[U_1|Uc] = \sigma_{1c}Uc \\
& E[U_0|Uc] = \sigma_{0c}Uc
\end{aligned}
$$

Also,

$$
\begin{aligned}
E[y_1|Xa, Xb, \mu_c(Xa_i, Z_i) = a, C = 1] &= E[y_1|Xa, Xb, Uc < a] \\
&= (Xa\;Xb)\beta_1 + E[U_1|Uc < a] \\
E[y_0|Xa, Xb, \mu_c(Xa_i, Z_i) = a, C = 0] &= E[y_0|Xa, Xb, Uc \geq a] \\
&= (Xa\;Xb)\beta_0 + E[U_0|Uc \geq a]
\end{aligned}
$$

Where

$$
\begin{aligned}
& E[U_1|Uc \leq a] = \sigma_{1c}E[Uc|Uc \leq a] \\
& E[U_0|Uc > a] = \sigma_{0c}E[Uc|Uc > a]
\end{aligned}
$$

$$
\begin{aligned}
E[Uc|Uc \leq \tilde{\mu_c}] &= \int_{x-\infty}^{a} ucf_{uc}(uc|uc \leq a)duc \\
&= \int_{-\infty}^{a} ucf_{uc}(uc)/P(uc \leq a)duc \\
&= \int_{-\infty}^{a} uc\phi(uc)/\Phi(a)duc \\
&= -\phi(a)/\Phi(a) \\
\Rightarrow E[U_1|Uc \leq a] &= \sigma_{1c} * (-\phi(a)/\Phi(a))
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
E[Uc|Uc > a] &= \int_{a}^{\infty} ucf_{uc}(uc|uc > a)duc \\
&= \int_{a}^{\infty} ucf_{uc}(uc)/P(uc > a)duc \\
&= \int_{a}^{\infty} uc\phi(uc)/(1-\Phi(a))duc \\
&= \phi(a)/(1-\Phi(a)) \\
\Rightarrow E[U_0|Uc \geq a] &= \sigma_{0c}(\phi(a)/\,(1-\Phi(a)))
\end{aligned}
$$

Note that $P \equiv P(C=1|Xa, Xb, Z) = \Phi(\mu_c(Xa, Z))$. And since conditioning on $\mu_c(Xa_i, Z_i) = a$ is equivalent to conditioning on $\Phi(\mu_c(Xa, Z)) = \Phi(a)$,

$$
\begin{aligned}
&E[y_1|Xa, Xb, P = \Phi(a), C = 1] = (Xa\;Xb)\beta_1 - \sigma_{1c}(\phi(\Phi^{-1}(P))/P) \\
&E[y_0|Xa, Xb, P = \Phi(a), C = 0] = (Xa\;Xb)\beta_0 + \sigma_{0c}(\phi(\Phi^{-1}(P))/(1-P))
\end{aligned}
$$

From a linear regression of the observed outcome for those who graduated from college on $(Xa, Xb, \phi(\Phi^{-1}(\hat{P}))/\hat{P})$, (β_1, σ_{1c}) can be estimated. From a linear regression of the observed outcome for those who did not graduate from college on $(Xa, Xb, \phi(\Phi^{-1}(\hat{P}))/(1-\hat{P}))$, (β_0, σ_{0c}) can be estimated. We can therefore estimate

$$
M\hat{T}E(Xa, Xb, Uc) = (Xa\;Xb)(\hat{\beta_1} - \hat{\beta_0}) + \hat{\sigma_{1c}}Uc - \hat{\sigma_{0c}}Uc
$$

or, defining over the unit interval for comparability,

$$
M\hat{T}E(Xa, Xb, V = \Phi(Uc)) = (Xa\;Xb)(\hat{\beta_1} - \hat{\beta_0}) + \hat{\sigma_{1c}}\Phi^{-1}(V) - \hat{\sigma_{0c}}\Phi^{-1}(V)
$$

As before, the average $MTE(V = \Phi(Uc))$ can be estimated:

$$
\begin{aligned}
&MTE(V = \hat{p}) = E_{Xa,Xb|V}[E[Y_1 - Y_0|Xa, Xb, V = \Phi(Uc)]] \\
&\Rightarrow MTE(V = \Phi(Uc)) = E[Xa\;Xb|V = \Phi(Uc)](\beta_1 - \beta_0) + \sigma_{1c}\Phi^{-1}(V) - \sigma_{0c}\Phi^{-1}(V) \\
&\Rightarrow M\hat{T}E(V = \Phi(Uc)) = (\hat{\bar{X}}a\;\hat{\bar{X}}b)(\hat{\beta_1} - \hat{\beta_0}) + \hat{\sigma_{1c}}\Phi^{-1}(V) - \hat{\sigma_{0c}}\Phi^{-1}(V)
\end{aligned}
$$

3.4 Estimating the Weights

Heckman and Vytlacil (1999), Heckman and Vytlacil (2001b) and Heckman and Vytlacil (2005) show that various treatment parameters can be estimated as weighted averages of the MTE. In the following estimation of the ATE, TT and TUT are discussed.

$$\begin{aligned}
ATE(xa,xb) &= \int_0^1 MTE(xa,xb,v)dv \\
TT(xa,xb) &= \int_0^1 MTE(xa,xb,v) f_{V|Xa,Xb}(v|Xa,Xb,C=1)dv \\
&= \int_0^1 MTE(xa,xb,v)\left(\left(\int_v^1 f(p|Xa=xa,Xb=xb)dp\right)\frac{1}{E(P|Xa=xa,Xb=xb)}\right)dv \\
TUT(xa,xb) &= \int_0^1 MTE(xa,xb,v) f_{V|Xa,Xb}(v|Xa,Xb,C=0)dv \\
&= \int_0^1 MTE(xa,xb,v)\left(\left(\int_0^v f(p|Xa=xa,Xb=xb)dp\right)\frac{1}{E((1-P)|Xa=xa,Xb=xb)}\right)dv
\end{aligned}$$

Estimation of the weights also suffer from the curse of dimensionality, since the density of p conditional on (Xa,Xb) has to be estimated. However, Carneiro et al. (2011b) propose a simple simulation method for estimating the weights under the assumptions imposed in this section. Since V is assumed independent of (Xa, Xb, Z), Carneiro et al. (2011b) suggest drawing a large number (n) of V observations from a uniform distribution for each individual observation, and then evaluating the $MTE(Xa, Xb, V)$ for each of the n draws for that individual. To estimate $ATE(Xa, Xb)$ for an individual, the average over the n estimated $MTE(Xa, Xb, V)$ is taken, to estimate $TT(Xa, Xb)$, the average over the n estimated $MTE(Xa, Xb, V)$ where $V < P$ is taken and to estimate $TUT(Xa, Xb)$, the average over the n estimated $MTE(Xa, Xb, V)$ where $V \geq P$ is taken.[4] In addition, the average ATE in the population ($E[Y_1 - Y_0]$) can be estimated by averaging over $ATE(Xa, Xb)$ in the sample. The average TT in the population ($\bar{TT} = E[Y_1 - Y_0|C = 1]$) can be estimated by averaging over $TT(Xa, Xb)$ for those individuals in the sample who go to college, and the average TUT in the population ($\bar{TUT} = E[Y_1 - Y_0|C = 0]$) can be estimated by averaging over $TUT(Xa, Xb)$ for those individuals in the sample who do not go to college.

3.5 Selection into College

Selection into college is said to occur if the average gain from attending college is greater for those who actually attend college than for those who do not attend college. In addition, the total selection gain can be decomposed into a component due to observable characteristics and a component due to unobservable characteristics. Subtracting $TT(Xa, Xb)$ from $TUT(Xa, Xb)$ at a particular value of (Xa, Xb) gives the difference between the gain from unobservables for individuals with that value of (Xa, Xb) who choose to go to college and the gain from unobservables for individual with that value of (Xa, Xb) who choose not to go to college.

$$TT(Xa, Xb) - TUT(Xa, Xb) = E[U_1 - U_0|Xa, Xb, C = 1] - E[U_1 - U_0|Xa, Xb, C = 0]$$

The difference between the average TT in the population and the average TUT in the population gives the total selection gain, which can be decomposed into a component owing to observable characteristics and a component owing to unobservable characteristics.[5]

$$\begin{aligned}
\bar{TT} - \bar{TUT} &= (E[(\tilde{Xa}\,\tilde{Xb})|C = 1] - E[(\tilde{Xa}\,\tilde{Xb})|C = 0])(\tilde{B_1} - \tilde{B_0}) \\
&+ (E[U_1 - U_0|C = 1] - E[U_1 - U_0|C = 0])
\end{aligned}$$

Where $(E[(\tilde{Xa}\,\tilde{Xb})|C = 1] - E[(\tilde{Xa}\,\tilde{Xb})|C = 0])(\tilde{B_1} - \tilde{B_0})$ is the selection component due to observables and $(E[U_1 - U_0|C = 1] - E[U_1 - U_0|C = 0])$ is the component due to unobservables. The selection component due to observables can be estimated since $(\tilde{B_1} - \tilde{B_0})$ can be estimated using any of the estimation methods discussed in the above, and by subtracting $(\hat{E}[(\tilde{Xa}\,\tilde{Xb})|C = 1] - \hat{E}[(\tilde{Xa}\,\tilde{Xb})|C = 0])(\widehat{\tilde{B_1} - \tilde{B_0}})$ from $\hat{\bar{TT}} - \hat{\bar{TUT}}$ the selection component due to unobservables can also be estimated.

[4] Semi-parametric estimation of the weights is discussed in Carneiro et al. (2010) footnote 11.

[5] NOTE TO SELF: This is the same for normal/selection/spline models in contrast to the advantage discussion below. See PhD folder for cases and graphs

3.6 Advantage

Given the model framework and assumptions, it is possible to compare whether graduates have an advantage over non-graduates in the graduate labour market, and this advantage can be decomposed into a component explained by differences in observable characteristics and a component due to differences in unobservable characteristics. It is also possible to see whether one group has an advantage over the other in the non-graduate labour market, and to test whether there is a two-skill labour market.

The mean wages of graduates in the graduate labour market can be estimated using observed wages, $E[Y_1|C=1]$, and the component due to unobservables can be estimated using

$$E[Y_1|C=1] - E[Xa, Xb|C=1]B_1 = E[U_1|C=1]$$

The mean wages of non-graduates in the non-graduate labour market can also be estimated using observed wages $E[Y_0|C=0]$, and the component due to unobservables can be estimated using

$$E[Y_0|C=0] - E[Xa, Xb|C=0]B_0 = E[U_0|C=0]$$

The mean wages of graduates in the non-graduate labour market can be estimated from

$$\begin{aligned} E[Y_0|C=1] &= -(E[Y_1 - Y_0|C=1] - E[Y_1|C=1]) \\ &= -(T\bar{T} - E[Y_1|C=1]) \end{aligned}$$

And the component due to unobservable can be estimated using

$$E[Y_0|C=1] - E[Xa, Xb|C=1]B_0 = E[U_0|C=1]$$

The mean wages of non-graduates in the graduate labour market can be estimated from

$$\begin{aligned} E[Y_1|C=0] &= E[Y_0|C=0] + E[Y_1 - Y_0|C=0] \\ &= E[Y_0|C=0] + T\bar{U}T \end{aligned}$$

And the component due to unobservable can be estimated using

$$E[Y_1|C=0] - E[Xa, Xb|C=0]B_1 = E[U_1|C=0]$$

The above could also be estimated conditional on X, using non-parametric/parametric models to estimate $E[Y_1|Xa, Xb, C=1]$, $E[Y_0|Xa, Xb, C=0]$, from which $E[U_1|X, C=0]$ and $E[U_0|X, C=1]$ can be estimated. Alternatively, if you impose the assumption that is required to interpret OLS as discussed above, you can estimate the advantage from unobservables in the two labour markets at a particular value of X by subtracting $TT(Xa, Xb) - OLS(Xa, Xb) = E[U_0|Xa, Xb, C=0] - E[U_0|Xa, Xb, C=1]$, and by subtracting $TUT(Xa, Xb) - OLS(Xa, Xb) = E[U_1|Xa, Xb, C=0] - E[U_1|Xa, Xb, C=1]$. Getting values for mean X is more difficult using this approach.

Graduates are said to have an advantage in the graduate labour market if

$$E[Y_1|C=1] > E[Y_1|C=0]$$

with the advantage from observables being

$$E[Xa, Xb|C=1]B_1 - E[Xa, Xb|C=0]B_1$$

and the advantage from unobservables being

$$E[U_1|C=1] - E[U_1|C=0]$$

Graduates are said to have an advantage in the non-graduate labour market if

$$E[Y_0|C=1] > E[Y_0|C=0]$$

with the advantage from observables being

$$E[Xa, Xb|C = 1]B_0 - E[Xa, Xb|C = 0]B_0$$

and the advantage from unobservables being

$$E[U_0|C = 1] - E[U_0|C = 0]$$

If graduates have an advantage in the graduate labour market and non-graduate in the non-graduate labour market we say there is a two-skill labour market, with individuals selecting into the sector in which they have an advantage.[6]

[6] If using spline/Robinson's partially linear method for estimating the MTE model, the approach for estimating advantage (differences) is as above, however, it is not possible to estimate the levels of the mean unobservables. This is because the constants in the potential outcome models are not identified. For example, in the sieve case,
$E[Y_1|C = 1] - \theta_1 - E[X|C = 1]B_1 = E[U_1|C = 1] - \alpha_0 + \pi_1$,
$E[Y_1|C = 0] - \theta_1 - E[X|C = 0]B_1 = E[U_1|C = 0] - \alpha_0 + \pi_1$.
Therefore, comparing the difference between the two residuals;
$E[Y_1|C = 1] - \theta_1 - E[X|C = 1]B_1 - E[Y_1|C = 0] - \theta_1 - E[X|C = 0]B_1 =$
$E[U_1|C = 1] - E[U_1|C = 0]$
; the advantage of graduates in the graduate labour market owing to unobservables. For Robinson's partially linear approach,
$E[Y_1|C = 1] - E[X|C = 1]\tilde{B_1} = E[U_1|C = 1] + \alpha_1$,
$E[Y_1|C = 0] - E[X|C = 0]\tilde{B_1} = E[U_1|C = 0] + \alpha_0$.
Therefore, comparing the difference between the two residuals;
$E[Y_1|C = 1] - E[X|C = 1]\tilde{B_1} - E[Y_1|C = 0] - E[X|C = 0]\tilde{B_1} =$
$E[U_1|C = 1] - E[U_1|C = 0]$

4 Data

This analysis uses the UK National Child Development Study (NCDS)[7] to estimate the distribution of marginal treatment effects from higher education for males and females separately in the United Kingdom. The NCDS follows a group of people born in England, Scotland and Wales in a single week in March, 1958. The initial wave of the survey collected information from the mothers of babies born in that week. These children were subsequently followed in eight sweeps; in 1965 (aged 7), in 1969 (aged 11) in 1974 (aged 16), in 1981 (aged 23), in 1991 (aged 33), in 1999/2000 (aged 42), in 2004 (aged 46) and finally in 2008 (aged 50) (exam data was additionally collected in 1978 (aged 20)). The later life outcomes considered in this analysis are those reported in the 1991 when the respondents were aged 33, and additional results are presented from the 2008 wave when the respondents were aged 50.

The outcome variable of interest is gross hourly wage for employed individuals (self-employed individuals are excluded from the data as this information is not available for this group). The main explanatory variable is whether or not the individual has graduated from college (either a university/polytechnic) with a minimum of an undergraduate degree.

The Xa variables described in the previous section include number of siblings, ability (as measured by the Douglas general ability test score measured in 1969 when the individuals were aged 11), mothers' years of education and fathers' years of education. Missing variable dummies are also included for number of siblings, ability, and mothers' and fathers' education. Additional Xb variables are included in the specification with the second set of instrumental variables (distance to nearest university/polytechnic and region at age 16). The Xb variables for this specification include region at age 33 (or age 50 for the 2008 wave), and distance to nearest pre-existing college at age 33 (or age 50 for the 2008 wave). Cameron and Taber (2004) highlight the importance of controlling for current labour market conditions when past local labour market conditions are used as excluded instruments. This is because past local labour market conditions may be correlated with current local labour market conditions or other regional conditions which might impact upon later life outcomes. Similarly, as there is some concern that the distance to nearest university/polytechnic instrument might be proxying urban/rural status, distance to nearest pre-existing university/polytechnic at age 33 (50) is included as an additional Xb control variable.

In total, there are 18,558 individual who answer at least one of the first three rounds of the NCDS. Table 1 shows labour market participation figures for males and females in 1991 and 2008. In 1991, 74.6% of males and 61.2% of females were in employment (excluding self-employment). The corresponding figures for 2008 were 88.1% of males and 80.1% of females being employed.

Descriptives of the Xa, Xb and Z variables for these individuals are shown in Tables 2 and 3, with Table 2 showing the descriptives for the sample when the first set of instrumental variables are used (older siblings, financial shock and parental interest), and Table 3 showing the descriptives for the sample when the second set of instrumental variables are used (distance and region). Hourly gross income is also displayed in these tables. In 1991 approximately 11% of males and 19% of females report having graduated from college. Men and women who graduated from college have higher levels of mother's and father's education on average. They also have higher levels of ability as measured by the Douglas general ability test score at age 11. They have slightly fewer siblings on average. Later in life, graduates are more likely to be located in London and the South East as compared to non-graduates, and live about 1-2km closer to the nearest pre-existing college than non-graduates. Those who attended college were more likely to be living in London compared to non-graduates at age 16, and lived approximately 3-4km closer on average to their nearest university and polytechnic at age 16. Average hourly gross wage was £10.22 for males graduates compared to £7.08 for non-graduates, and £7.66 for females graduates compared to £4.72 for non-graduates.[8]

[7]University of London. Institute of Education. Centre for Longitudinal Studies, National Child Development Study [computer files]. Colchester, Essex: UK Data Archive [distributor]

[8]2.5% of observations have been trimmed from both tails of income for entire analysis

5 Results

The MTE distribution is estimated using the normal selection model and sieve estimation as discussed in the previous section. The propensity score is estimated in the first step. Table 4 tabulates the estimated model for the propensity score using the first set of instrumental variables (older siblings, financial shock and parental interest), and Table 5 tabulates the estimated model for the propensity score using the alternative set. Log odds are reported. The first set of IVs was highly significant for all groups with the exception of Males in 2008. The second set were significant for females in both time periods, but not for males. For the rest of the analysis, the paper uses the first set of instrumental variables, as they provide significant variation in college graduation for both males and females (although not significantly for males in 2008).

Figure 1 shows the estimated MTE curves estimated under the assumption of joint normality of the unobservables (in both the outcome equations and the selection into college equation), with homogeneity, and with linear, quadratic, linear with a median spline break and linear with quartile spline breaks for Males in 1991. The corresponding coefficients are shown in Table 6. Insignificant heterogeneity estimates were found for both the normal and linear models. While significant coefficients were estimated for the quadratic and linear spline models, all three of these models lead to implausibly high estimates of the return to college education over some regions and so it is assumed these models are not well-specified.

Table 7 tabulates the (β_0) estimates, and Table 8 the ($\beta_1 - \beta_0$) estimates for Males 1991. If significant ($\beta_1 - \beta_0$) estimates are found then this is evidence of heterogeneity in the returns from higher education from observable characteristics. There is evidence of significant heterogeneity here. Of particular note is the negative ability estimate, indicating that high ability is associated with lower returns from higher education participation. Table 7 shows the return to observed ability in the non-graduate labour market is positive, which implies that the return to ability in the graduate labour market must be lower than the return to observed ability in the non-graduate labour market. Further inspection of the results suggest that the returns to observed ability in the graduate labour market are close to zero.

Table 9 shows the estimates of whether graduates have an advantage in the graduate labour market, decomposing the advantage into a component due to observable and a component due to unobservable characteristics. Interestingly, graduates were not found to have an advantage over non-graduates in the graduate labour market. Table 10 shows whether graduates have an advantage in the non-graduate labour market. In this case, there was some evidence that graduates do better in the non-graduate labour market, although the results depend to some extent on the model specification.

Table 10 presents the selection estimates, again, decomposing the selection term into a component due to observable and a component due to unobservable characteristics. The total selection term measures whether the return to higher education is higher for graduates compared to non-graduates. Surprisingly, the return to higher education was found to be lower for graduates, with significant negative differences found for both the OLS and constant specification. A large proportion of this negative selection is explained by observed ability. This result is interpreted as follows; high ability individuals earn more in the non-graduate labour market compared to low ability individuals. However, high ability individuals earn the same in the graduate labour market compared to low ability individuals. High ability individual, who are more likely to graduate from college, are therefore giving up more (earnings in non-graduate labour market) to earn the same (in the graduate labour market), which explains this negative selection result.

Various explanations for this negative selection finding were considered. One possibility considered was a model of compensating differentials. A rational individual chooses to participate in higher education if benefits > costs. There might be non-monetary costs and benefits associated with higher education, in addition to monetary costs and benefits. Since this paper found evidence of higher monetary benefits for non-graduates, this reasoning implies that graduates have higher levels of non-monetary benefits, lower monetary costs or lower non-monetary costs. Compensating differentials in the graduate labour market could explain why graduates might have higher non-monetary benefits of higher education participation. Suppose high ability graduates are maximising their utility by trading off some potential salary for nice jobs. Low ability individuals might not have this same option due to not meeting the threshold ability levels for some of these nice/high-prestige jobs. This could imply a higher non-monetary benefit of higher education for high versus low ability individuals. This potential explanation was tested by regressing positive job characteristics on the set of control variables, and the set of control variables interacted with graduate status. The job characteristics considered were whether or not the individual has firm shares/company car/travel benefits/subsidised meals/medical insurance/pension/childcare/discounts/other benefits provided by their work. Additionally, the total number of benefits was also regressed on the set of controls and interactions. Results of this analysis are shown in Table 12 and Table 13. No evidence was found of higher ability graduates having higher levels of positive job characteristics than graduates with lower ability.

Remarkably similar results were found for females in 1991, and for both genders in the 2008 wave of data. (Results shown in Figures 2-4 and Tables 14-33).

6 Discussions and Conclusion

This paper exploits a rich longitudinal data set to answer three related questions. Firstly, is there any evidence of heterogeneity in the returns from higher education, due to either unobservable or observable characteristics? Secondly, if heterogeneous returns exist, are individuals selecting into higher education based on this heterogeneity? Finally, do one group (graduates/non-graduates) have an advantage over the other in the graduate/non-graduate labour market?

The results presented in the previous section are somewhat surprising. The empirical evidence suggests that there is significant heterogeneity in the returns to education based on observable characteristics, in particular, with high ability individuals receiving a lower return to higher education than lower ability individuals. Since higher education participation is higher among high ability students, this creates negative selection into higher education, with individuals who attend receiving lower returns than the hypothetical returns the lower ability group would have received had they graduated. This finding is largely driven by the fact that there are positive returns to ability in the non-graduate labour market but low/negative returns to ability in the graduate labour market, implying that high ability graduates give up more (in the non-graduate labour market) to receive the same (in the graduate labour market), generating a lower return. There is no strong evidence that graduates have an advantage over non-graduates in either the graduate/non-graduate labour market. In addition, there is little evidence of heterogeneity in the returns to unobservables characteristics.

A rational individual chooses to participate in higher education if monetary & non-monetary benefits > monetary & non-monetary costs. This paper found evidence of higher monetary benefits for non-graduates, which implies that in a rational model, graduates must have higher levels of non-monetary benefits, lower monetary costs or lower non-monetary costs to explain why they attend college while non-graduates who would receive higher monetary returns do not attend. Compensating differentials in the graduate labour market was ruled out, as no evidence was found that high ability graduates have nicer job attributes than lower ability graduates. Alternative explanations for negative selection include lower monetary costs; perhaps non-graduates have higher opportunity costs than graduates. This could be the case if there was a two-skill labour market, however the empirical results (graduates have higher potential earnings in the non-graduate labour market) suggest that this is not the case. High ability individuals could have lower non-monetary costs, for instance, if they have a lower effort cost, or if there is a higher probability of course failure for low ability students. Finally, another possible explanation for this negative selection is that low ability students are facing barriers to entry. This could be indirectly though credit constraints for instance, or directly, though stringent college admission procedures.

Table 1: Labour Market Participation

	1991				2008			
	Male		Female		Male		Female	
Employed	90.51%		68.01%		68.23%		73.07%	
Employed excl. self employed	74.64%		61.24%		88.12%		80.66%	
	No HE	HE	No HE	HE	No HE	HE	No HE	HE
Employed	89.66%	95.71%	66.82%	77.58%	86.86%	93.26%	79.08%	87.38%
Employed excl. self employed	73.29%	82.83%	60.33%	68.52%	66.77%	74.18%	72.12%	77.09%

Table 2: Moffitt Descriptives

		Male 1991		Male 2008		Female 1991		Female 2008	
	College Graduate	0.11		0.16		0.19		0.23	
		HE	No HE	HE	No HE	HE	No HE	HE	No HE
Xa:	Ability	58.44	42.21	56.01	42.28	60.41	44.70	56.23	45.18
	Mother ed yrs	15.86	14.80	15.40	14.77	16.31	14.84	15.76	14.84
	Mother ed yrs m	0.02	0.02	0.03	0.02	0.01	0.02	0.03	0.01
	Father ed yrs	16.44	14.76	16.06	14.78	16.36	14.83	15.93	14.81
	Father ed yrs m	0.03	0.04	0.04	0.04	0.02	0.05	0.06	0.05
	Siblings	1.89	2.35	2.04	2.40	1.88	2.38	1.98	2.32
	North West 16	0.10	0.10	0.12	0.11	0.09	0.12	0.12	0.12
	Yorkshire and Humber 16	0.08	0.09	0.10	0.09	0.08	0.08	0.08	0.09
	East Midlands 16	0.06	0.08	0.06	0.08	0.05	0.08	0.04	0.08
	West Midlands 16	0.12	0.10	0.12	0.09	0.08	0.10	0.07	0.11
	Eastern 16	0.07	0.10	0.09	0.11	0.11	0.10	0.10	0.10
	London 16	0.15	0.09	0.12	0.09	0.20	0.09	0.17	0.08
	South East 16	0.15	0.13	0.10	0.14	0.18	0.12	0.15	0.13
	South West 16	0.09	0.08	0.09	0.08	0.07	0.09	0.08	0.09
	Wales 16	0.04	0.06	0.04	0.05	0.05	0.06	0.03	0.05
	Scotland 16	0.09	0.10	0.09	0.11	0.05	0.10	0.10	0.10
	Region m 16	0.15	0.15	0.17	0.16	0.12	0.13	0.15	0.13
Z:	Financial Shock	0.04	0.15	0.07	0.14	0.06	0.16	0.07	0.15
	Parental Interest	0.76	0.41	0.66	0.42	0.85	0.40	0.75	0.40
	Older Siblings	0.83	1.15	0.92	1.16	0.84	1.16	0.88	1.11
Y:	Hourly income	10.22	7.08	23.39	15.72	7.66	4.72	16.23	10.41
	N	2416		1761		2517		1969	

Notes

1. Individuals with non-reported income/college have been dropped from analysis. Individuals with non-reported values of the instrumental variables have also been dropped (financial shock, parental interest and older siblings in this case). In addition, individuals with income outside the 2.5th and 97.5th quantiles have been dropped from the analysis.

2: Very few individuals had missing ability once the above sample had been dropped. Therefore, these individuals were also dropped as keeping them in the sample led to perfect collinearity in the estimation.

Table 3: Distreg Descriptives

		Male 1991		Male 2008		Female 1991		Female 2008	
	College Graduate	0.12		0.16		0.20		0.21	
		HE	No HE	HE	No HE	HE	No HE	HE	No HE
Xa:	Ability	58.54	42.27	56.72	42.45	60.45	44.94	56.80	45.30
	Ability m	0.13	0.13	0.11	0.12	0.16	0.13	0.13	0.12
	Mother ed yrs	15.98	14.78	15.44	14.74	16.29	14.83	15.80	14.87
	Mother ed yrs m	0.22	0.24	0.22	0.24	0.27	0.24	0.26	0.23
	Father ed yrs	16.51	14.75	16.13	14.74	16.34	14.85	15.89	14.84
	Father ed yrs m	0.23	0.26	0.23	0.25	0.29	0.26	0.28	0.25
	Siblings	1.90	2.32	2.01	2.40	1.81	2.36	2.04	2.31
	Siblings m	0.21	0.25	0.22	0.25	0.28	0.23	0.26	0.23
Xb:	North West 33	0.08	0.11	0.11	0.12	0.13	0.12	0.12	0.12
	Yorkshire and Humber 33	0.07	0.10	0.07	0.09	0.08	0.09	0.11	0.09
	East Midlands 33	0.06	0.09	0.07	0.09	0.04	0.07	0.04	0.08
	West Midlands 33	0.09	0.10	0.09	0.10	0.08	0.10	0.07	0.10
	Eastern 33	0.10	0.10	0.11	0.12	0.12	0.10	0.12	0.11
	London 33	0.17	0.07	0.09	0.05	0.18	0.08	0.09	0.06
	South East 33	0.17	0.15	0.16	0.15	0.19	0.14	0.17	0.14
	South West 33	0.10	0.07	0.11	0.07	0.06	0.09	0.10	0.10
	Wales 33	0.06	0.05	0.05	0.05	0.03	0.05	0.05	0.05
	Scotland 33	0.07	0.08	0.10	0.10	0.08	0.09	0.09	0.09
	Distance Uni 33	19.70	21.91	23.32	23.90	18.25	22.11	22.22	23.86
	Distance Poly 33	24.04	25.66	27.34	28.21	21.36	26.27	27.65	28.48
Z:	North West 16	0.11	0.12	0.13	0.12	0.11	0.13	0.12	0.13
	Yorkshire and Humber 16	0.08	0.10	0.10	0.09	0.07	0.09	0.09	0.09
	East Midlands 16	0.06	0.09	0.06	0.08	0.05	0.07	0.04	0.07
	West Midlands 16	0.09	0.10	0.10	0.09	0.08	0.10	0.06	0.10
	Eastern 16	0.09	0.09	0.08	0.10	0.09	0.09	0.09	0.10
	London 16	0.15	0.10	0.13	0.10	0.23	0.11	0.19	0.10
	South East 16	0.17	0.13	0.13	0.14	0.18	0.12	0.14	0.13
	South West 16	0.09	0.07	0.09	0.07	0.06	0.08	0.08	0.08
	Wales 16	0.04	0.05	0.04	0.05	0.04	0.05	0.03	0.05
	Scotland 16	0.08	0.08	0.08	0.10	0.07	0.09	0.09	0.09
	Distance Uni 16	22.70	25.06	23.16	25.54	19.82	25.17	22.39	25.56
	Distance Poly 16	21.77	24.14	21.89	24.59	19.12	24.80	21.89	25.06
Y:	Hourly income	10.21	7.12	23.57	15.43	7.85	4.75	16.65	10.40
	N	3073		2222		3276		2502	

Notes
1. Individuals with non-reported income/college have been dropped from analysis. Individuals with non-reported values of the instrumental variables have also been dropped (region at age 16 and distance to nearest uni/poly at age 16), individuals with missing Xb variables (distance later in life and distance to nearest pre-existing uni/poly) were also dropped from the analysis. In addition, individuals with income outside the 2.5th and 97.5th quantiles have been dropped from the analysis.

Table 4: Propensity Score Estimates - Moffitt Instruments

	Male 1991	Female 1991	Male 2008	Female 2008
Financial	-0.447^{**}	-0.204	-0.190	-0.270^{*}
	(0.154)	(0.157)	(0.131)	(0.129)
Parental int	0.337^{***}	0.615^{***}	0.119	0.456^{***}
	(0.079)	(0.096)	(0.081)	(0.080)
Older sib	0.015	0.025	-0.017	-0.042
	(0.042)	(0.047)	(0.041)	(0.040)
Constant	-6.004^{***}	-6.652^{***}	-4.257^{***}	-4.288^{***}
	(0.405)	(0.456)	(0.428)	(0.395)
Ability	0.043^{***}	0.044^{***}	0.037^{***}	0.026^{***}
	(0.003)	(0.004)	(0.003)	(0.003)
Mother ed yrs	0.061^{*}	0.121^{***}	0.012	0.088^{***}
	(0.026)	(0.025)	(0.030)	(0.025)
Mother ed yrs m	0.097	0.480	0.191	0.694^{*}
	(0.281)	(0.356)	(0.272)	(0.296)
Father ed yrs	0.111^{***}	0.050^{*}	0.104^{***}	0.054^{**}
	(0.019)	(0.021)	(0.021)	(0.021)
Father ed yrs m	0.140	-0.612^{*}	0.137	0.075
	(0.225)	(0.272)	(0.207)	(0.183)
Siblings	-0.031	-0.020	0.007	0.032
	(0.034)	(0.038)	(0.033)	(0.031)
North West	0.241	-0.016	0.001	-0.331
	(0.205)	(0.252)	(0.197)	(0.191)
Yorkshire and Humber	0.173	0.154	-0.003	-0.371
	(0.209)	(0.261)	(0.204)	(0.206)
East Midlands	-0.106	-0.095	-0.318	-0.646^{**}
	(0.221)	(0.275)	(0.221)	(0.229)
West Midlands	0.237	0.065	0.134	-0.519^{*}
	(0.201)	(0.260)	(0.201)	(0.205)
Eastern	-0.282	-0.053	-0.306	-0.493^{*}
	(0.214)	(0.253)	(0.205)	(0.198)
London	0.229	0.376	-0.062	-0.026
	(0.197)	(0.240)	(0.203)	(0.193)
South East	0.058	0.201	-0.399^{*}	-0.405^{*}
	(0.192)	(0.242)	(0.199)	(0.189)
South West	0.102	0.002	-0.031	-0.599^{**}
	(0.208)	(0.265)	(0.211)	(0.208)
Wales	-0.172	0.013	-0.340	-0.653^{**}
	(0.241)	(0.288)	(0.247)	(0.247)
Scotland	0.105	-0.030	-0.201	-0.205
	(0.205)	(0.266)	(0.206)	(0.197)
Region m	0.166	0.181	-0.055	-0.245
	(0.187)	(0.243)	(0.183)	(0.183)
N	2416	2517	1761	1969
McFadden R-sq.	0.529	0.649	0.372	0.422
P-Value IV	0.000^{***}	0.000^{***}	0.2173	0.000^{***}

Table 5: Propensity Score Estimates - Distance Region Instruments

	Male 1991	Female 1991	Male 2008	Female 2008
Distance Uni 16	−0.002	−0.003	−0.001	−0.001
	(0.002)	(0.002)	(0.002)	(0.002)
Distance Poly 16	−0.000	−0.004	−0.001	−0.001
	(0.002)	(0.002)	(0.002)	(0.002)
North West	0.173	0.262	0.036	−0.168
	(0.165)	(0.202)	(0.164)	(0.158)
Yorkshire and Humber	0.207	0.261	0.021	−0.124
	(0.173)	(0.214)	(0.174)	(0.168)
East Midlands	0.034	0.148	−0.246	-0.460^{*}
	(0.180)	(0.226)	(0.185)	(0.190)
West Midlands	0.167	0.319	0.094	−0.317
	(0.170)	(0.210)	(0.170)	(0.172)
Eastern	0.043	0.366	−0.181	−0.268
	(0.178)	(0.215)	(0.181)	(0.173)
London	0.245	0.558^{**}	0.024	0.061
	(0.164)	(0.196)	(0.166)	(0.157)
South East	0.265	0.517^{**}	−0.095	−0.213
	(0.159)	(0.198)	(0.163)	(0.158)
South West	0.248	0.245	0.172	−0.205
	(0.177)	(0.223)	(0.182)	(0.176)
Wales	0.049	0.148	−0.207	-0.455^{*}
	(0.200)	(0.243)	(0.213)	(0.206)
Scotland	0.265	0.458^{*}	−0.081	−0.013
	(0.177)	(0.215)	(0.180)	(0.169)
Constant	-6.630^{***}	-7.294^{***}	-4.826^{***}	-5.209^{***}
	(0.389)	(0.434)	(0.418)	(0.395)
Ability	0.045^{***}	0.051^{***}	0.039^{***}	0.036^{***}
	(0.003)	(0.004)	(0.003)	(0.003)
Ability m	0.300^{***}	0.634^{***}	0.166	0.284^{**}
	(0.091)	(0.099)	(0.100)	(0.095)
Mother ed yrs	0.097^{***}	0.155^{***}	0.021	0.112^{***}
	(0.025)	(0.024)	(0.030)	(0.024)
Mother ed yrs m	0.109	0.241	−0.061	0.382
	(0.241)	(0.257)	(0.245)	(0.233)
Father ed yrs	0.122^{***}	0.061^{**}	0.127^{***}	0.069^{***}
	(0.019)	(0.020)	(0.021)	(0.021)
Father ed yrs m	0.074	−0.288	0.145	−0.110
	(0.211)	(0.202)	(0.209)	(0.176)
Siblings	−0.048	-0.071^{*}	−0.024	−0.003
	(0.025)	(0.029)	(0.025)	(0.024)
Siblings m	−0.282	0.112	−0.218	−0.141
	(0.199)	(0.218)	(0.198)	(0.199)
N	3073	3276	2222	2502
McFadden R-sq.	0.509	0.597	0.383	0.387
P-Value IV	0.529	0.000^{***}	0.247	0.002^{**}

Figure 1: MTE estimates - 1991 - Male

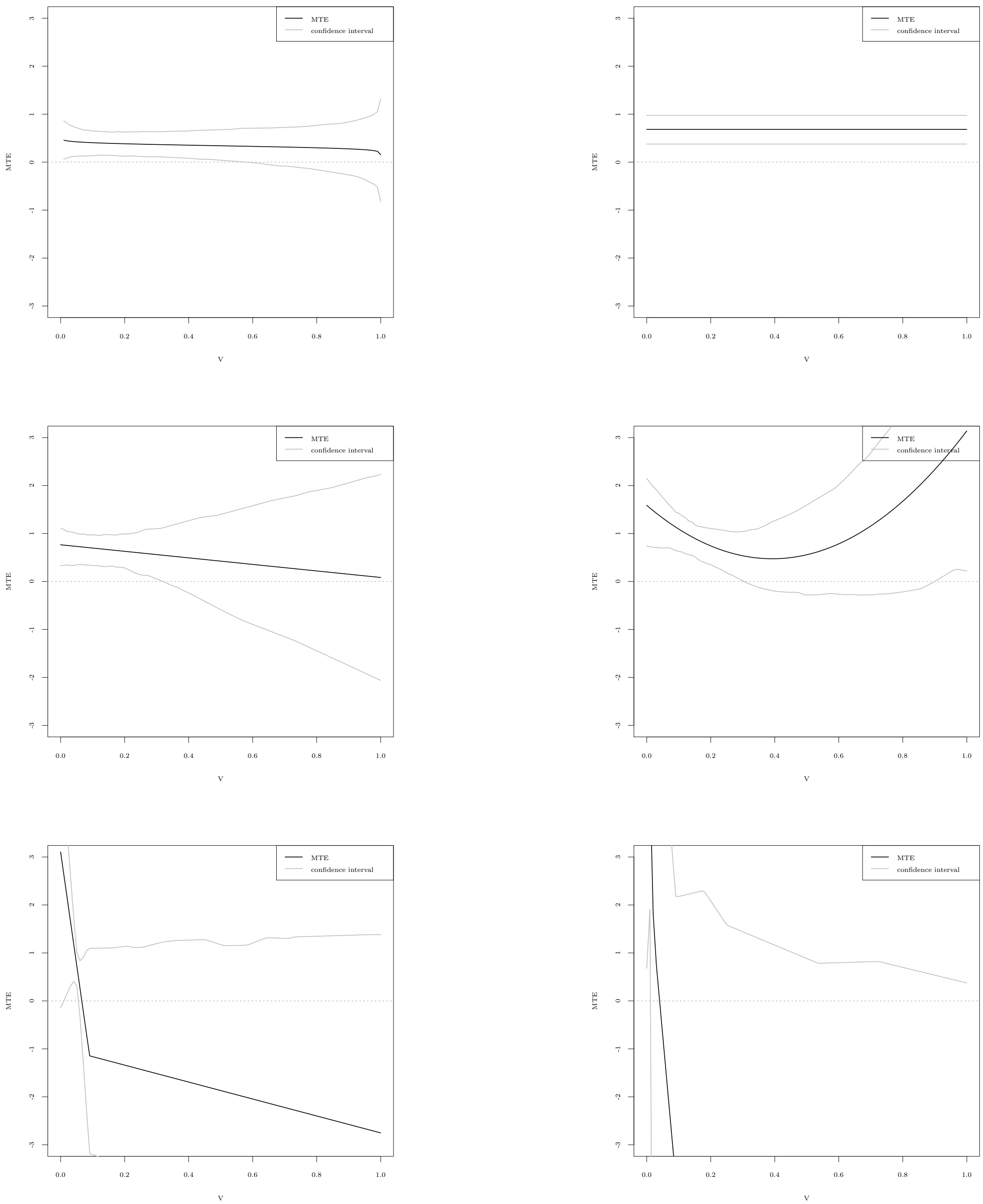

Table 6: Heterogeneity from Unobservables - 1991 - Male

	Normal	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
NormalCoef	−0.050					
Linear			−0.682	−5.674**	−46.906*	−278.605
Quadratic				7.225**		
MedianBreak					45.139*	63.944**
LowerQuartileBreak						204.634
UpperQuartileBreak						6.975**

Table 7: B0 estimates - 1991 - Male

	OLS	*Normal*	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
Cons	1.225***	1.305***	1.317***	1.353***	1.352***	1.423***	1.543***
Ability	0.007***	0.006***	0.006***	0.006***	0.004***	0.005***	0.003**
Motheredyrs	0.003	0.001	−0.003	−0.004	−0.003	−0.005	−0.008
Motheredyrsm	0.006	0.003	−0.011	−0.013	−0.022	−0.017	−0.023
Fatheredyrs	0.024***	0.021***	0.022**	0.020**	0.021**	0.016**	0.012
Fatheredyrsm	−0.035	−0.035	−0.054	−0.055	−0.057	−0.055	−0.057
Siblings	−0.007	−0.007*	−0.005	−0.005	−0.002	−0.003	−0.001
*NorthWest*16	−0.012	−0.016	−0.005	−0.007	−0.013	−0.015	−0.024
*YorkshireandHumber*16	−0.103**	−0.106***	−0.058	−0.060	−0.066*	−0.062	−0.068
*EastMidlands*16	0.014	0.014	0.062	0.062	0.062	0.066	0.071
*WestMidlands*16	−0.037	−0.040	−0.057	−0.060	−0.066*	−0.066*	−0.076*
*Eastern*16	0.060	0.064*	0.113**	0.114***	0.118**	0.122***	0.134**
*London*16	0.170***	0.166***	0.164***	0.163***	0.154***	0.156**	0.146***
*SouthEast*16	0.102**	0.101**	0.171***	0.170***	0.163***	0.171***	0.167***
*SouthWest*16	−0.056	−0.058*	−0.012	−0.013	−0.024	−0.014	−0.018
*Wales*16	−0.095**	−0.093**	−0.098*	−0.097**	−0.100*	−0.090*	−0.083
*Scotland*16	−0.083**	−0.084**	−0.033	−0.033	−0.036	−0.036	−0.040
*Regionm*16	−0.080**	−0.082**	−0.033	−0.035	−0.043	−0.038	−0.044

Table 8: B1-B0 estimates - 1991 - Male

	OLS	*Normal*	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
Cons	0.838***	0.782	2.034***	1.627*	3.095***	3.408**	7.029**
Ability	−0.007***	−0.006**	−0.018***	−0.013*	−0.015*	−0.009	−0.001
Motheredyrs	−0.016	−0.014	−0.009	−0.003	−0.021	0.005	0.014
Motheredyrsm	−0.051	−0.049	0.022	0.038	0.038	0.053	0.075
Fatheredyrs	0.002	0.005	−0.010	0.003	−0.020	0.020	0.036
Fatheredyrsm	0.143	0.142	0.298*	0.311*	0.315*	0.320*	0.331*
Siblings	−0.009	−0.009	−0.007	−0.009	−0.011	−0.015	−0.024
NorthWest16	−0.067	−0.064	−0.246	−0.219	−0.225	−0.186	−0.151
YorkshireandHumber16	0.072	0.074	−0.291	−0.273*	−0.244	−0.261	−0.243
EastMidlands16	−0.095	−0.095	−0.419*	−0.428*	−0.400*	−0.443**	−0.469*
WestMidlands16	0.033	0.035	−0.001	0.028	0.028	0.060	0.090
Eastern16	−0.073	−0.075	−0.431*	−0.456**	−0.387*	−0.493**	−0.546*
London16	−0.014	−0.010	−0.169	−0.145	−0.126	−0.120	−0.092
SouthEast16	−0.064	−0.064	−0.546***	−0.540***	−0.507**	−0.540***	−0.536***
SouthWest16	0.063	0.065	−0.304	−0.295	−0.245	−0.292	−0.287
Wales16	0.021	0.019	0.130	0.117	0.174	0.092	0.064
Scotland16	0.066	0.067	−0.329	−0.322	−0.315	−0.312	−0.302
Regionm16	0.111	0.112*	−0.276	−0.255	−0.227	−0.239	−0.220

Table 9: Graduate Advantage in Graduate Labour Market - 1991 - Male

	OLS	*Normal*	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
TotalAdvgrad	0.048*	0.038	−0.348**	−0.032	−0.856	2.288	6.460*
AdvgradObs	0.048*	0.045	−0.167**	−0.074	−0.176	0.017	0.142
AdvgradUnobs	0.000***	−0.007	−0.181**	0.042	−0.680	2.271	6.318*
Decomposition							
Cons	–	–	–	–	–	–	–
Ability	2.85%	−2.79%	111.42%**	165.57%	95.77%*	−367.70%	20.95%
Motheredyrs	−27.85%	−30.57%	7.72%	9.48%	14.24%	−1.74%	4.68%
Motheredyrsm	−0.40%	−0.45%	−0.03%	−0.15%	−0.04%	0.92%	0.16%
Fatheredyrs	87.27%	92.96%	−11.14%	−51.83%	−1.81%	343.69%	55.14%
Fatheredyrsm	−1.43%	−1.54%	0.93%	2.24%	0.94%	−10.00%	−1.24%
Siblings	14.93%	15.97%	−3.17%	−8.95%	−3.58%	49.63%	8.07%
*NorthWest*16	−0.56%	−0.61%	0.51%	1.05%	0.46%	−4.04%	−0.42%
*YorkshireandHumber*16	0.61%	0.67%	−2.03%	−4.40%	−1.71%	18.45%	2.13%
*EastMidlands*16	3.60%	3.88%	−4.59%	−10.66%	−4.12%	47.57%	6.02%
*WestMidlands*16	−0.09%	−0.12%	0.46%	0.56%	0.28%	−0.50%	0.13%
*Eastern*16	0.57%	0.57%	−4.23%	−10.31%	−3.40%	48.44%	6.45%
*London*16	16.52%	17.75%	0.15%	−1.24%	−0.81%	10.90%	1.96%
*SouthEast*16	1.24%	1.33%	3.59%	8.04%	3.13%	−34.74%	−4.16%
*SouthWest*16	0.04%	0.05%	0.61%	1.35%	0.49%	−5.77%	−0.69%
*Wales*16	2.43%	2.61%	0.30%	0.43%	0.67%	−0.25%	0.21%
*Scotland*16	0.14%	0.16%	−0.90%	−2.00%	−0.83%	8.46%	0.99%
*Regionm*16	0.13%	0.14%	0.38%	0.81%	0.31%	−3.34%	−0.38%

Table 10: Graduate Advantage in Non-Graduate Labour Market - 1991 - Male

	OLS	*Normal*	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
TotalAdvnongrad	0.170***	0.080	−0.037	−0.078	−0.360*	0.984*	3.334*
AdvnongradObs	0.170***	0.153***	0.143***	0.136***	0.109***	0.112***	0.066*
AdvnongradUnobs	–	−0.073	−0.181**	−0.213*	−0.469***	0.872	3.268*
Decomposition							
Cons	–	0.000	–	–	–	–	–
Ability	67.25%***	68.14%***	70.07%***	70.62%***	62.23%***	72.32%***	71.52%*
Motheredyrs	1.73%	0.81%	−2.49%	−2.90%	−2.96%	−4.62%	−12.23%
Motheredyrsm	0.01%	0.01%	−0.03%	−0.04%	−0.09%	−0.07%	−0.15%
Fatheredyrs	22.64%***	22.14%***	24.56%**	24.13%**	32.25%**	23.73%**	29.86%
Fatheredyrsm	0.13%	0.15%	0.24%	0.26%	0.34%	0.32%	0.55%
Siblings	1.92%*	2.01%*	1.59%	1.64%	1.03%	1.41%	0.51%
NorthWest16	−0.02%	−0.04%	−0.01%	−0.02%	−0.04%	−0.04%	−0.13%
YorkshireandHumber16	0.59%	0.67%	0.40%	0.43%	0.59%	0.54%	0.99%
EastMidlands16	−0.17%	−0.20%	−0.92%	−0.98%	−1.23%	−1.26%	−2.30%
WestMidlands16	−0.28%	−0.34%	−0.52%	−0.57%	−0.79%	−0.77%	−1.49%
Eastern16	−0.79%	−0.92%*	−1.75%*	−1.88%**	−2.42%*	−2.43%**	−4.50%
London16	5.14%***	5.55%***	5.87%***	6.15%***	7.26%***	7.15%**	11.32%*
SouthEast16	0.96%	1.06%	1.90%	2.01%	2.40%	2.44%	4.04%
SouthWest16	−0.11%	−0.12%	−0.03%	−0.03%	−0.07%	−0.04%	−0.09%
Wales16	0.89%*	0.96%*	1.08%	1.13%	1.45%	1.27%	1.98%
Scotland16	0.20%	0.23%	0.10%	0.10%	0.14%	0.13%	0.25%
Regionm16	−0.10%	−0.11%	−0.05%	−0.05%	−0.08%	−0.07%	−0.14%

Table 11: Selection - 1991 - Male

	OLS	*Normal*	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
TotalSelection	−0.121***	−0.042	−0.310***	0.046	−0.496	1.304	3.125**
SelectionObs	−0.121***	−0.108	−0.310***	−0.209	−0.284*	−0.095	0.076
SelectionUnobs	0.000***	0.066	0.000	0.255	−0.211	1.399*	3.050**
Decomposition							
Cons	–	–	–	–	–	–	–
Ability	92.94%	97.57%	92.32%***	104.02%*	82.96%**	151.13%	−23.39%
Motheredyrs	13.53%	13.83%	3.00%	1.45%	7.67%	−5.14%	19.51%
Motheredyrsm	0.18%	0.20%	−0.03%	−0.08%	−0.06%	−0.24%	0.43%
Fatheredyrs	−3.15%	−7.24%	5.35%	−2.59%	11.19%	−33.58%	77.29%
Fatheredyrsm	0.76%	0.84%	0.62%	0.96%	0.71%	2.16%	−2.81%
Siblings	−3.27%	−3.78%	−0.97%	−2.08%	−1.82%	−7.23%	14.70%
*NorthWest*16	0.19%	0.20%	0.27%	0.36%	0.27%	0.67%	−0.69%
*YorkshireandHumber*16	0.58%	0.67%	−0.91%	−1.27%	−0.83%	−2.67%	3.13%
*EastMidlands*16	−1.68%	−1.89%	−2.89%	−4.39%	−3.02%	−10.01%	13.32%
*WestMidlands*16	−0.36%	−0.43%	0.00%	−0.17%	−0.13%	−0.82%	1.55%
*Eastern*16	−1.33%	−1.54%	−3.08%	−4.84%	−3.03%	−11.54%	16.05%
*London*16	0.61%	0.49%	2.79%	3.55%	2.27%	6.47%	−6.24%
*SouthEast*16	0.85%	0.95%	2.81%	4.13%	2.85%	9.10%	−11.34%
*SouthWest*16	−0.17%	−0.19%	0.31%	0.45%	0.28%	0.99%	−1.22%
*Wales*16	0.27%	0.28%	0.66%	0.89%	0.97%	1.54%	−1.34%
*Scotland*16	0.23%	0.26%	−0.44%	−0.64%	−0.46%	−1.36%	1.65%
*Regionm*16	−0.19%	−0.21%	0.18%	0.25%	0.16%	0.52%	−0.60%

Table 12: Nicejobs - 1991 - Male

	Benefits	Firmshares	Companycar	Travelben	Subsmeals
Constant	1.217^{*}	0.030	−0.180	0.152	0.312^{*}
	(0.494)	(0.144)	(0.138)	(0.139)	(0.149)
Ability	0.018^{***}	0.002^{**}	0.004^{***}	0.003^{***}	0.000
	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)
Mother ed yrs	0.023	0.012	−0.008	−0.009	0.011
	(0.034)	(0.010)	(0.010)	(0.010)	(0.010)
Mother ed yrs m	−0.492	−0.051	−0.096	−0.038	−0.063
	(0.299)	(0.087)	(0.084)	(0.084)	(0.090)
Father ed yrs	0.018	−0.003	0.023^{**}	0.010	−0.014
	(0.027)	(0.008)	(0.008)	(0.008)	(0.008)
Father ed yrs m	0.076	−0.006	0.003	0.124^{*}	−0.040
	(0.214)	(0.062)	(0.060)	(0.060)	(0.064)
Siblings	−0.017	0.002	0.003	−0.010	0.004
	(0.022)	(0.006)	(0.006)	(0.006)	(0.007)
North West	−0.041	0.007	0.086	−0.024	0.006
	(0.199)	(0.058)	(0.056)	(0.056)	(0.060)
Yorkshire and Humber	−0.067	0.040	−0.022	−0.033	−0.009
	(0.202)	(0.059)	(0.056)	(0.057)	(0.061)
East Midlands	0.212	0.065	0.040	−0.013	0.062
	(0.202)	(0.059)	(0.056)	(0.057)	(0.061)
West Midlands	0.195	0.074	0.057	−0.021	0.036
	(0.200)	(0.058)	(0.056)	(0.056)	(0.060)
Eastern	0.394^{*}	0.076	0.050	−0.003	0.103
	(0.199)	(0.058)	(0.056)	(0.056)	(0.060)
London	0.568^{**}	0.098	0.106	0.056	0.134^{*}
	(0.199)	(0.058)	(0.056)	(0.056)	(0.060)
South East	0.424^{*}	0.112^{*}	0.040	0.034	0.125^{*}
	(0.189)	(0.055)	(0.053)	(0.053)	(0.057)
South West	−0.028	0.099	−0.004	−0.061	0.005
	(0.206)	(0.060)	(0.057)	(0.058)	(0.062)
Wales	−0.091	0.047	0.013	−0.010	0.026
	(0.222)	(0.065)	(0.062)	(0.062)	(0.067)
Scotland	0.210	0.032	0.032	0.053	0.078
	(0.200)	(0.058)	(0.056)	(0.056)	(0.060)
Region m	0.102	0.022	0.015	0.007	0.024
	(0.182)	(0.053)	(0.051)	(0.051)	(0.055)
Constant * College	1.294	−0.140	0.137	0.637^{*}	0.117
	(0.968)	(0.282)	(0.270)	(0.272)	(0.291)
Ability * College	−0.007	0.004	−0.001	-0.006^{*}	−0.001
	(0.008)	(0.002)	(0.002)	(0.002)	(0.003)
Mother ed yrs * College	−0.028	−0.026	0.023	0.001	−0.022
	(0.058)	(0.017)	(0.016)	(0.016)	(0.018)
Mother ed yrs m * College	1.109	−0.014	0.394	0.228	0.160
	(0.767)	(0.224)	(0.214)	(0.216)	(0.231)
Father ed yrs * College	0.012	0.016	−0.019	−0.017	0.030^{*}
	(0.045)	(0.013)	(0.013)	(0.013)	(0.014)
Father ed yrs m * College	−0.021	0.263	−0.244	−0.194	−0.022
	(0.634)	(0.185)	(0.177)	(0.178)	(0.191)
Siblings * College	−0.075	0.008	−0.011	−0.002	−0.003
	(0.065)	(0.019)	(0.018)	(0.018)	(0.020)
North West * College	−0.473	0.029	−0.091	−0.185	−0.154
	(0.512)	(0.149)	(0.143)	(0.144)	(0.154)
Yorkshire and Humber * College	−0.132	−0.023	0.057	0.012	−0.164
	(0.534)	(0.156)	(0.149)	(0.150)	(0.161)
East Midlands * College	−0.109	0.075	−0.005	−0.049	−0.145
	(0.545)	(0.159)	(0.152)	(0.153)	(0.164)
West Midlands * College	-1.039^{*}	−0.119	−0.199	−0.125	−0.168
	(0.503)	(0.146)	(0.140)	(0.141)	(0.151)
Eastern * College	−0.487	−0.013	0.117	0.044	-0.350^{*}
	(0.539)	(0.157)	(0.151)	(0.151)	(0.162)
London * College	−0.929	0.015	−0.153	−0.179	-0.306^{*}
	(0.487)	(0.142)	(0.136)	(0.137)	(0.147)
South East * College	−0.545	−0.055	−0.006	−0.108	−0.176
	(0.486)	(0.142)	(0.136)	(0.137)	(0.146)
South West * College	−0.236	−0.069	−0.003	0.116	−0.036
	(0.526)	(0.153)	(0.147)	(0.148)	(0.158)
Wales * College	−0.924	−0.153	−0.140	0.048	−0.188
	(0.630)	(0.184)	(0.176)	(0.177)	(0.190)
Scotland * College	−0.565	0.043	−0.130	−0.123	−0.035
	(0.521)	(0.152)	(0.145)	(0.146)	(0.157)
Region m * College	−0.459	0.048	−0.091	−0.020	−0.108
	(0.472)	(0.138)	(0.132)	(0.133)	(0.142)
N	2259	2259	2259	2259	2259

Table 13: Nicejobs - 1991 - Male

	Medicalins	Pension	Childcare	Discounts	Otherben
Constant	-0.387^{**}	0.730^{***}	0.021	0.558^{***}	−0.018
	(0.133)	(0.134)	(0.032)	(0.151)	(0.127)
Ability	0.003^{***}	0.003^{***}	0.000	0.001	0.002^{***}
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Mother ed yrs	0.022^{*}	0.001	−0.002	−0.002	−0.002
	(0.009)	(0.009)	(0.002)	(0.011)	(0.009)
Mother ed yrs m	−0.111	0.096	−0.005	−0.064	-0.160^{*}
	(0.080)	(0.081)	(0.019)	(0.092)	(0.077)
Father ed yrs	0.008	−0.001	0.000	−0.015	0.009
	(0.007)	(0.007)	(0.002)	(0.008)	(0.007)
Father ed yrs m	−0.050	−0.017	−0.009	0.069	0.001
	(0.057)	(0.058)	(0.014)	(0.065)	(0.055)
Siblings	−0.005	-0.015^{*}	−0.001	0.004	0.001
	(0.006)	(0.006)	(0.001)	(0.007)	(0.006)
North West	−0.034	-0.115^{*}	0.007	0.034	−0.008
	(0.053)	(0.054)	(0.013)	(0.061)	(0.051)
Yorkshire and Humber	0.011	−0.097	0.007	0.042	−0.006
	(0.054)	(0.055)	(0.013)	(0.062)	(0.052)
East Midlands	0.049	−0.066	0.014	0.041	0.018
	(0.054)	(0.055)	(0.013)	(0.062)	(0.052)
West Midlands	0.054	−0.078	0.001	0.081	−0.009
	(0.054)	(0.054)	(0.013)	(0.061)	(0.051)
Eastern	0.083	−0.036	0.013	0.078	0.030
	(0.053)	(0.054)	(0.013)	(0.061)	(0.051)
London	0.129^{*}	−0.018	0.007	0.064	−0.007
	(0.053)	(0.054)	(0.013)	(0.061)	(0.051)
South East	0.084	−0.071	0.006	0.049	0.047
	(0.051)	(0.051)	(0.012)	(0.058)	(0.048)
South West	−0.030	-0.154^{**}	0.008	0.124	−0.017
	(0.055)	(0.056)	(0.013)	(0.063)	(0.053)
Wales	−0.017	-0.186^{**}	0.021	0.024	−0.008
	(0.060)	(0.060)	(0.014)	(0.068)	(0.057)
Scotland	−0.020	−0.063	0.007	0.067	0.024
	(0.054)	(0.054)	(0.013)	(0.061)	(0.051)
Region m	0.035	−0.082	0.020	0.062	−0.001
	(0.049)	(0.049)	(0.012)	(0.056)	(0.047)
Contant * College	0.607^{*}	0.322	−0.064	−0.343	0.022
	(0.260)	(0.263)	(0.062)	(0.297)	(0.248)
Ability * College	−0.001	−0.004	0.001	0.002	−0.001
	(0.002)	(0.002)	(0.001)	(0.003)	(0.002)
Mother ed yrs * College	−0.019	−0.006	0.004	0.008	0.007
	(0.016)	(0.016)	(0.004)	(0.018)	(0.015)
Mother ed yrs m * College	0.257	0.077	−0.016	0.058	−0.035
	(0.206)	(0.209)	(0.049)	(0.235)	(0.197)
Father ed yrs * College	−0.004	0.000	−0.002	0.016	−0.009
	(0.012)	(0.012)	(0.003)	(0.014)	(0.012)
Father ed yrs m * College	−0.013	−0.135	0.007	0.188	0.128
	(0.170)	(0.173)	(0.040)	(0.194)	(0.163)
Siblings * College	-0.049^{**}	0.006	−0.004	−0.006	−0.014
	(0.018)	(0.018)	(0.004)	(0.020)	(0.017)
North West * College	−0.063	0.077	0.027	−0.194	0.081
	(0.137)	(0.139)	(0.033)	(0.157)	(0.131)
Yorkshire and Humber * College	0.001	0.173	0.039	-0.343^{*}	0.115
	(0.143)	(0.145)	(0.034)	(0.164)	(0.137)
East Midlands * College	−0.104	0.050	−0.012	−0.038	0.119
	(0.146)	(0.148)	(0.035)	(0.167)	(0.140)
West Midlands * College	−0.089	0.037	0.001	-0.398^{**}	0.022
	(0.135)	(0.137)	(0.032)	(0.154)	(0.129)
Eastern * College	0.023	0.006	−0.015	-0.365^{*}	0.067
	(0.145)	(0.147)	(0.034)	(0.165)	(0.138)
London * College	−0.125	−0.107	−0.005	−0.249	0.180
	(0.131)	(0.132)	(0.031)	(0.149)	(0.125)
South East * College	−0.031	0.065	0.038	-0.304^{*}	0.031
	(0.130)	(0.132)	(0.031)	(0.149)	(0.125)
South West * College	−0.183	0.101	−0.003	-0.352^{*}	0.192
	(0.141)	(0.143)	(0.034)	(0.161)	(0.135)
Wales * College	−0.187	0.020	−0.014	−0.369	0.058
	(0.169)	(0.171)	(0.040)	(0.193)	(0.162)
Scotland * College	−0.207	0.021	0.032	−0.243	0.079
	(0.140)	(0.142)	(0.033)	(0.160)	(0.134)
Region m * College	−0.012	0.002	0.003	−0.281	0.001
	(0.127)	(0.128)	(0.030)	(0.145)	(0.121)
N	2259	2259	2259	2259	2259

Figure 2: MTE estimates - 1991 - Female

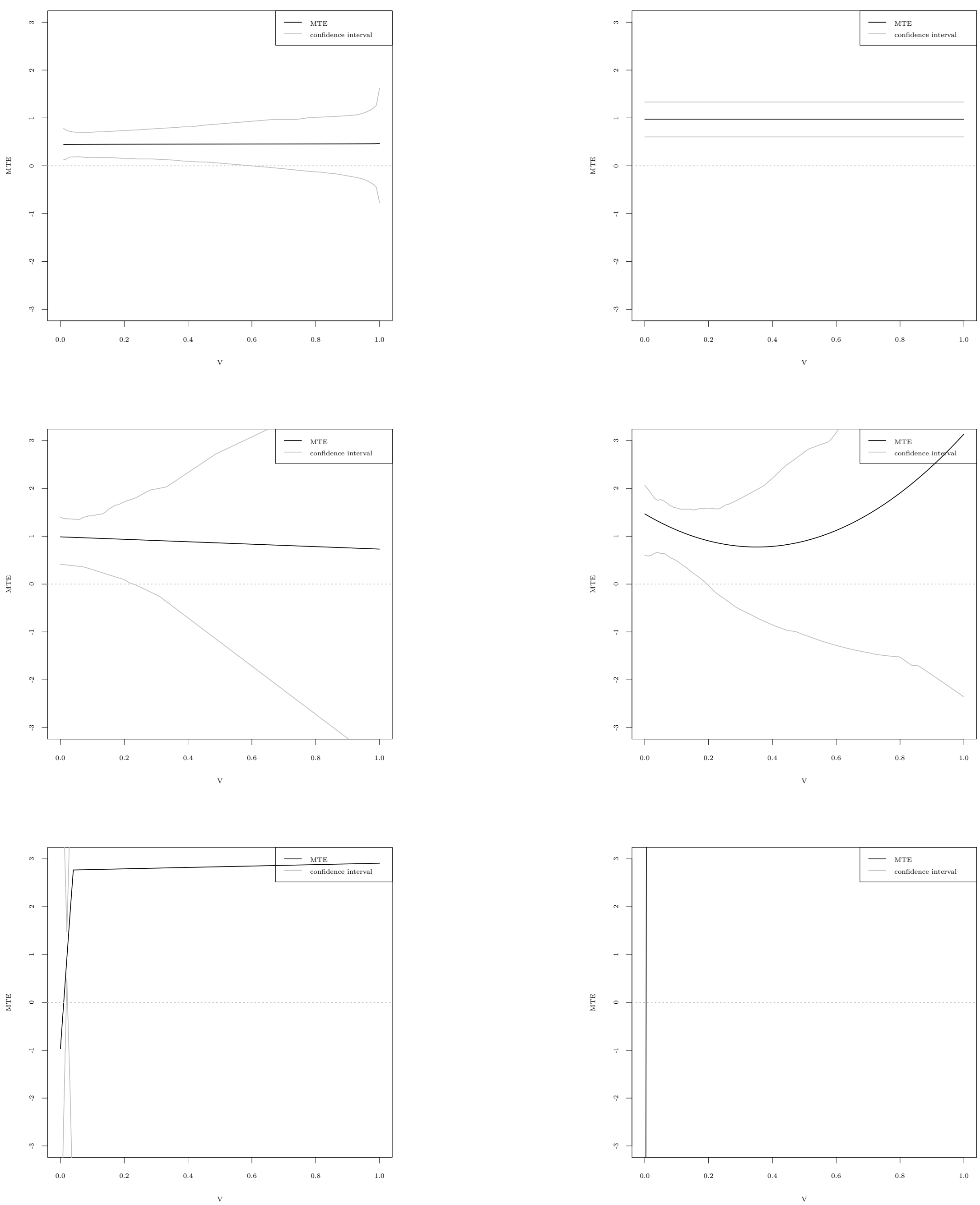

Table 14: Heterogeneity from Unobservables - 1991 - Female

	Normal	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
NormalCoef	0.004					
Linear			−0.255	−3.942	93.582	6437.014
Quadratic				5.609*		
MedianBreak					−93.436	1.187
LowerQuartileBreak						−6463.088
UpperQuartileBreak						22.931*

Table 15: B0 estimates - 1991 - Female

	OLS	*Normal*	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
Cons	0.620***	0.608***	0.861***	0.870***	0.868***	0.850***	1.037***
Ability	0.008***	0.008***	0.008***	0.008***	0.007***	0.008***	0.006***
Motheredyrs	0.018**	0.019**	0.003	0.002	0.004	0.004	−0.002
Motheredyrsm	−0.110	−0.109*	−0.068	−0.069	−0.078	−0.068	−0.088
Fatheredyrs	0.015**	0.015**	0.012*	0.012**	0.012*	0.012*	0.009
Fatheredyrsm	0.006	0.005	0.042	0.043	0.052	0.040	0.065
Siblings	−0.022***	−0.022***	−0.020***	−0.020***	−0.019***	−0.020***	−0.019***
*NorthWest*16	0.057	0.057**	0.012	0.012	0.011	0.011	0.011
*YorkshireandHumber*16	0.008	0.008	−0.000	−0.000	−0.002	−0.000	−0.006
*EastMidlands*16	0.027	0.026	0.039	0.039	0.037	0.038	0.037
*WestMidlands*16	0.069	0.069**	0.078*	0.078*	0.076*	0.078*	0.073*
*Eastern*16	0.051	0.050	0.067*	0.067*	0.064*	0.066*	0.067
*London*16	0.312***	0.313***	0.324***	0.324***	0.319***	0.324***	0.309***
*SouthEast*16	0.064	0.064*	0.036	0.035	0.031	0.036	0.028
*SouthWest*16	−0.051	−0.051	−0.025	−0.025	−0.028	−0.025	−0.027
*Wales*16	−0.010	−0.010	−0.013	−0.013	−0.017	−0.014	−0.017
*Scotland*16	0.032	0.032	0.066*	0.067*	0.065*	0.066	0.069
*Regionm*16	0.059	0.059*	0.091**	0.091**	0.088**	0.091**	0.085**

Table 16: B1-B0 estimates - 1991 - Female

	OLS	*Normal*	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
Cons	1.326***	1.302*	2.691***	2.505*	3.231**	0.747	−18.878
Ability	−0.006**	−0.006	−0.030***	−0.028**	−0.027*	−0.029**	−0.014
Motheredyrs	−0.025*	−0.025	0.010	0.015	0.003	0.009	0.044
Motheredyrsm	0.285	0.285	−0.085	−0.067	−0.044	−0.083	0.057
Fatheredyrs	−0.008	−0.008	−0.015	−0.013	−0.018	−0.015	0.003
Fatheredyrsm	0.167	0.166	−0.261	−0.283	−0.314	−0.261	−0.411
Siblings	0.021	0.021*	0.020	0.019	0.017	0.019	0.011
NorthWest16	−0.105	−0.105	0.400	0.397	0.389	0.401	0.409
YorkshireandHumber16	−0.006	−0.006	−0.037	−0.034	−0.040	−0.035	−0.001
EastMidlands16	−0.232	−0.232*	−0.402	−0.406	−0.381	−0.399	−0.412
WestMidlands16	0.001	0.002	−0.145	−0.144	−0.146	−0.147	−0.125
Eastern16	−0.301*	−0.300**	−0.514*	−0.517*	−0.484*	−0.510*	−0.526
London16	−0.379**	−0.378**	−0.617***	−0.604*	−0.610**	−0.613**	−0.514*
SouthEast16	−0.225	−0.224*	−0.178	−0.170	−0.157	−0.174	−0.120
SouthWest16	−0.222	−0.221	−0.556*	−0.558*	−0.535*	−0.556*	−0.562*
Wales16	0.015	0.015	−0.019	−0.019	0.004	−0.015	−0.004
Scotland16	−0.042	−0.042	−0.589*	−0.593*	−0.579	−0.588*	−0.611
Regionm16	−0.192	−0.191*	−0.645**	−0.641**	−0.631**	−0.643**	−0.593*

Table 17: Graduate Advantage in Graduate Labour Market - 1991 - Female

	OLS	*Normal*	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
TotalAdvgrad	0.019	0.033	−0.522**	−0.392	−0.942	−2.364	−20.149
AdvgradObs	0.019	0.024	−0.325**	−0.285	−0.315	−0.318	−0.038
AdvgradUnobs	0.000***	0.009	−0.197**	−0.107	−0.627	−2.046	−20.111
Decomposition							
Cons	–	–	–	–	–	–	–
Ability	163.44%	143.25%	105.58%**	111.06%	102.24%	106.03%	369.79%
Motheredyrs	−53.94%	−39.10%	−5.67%	−9.07%	−2.90%	−5.71%	−161.92%
Motheredyrsm	−1.25%	−0.98%	−0.06%	−0.06%	−0.05%	−0.06%	−0.11%
Fatheredyrs	54.71%	44.75%	1.43%	0.47%	2.80%	1.32%	−46.31%
Fatheredyrsm	−23.29%	−18.09%	−1.69%	−2.12%	−2.10%	−1.75%	−23.22%
Siblings	0.34%	0.56%	−0.05%	−0.12%	−0.38%	−0.12%	−11.04%
NorthWest16	5.24%	4.04%	2.56%	2.90%	2.57%	2.62%	22.64%
YorkshireandHumber16	0.00%	0.01%	0.01%	0.01%	0.01%	0.01%	0.01%
EastMidlands16	21.33%	16.75%	−2.18%	−2.51%	−2.13%	−2.21%	−19.46%
WestMidlands16	−7.62%	−6.03%	−0.42%	−0.47%	−0.45%	−0.44%	−2.77%
Eastern16	−4.38%	−3.43%	0.45%	0.52%	0.44%	0.46%	3.99%
London16	−35.17%	−26.82%	8.81%	9.58%	9.02%	8.85%	53.29%
SouthEast16	−50.83%	−39.49%	2.57%	2.78%	2.35%	2.56%	14.35%
SouthWest16	22.70%	17.78%	−2.78%	−3.18%	−2.79%	−2.85%	−24.38%
Wales16	−0.18%	−0.15%	−0.08%	−0.09%	−0.03%	−0.07%	−0.43%
Scotland16	2.41%	1.92%	−6.93%	−7.95%	−7.04%	−7.08%	−62.15%
Regionm16	6.47%	5.02%	−1.55%	−1.75%	−1.57%	−1.57%	−12.28%

Table 18: Graduate Advantage in Non-Graduate Labour Market - 1991 - Female

	OLS	*Normal*	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
TotalAdvnongrad	0.210^{***}	0.227^{**}	-0.017	-0.025	-0.175	-1.355	-20.644
AdvnongradObs	0.210^{***}	0.213^{***}	0.181^{***}	0.179^{***}	0.169^{***}	0.185^{***}	0.133^{***}
AdvnongradUnobs	–	0.014	-0.197^{**}	-0.204^{*}	-0.344^{***}	-1.540	-20.777
Decomposition							
Cons	–	0.000	–	–	–	–	–
Ability	$57.35\%^{***}$	$57.30\%^{***}$	$67.43\%^{***}$	$67.61\%^{***}$	$65.29\%^{***}$	$67.36\%^{***}$	$66.62\%^{***}$
Motheredyrs	$12.60\%^{**}$	$12.76\%^{**}$	2.16%	1.90%	3.08%	2.76%	−2.44%
Motheredyrsm	0.07%	0.07%	0.05%	0.05%	0.06%	0.05%	0.09%
Fatheredyrs	$10.56\%^{***}$	$10.58\%^{**}$	$9.79\%^{*}$	$9.74\%^{**}$	$10.26\%^{*}$	$9.64\%^{*}$	9.89%
Fatheredyrsm	−0.08%	−0.06%	−0.59%	−0.60%	−0.77%	−0.54%	−1.22%
Siblings	$5.13\%^{***}$	$5.08\%^{***}$	$5.54\%^{***}$	$5.60\%^{***}$	$5.68\%^{***}$	$5.44\%^{***}$	$7.11\%^{***}$
NorthWest16	−0.55%	−0.54%	−0.14%	−0.14%	−0.13%	−0.12%	−0.17%
YorkshireandHumber16	0.00%	0.00%	−0.00%	−0.00%	−0.00%	−0.00%	−0.00%
EastMidlands16	−0.25%	−0.24%	−0.42%	−0.42%	−0.42%	−0.40%	−0.54%
WestMidlands16	−0.66%	−0.66%	−0.87%	−0.88%	−0.91%	−0.86%	−1.12%
Eastern16	0.08%	0.08%	0.12%	0.12%	0.12%	0.12%	0.17%
London16	$14.45\%^{***}$	$14.34\%^{***}$	$17.53\%^{***}$	$17.65\%^{***}$	$18.45\%^{***}$	$17.13\%^{***}$	$22.77\%^{***}$
SouthEast16	$1.79\%^{*}$	$1.78\%^{*}$	1.16%	1.16%	1.08%	1.14%	1.25%
SouthWest16	0.38%	0.37%	0.22%	0.22%	0.26%	0.21%	0.31%
Wales16	0.04%	0.04%	0.06%	0.06%	0.08%	0.06%	0.10%
Scotland16	−0.65%	−0.64%	$-1.58\%^{*}$	$-1.60\%^{*}$	$-1.66\%^{*}$	−1.54%	−2.24%
Regionm16	−0.25%	−0.25%	−0.46%	−0.46%	−0.47%	−0.45%	−0.58%

Table 19: Selection - 1991 - Female

	OLS	*Normal*	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
TotalSelection	−0.192***	−0.194	−0.506***	−0.367	−0.766	−1.009	0.496
SelectionObs	−0.192***	−0.189*	−0.506***	−0.464	−0.483	−0.503	−0.170
SelectionUnobs	0.000***	−0.005	0.000	0.097	−0.283	−0.506	0.666
Decomposition							
Cons	–	–	–	–	–	–	–
Ability	47.00%	46.46%	91.96%**	94.31%	89.34%*	91.81%*	133.52%
Motheredyrs	19.09%	19.30%	−2.87%	−4.85%	−0.81%	−2.59%	−37.63%
Motheredyrsm	0.20%	0.20%	−0.02%	−0.02%	−0.01%	−0.02%	0.04%
Fatheredyrs	6.25%	6.28%	4.42%	4.04%	5.41%	4.38%	−2.51%
Fatheredyrsm	2.19%	2.21%	−1.30%	−1.53%	−1.64%	−1.31%	−6.08%
Siblings	5.60%	5.65%	1.94%	2.08%	1.74%	1.92%	3.10%
NorthWest16	−1.11%	−1.12%	1.60%	1.73%	1.63%	1.61%	4.87%
YorkshireandHumber16	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
EastMidlands16	−2.35%	−2.38%	−1.55%	−1.70%	−1.53%	−1.55%	−4.71%
WestMidlands16	0.02%	0.02%	−0.58%	−0.63%	−0.61%	−0.59%	−1.49%
Eastern16	0.51%	0.52%	0.33%	0.36%	0.33%	0.33%	1.01%
London16	19.29%	19.53%	11.92%**	12.69%	12.32%	11.90%	29.50%
SouthEast16	6.92%	6.98%	2.07%	2.15%	1.91%	2.04%	4.14%
SouthWest16	−1.80%	−1.82%	−1.71%	−1.87%	−1.72%	−1.72%	−5.13%
Wales16	0.06%	0.06%	−0.03%	−0.03%	0.01%	−0.02%	−0.02%
Scotland16	−0.95%	−0.97%	−5.02%	−5.50%	−5.16%	−5.04%	−15.47%
Regionm16	−0.91%	−0.92%	−1.16%	−1.25%	−1.18%	−1.16%	−3.16%

Table 20: Nicejobs - 1991 - Female

	Benefits	Firmshares	Companycar	Travelben	Subsmeals
Constant	1.053*	0.374**	−0.074	−0.355**	0.371*
	(0.414)	(0.127)	(0.079)	(0.122)	(0.155)
Ability	0.009***	0.002**	0.000	0.001	−0.002
	(0.002)	(0.001)	(0.000)	(0.001)	(0.001)
Mother ed yrs	0.004	−0.016	0.011*	0.014	0.005
	(0.027)	(0.008)	(0.005)	(0.008)	(0.010)
Mother ed yrs m	−0.295	−0.006	−0.023	−0.118	−0.027
	(0.262)	(0.080)	(0.050)	(0.077)	(0.098)
Father ed yrs	0.034	0.001	−0.003	0.018**	−0.005
	(0.022)	(0.007)	(0.004)	(0.006)	(0.008)
Father ed yrs m	0.006	0.027	−0.010	0.046	0.105
	(0.149)	(0.046)	(0.028)	(0.044)	(0.056)
Siblings	−0.045*	−0.009	−0.006	−0.012*	0.012
	(0.019)	(0.006)	(0.004)	(0.006)	(0.007)
North West	0.057	−0.063	0.055	0.052	0.009
	(0.172)	(0.053)	(0.033)	(0.051)	(0.064)
Yorkshire and Humber	0.004	−0.069	−0.008	−0.015	0.023
	(0.184)	(0.057)	(0.035)	(0.054)	(0.069)
East Midlands	0.003	−0.014	0.014	0.019	0.046
	(0.187)	(0.057)	(0.036)	(0.055)	(0.070)
West Midlands	0.294	0.024	0.036	0.062	−0.001
	(0.176)	(0.054)	(0.034)	(0.052)	(0.065)
Eastern	0.263	0.047	0.035	0.007	0.021
	(0.175)	(0.054)	(0.033)	(0.052)	(0.065)
London	0.427*	0.014	0.016	0.086	0.085
	(0.176)	(0.054)	(0.034)	(0.052)	(0.066)
South East	0.217	0.028	0.015	0.020	0.067
	(0.173)	(0.053)	(0.033)	(0.051)	(0.064)
South West	0.165	0.012	−0.012	0.031	0.043
	(0.189)	(0.058)	(0.036)	(0.056)	(0.070)
Wales	0.038	−0.037	−0.001	0.086	−0.039
	(0.204)	(0.063)	(0.039)	(0.060)	(0.076)
Scotland	−0.142	−0.099	0.033	0.058	−0.013
	(0.177)	(0.055)	(0.034)	(0.052)	(0.066)
Region m	0.097	−0.008	0.020	0.083	0.003
	(0.167)	(0.051)	(0.032)	(0.049)	(0.062)
Constant * College	−0.212	−0.460	0.276	0.650*	−0.532
	(0.933)	(0.287)	(0.178)	(0.275)	(0.348)
Ability * College	0.010	−0.000	0.003	−0.003	0.001
	(0.009)	(0.003)	(0.002)	(0.003)	(0.003)
Mother ed yrs * College	0.053	0.024	−0.014	−0.021	0.026
	(0.050)	(0.015)	(0.010)	(0.015)	(0.019)
Mother ed yrs m * College	−0.351	0.242	−0.162	−0.161	−0.096
	(0.781)	(0.240)	(0.149)	(0.230)	(0.291)
Father ed yrs * College	−0.052	−0.004	0.002	−0.003	−0.003
	(0.043)	(0.013)	(0.008)	(0.013)	(0.016)
Father ed yrs m * College	0.794	−0.250	0.195	0.419*	−0.015
	(0.677)	(0.208)	(0.129)	(0.199)	(0.253)
Siblings * College	0.007	0.009	−0.017	−0.005	−0.010
	(0.065)	(0.020)	(0.012)	(0.019)	(0.024)
North West * College	−0.258	0.059	−0.262**	0.109	0.114
	(0.526)	(0.162)	(0.101)	(0.155)	(0.196)
Yorkshire and Humber * College	−1.058	−0.024	−0.120	0.029	−0.076
	(0.541)	(0.166)	(0.103)	(0.159)	(0.202)
East Midlands * College	−0.626	−0.003	−0.149	−0.195	−0.026
	(0.579)	(0.178)	(0.111)	(0.170)	(0.216)
West Midlands * College	−0.640	0.028	−0.286**	−0.045	0.203
	(0.530)	(0.163)	(0.101)	(0.156)	(0.198)
Eastern * College	−0.765	0.010	−0.275**	0.073	0.227
	(0.536)	(0.165)	(0.102)	(0.158)	(0.200)
London * College	−0.989*	0.084	−0.186*	−0.154	0.042
	(0.495)	(0.152)	(0.095)	(0.146)	(0.185)
South East * College	−0.788	0.051	−0.219*	−0.104	0.080
	(0.498)	(0.153)	(0.095)	(0.146)	(0.186)
South West * College	−0.340	0.017	−0.083	0.168	0.114
	(0.556)	(0.171)	(0.106)	(0.164)	(0.207)
Wales * College	−0.535	0.031	−0.328**	0.191	0.224
	(0.614)	(0.189)	(0.117)	(0.181)	(0.229)
Scotland * College	−0.782	0.016	−0.306**	0.040	−0.118
	(0.572)	(0.176)	(0.109)	(0.168)	(0.213)
Region m * College	−0.802	−0.040	−0.189	−0.106	0.016
	(0.511)	(0.157)	(0.098)	(0.150)	(0.191)
N	1929	1929	1928	1927	1927

Table 21: Nicejobs - 1991 - Female

	Medicalins	Pension	Childcare	Discounts	Otherben
Constant	−0.153	0.306*	−0.069	0.662***	0.025
	(0.102)	(0.156)	(0.050)	(0.155)	(0.121)
Ability	0.001*	0.005***	0.000	−0.001	0.002**
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Mother ed yrs	0.008	0.012	0.001	−0.018	−0.016*
	(0.007)	(0.010)	(0.003)	(0.010)	(0.008)
Mother ed yrs m	−0.044	0.041	0.051	−0.106	−0.062
	(0.065)	(0.098)	(0.031)	(0.097)	(0.076)
Father ed yrs	0.004	−0.007	0.004	0.007	0.015*
	(0.005)	(0.008)	(0.003)	(0.008)	(0.006)
Father ed yrs m	−0.060	−0.078	−0.008	0.011	−0.028
	(0.037)	(0.056)	(0.018)	(0.056)	(0.044)
Siblings	−0.008	−0.017*	−0.000	0.002	−0.005
	(0.005)	(0.007)	(0.002)	(0.007)	(0.005)
North West	0.040	−0.085	−0.001	−0.010	0.061
	(0.042)	(0.065)	(0.021)	(0.064)	(0.050)
Yorkshire and Humber	0.009	−0.061	0.044*	−0.023	0.105
	(0.045)	(0.069)	(0.022)	(0.069)	(0.054)
East Midlands	0.025	−0.102	0.005	−0.018	0.030
	(0.046)	(0.070)	(0.022)	(0.070)	(0.055)
West Midlands	0.075	0.001	0.020	−0.059	0.137**
	(0.043)	(0.066)	(0.021)	(0.065)	(0.051)
Eastern	0.087*	−0.063	0.029	−0.051	0.143**
	(0.043)	(0.066)	(0.021)	(0.065)	(0.051)
London	0.157***	0.029	0.011	−0.094	0.125*
	(0.043)	(0.066)	(0.021)	(0.066)	(0.051)
South East	0.100*	−0.066	−0.004	0.006	0.053
	(0.043)	(0.065)	(0.021)	(0.064)	(0.050)
South West	0.036	−0.056	0.005	−0.001	0.108
	(0.047)	(0.071)	(0.023)	(0.070)	(0.055)
Wales	0.061	−0.064	−0.001	−0.006	0.039
	(0.050)	(0.077)	(0.024)	(0.076)	(0.060)
Scotland	0.011	−0.065	0.007	−0.109	0.036
	(0.044)	(0.067)	(0.021)	(0.066)	(0.052)
Region m	0.049	−0.070	0.004	−0.025	0.044
	(0.041)	(0.063)	(0.020)	(0.062)	(0.049)
Constant * College	0.231	0.343	0.031	−0.498	−0.288
	(0.230)	(0.351)	(0.112)	(0.348)	(0.273)
Ability * College	0.003	−0.006	0.002*	0.007*	0.002
	(0.002)	(0.004)	(0.001)	(0.003)	(0.003)
Mother ed yrs * College	0.007	−0.008	−0.012*	0.021	0.033*
	(0.012)	(0.019)	(0.006)	(0.019)	(0.015)
Mother ed yrs m * College	0.035	−0.176	−0.117	0.361	−0.279
	(0.193)	(0.294)	(0.094)	(0.291)	(0.228)
Father ed yrs * College	−0.006	0.005	−0.001	−0.019	−0.023
	(0.011)	(0.016)	(0.005)	(0.016)	(0.013)
Father ed yrs m * College	0.305	−0.305	−0.006	0.047	0.402*
	(0.167)	(0.254)	(0.081)	(0.252)	(0.198)
Siblings * College	−0.020	0.072**	0.001	−0.004	−0.020
	(0.016)	(0.025)	(0.008)	(0.024)	(0.019)
North West * College	−0.379**	0.133	0.108	−0.162	0.023
	(0.130)	(0.198)	(0.063)	(0.196)	(0.154)
Yorkshire and Humber * College	−0.399**	0.021	−0.026	−0.324	−0.140
	(0.134)	(0.204)	(0.065)	(0.202)	(0.158)
East Midlands * College	−0.296*	0.174	0.011	−0.098	−0.045
	(0.143)	(0.218)	(0.069)	(0.216)	(0.169)
West Midlands * College	−0.491***	0.139	−0.014	−0.038	−0.136
	(0.131)	(0.199)	(0.064)	(0.197)	(0.155)
Eastern * College	−0.548***	−0.037	0.040	−0.156	−0.090
	(0.132)	(0.202)	(0.064)	(0.200)	(0.157)
London * College	−0.526***	−0.062	0.012	−0.088	−0.112
	(0.122)	(0.186)	(0.059)	(0.184)	(0.145)
South East * College	−0.479***	0.050	0.050	−0.233	0.014
	(0.123)	(0.187)	(0.060)	(0.185)	(0.145)
South West * College	−0.407**	−0.049	0.138*	−0.165	−0.072
	(0.137)	(0.209)	(0.067)	(0.207)	(0.162)
Wales * College	−0.609***	0.329	0.106	−0.298	−0.182
	(0.151)	(0.231)	(0.074)	(0.229)	(0.179)
Scotland * College	−0.483***	0.150	0.021	−0.071	−0.031
	(0.141)	(0.215)	(0.069)	(0.213)	(0.167)
Region m * College	−0.360**	−0.027	0.023	−0.135	0.011
	(0.126)	(0.192)	(0.061)	(0.190)	(0.149)
N	1927	1928	1927	1927	1927

Figure 3: MTE estimates - 2008 - Male

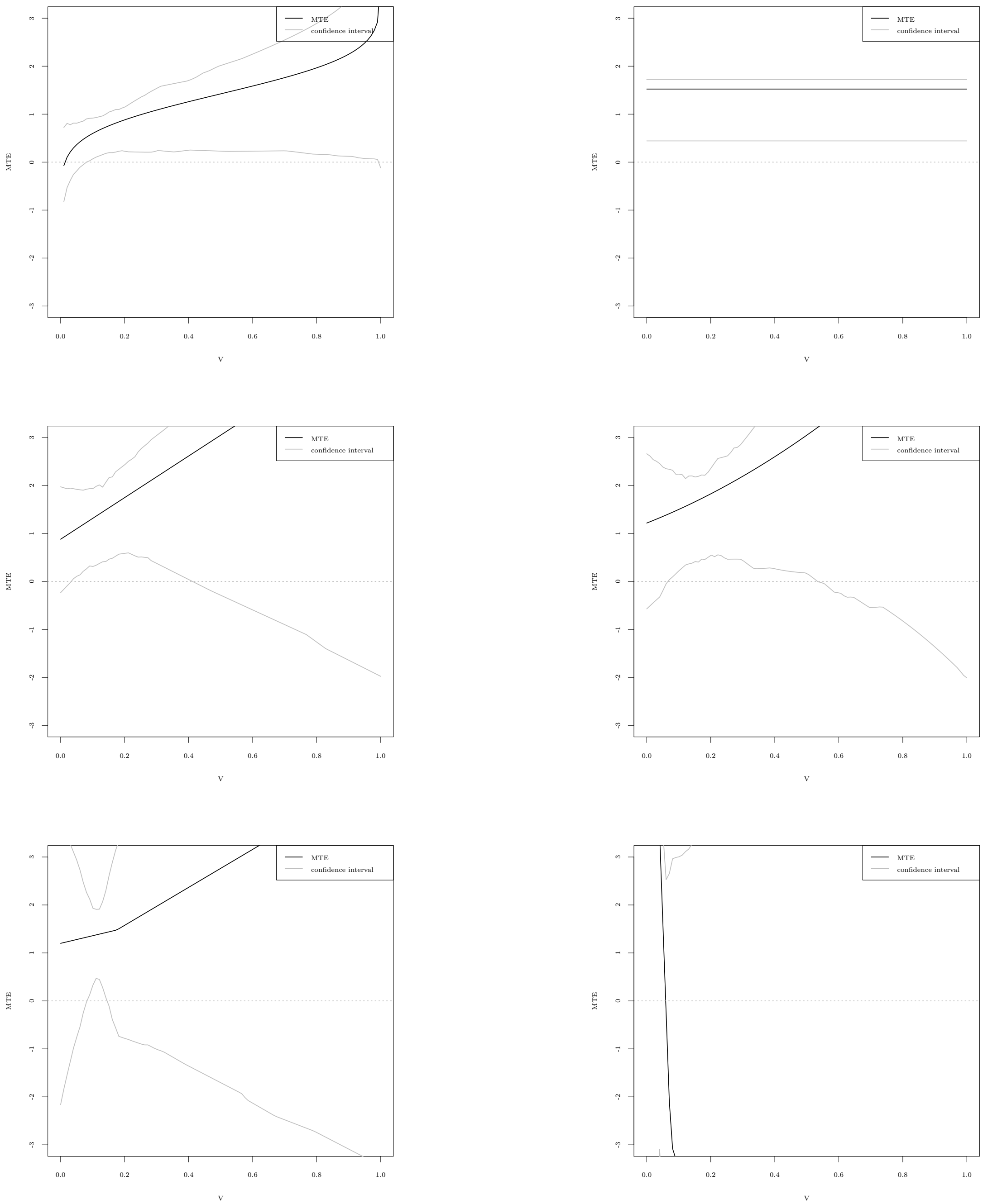

Table 22: Heterogeneity from Unobservables - 2008 - Male

	Normal	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
NormalCoef	0.645					
Linear			4.338	2.632	1.586	−184.367
Quadratic				2.063		
MedianBreak					2.372	15.008
LowerQuartileBreak						162.851
UpperQuartileBreak						6.441

Table 23: B0 estimates - 2008 - Male

	OLS	*Normal*	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
Cons	1.921***	2.013***	2.390***	2.215***	2.242***	2.239***	2.375***
Ability	0.007***	0.006***	0.001*	0.003*	0.002	0.003	−0.001
Motheredyrs	0.010	0.009	0.009	0.009	0.009	0.009	0.008
Motheredyrsm	0.133	0.125*	0.075	0.086	0.082	0.084	0.052
Fatheredyrs	0.020**	0.016*	−0.012	−0.004	−0.005	−0.005	−0.020
Fatheredyrsm	0.027	0.026	0.051	0.059	0.056	0.059	0.056
Siblings	−0.003	−0.002	−0.004	−0.003	−0.003	−0.003	−0.002
*NorthWest*16	−0.109*	−0.110**	−0.130*	−0.128*	−0.129*	−0.129*	−0.124
*YorkshireandHumber*16	−0.161**	−0.162**	−0.023	−0.022	−0.024	−0.022	−0.020
*EastMidlands*16	−0.057	−0.050	0.079	0.063	0.069	0.066	0.108
*WestMidlands*16	−0.113*	−0.117**	−0.040	−0.027	−0.031	−0.029	−0.044
*Eastern*16	0.098	0.105*	0.196**	0.183**	0.188**	0.186**	0.227*
*London*16	0.126**	0.126**	0.271***	0.272***	0.270**	0.272***	0.285***
*SouthEast*16	0.068	0.079	0.219***	0.200***	0.207**	0.204**	0.255*
*SouthWest*16	−0.036	−0.036	0.018	0.019	0.018	0.019	0.021
*Wales*16	−0.220***	−0.210***	−0.103*	−0.119*	−0.113*	−0.115	−0.064
*Scotland*16	−0.079	−0.073	0.044	0.033	0.037	0.035	0.063
*Regionm*16	−0.022	−0.021	0.082	0.082	0.082	0.083	0.097

Table 24: B1-B0 estimates - 2008 - Male

	OLS	*Normal*	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
Cons	0.570**	3.250**	3.136***	5.045**	5.382**	5.192**	13.365*
Ability	−0.004**	−0.023**	−0.023***	−0.050*	−0.049*	−0.048	−0.032
Motheredyrs	−0.018	−0.025	−0.033	−0.044	−0.045	−0.043	−0.037
Motheredyrsm	−0.194	−0.279	−0.128	−0.291	−0.287	−0.281	−0.171
Fatheredyrs	0.010	−0.036	0.006	−0.077	−0.080	−0.072	−0.018
Fatheredyrsm	0.031	−0.026	−0.197	−0.283	−0.277	−0.279	−0.254
Siblings	−0.003	0.005	0.034	0.037	0.036	0.037	0.034
NorthWest16	0.089	0.090	0.098	0.080	0.085	0.082	0.070
YorkshireandHumber16	0.215	0.217*	−0.401*	−0.406	−0.396	−0.407	−0.419
EastMidlands16	−0.142	0.010	−0.410*	−0.190	−0.186	−0.204	−0.363
WestMidlands16	0.130	0.071	−0.350*	−0.485	−0.476	−0.476	−0.420
Eastern16	−0.067	0.075	−0.189	0.007	0.013	−0.007	−0.161
London16	0.026	0.037	−0.545**	−0.532*	−0.517*	−0.534*	−0.581*
SouthEast16	0.118	0.307*	−0.132	0.136	0.140	0.118	−0.079
SouthWest16	−0.090	−0.075	−0.318	−0.311	−0.304	−0.311	−0.327
Wales16	0.035	0.190	−0.090	0.125	0.129	0.110	−0.074
Scotland16	0.255*	0.341**	−0.098	0.038	0.038	0.029	−0.071
Regionm16	0.137	0.162	−0.295	−0.279	−0.269	−0.281	−0.328

Table 25: Graduate Advantage in Graduate Labour Market - 2008 - Male

	OLS	*Normal*	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
TotalAdvgrad	0.080^{***}	-1.294^{*}	-1.184^{**}	-3.217^{*}	-3.392	-2.895	6.219
AdvgradObs	0.080^{***}	-0.282	-0.354^{*}	-0.811^{*}	-0.815^{*}	-0.786	-0.562
AdvgradUnobs	0.000^{***}	-1.012^{*}	-0.830^{**}	-2.406	-2.577	-2.109	6.781
Decomposition							
Cons	–	–	–	–	–	–	–
Ability	$56.93\%^{**}$	$82.27\%^{**}$	$84.45\%^{**}$	$78.15\%^{**}$	$77.69\%^{**}$	$78.37\%^{**}$	$80.72\%^{**}$
Motheredyrs	−5.79%	3.38%	4.19%	2.65%	2.69%	2.67%	3.23%
Motheredyrsm	−0.46%	0.33%	0.09%	0.15%	0.15%	0.15%	0.13%
Fatheredyrs	$45.66\%^{***}$	8.67%	2.05%	12.35%	12.84%	12.05%	8.36%
Fatheredyrsm	0.37%	−0.00%	0.21%	0.14%	0.14%	0.14%	0.18%
Siblings	2.59%	0.30%	3.09%	1.50%	1.44%	1.53%	2.05%
NorthWest16	−0.31%	0.09%	0.11%	0.07%	0.07%	0.07%	0.12%
YorkshireandHumber16	0.72%	−0.21%	1.30%	0.57%	0.56%	0.59%	0.85%
EastMidlands16	3.96%	−0.23%	−1.50%	−0.25%	−0.23%	−0.28%	−0.73%
WestMidlands16	0.52%	0.38%	2.59%	1.48%	1.46%	1.51%	1.94%
Eastern16	−0.49%	0.82%	0.03%	0.30%	0.32%	0.29%	0.15%
London16	4.87%	−1.48%	1.99%	0.82%	0.78%	0.86%	1.35%
SouthEast16	$-7.18\%^{*}$	4.23%	0.76%	1.28%	1.32%	1.27%	0.97%
SouthWest16	−1.73%	0.44%	0.94%	0.40%	0.39%	0.41%	0.60%
Wales16	2.59%	−0.08%	−0.61%	0.01%	0.02%	−0.01%	−0.28%
Scotland16	−3.60%	1.56%	−0.25%	0.14%	0.15%	0.13%	−0.02%
Regionm16	1.35%	−0.47%	0.56%	0.23%	0.21%	0.24%	0.39%

Table 26: Graduate Advantage in Non-Graduate Labour Market - 2008 - Male

	OLS	*Normal*	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
TotalAdvnongrad	0.133***	−0.001	−0.823*	−0.610	−0.761	−0.599	3.151
AdvnongradObs	0.133***	0.111***	0.006	0.047	0.031	0.040	−0.042
AdvnongradUnobs	–	−0.113	−0.830**	−0.657*	−0.792*	−0.639*	3.193
Decomposition							
Cons	–	0.000	–	–	–	–	–
Ability	76.39%***	77.85%***	291.91%	99.75%**	106.86%*	105.47%*	38.71%*
Motheredyrs	4.73%	5.13%	86.73%	12.21%	17.88%	14.10%	−11.27%
Motheredyrsm	0.62%	0.69%	7.21%	1.12%	1.59%	1.29%	−0.76%
Fatheredyrs	18.83%**	17.81%*	−232.32%	−9.53%	−17.93%	−15.73%	58.91%
Fatheredyrsm	0.10%	0.12%	4.12%	0.64%	0.90%	0.75%	−0.68%
Siblings	0.68%	0.78%	20.65%	2.64%	3.25%	3.04%	−1.87%
*NorthWest*16	−1.01%	−1.22%	−25.16%	−3.35%	−5.08%	−3.99%	3.62%
*YorkshireandHumber*16	−1.32%	−1.58%	−3.89%	−0.50%	−0.82%	−0.59%	0.52%
*EastMidlands*16	0.69%	0.71%	−19.88%	−2.13%	−3.51%	−2.65%	4.09%
*WestMidlands*16	−1.99%	−2.47%*	−14.82%	−1.32%	−2.28%	−1.69%	2.43%
*Eastern*16	−0.95%	−1.21%	−39.28%	−4.95%	−7.68%	−5.98%	6.90%
*London*16	2.44%	2.91%	108.93%	14.71%	22.01%	17.48%	−17.32%
*SouthEast*16	−1.60%	−2.20%	−106.23%	−13.03%	−20.32%	−15.78%	18.67%
*SouthWest*16	−0.30%	−0.36%	3.19%	0.44%	0.62%	0.52%	−0.55%
*Wales*16	1.87%	2.12%	18.15%	2.83%	4.03%	3.25%	−1.70%
*Scotland*16	0.97%	1.08%	−11.40%	−1.15%	−1.95%	−1.44%	2.45%
*Regionm*16	−0.15%	−0.17%	12.09%	1.63%	2.43%	1.95%	−2.15%

Table 27: Selection - 2008 - Male

	OLS	*Normal*	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
TotalSelection	−0.052**	−1.293*	−0.360***	−2.607	−2.631	−2.296	3.068
SelectionObs	−0.052**	−0.394**	−0.360***	−0.858*	−0.846*	−0.826	−0.520
SelectionUnobs	0.000***	−0.899	−0.000	−1.749	−1.785	−1.470	3.588
Decomposition							
Cons	–	–	–	–	–	–	–
Ability	106.38%**	81.02%**	88.13%***	79.35%***	78.77%**	79.68%**	84.13%*
Motheredyrs	20.94%	3.87%	5.65%	3.17%	3.25%	3.23%	4.40%
Motheredyrsm	2.27%	0.43%	0.22%	0.21%	0.21%	0.21%	0.20%
Fatheredyrs	−22.51%	11.26%	−2.10%	11.14%	11.70%	10.71%	4.26%
Fatheredyrsm	−0.30%	0.03%	0.28%	0.17%	0.17%	0.17%	0.25%
Siblings	−2.27%	0.43%	3.40%	1.56%	1.51%	1.60%	2.37%
NorthWest16	−2.10%	−0.28%	−0.34%	−0.12%	−0.12%	−0.12%	−0.17%
YorkshireandHumber16	−4.47%	−0.60%	1.21%	0.51%	0.51%	0.54%	0.88%
EastMidlands16	−4.34%	0.04%	−1.82%	−0.35%	−0.35%	−0.40%	−1.12%
WestMidlands16	−5.85%	−0.42%	2.28%	1.33%	1.32%	1.35%	1.90%
Eastern16	−1.65%	0.24%	−0.67%	0.01%	0.02%	−0.01%	−0.40%
London16	−1.30%	−0.24%	3.88%	1.59%	1.57%	1.66%	2.87%
SouthEast16	7.00%	2.42%*	−1.13%	0.49%	0.51%	0.44%	−0.47%
SouthWest16	1.91%	0.21%	0.98%	0.40%	0.40%	0.42%	0.70%
Wales16	0.75%	0.54%	−0.28%	0.16%	0.17%	0.15%	−0.16%
Scotland16	8.01%	1.42%	−0.45%	0.07%	0.07%	0.06%	−0.23%
Regionm16	−2.46%	−0.39%	0.77%	0.30%	0.30%	0.32%	0.59%

Figure 4: MTE estimates - 2008 - Female

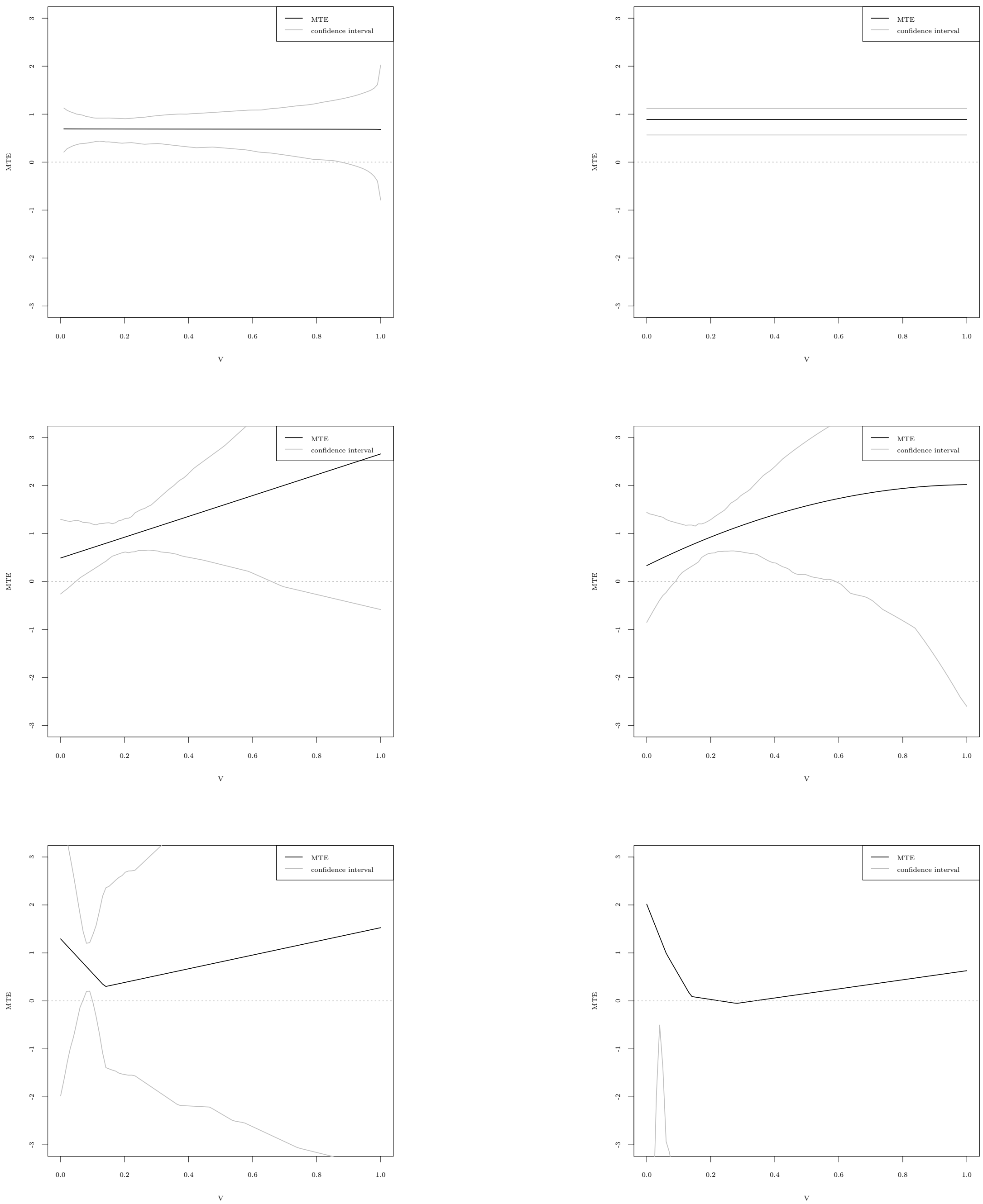

Table 28: Heterogeneity from Unobservables - 2008 - Female

	Normal	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
NormalCoef	−0.002					
Linear			2.170	3.299	−7.184	−16.822
Quadratic				−1.610		
MedianBreak					8.611	10.499
LowerQuartileBreak						5.299
UpperQuartileBreak						1.972

Table 29: B0 estimates - 2008 - Female

	OLS	*Normal*	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
Cons	1.387***	1.561***	1.670***	1.509***	1.524***	1.548***	1.573***
Ability	0.006***	0.005***	0.004***	0.005***	0.005***	0.005**	0.004*
Motheredyrs	0.026***	0.020**	0.020*	0.026*	0.026*	0.024*	0.022*
Motheredyrsm	−0.055	−0.096	0.112	0.163	0.163	0.146	0.132
Fatheredyrs	0.012	0.007	−0.001	0.003	0.002	0.002	0.001
Fatheredyrsm	−0.001	0.003	0.038	0.040	0.040	0.041	0.042
Siblings	−0.005	−0.005	−0.006	−0.006	−0.007	−0.006	−0.005
*NorthWest*16	0.088*	0.101**	0.040	0.025	0.022	0.029	0.037
*YorkshireandHumber*16	0.015	0.029	0.059	0.042	0.038	0.046	0.054
*EastMidlands*16	0.028	0.053	0.044	0.019	0.014	0.028	0.042
*WestMidlands*16	−0.005	0.017	−0.076	−0.099	−0.103	−0.091	−0.080
*Eastern*16	0.034	0.054	0.084	0.063	0.060	0.071	0.081
*London*16	0.215***	0.208***	0.275***	0.279***	0.276***	0.276***	0.279***
*SouthEast*16	0.068	0.084*	0.122*	0.104	0.101	0.109	0.118
*SouthWest*16	−0.023	0.002	0.097	0.070	0.066	0.080	0.093
*Wales*16	0.020	0.048	0.036	0.006	0.002	0.016	0.030
*Scotland*16	0.066	0.074	0.109	0.099	0.097	0.099	0.104
*Regionm*16	0.044	0.052	0.143**	0.131**	0.128*	0.133*	0.139*

Table 30: B1-B0 estimates - 2008 - Female

	OLS	*Normal*	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
Cons	1.359***	1.761***	1.864***	2.691***	2.481**	3.173**	3.671
Ability	−0.004*	−0.006*	−0.008*	−0.017*	−0.017*	−0.014*	−0.012
Motheredyrs	−0.039***	−0.042***	−0.052	−0.089*	−0.087*	−0.079	−0.073
Motheredyrsm	−0.161	−0.186	−0.864**	−1.164**	−1.147**	−1.085**	−1.038**
Fatheredyrs	−0.013	−0.014	0.008	−0.016	−0.013	−0.010	−0.007
Fatheredyrsm	0.084	0.075	0.014	−0.006	−0.008	−0.007	−0.009
Siblings	0.012	0.013	0.027	0.024	0.027	0.025	0.023
*NorthWest*16	0.077	0.093	0.503	0.607	0.618**	0.585*	0.558*
*YorkshireandHumber*16	0.148	0.173	0.094	0.220	0.229	0.194	0.167
*EastMidlands*16	−0.091	−0.059	0.240	0.433	0.445	0.382	0.327
*WestMidlands*16	0.187	0.212*	1.061**	1.228**	1.238***	1.189***	1.147***
*Eastern*16	0.010	0.033	−0.062	0.093	0.098	0.051	0.013
*London*16	−0.149	−0.141	−0.411	−0.426	−0.416	−0.420	−0.428
*SouthEast*16	−0.125	−0.105	−0.219	−0.088	−0.084	−0.120	−0.152
*SouthWest*16	−0.079	−0.048	−0.559*	−0.365	−0.359	−0.417	−0.466
*Wales*16	0.069	0.111	0.369	0.579	0.587	0.525	0.468
*Scotland*16	0.030	0.040	−0.119	−0.058	−0.054	−0.068	−0.083
*Regionm*16	−0.008	0.008	−0.393	−0.312	−0.302	−0.328	−0.349

Table 31: Graduate Advantage in Graduate Labour Market - 2008 - Female

	OLS	*Normal*	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
TotalAdvgrad	0.006	−0.261	−0.479***	−1.368	−1.219	−0.557	0.110
AdvgradObs	0.006	−0.053	−0.143**	−0.296**	−0.298**	−0.259	−0.226
AdvgradUnobs	0.000***	−0.208	−0.335***	−1.072	−0.921	−0.298	0.335
Decomposition							
Cons	–	–	–	–	–	–	–
Ability	401.24%	22.22%	29.10%	42.91%*	44.45%*	41.04%*	36.92%
Motheredyrs	−188.33%	38.34%	20.41%	19.10%	18.37%*	19.23%	20.28%
Motheredyrsm	−55.26%	8.75%	8.57%*	5.52%*	5.39%*	5.93%	6.55%
Fatheredyrs	−15.26%	14.96%	−5.05%	4.67%	3.92%	3.41%	3.00%
Fatheredyrsm	14.18%	−1.62%	−0.40%	−0.12%	−0.12%	−0.15%	−0.16%
Siblings	−35.15%	5.27%	4.90%	2.05%	2.35%	2.46%	2.58%
NorthWest16	−20.47%	2.90%	3.00%	1.69%	1.70%	1.88%	2.08%
YorkshireandHumber16	−16.50%	2.47%	0.69%	0.57%	0.58%	0.60%	0.63%
EastMidlands16	30.40%	−0.32%	6.09%	4.69%	4.72%	4.86%	5.01%
WestMidlands16	−93.59%	14.18%	22.51%*	12.49%*	12.46%*	13.88%	15.44%
Eastern16	−4.57%	1.08%	0.10%	0.34%	0.35%	0.31%	0.27%
London16	77.34%	−9.55%	7.11%	3.72%	3.51%	4.15%	4.94%
SouthEast16	−13.65%	0.61%	1.02%	−0.08%	−0.08%	0.06%	0.23%
SouthWest16	27.79%	−1.50%	−5.59%	−1.73%	−1.70%	−2.25%	−2.86%
Wales16	−22.54%	4.83%	4.54%	3.17%	3.17%	3.35%	3.54%
Scotland16	4.85%	−0.69%	0.02%	−0.04%	−0.05%	−0.04%	−0.03%
Regionm16	9.53%	−1.93%	2.97%	1.04%	1.00%	1.28%	1.58%

Table 32: Graduate Advantage in Non-Graduate Labour Market - 2008 - Female

	OLS	*Normal*	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
TotalAdvnongrad	0.121***	−0.120	−0.248*	−0.060	−0.005	0.055	0.111
AdvnongradObs	0.121***	0.091***	0.088***	0.114**	0.117***	0.104**	0.093**
AdvnongradUnobs	–	−0.211*	−0.335***	−0.174	−0.122	−0.049	0.018
Decomposition							
Cons	–	0.000	–	–	–	–	–
Ability	55.91%***	57.41%***	50.19%***	49.98%**	51.51%***	49.78%**	47.73%**
Motheredyrs	19.35%***	19.52%**	20.05%*	20.47%*	19.80%*	20.48%*	21.24%*
Motheredyrsm	−0.74%	−1.72%	2.07%	2.34%	2.28%	2.28%	2.31%
Fatheredyrs	10.44%*	8.04%	−1.17%	2.81%	2.21%	1.78%	1.25%
Fatheredyrsm	−0.01%	0.04%	0.47%	0.38%	0.37%	0.43%	0.50%
Siblings	1.49%	1.76%	2.49%	1.79%	1.89%	1.90%	1.95%
*NorthWest*16	−0.57%	−0.87%	−0.36%	−0.17%	−0.15%	−0.22%	−0.31%
*YorkshireandHumber*16	−0.08%	−0.20%	−0.43%	−0.24%	−0.21%	−0.28%	−0.37%
*EastMidlands*16	−0.71%	−1.79%	−1.54%	−0.50%	−0.36%	−0.83%	−1.37%
*WestMidlands*16	0.13%	−0.60%	2.81%	2.83%	2.89%	2.86%	2.81%
*Eastern*16	−0.19%	−0.39%	−0.63%	−0.36%	−0.34%	−0.45%	−0.57%
*London*16	13.29%***	17.06%***	23.35%***	18.23%**	17.67%***	19.78%**	22.28%**
*SouthEast*16	0.85%	1.39%	2.10%	1.37%	1.31%	1.58%	1.91%
*SouthWest*16	0.33%	−0.04%	−1.90%	−1.06%	−0.99%	−1.33%	−1.72%
*Wales*16	−0.27%	−0.84%	−0.66%	−0.09%	−0.02%	−0.25%	−0.52%
*Scotland*16	0.17%	0.26%	0.40%	0.28%	0.27%	0.31%	0.36%
*Regionm*16	0.61%	0.96%	2.77%	1.95%	1.87%	2.17%	2.54%

Table 33: Selection - 2008 - Female

	OLS	*Normal*	*Constant*	*Linear*	*Cubic*	*MedianBreak*	*QuartileBreaks*
TotalSelection	−0.115***	−0.141	−0.231***	−1.308	−1.213	−0.612	−0.002
SelectionObs	−0.115***	−0.144**	−0.231***	−0.410**	−0.414**	−0.363*	−0.319*
SelectionUnobs	0.000***	0.003	−0.000	−0.898	−0.799	−0.249	0.317
Decomposition							
Cons	–	–	–	–	–	–	–
Ability	36.74%**	44.52%**	37.13%*	44.88%**	46.43%*	43.55%**	40.08%
Motheredyrs	30.88%**	26.41%**	20.27%	19.48%*	18.77%*	19.59%	20.56%
Motheredyrsm	2.29%	2.11%	6.10%*	4.63%*	4.52%**	4.88%*	5.31%
Fatheredyrs	11.87%	10.57%	−3.57%	4.15%	3.44%	2.94%	2.49%
Fatheredyrsm	−0.80%	−0.57%	−0.07%	0.02%	0.02%	0.02%	0.03%
Siblings	3.52%	3.05%	3.98%	1.98%	2.22%	2.30%	2.39%
NorthWest16	0.53%	0.51%	1.72%	1.17%	1.18%	1.27%	1.38%
YorkshireandHumber16	0.83%	0.78%	0.26%	0.35%	0.36%	0.35%	0.34%
EastMidlands16	−2.44%	−1.25%	3.18%	3.24%	3.29%	3.22%	3.14%
WestMidlands16	5.33%*	4.81%	15.02%**	9.80%*	9.77%**	10.71%*	11.74%*
Eastern16	0.06%	0.15%	−0.18%	0.15%	0.16%	0.09%	0.03%
London16	9.73%*	7.31%	13.29%	7.76%	7.49%	8.64%	10.01%
SouthEast16	1.65%	1.11%	1.43%	0.33%	0.31%	0.50%	0.72%
SouthWest16	−1.19%	−0.57%	−4.19%	−1.54%	−1.50%	−1.99%	−2.53%
Wales16	0.97%	1.24%	2.56%	2.26%	2.27%	2.32%	2.35%
Scotland16	−0.09%	−0.09%	0.17%	0.05%	0.04%	0.06%	0.08%
Regionm16	0.12%	−0.10%	2.90%	1.30%	1.24%	1.54%	1.86%

References

Belloni A, Chernozhukov V, Fernandez-Val I. 2011. Conditional quantile processes based on series or many regressors. *Working PAPER* .

Björklund A, Moffitt R. 1987. The estimation of wage gains and welfare gains in self-selection models. *The Review of Economics and Statistics* **69**: 42–49.

Blundell R, Dearden L, Sianesi B. 2005. Evaluating the effect of education on earnings: models, methods and results from the National Child Development Survey. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **168**: 473–512. ISSN 0964-1998.

Brinch CN, Mogstad M, Wiswall M. 2012. Beyond LATE with a discrete instrument. Heterogeneity in the quantity-quality interaction of children. *Statistics Norway Research Department Discussion Paper* .

Cameron S, Heckman J. 1998. Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of American males. *Journal of Political Economy* **106**: 262–333.

Cameron SV, Taber C. 2004. Estimation of Educational Borrowing Constraints Using Returns to Schooling. *Journal of Political Economy* **112**: 132–182. ISSN 0022-3808.

Card D. 1995. Using geographic variation in college proximity to estimate the return to schooling. In Christofides L, Grant E, Swidinsky R (eds.) *Aspects of Labor Market Behaviour: Essays in Honour of John Vanderkamp*. Toronto: University of Toronto Press.

Carneiro P, Heckman J, Vytlacil E. 2010. Evaluating Marginal Policy Changes and the Average Effect of Treatment for Individuals at the Margin. *Econometrica* **78**: 377–394. ISSN 0012-9682.

Carneiro P, Heckman J, Vytlacil E. 2011a. Estimating marginal returns to education. *American Economic Review* **101**: 2754–2781.

Carneiro P, Lee S. 2009. Estimating distributions of potential outcomes using local instrumental variables with an application to changes in college enrollment and wage inequality. *Journal of Econometrics* **149**: 191–208. ISSN 03044076.

Carneiro P, Lokshin M, Ridao-Cano C, Umapathi N. 2011b. Average and marginal returns to upper secondary schooling in Indonesia. *Cemmap Working Papers* .

Chetty R. 2009. Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods. *Annual Review of Economics* **1**: 451–488. ISSN 1941-1383.

Heckman J, Ichimura H, Smith J, Todd P. 1998. Characterizing selection bias using experimental data. *Econometrica* **66**: 1017–1098.

Heckman J, Urzua S, Vytlacil E. 2006. Understanding instrumental variables in models with essential heterogeneity. *The review of economics and statistics* **88**: 389–432.

Heckman J, Vytlacil E. 1999. Local instrumental variables and latent variable models for identifying and bounding treatment effects. *Proceedings of the National Academy of Sciences* **96**: 4730–4734.

Heckman J, Vytlacil E. 2001a. Instrumental Variables, Selection Models, and Tight Bounds on the Average Treatment Effect. In Lechner M, Pfeiffer F (eds.) *Econometric evaluation of labour market policies*. New York: Springer, 1–15.

Heckman J, Vytlacil E. 2001b. Local instrumental variables. In Amemiya T, Morimun K, Powell J (eds.) *Nonlinear statistical modeling: proceedings of the thirteenth International Symposium in Economic Theory and Econometrics: essays in honor of Takeshi Amemiya*, volume 252, chapter Chapter 1. Cambridge University Press, 1–46.

Heckman J, Vytlacil E. 2001c. Policy-Relevant treatment effects. *American Economic Review* **91**: 107–111.

Heckman J, Vytlacil E. 2005. Structural Equations, Treatment Effects and Econometric Policy Evaluation. *Econometrica* **73**: 669–738.

Heckman J, Vytlacil E. 2007. Econometric evaluation of social programs, part II: Using the marginal treatment effect to organise alternative econometric estimators to evaluation social programs, and to forecast their effects in new environments. *Handbook of Econometrics* **Chapter 71**.

Heckman JJ. 1979. Sample selection bias as a specification error. *Econometrica* **47**: 153–161.

Moffitt R. 2008. Estimating marginal treatment effects in heterogeneous populations. *Annals of Economics and Statistics* : 239–261.