DISCUSSION PAPER SERIES

No. 6266

INCREASING RETURNS TO EDUCATION: THEORY AND EVIDENCE

Alison L Booth, Melvyn G Coles
and Xiaodong Gong

LABOUR ECONOMICS

ISSN 0265-8003

INCREASING RETURNS TO EDUCATION: THEORY AND EVIDENCE

Alison L Booth, University of Essex and CEPR
Melvyn G Coles, ICREA and IAE
Xiaodong Gong, Australian National University (ANU)

Discussion Paper No. 6266
May 2007

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 6266

May 2007

ABSTRACT

Increasing Returns to Education: Theory and Evidence*

We model educational investment and labour supply in a competitive economy with home and market production. Heterogeneous workers are assumed to have different productivities both at home and in the workplace. Following Rosen (1983), we show that there are private increasing returns to education at the labour market participation margin. We show that these depend directly on the elasticity of labour supply with respect to wages. Thus the increasing returns to education problem will be most relevant for women or other types with large enough home productivity. We estimate a three equation recursive model of working hours, wages and years of schooling, and find empirical support for the main predictions of the model.

JEL Classification: J22, J24 and J31
Keywords: home production, labour supply and returns to education

Alison L Booth
Department of Economics
University of Essex
Wivenhoe Park
Colchester
CO4 3SQ
Email: albooth@essex.ac.uk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=107752

Melvyn G Coles
Institut d'Anàlisi Econòmica
Campus UAB
08139 Bellaterra
SPAIN
Email: melvyn.coles@uab.es

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=116675

Xiaodong Gong
Economics Program
Research School of Social Sciences
H C Coombs Bldg
The Australian National University
Canberra ACT 0200
Australia
Email: xiaodong.gong@anu.edu.au

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=166336

* We are grateful to Olivier Deschênes and participants of various seminars and conferences for their helpful comments, to the ARC for financial support under a Discovery Project Grant, and to Hiau Joo Kee and Margi Wood for excellent data assistance. Melvyn Coles also thanks the Barcelona Economics Program of CREA for support.

Submitted 10 April 2007

1 Introduction

This paper introduces home production into the competitive labor market model and shows how this can affect educational choice as well as labor supply to the market sector. This approach provides a richer framework for modeling the returns to education than is usually found in the literature. We demonstrate that, for some sectors of the population, there are increasing returns to education even in a competitive labor market. This prediction is borne out by the findings of our econometric model, which have important implications for estimation of the true returns to education.

The traditional approach to measuring the returns to human capital estimates how wage rates change with an increase in education or training (see Ashenfelter and Rouse (1998) and Card (1999) for extensive discussion and surveys of the returns to education literature). But the wage effect does not describe the full return to education. For example a degree in chemical engineering is only valuable if the student subsequently participates in the labor market. The private return to that qualification is zero if the individual instead uses his/her time in household production.[1] Optimal human capital investment not only depends on wage effects but also on expectations of future labor supply, which are in turn responsive to expectations of future home productivity.

A related insight was developed in an important paper by Rosen (1983). He showed that returns to human capital investment are increasing in subsequent utilization rates because investment costs are independent of how acquired skills are employed. Our approach is complementary to Rosen's important - and often overlooked - insight. In our paper we show that there are increasing returns at the labor market participation, and demonstrate that these depend directly on the elasticity of labor supply with respect to wages.[2] There are two reasons for the increasing returns result. First, there is the participation effect. Workers with greater workplace skills receive better wage offers and so are more likely to participate in the workplace, and a higher participation probability raises the ex ante expected marginal return to human capital investment. Second, there is an increasing labor supply effect that arises because more educated workers may find it worthwhile to work longer hours. This further increases the marginal return to education.

[1] In the theory below we assume that higher education does not increase home productivity. However all our results also carry through if we allow such qualifications to increase home productivity, provided that they increase home productivity less than market productivity.

[2] We use the term 'increasing returns' in the same sense as Rosen (1983), to refer to the private returns to human capital investment.

A central feature of our model is that workers are heterogeneous, having different productivities both in home production and in the workplace.[3] A critical insight is that, when taking labor market participation choice into account, there may be increasing returns to education that vary across workers. For example, workers with low market sector productivity do not participate in the labor market but instead engage in home production. Consequently they have a zero marginal return to education *ex ante*. Since higher productivity workers choose positive labor supply *ex post,* they have a positive marginal return to education *ex ante,* and so the marginal return to education must be increasing over some range.

In three companion papers we examine how such increasing returns interact with various market distortions such as (a) an imperfectly competitive labor market (Booth and Coles, 2005a); (b) endogenous household formation with matching frictions in the marriage market (Booth and Coles, 2005b); and (c) government tax policy, where increasing returns to education cause large substitution effects, and hence large deadweight losses around the non-participant margin (Booth and Coles, 2006). Our aim in the present paper is threefold. First, we illustrate why we should expect increasing returns to education. Second, we show that it is women at the non-participation margin who are most likely to face increasing returns. Third, we identify such effects using individual-level data.

An important theoretical contribution of our paper is that we show that the presence (or absence) of increasing returns to education is closely related to the elasticity of labor supply with respect to wages. Specifically we show that the marginal return to education is increasing in education where the elasticity of labor supply with respect to wages is more than one. Conversely there are decreasing returns if this elasticity is negative (and it is ambiguous otherwise). As this labor supply elasticity is sensitive to home productivity, it should not be surprising that different individuals with different home productivities face different investment margins.

The paper is set out as follows. Section 2 provides some background discussion, while Section 3 sets out the basics of the model. Sections 4 and 5 establish optimal labor supply and optimal education respectively. Section 6 develops an estimation strategy for investigating whether or not there are increasing returns to education, and also describes the data source. Section 7 presents and discusses the empirical estimates. The final section concludes.

[3] Rios-Rull (1993) models skills acquisition in a competitive economy with home and market production. In his model agents are ex ante homogeneous but become heterogeneous through educational investments.

2 Background

The insight for increasing returns to education can be simply demonstrated using the standard textbook model:

$$\max_{C,l} U(C, 1-l) \text{ subject to } pC \leq M + wl$$

where U is an increasing concave utility function, C is consumption, l is labor supply, M is non-labor income, p is the price level, and w the hourly wage rate. Suppose in an earlier period the worker can invest in education e, and so earn a higher wage in the labor market, given by $w = \widehat{w}(a, e)$ where a is the worker's underlying ability. This reflects the Mincerian human capital literature in which education yields higher wage rates. The augmented optimization problem is then

$$\max_{e,C,l} U(C, 1-l) \text{ subject to } pC \leq M + \widehat{w}(a, e)l - ce$$

where c is the cost of education. Notice that this programming problem is not concave. If the worker doubles investment e and doubles labor supply l, then market sector earnings $\widehat{w}l$ more than double. It is this simple non-convexity which generates increasing returns to education. For example, a worker who does not participate *ex post* (i.e. chooses $l = 0$) has a zero return to education e (as $\widehat{w}(a, e)l = 0$). In contrast a worker who supplies more labor obtains a higher return to education; that is, the return to education increases as labor supply increases.

There is a large literature estimating how wages increase with education and that identifies $\partial\widehat{w}/\partial e$. However, this does not describe the *full* marginal return to education. Instead when consumption and labor supply are chosen optimally, the following establishes (using the Envelope Theorem) that the *ex ante* marginal return to education is

$$MR = \frac{\partial \widehat{w}}{\partial e} l^* \frac{\partial U}{\partial C}$$

where l^* is the *ex post* optimal labor supply rule (given education choice e). As noted above, the worker who does not participate *ex post* in the market sector ($l^* = 0$) has a zero marginal return to education, and the marginal return to education increases directly with expected labor supply. Assuming for simplicity that $\widehat{w}$ is linearly increasing in e, we show (i) there are always increasing marginal returns to education at the non-participant margin (where $l^* = 0$ and the return to education is zero); (ii) there are increasing returns to education where the (endogenous) elasticity of labor supply with respect to wages exceeds one; and (iii) there are decreasing returns where the labor supply elasticity is negative.

This simple example illustrates how increasing returns to education can arise even in the canonical textbook model of labor supply. In other words, increasing returns to education are not an artifact of our modeling approach. In the remainder of the paper we focus on extending this basic framework by specifying a utility function $U(.)$ consistent with a model of home production, and allowing for heterogeneity in home productivity b as well as in innate ability a. We wish to introduce the assumption that different workers have different expected home productivities not only because it is realistic, although that is clearly an advantage.[4] But more importantly, with this assumption we can demonstrate how the impact of home productivity on optimal educational choice and market sector labor supply varies across individuals.[5]

3 The Model

We consider a two period framework where a representative worker is born with ability a and has expectations of future home productivity b. In the first period, the worker can invest in e units of workplace human capital which - consistent with the human capital literature - we refer to as education. Let $\alpha = \alpha(a, e)$ denote the worker's marginal product in the market sector in the second period and $\alpha(.)$ is an increasing function. In the second period, home productivity b is realised. The worker then has a unit time endowment where $l \in [0, 1]$ is time spent working in the labor market and $h = 1 - l$ is time spent on home activities. Traditionally h might be interpreted as leisure, but here we think of it as time spent raising children and other domestic activities.[6] Of course some individuals have no children and for them h might be pure leisure. We interpret these latter types as having relatively low home productivity.

[4] High returns to home productivity might be realized by those involved with care of young children or elderly parents, or for individuals with a taste for leisure or for home renovations, or for those with a strong aversion to workplace employment.

[5] Our model has some superficial similarities to Acemoglu (1996). In that model, all firms' have constant returns to scale production functions and an interaction between *ex ante* human capital investments and bilateral search, resulting in *social* increasing returns to average human capital.

[6] Theoretical contributions to the literature considering the allocation of time to home production - albeit in a different context to ours - include *inter alia* Becker (1965), Gronau (1977), and Apps and Rees (1997).

The labor market is competitive. Given productivity $\alpha(.)$, the competitive wage rate is αw. With $w > 0$, a worker who supplies l hours to the market sector enjoys earnings $y = w\alpha l$. Note that doubling second period labor supply l and doubling e implies second period earnings more than double. Hence there are joint increasing returns to education and labor supply in earnings. This specification is crucial to our results.[7] An important issue is that first order conditions are not sufficient to characterize the optimum. To keep the discussion of second order conditions manageable, the model used is relatively straightforward.

Assume productivity α is uni-dimensional, where a worker with ability a who invests in $e \geq 0$ units of education has second period workplace productivity $\alpha = a + e$. The competitive wage rate is therefore $\widehat{w}(.) = (a+e)w$, and $w > 0$ describes the Mincerian rate of return to education (i.e. $\partial\widehat{w}/\partial e = w$) and is assumed the same for all. Assume the cost of attaining education level e is $c_a e$ where $c_a > 0$ depends on the ability of the worker. For ease of exposition c_a is a disutility cost and represents the pain of studying to pass exams rather than a financial cost. Further, assume higher ability types have lower education costs; that is $a > a'$ implies $c_a \leq c_{a'}$.[8]

Young people typically make their human capital investments prior to meeting their future partners, and before raising families. We therefore assume education choice is made in the first period given expectations of second period home productivity. This timing is not critical to the results - there are increasing returns regardless of whether education and labor supply are chosen simultaneously or sequentially. This sequencing implies the second period labor supply choice (choosing l, given e fixed) has well-behaved second order properties. But given the optimal labor supply rule in the second period, we shall characterize conditions which imply increasing returns to education in the first period.

[7] An alternative, more abstract, approach might simply posit a gross earnings function $y_G = y(a, e, l)$ with constant (or decreasing) returns to e and l. Such a specification is convenient (second order conditions are automatically satisfied) rather than compelling. Booth and Coles (2005b) consider a related model of partnership, in which part-time employment is never optimal and workers ex post make a simple labor market participation choice. When making their ex ante education investments, workers anticipate a participation probability $l \in [0, 1]$ (where future participation depends on possibly random home and workplace productivities). Expected future earnings are then $l\widehat{w}(a, e)$, where $\widehat{w}(a, e)$ is the market wage rate, and implies joint increasing returns to l and e. Indeed if a worker has average participation rate l over a working lifetime, then average lifetime earnings are $l\widehat{w}(a, e)$.

[8] A more general specification might instead assume $\widehat{w}(.)$ is non-linear and that the cost of education e is $\widehat{c}_a(e)$ where $\widehat{c}_a(.)$ is an increasing and strictly convex function. Such extensions are qualitatively unimportant. Given any educational investment k, and hence corresponding educational attainment $e = \widehat{c}_a^{-1}(k)$, second period gross earnings $y_G = l\widehat{w}(a, \widehat{c}_a^{-1}(k))$ continue to imply increasing returns to l and k. Assuming $\widehat{w}(.)$ is linear is a useful simplification which implies w can be interpreted as the return to education. Of course linear returns and costs could potentially imply an individual makes an unboundedly large investment. A strictly concave utility function, however, ensures this is not optimal.

Booth and Coles (2005b) consider an equilibrium matching framework where realized home productivity is an endogenous outcome, determined as part of an optimal joint labor supply decision. Realised home productivity is then random, depending on the matching and fertility outcome. However, our focus here is on educational choice and labor supply, and we simplify by assuming b is known in the first period; that is, agents have perfect foresight. Of course different individuals might expect different future home productivities.

A useful simplifying assumption is that human capital investment e does not affect second period home productivity b. While restrictive, it is easy to show that this assumption does not affect the qualitative results obtained here.[9] As the scope for investment in home productivity seems significantly less than the scope for investment in workplace human capital, we therefore simplify by assuming there is no investment in home productivity, and so take b as exogenous.

Labor market earnings are

$$y = \alpha w l$$

Given income y, the worker consumes $C = y/p$ units of a consumption good, where we normalise $p = 1$.

Utility in the second period through consuming C units of the consumption good and enjoying h units of home production is:

$$U_2(C, h) = u(C) + bx(h)$$

where u, x are increasing, concave and twice differentiable functions. Note that an increase in parameter b implies a greater marginal return to home production. As in the traditional labor supply literature we assume x is strictly concave; i.e. there are diminishing marginal returns to leisure, or home production in this case. It should be noted, however, that assuming labor income is linear in l and home production is concave in h is a *normalisation.* If we instead assumed diminishing marginal returns in the market sector, so that labor earnings $y = \alpha w \phi(l)$ where $\phi(.)$ is

[9] For example, suppose instead we allowed a worker also to make home productivity investments $\widetilde{e}$ and so have home productivity parameter $b = b(a, \widetilde{e})$ in the second period. We could then specify a home production function $x(b, h)$ where $h = 1 - l$ is time spent in home production and $x(.)$ is increasing in b, h. In a perfectly symmetric framework, we would also have joint increasing returns to $\widetilde{e}$ and h in home production. As is the case in the model, increasing returns imply workers specialize - a worker either participates mainly in the labour market (choosing e large and $\widetilde{e}$ small) or mainly in the home (choosing e small and $\widetilde{e}$ large). The underlying insight however is unchanged: there are large substitution effects for some workers. For a given individual, the option to invest in home productivity makes substitution to home production less costly. Nevertheless for the marginal participant, the substitution effect remains large.

an increasing but strictly concave function, we could simply redefine the model in terms of effective labor units. Specifically we would define effective labor supply as $\widehat{l} = \phi(l)$, earnings would then have the functional form $y = \alpha w \widehat{l}$ as assumed above, and home production becomes $x(1 - \phi^{-1}(\widehat{l}))$ which is a concave function of $\widehat{l}$. Diminishing marginal returns to labor l in the market sector implies the same underlying 'leisure/labor' trade-off, but described in effective units.[10]

As the budget constraint implies $C = y$ and the time constraint implies $h = 1 - l$, the worker chooses $l \in [0, 1]$ in the second period to solve

$$\max_{l \in [0,1]} u(\alpha w l) + bx(1 - l). \tag{1}$$

As this objective function is concave in l, standard first order conditions are sufficient to describe a maximum. Claim 1 below describes those conditions. Expected second period utility is

$$U_2^*(\alpha, b) = \left[\max_{l \in [0,1]} u(\alpha w l) + bx(1 - l)\right].$$

Given productivity parameters a and b, the worker in the first period chooses education to solve

$$\max_{\alpha \geq a} \left[U_2^*(\alpha, .) - c_a[\alpha - a]\right]. \tag{2}$$

This objective function is not concave in α due to the increasing returns feature. The next section considers the worker's optimization problem recursively, starting in the second period.

4 Optimal Labor Supply

We solve the worker's decision problem recursively, starting in the second period. Given second period productivity parameters (α, b), the worker's optimal second period labor supply choice, properly denoted $l^*(\alpha, b)$, solves (1). As there may be corner solutions, define the functions

[10] Though note that, if we have diminishing returns to labor *and* to education, we could have an earnings function of the form $y = A(a + e)^\gamma l^\delta$ and $\gamma + \delta \leq 1$ implies decreasing returns.

$$b_{PT}(\alpha) = \alpha w u'(0)/x'(1),$$

$$b_{FT}(\alpha) = \alpha w u'(\alpha)/x'(0),$$

and note that concavity of u and x implies $b_{PT} > b_{FT}$. Figures 1a and 1b below plot these functions when $u(.)$ exhibits constant relative risk aversion (CRRA).

As the objective function in (1) is concave, the Kuhn-Tucker first order conditions fully characterise $l^*(.)$. Claim 1 now describes those conditions.

Claim 1. Optimal Second Period Labor Supply.

Given $\alpha, b \geq 0$, optimality implies:

(i) $l^* = 0$ if $b > b_{PT}$;

(ii) $l^* = 1$ if $b < b_{FT}$;

(iii) otherwise l^* is described by the first order condition

$$bx'(1 - l^*) = \alpha w u'(\alpha w l^*). \tag{3}$$

Claim 1 describes the Kuhn-Tucker conditions implied by (1). People with very high home productivity, $b > b_{PT}$, do not participate in the labor market; they choose $l^* = 0$. Conversely people with very low home productivity, $b < b_{FT}$, participate in full time employment; they choose $l^* = 1$. In the intermediate region where $b \in (b_{FT}, b_{PT})$, optimal labor supply implies $l^* \in (0, 1)$ and (3) describes the optimal trade-off between home production and employment in the market sector. Note that $l^* \in (0, 1)$ can be interpreted in two ways. One is that it describes part-time employment, which is common among partnered women with children in many OECD countries (see for e xample Jaumotte, 2003). Alternatively it can be interpreted as the proportion of time spent in full-time employment. For example if there are fixed commuting costs, part-time employment may not be efficient. In that case the worker might instead convexify by spending a proportion of time l^* in full-time employment, $1 - l^*$ in non-participation and l^* then describes a worker's average participation rate. The taxonomy used here is that the interval $b \in (b_{FT}, b_{PT})$ is the part-time region, the region $b \leq b_{FT}$ is the full time region while $b \geq b_{PT}$ is the non-participant region.

The optimal education choice e depends on how l^* varies with productivity parameters (α, b). Standard comparative statics establish that labor supply l^* is strictly decreasing in home productivity in the part-time region (and is constant in the constrained regions). If the worker is risk neutral, the second period utility function U_2 is quasi-linear and labor supply unambiguously increases with α. Risk aversion is more complicated as there are additional income effects. Suppose for example a constant relative risk aversion (CRRA) utility function, $u(C) = C^{1-\sigma}/(1-\sigma)$ where $\sigma \geq 0$ is the degree of relative risk aversion. Claim 2 describes how l^* varies with α in this case.

Claim 2. Optimal Labor Supply with CRRA.
(i) If $\sigma < 1$ then l^* is strictly increasing in α for all $b \in (b_{FT}, b_{PT})$. Further b_{PT}, b_{FT} are strictly increasing in α.
(ii) if $\sigma > 1$ then

(a) for low productivities $\alpha < 1/[(\sigma - 1)w]$, l^* and b_{FT} are both increasing in α;

(b) for $\alpha > 1/[(\sigma - 1)w]$, b_{FT} is decreasing in α. Further, a $b^c \in (b_{FT}, b_{PT})$ exists where l^* is strictly increasing in α for $b \in (b^c, b_{PT})$ and strictly decreasing in α for $b \in (b_{FT}, b^c]$.

Proof is in the Appendix.

Figures 1a and 1b depict these two cases. Figure 1a describes the thresholds b_{PT} and b_{FT} for low levels of risk aversion, $\sigma < 1$. Claim 2 implies labor supply is always increasing in α. Further, α high enough implies the worker takes full time employment $l^* = 1$. Figure 1b holds when there is high risk aversion, $\sigma > 1$. Note that l^* is decreasing in α for α high enough - high risk aversion implies the shadow value of consumption becomes very small at high income levels and the worker instead consumes more 'leisure' (home production). Standard comparative statics establish that b^c, as drawn in Figure 1b, is strictly increasing in α.

Figures 1a and 1b here.

Given this characterization of l^* we now consider the optimal education choice in the first period.

5 Optimal Education

Claim 1 describes the optimal labor supply rule, $l^*(.)$, in the second period. A worker who invests to productivity level $\alpha \geq a$ in the first period obtains expected utility

$$U_1(\alpha, .) \equiv [u(\alpha w l^*) + bx(1 - l^*)] - c_a[\alpha - a].$$

This objective function is not concave in α. A most important object for what follows is

$$MR = wl^* u'(C). \tag{4}$$

where $C = \alpha w l^*$. Totally differentiating U_1 with respect to α, noting that l^* is chosen optimally, the Envelope Theorem implies

$$\frac{dU_1}{d\alpha} = MR - c_a$$

Hence MR describes the worker's marginal return to education.

First consider the simplest case, that workers are risk neutral and so without further loss of generality $u(C) = C$. Then $MR = wl^*$. Thus the marginal return to education is the Mincer rate of return multiplied by expected labor supply. Since Claim 2 implies l^* is increasing in α (strictly in the part-time region), MR is an increasing function of α. That is, risk neutrality guarantees there are increasing returns to education. The reason is simple - very low α workers who do not participate in the labor market have a zero marginal return to workplace capital investment. In contrast, very high productivity workers who choose $l^* = 1$ have the highest return. Increasing returns then occur as labor supply is increasing in productivity.

The case with strictly risk averse workers is more complicated because the marginal return to education depends on the marginal utility of consumption. As Figure 1b demonstrates, it is possible that labor supply decreases with productivity. We shall say that there are strictly increasing (decreasing) returns to education whenever MR is strictly increasing (decreasing) in α.

Theorem 1. Returns to education.

For any concave utility function $u(.)$:

(a) there are strictly increasing returns to education when the elasticity of labor supply with respect to wages is more than one; i.e. when

$$\frac{\alpha}{l^*}\frac{\partial l^*}{\partial \alpha} \geq 1, \tag{5}$$

(b) there are decreasing returns to education when the elasticity of labor supply with respect to wages is negative; i.e. when

$$\frac{\alpha}{l^*}\frac{\partial l^*}{\partial \alpha} \leq 0. \tag{6}$$

Proof is in the Appendix.

Theorem 1 establishes that increasing returns to education depend directly on the elasticity of labor supply with respect to wages. The underlying insight is that there are only joint increasing returns to education *and* labor supply. If for example labor supply were exogenously fixed, we would have decreasing returns to education. Part (a) establishes that a labor supply elasticity exceeding one is sufficient to imply increasing returns to education.[11] In that case, given an increase in education, the subsequent increase in labor supply further increases the marginal return to education. Conversely part (b) establishes that, if labor supply does not increase with wages, then there must be decreasing returns to education.

We now simplify further by assuming a CRRA utility function with $\sigma \leq 1$.[12] Noting that $b = b_{PT}(\alpha)$ is linear in α, we can invert this function and so define $\alpha_{PT} = (b_{PT})^{-1}(b)$ where

$$\alpha_{PT} = bx'(1)/[wu'(0)].$$

Note that $\sigma < 1$ and Claim 2 imply $b = b_{FT}(\alpha)$ is a strictly increasing function. Hence its inverse function is also well-defined. Hence we can define $\alpha_{FT} = (b_{FT})^{-1}(b)$ and $\alpha_{FT}(.)$ is also an

[11] There is a large literature estimating male and female own-price elasticities (see for example Pencavel (2002) and references therein). They are typically small and positive for men, suggesting the increasing returns to education problem is not relevant for them (they anticipate $l^* = 1$ and there are diminishing returns). However the empirical elasticities are larger for women, for whom increasing returns are more likely to apply. (It has recently been argued that estimates of intertemporal substitution elasticities are biased towards zero since they are based on the false assumption of no learning on the job; see for example Imai and Keane (2004). Empirical work estimating own-price elasticities once on-the-job learning is taken into account find the elasticities to be much larger. For instance, Imai and Keane (2004) find substitution elasticities of over 3.)

[12] The results are qualitatively identical with $\sigma > 1$ but the exposition is more complicated as Figure 1b implies the full-time region may not exist (e.g. when b is large). The properties of MR with $\sigma > 1$ are identical to the case $\sigma \in (0, 1)$ as drawn in Figure 2; there are zero returns for α in the non-participation region, increasing marginal returns in the early part of the part-time region (as returns become strictly positive) and decreasing marginal returns for large enough α (as labor supply is then decreasing with productivity - see Figure 1b) but the full-time region may not exist.

increasing function. In Figure 1a, $\alpha = \alpha_{PT}(b)$ corresponds to the locus labelled b_{PT} while $\alpha_{FT}(b)$ corresponds to the locus labelled b_{FT}. Figure 1a and (4) now imply $MR = MR(\alpha, b)$ where

$$\begin{aligned} MR &= 0 \text{ if } \alpha \leq \alpha_{PT}(b) \\ &= wl^* u'(\alpha w l^*) \text{ if } \alpha \in (\alpha_{PT}(b), \alpha_{FT}(b)) \\ &= wu'(\alpha w) \text{ if } \alpha \geq \alpha_{FT}(b). \end{aligned} \tag{7}$$

Figure 2 below graphs MR by productivity α, given $\sigma \leq 1$ and b fixed. Note that $\alpha \leq \alpha_{PT}(b)$ and Claim 1 imply $l^* = 0$ and so $MR = 0$; the marginal return to education is zero in the non-participant region. Note that $\alpha \geq \alpha_{FT}(b)$ and Claim 1 imply $l^* = 1$ and (4) then implies $MR = wu'(\alpha w)$. Consistent with Theorem 1, note there are strictly diminishing marginal returns to education in the full time region (where $\partial l^*/\partial \alpha = 0$). In the part-time region $\alpha \in (\alpha_{PT}, \alpha_{FT})$, MR depends on expected labor supply l^*. There are increasing returns to education for α around the non-participant margin, $\alpha = \alpha_{PT}$, because returns become strictly positive at that point. However, as earnings increase with α, the marginal utility of consumption decreases. Thus it is not necessarily the case that MR is increasing over the entire part-time region. Theorem 1(a) establishes there are increasing returns while the labor supply elasticity exceeds one. For ease of exposition, we shall assume MR is single peaked in this region. Although MR is continuous in α (as labor supply is continuous) its slope is not continuous at the margins α_{PT}, α_{FT} as $\partial l^*/\partial \alpha$ is constrained equal to zero outside the part-time region.[13]

Figure 2 here.

Given this characterization of $MR(.)$, we can now describe the optimal education decision of a worker given ability a and expected home productivity b. Recall that the worker's first period problem is

$$\max_{\alpha \geq a} [U_2^*(\alpha, b) - c_a[\alpha - a]]$$

where $MR \equiv \partial U_2^*/\partial \alpha$. The necessary conditions for optimality imply either a corner solution

(i) $\alpha = a$ and $MR(a, b) \leq c_a$;

or an interior optimum

(ii) $\alpha = \alpha^*(a, b)$ where $MR(\alpha^*, b) = c_a$.

[13]See the Technical Appendix which describes the slope of MR.

Assuming MR is single-peaked as drawn in Figure 2 implies there are only two candidate optima. A local maximum occurs where $MR(\alpha, b) = c_a$ on the decreasing portion of the marginal revenue curve and we let $\alpha^*(a, b)$ denote that solution. The other candidate maximum is that the worker chooses zero education where such a choice is optimal only if $MR(a, b) \leq c_a$.

Consider then workers with ability a and home productivity b satisfying $MR(a, b) < c_a$; e.g. workers with $a < \alpha_{PT}(b)$ for whom $MR = 0$. With increasing returns to education, these workers compare the value of no education, $\alpha = a$, against educating up to $\alpha = \alpha^*(a, b)$. Define

$$V(a, b) = \int_a^{\alpha^*} (MR(\alpha, b) - c_a)\, d\alpha.$$

If $V > 0$ the optimal education choice implies $\alpha = \alpha^*(a, .)$ as it generates positive value relative to no education. The converse is implied by $V < 0$; the worker is better off choosing no education $\alpha = a$. The optimal investment choice therefore depends on the sign of V.

Figure 2 depicts the case when a critical ability a^c exists where $V(a^c, b) = 0$. A worker with ability $a = a^c$ and home productivity b is indifferent between no education and education to $\alpha^c = \alpha^*(a^c)$, where indifference requires that the two shaded areas are equal. Note that a^c can only occur on the increasing portion of the marginal revenue curve, and so $a^c < \alpha_{FT}$. Hence the marginal investor a^c who chooses not to invest to α^* will ex-post choose $l^* < 1$.

Assuming such a critical ability worker exists, consider $V(a, b)$ for $a < a^c$. Differentiation of V with respect to a implies

$$\frac{\partial V}{\partial a} = (c_a - MR(a, b)) - \frac{dc_a}{da}[\alpha^* - a].$$

As c_a is non-increasing in a by assumption and there are increasing marginal returns for $a \leq a^c$, this implies $\partial V / \partial a > 0$ for $a < a^c$. As $V(a^c, .) = 0$ we then have $V < 0$ for $a < a^c$ and so these workers strictly prefer no education. Further we have $\frac{\partial V}{\partial a} > 0$ at $a = a^c$ and so $V > 0$ for $a > a^c$. Hence we have established the following.

Proposition 1. If an ability $a^c < \alpha_{FT}$ exists where $V(a^c, b) = 0$ then

(i) workers with ability $a < a^c$ choose $\alpha = a$ (no education)

(ii) workers with ability $a > a^c$ choose $\alpha = \alpha^*(a) \gg a$.

Proposition 1 establishes that, with increasing returns to education, investment choices may be discontinuous in ability. Low ability types with $a < a^c$ choose no education and, as $a^c < \alpha_{FT}$, these workers either do not participate in the labor market, or only take part-time employment. Workers with sufficiently high ability however choose investment $\alpha^* > a$ and, if $\alpha^* > \alpha_{FT}$ as drawn in Figure 2, participate *ex post* in full time employment. Of course it is the switch to full time employment which makes the *ex ante* education decision worthwhile.

We refer to workers with abilities satisfying $a \leq a^c(b)$ as home specialists. Such workers do not invest in workplace human capital and so have relatively low labor supplies $l^* < 1$. Proposition 2 shows how this critical ability depends on home productivity.

Proposition 2. Home specialists.

$a^c(b)$ is increasing in b.

Proof is in the Appendix.

Home specialists compare the payoff of choosing no education with investing up to productivity $\alpha = \alpha^* \gg a$. An increase in home productivity increases the opportunity cost of working in the market sector and so lowers the relative return to education. Hence workers with greater home productivity are more likely to be home specialists.[14]

6 The Empirical Specification and the Data

Our theory demonstrated that there are increasing returns to education at the labor market participation margin (see Figure 2). Moreover, Theorem 1 showed that the presence (or absence) of increasing returns to education is directly related to the elasticity of labor supply with respect to wages. Proposition 1 established that, with increasing returns to education, the optimal education decision of individuals with abilities a satisfying $MR(a, b_0) < c_a$ is non-marginal. Specifically an ability threshold a^c exists where lower ability types do not invest in education. Proposition 2 demonstrated that this threshold is increasing in expected home productivity, as that determines the payoff to non-participation in the market sector. If women have higher expected home productivities than men, as many believe, this suggests that increasing returns to education will most

[14]It is also possible to show that the critical ability level depends on initial endowed wealth W_o. Suppose each individual faces a budget constraint $C = \alpha wl + W_o$.All of the previous results go through unchanged and the investment margin becomes $a = a^c(b, W_o)$, with $\partial a^c/\partial W_o > 0$. Richer types are less likely to invest in education since their endowed wealth allows them to reduce their participation.

affect female educational choices. We next turn to our empirical model of working hours, wages and years of schooling in order to further investigate increasing returns to education.

We do not test directly for increasing returns to education. That would be akin to testing whether terms like pC and wl should appear in a budget constraint. In an ideal world we would identify the investment threshold $a = a^c(b, W_0)$. Unfortunately ability a, expected future home productivity b (when young) and (endowed) initial wealth W_0 are all unobserved by the econometrician (although to some extent family background variables are likely to pick up some of the endowed wealth effect). Instead we try to identify the marginal return to education and determine whether this is increasing or decreasing for different individuals.

As men and women typically have - or anticipate having - different responsibilities for child care, they have different expected home productivities when young. They therefore make different educational choices when young and have different participation rates in the labor market when older. For this reason we estimate the returns to education separately by gender. As individuals with different abilities face different returns to education, we also split the sample into those who choose low educational investment (that is, those who complete schooling by the age of 18 and who are likely to have $a < a^c$) and those who choose to carry on in further or higher education beyond the age of 18. For each subsample, we estimate the following econometric structure. We also estimate the model on both educational groups combined.

6.1 A Joint Model of Labor Supply, Education and Wages

The theory implies that *ex post* labor supply is

$$l_i = \max[0, l_i^*]$$

where equation (3) implies that, when $l_i > 0$, it is given by

$$l_i^* = l^*(a_i, e_i, b_i, .);$$

The optimal *ex post* labor supply choice reflects productivities in the home and in the workplace, given by b_i and a_i respectively, and e_i denotes educational investment. We shall adopt the econometric structure:

$$l_i^* = x_i'\beta^l + e_i\gamma^l + \varepsilon_i \tag{8}$$

with

$$\begin{aligned} l_i &= 0 \text{ if } x_i'\beta^l + e_i\gamma^l + \varepsilon_i \leq 0 \\ &= l_i^* \text{ if } x_i'\beta^l + e_i\gamma^l + \varepsilon_i > 0 \end{aligned}$$

where x_i is a vector of demographic variables including, age, current health status, family structure (number of children) and the like. We use the number of children and health status as exclusion restriction from other equations. These are standard exclusion restrictions in labor supply models. Note γ^l describes the expected impact of education choice on *ex post* labor supply (when labor supply is positive). ε_i captures unobservable factors such as underlying ability a_i.

Individual i makes an education choice before marriage and fertility are realized; that is, education e_i depends on expectations of future home productivity and is based on first period information. Let m_i denote family background variables (such as own sex, parental socio-economic variables, own birth-order and the like), which an individual knows when young and making education investments.[15] Birth order and other family background variables serve as exclusion restrictions to identify the equation. Birth order in particular can be argued as affecting educational choices of the parents, and is not obviously related to individuals' ability and hence labor supply or wage in later years. Thus ex-ante, each individual i makes an initial education choice

$$e_i = m_i'\beta^e + \phi_i \tag{9}$$

where ϕ_i denotes the error term capturing ability variables known to the individual but not observed by the econometrician. The error terms ε_i and ϕ_i are therefore correlated because of unobserved ability that affects both educational investment and labor supply.

Given e_i described by (9), then substituting out e_i in (8) implies optimal labor supply, in

[15] Black, Devereux and Salvanes (2005) show the importance of family size and birth order in affecting educational attainment. See also Booth and Kee (2005), who show using the BHPS data the importance of family background characteristics in affecting childrens subsequent educational attainment. We incorporate those variables in m.

reduced form, is given by

$$l_i^* = x_i'\beta^l + \gamma^l[m_i'\beta^e] + \tilde{\varepsilon}_i \tag{10}$$

where $\tilde{\varepsilon}_i = \gamma^l\phi_i + \varepsilon_i$.

Next consider the econometric specification of the wage function. An individual's market sector hourly wage remuneration is given by $\widehat{w}_i = (a_i + e_i)w$ which we model as:

$$\widehat{w}_i = z_i'\beta^w + e_i\gamma^w + \mu_i \tag{11}$$

and note that γ^w corresponds to the Mincerian rate of return to education (i.e. γ^w identifies w). The vector z_i consists of independent variables such as the individual's ethnicity and age (as a proxy for experience). As a competitive labor market implies this wage rate depends only on productivity parameters, it must be independent of family structure when young. Thus we exclude variables m_i from z_i. The error term μ_i (capturing workplace ability as well as random shocks affecting productivity) is also correlated with ϕ_i and ε_i.

Again given e_i described by (9), substituting out e_i implies a reduced form Mincer wage equation:

$$\widehat{w}_i = z_i'\beta^w + \gamma^w[m_i'\beta^e] + \tilde{\mu}_i \tag{12}$$

where $\tilde{\mu}_i = \gamma^w\phi_i + \mu_i$.

Thus we obtain a reduced form econometric structure for $(e_i, l_i, \widehat{w}_i)$ which depends on the underlying explanatory variables m_i, x_i, z_i.[16] Assuming the three error terms $(\varepsilon_i, \mu_i, \phi_i)$ follow a joint normal distribution which is independent of the explanatory variables, m_i, x_i, z_i;i.e.

$$\begin{pmatrix} \varepsilon_i \\ \mu_i \\ \phi_i \end{pmatrix} = N(0, \Sigma), \tag{13}$$

[16]Our model is similar to a Heckman-type model. In this framework, we endogenize education and allow unobserved ability to affect education and hence labour supply subsequently.

implies the composite error terms are also normally distributed

$$\begin{pmatrix} \tilde{\varepsilon_i} \\ \tilde{\mu_i} \\ \phi_i \end{pmatrix} = \begin{pmatrix} \gamma^l \phi_i + \varepsilon_i \\ \gamma^w \phi_i + \mu_i \\ \phi_i \end{pmatrix} = N(0, \tilde{\Sigma}), \tag{14}$$

where $\tilde{\Sigma} = A \Sigma A'$ and $A = \begin{pmatrix} 1 & 0 & \gamma^h \\ 0 & 1 & \gamma^w \\ 0 & 0 & 1 \end{pmatrix}$.

We also take into account the sample selection problem, that when individual i is not working ($l_i = 0$) then individual i's wage rate, denoted w_i, is not observed; i.e.,

$$\begin{aligned} w_i &= \text{missing if } x_i'\beta^l + e_i\gamma^l + \varepsilon_i \leq 0 \\ &= \widehat{w}_i \text{ if } x_i'\beta^l + e_i\gamma^l + \varepsilon_i > 0 \end{aligned}$$

These equations are estimated jointly using maximum likelihood. Letting f be the joint normal density function of the three composite error terms, we can write the likelihood contribution of each individual as

- if the individual is working, $l_i > 0$ and $w_i > 0$,

$$lik_i = f(\tilde{\varepsilon_i}, \tilde{\mu_i}, \phi_i)$$

  where $\tilde{\varepsilon}_i = l_i - (x_i'\beta^l + m_i'(\gamma^l\beta^e))$, $\tilde{\mu_i} = w_i - (z_i'\beta^w + m_i'(\gamma^w\beta^e))$, and $\phi_i = e_i - m_i'\beta^e$;

- if she is not working, $l_i = 0$ and w_i is missing, and hence $\tilde{\varepsilon}_i$ and $\tilde{\mu_i}$ need to be integrated out, and

$$glik_i = \int_{-\infty}^{x_i'\beta^l + m_i'(\gamma^l\beta^e)} \int_{-\infty}^{+\infty} f(\tilde{\varepsilon_i}, \tilde{\mu_i}, \phi_i) \mathrm{d}\tilde{\mu_i} \mathrm{d}\tilde{\varepsilon_i}$$

  where $\phi_i = e_i - m_i'\beta^e$

Thus the liklihood function is

$$lik = \prod_i lik_i^{d_i} glik_i^{(1-d_i)},$$

where d_i is an indicator which takes the value one if the individual is working and zero otherwise.

6.2 Data

Our data are from wave 13 of the British Household Panel Survey (BHPS), conducted in 2003-4. The BHPS is a nationally representative random-sample survey of private households in Britain. The survey provides detailed information for individuals on hours worked, earnings, education and other demographics. In wave 13, additional information was collected about family background and the childhood home (for example, sibling numbers, birth order, parents' education, and the number of books at home when the respondent was a child). These variables are used as instruments for the education equation. Although the BHPS reports each individual's highest educational qualification and not years of education, we used information about the British education system to calculate the equivalent years of schooling for each qualification (for details of the British education system, see the summary in Booth and Kee, 2005). For variable definitions and sample statistics see Appendix B. Our sample comprises 3839 women and 2665 men aged 25 to 60 years.

In Figure 3A, we plot, for women, the mean hours of work (conditional on participation), labor force participation and the natural logarithm of hourly wage (workers only) by years of schooling. The figure shows that more highly educated women typically earn more, participate more and, conditional on participation, work a longer working week. Labor force participation is about 50% for women with only 10 years of education but is about 83% for those with 17 years. Conditional on participation, women with 10 years of education work on average 28 hours while those with 17 years education work about 37 hours. In Figure 3B, the same information is summarised for men.[17] The average wage of the men who received 17 years of education was about 70% higher than those who only received 10 years of education. As schooling increases from 10 to 17 years, participation rates increase from 63% to 88% but hours worked (conditional on participation) decrease from 45 hours to about 42 hours. Note the selection effect tends to bias downwards the effect of education on wages - potentially low paid individuals choose not to work and so their wage information is missing from the above. Of course, unobserved ability tends to bias upwards the (Mincer) wage effect. The estimation procedure takes into account both this selection effect

[17]Since only three men received 14 years of education, they are excluded from the graph. The strange blip in wage rates for women occurs at the same 14 year point.

and unobserved ability.

7 Empirical Results

7.1 Estimates of the Key Parameters

To see whether or not the education effects vary with education level, we estimate the model separately for two groups - more highly educated and less highly educated. The more highly educated types are those with 14 or more years of education, which typically means investment at college/university level. The less highly educated are those with fewer than 14 years of education, and who have typically completed their education by the age of 18. We also estimate the model for these two educational groups combined. The estimates of the key parameters of interest are presented as Model A of Table 1A for women and Table 1B for men.

As a sensitivity test, we estimated three additional models progressively imposing restrictions on the error structures. The key estimates from the restricted models are presented in the lower three panels of Tables 1A and 1B. The unrestricted results (Model A) summarise the principal results of the full model reported in Appendix C, and suggest that correlations between the unobservables play significant roles. Not surprisingly, the error terms of hours worked and wages are positively correlated - individuals with higher workplace productivity also work more hours. It is interesting, however, that for the separate subsamples, the correlations between the error terms of education with labor supply and with wages are statistically significant and negative for the less educated group. This cannot be explained by unobserved ability; individuals with higher (unobserved) ability will tend to invest in more education and participate more in the labor market, thus yielding a positive correlation. Instead this negative covariance may be due to unobserved wealth or "class" effects. In the U.K., children from middle class backgrounds are more likely to attend university than are children from working class backgrounds. The latter instead tend to leave school at the minimum school-leaving age of 16 years after which a minority might engage in apprenticeships and the like. Our ML estimates of years of schooling (see Tables C3A and C3B) control for observable family background characteristics such as parental education levels, but are unable to control for unobservables that might differ across different groups or classes in society.[18]

[18] Examples of unobservables likely to affect working class children differently from middle class children are

Model B restricts σ_{le}=0, Model C restricts $\sigma_{le} = \sigma_{we} = 0$, and in Model D no correlations are allowed. The preferred specification is Model A. Nonetheless, it is interesting to note that, when correlations are disallowed in Model D, the estimates of the education effects on hours and wages are larger for the more educated group and smaller for the less educated group, in contrast to the results from Model A. This is not surprising since, according to the *unrestricted* model, the correlation between education and hours and wages is negative for the less educated group and positive for the more highly educated group. When correlations are restricted to zero, these impacts are attributed to the coefficients of the corresponding variables.

Table 1A about here

The two main parameters of interest are γ^l and γ^w, where γ^l describes the impact of education on ex post labor supply, and γ^w describes the Mincerian rate of wage returns to education. Model A in Table 1A shows that γ^w is significantly positive both for more highly educated women and for the less educated women. For more highly educated women, the estimated wage return to education is about 17% per year. For less educated women they are much higher, at about 28%. When the two groups are combined, the estimate is downwardly biased to about 9%.

Recall that the theory section implies the marginal return to education is:

$$MR = \gamma^w l^* u'(C). \tag{15}$$

If workers are risk neutral, then $MR = \gamma^w l^*$. For given γ^w, the slope of MR is then

$$\frac{\partial}{\partial e} MR = \gamma^w \gamma^l.$$

Model A in Table 1A finds that γ^l is not statistically significant for more highly educated women but it is statistically significant, positive and large for the less educated women, for whom one more year of education increases average hours worked by about 7 hours. Thus for the less educated women the slope of MR is $\gamma^w\gamma^l$ =0.278×7.43=2.067. For the more educated women it is 0.173×0.493 =

heterogeneity in discount rates, a distaste for schooling, social custom factors or simply lack of information about the potentially high returns to schooling. Notice that the negative covariance also appears for men (see Table 1B). These wealth or "class" effects also potentially explain why wages and education may be negatively correlated, $\sigma_{we} < 0$: students in continuing education are over-represented by the middle classes. To the extent that ability is the same across the classes, some lower ability middle class students are continuing in school, while higher ability children from working class backgrounds are leaving school at 16 years of age.

0.08. The large labor supply response of less educated women to investment in education implies they face strong increasing returns to education. In contrast, γ^l is not statistically significant from zero for the more educated women, which suggests constant MR at high levels of education.

We also estimated the three equation recursive system separately for men. The principal results are reported as Model A in Table 1B (the progressively more restricted models are presented in the lower panels of the table). Once again, the preferred specification is Model A, the unrestricted model. The γ^w are significantly positive for both less educated and more highly educated men.

The estimates find that the response of labor supply to education (γ^l) is large and positive for the less educated men, and negative for the highly educated men. Thus the estimates again imply strong increasing returns to education for the less educated men and decreasing returns for the more educated men.

Table 1B about here.

Although the main parameters of interest are γ^l and γ^w, the reader may also be interested in seeing the estimated coefficients for the other explanatory variables included in the ML model. These are reported in full in Appendix C (the definitions of the variables used are given in Table B). The estimates of the hours equation are basically in line with the literature. For example, the number of young children reduces women's labor supply as also does higher non-labor income. Higher non-labor income also reduces male labor supply. Better health increases both male and female labor supply and this effect is typically larger for the less educated group. The wage equation estimates in Table C2A show that the impact of age (proxying experience) on wages differs across the two groups of women - it has a quadratic pattern for the more highly educated group, but not for the other group. In contrast, for men the impact is the usual quadratic.

The estimated coefficients of the education equation are presented in Table C3A. Observable family background strongly affects educational attainment. For both educational groups for both sexes, parental education plays a significant role, as does family size. Overall, the effects are larger for the less educated group, especially the impact of maternal education. Growing up in a family in which both biological parents are not present when the respondent was aged up to 16 years (*prntmiss)*, or in a family with more siblings, or with fewer books, especially reduces less educated individuals' years of schooling. However, being first-born in the family increases years of schooling for more highly educated women.

7.2 Calculating the Elasticities

How do the empirical estimates for women and men relate to the theoretical restrictions on wage elasticities derived in Theorem 1? Table 2 gives the sample means - of years of education, hours of labor, and hourly wage rates - for each of our four groups of interest: women and men with low and high levels of schooling respectively. The last row of the table shows the wage elasticities calculated from our estimates. Less educated women - those with fewer than 14 years of schooling - have an estimated wage elasticity of 1.53.[19] Men in the less educated group have an estimated wage elasticity of 1.21. From Theorem 1(a) it can therefore be seen that less educated women and men are characterized by *increasing* returns to education regardless of the assumed curvature on the utility function $u(.)$.

Now consider more highly educated individuals, those with 14 or more years of schooling. The last row of Table 2 reveals that women with 14 or more years of education have a wage elasticity of 0.13, while men in that group have a wage elasticity of -0.31. Thus from Theorem 1(b) more highly educated men are necessarily characterized by *decreasing* returns to education.

Table 2 here.

In summary, our findings for both women and men confirm the predictions of the theoretical model, that the returns to education are increasing for less educated women and men once labor supply behavior is taken into account. However, for more highly educated men, whose labor supply is unaffected by years of education, the empirical returns to education are decreasing. For more highly educated women they appear to be constant.

8 Conclusion

In this paper we modelled educational investment and hours of work in a competitive labor market in which heterogeneous workers have different productivities, both at home and in the workplace. We showed that there are increasing returns to education at the participation margin. These arise for two reasons. First, workers with greater workplace skills receive better wage offers and so are more likely to participate in the workplace, and a higher participation probability raises the *ex*

[19] The elasticity is calculated as $\frac{w}{l}\gamma^l$,where γ^l=7.43 (see Panel A in Table 4) and $\frac{w}{l}$ is the ratio of average wages (6.15) to average hours (29.88). These averages for each group are given in Table 5. This yields an elasticity for the less educated women of 7.43(6.15/29.88)=1.53. Unfortunately we cannot with our data calculate the longer run elasticities.

ante expected marginal return to human capital investment. Second, there is an increasing labor supply effect that arises because more educated workers may find it worthwhile to work longer hours. This further increases the marginal return to education. Those individuals most likely to be affected in this way are those types with large enough home productivity, who may be either involved in home or black market production, or may be characterized by a strong preference for other non-market sector activities.

Our model demonstrated how the importance of increasing returns to education varies across individuals. An important theoretical contribution of our paper is that we show that the presence (or absence) of increasing returns to education is closely related to the elasticity of labor supply with respect to wages. Specifically we show that the marginal return to education is increasing in education where the elasticity of labor supply with respect to wages is more than one. Conversely there are decreasing returns if this elasticity is negative (and it is ambiguous otherwise). Because this labor supply elasticity is sensitive to home productivity, it is not surprising that different individuals with different home productivities face different investment margins. As an illustration, we estimated a three-equation recursive model of working hours, wages and years of schooling using new data for Britain. We found empirical support for the main predictions of the model with regard to increasing returns to education once labor supply behavior is taken into account.

References

[1] Acemoglu, Daron (1996) "A Microfoundation for Social Increasing Returns in Human Capital Accumulation". *The Quarterly Journal of Economics*, Vol. 111, No. 3. (Aug., 1996), 779-804.

[2] Apps Patricia F and Ray Rees (1999) 'Individual vs. Joint Taxation in Models with Household Production', *Journal of Political Economy,* 107, 393-403.

[3] Ashenfelter, Orley and Cecilia Rouse (1998). "Income, Schooling, and Ability: Evidence from a New

[4] Sample of Identical Twins". *The Quarterly Journal of Economics*, Vol. 113, No. 1. (Feb., 1998), pp. 253-284.

[5] Becker, Gary S (1965). 'A Theory of the Allocation of Time', *The Economic Journal*, 75, pp. 493-517.

[6] Black, Sandra E, Paul J Devereux and Kjell G. Salvanes (2005). "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education". *Quarterly Journal of Economics,* 120(2), May.

[7] Booth Alison L and Melvyn G Coles (2005a). "The Under-participation Trap". . CEPR Discussion Paper No.. Forthcoming European Economic Review, 2007.

[8] Booth Alison L and Melvyn G Coles (2005b). "Education, Matching and the Allocative Value of Romance". CEPR Discussion Paper No. 5099. June.

[9] Booth Alison L and Melvyn G Coles (2006). "Modeling the Impact of Fiscal Policy on Labor Supply and Education in an Economy with Household and Market Production". Mimeo, University of Essex.

[10] Booth, Alison L. and H. Joo Kee (2005). 'Birth Order Matters: The Effect of Family Size and Birth Order on Educational Attainment', IZA Discussion Paper No. 1713, August.

[11] Card, David (1999). "The Causal Effect of Education on Earnings", Chapter 30 pp. 1801-1863 in *Handbook of Labor Economics*, Vol. 3A, (eds) Ashenfelter, Orley and David Card, Elsevier.

[12] Gomme, Paul, Richard Rogerson, Peter Rupert and Randall Wright (2004) "The Business Cycle and the Life Cycle". Mimeo.

[13] Gronau, R (1977). "Leisure, Home Production and Work – the Theory of the Allocation of Time Revisited", *Journal of Political Economy*, 85(6), pp 1099-1124, December.

[14] Imai, Susumi and Michael P. Keane (2004). "Intertemporal Labor Supply and Human Capital Accumulation." *International Economic Review*,45 (2), 601-641, May.

[15] Jaumotte, Florence (2003) "Female Labour Force Participation: Past Trends and Main Determinants in OECD Countries". Economics Dept. Working Paper No. 376, Paris: OECD.

[16] Pencavel, John (2002). "A Cohort Analysis of the Association between Work Hours and Wages among Men." *Journal of Human Resources,* 37 (2), 251-274. Spring.

[17] Rios-Rull, Jose-Victor (1993) "Working in the Market, Working at Home and the Acquisition of Skills: A General-Equilibrium Approach". *The American Economic Review,* 83(4), 893-907. September.

[18] Rosen, Sherwin (1983) "Specialization and Human Capital", *Journal of Labor Economics*, 1(1), January, 43-49.

Figures

Figure 1: Labor Supply

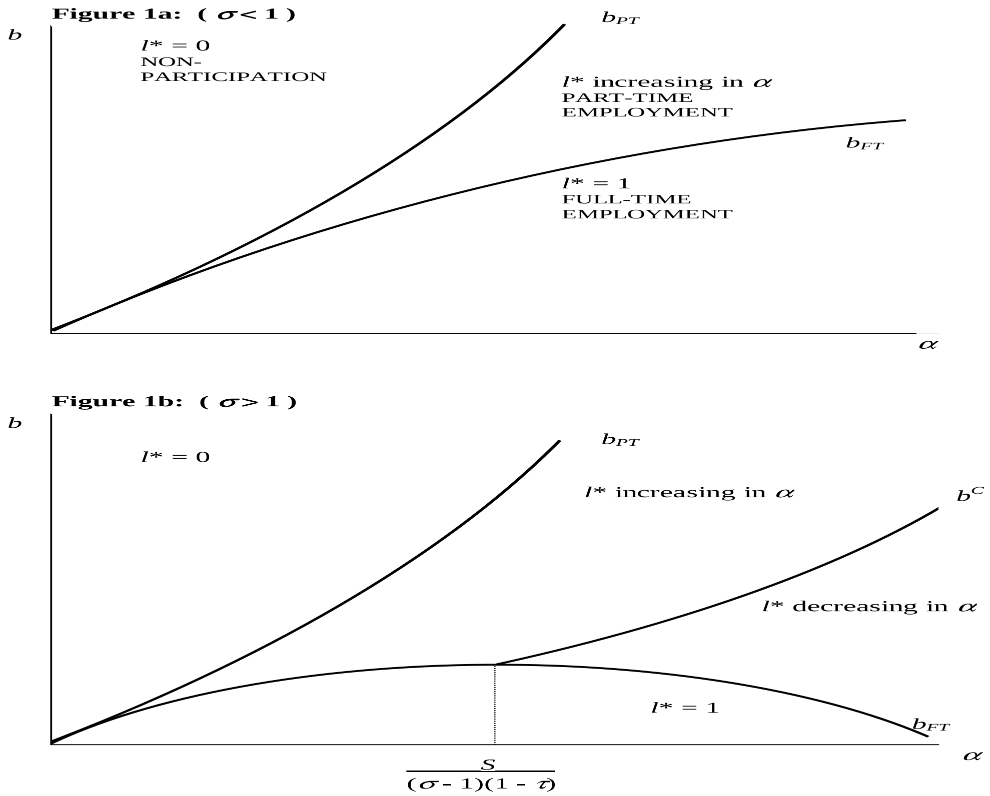


Figure 2: Optimal Education Choice (b_0 given)

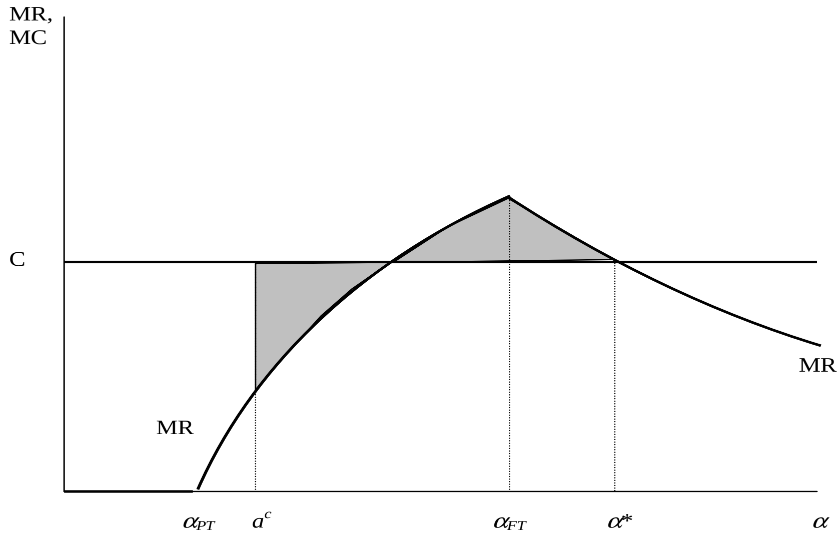


Figure 3A. Hours, LFP, and Wage by education (women)

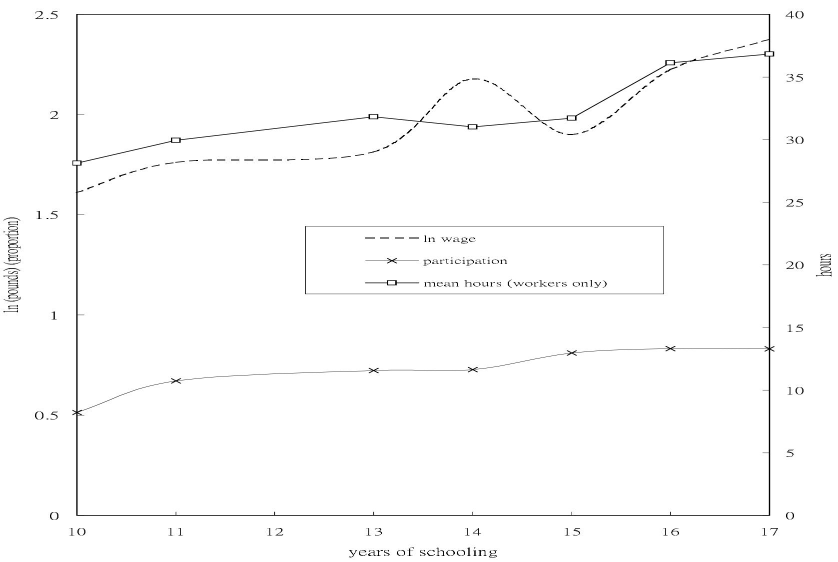
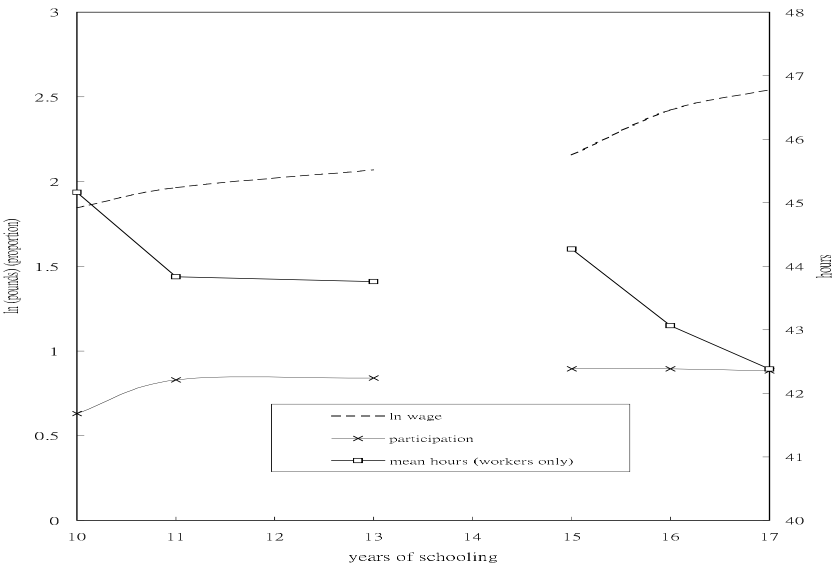


Figure 3B. Hours, LFP, and Wage by education (men)



Tables

Table 1A. Key parameter estimates (Women)

	All		low edu.		high edu.	
Model A, unrestricted model						
γ^l	0.787	(1.11)	7.435	(4.04)	0.493	(0.18)
γ^w	0.091	(7.04)	0.278	(4.44)	0.173	(1.88)
$\sigma 2_l$	289.714	(27.57)	361.405	(13.24)	274.144	(17.94)
$\sigma 2_w$	0.298	(34.42)	0.319	(11.90)	0.363	(19.52)
$\sigma 2_e$	5.124	(19.74)	1.098	(18.29)	0.443	(25.06)
σ_{lw}	3.580	(14.24)	5.473	(7.58)	4.062	(8.91)
σ_{le}	6.039	(3.00)	-4.669	(-2.26)	1.835	(1.75)
σ_{we}	0.144	(2.03)	-0.167	(-2.37)	0.045	(1.08)
Model B, restricted model, $\sigma_{le} = 0$						
γ^l	1.349	(1.23)	7.495	(12.14)	1.192	(0.48)
γ^w	0.118	(9.23)	0.242	(4.86)	0.182	(2.00)
$\sigma 2_l$	275.459	(30.71)	375.553	(18.57)	251.715	(23.41)
$\sigma 2_w$	0.308	(38.32)	0.282	(20.90)	0.367	(22.35)
$\sigma 2_e$	5.072	(19.97)	1.151	(16.55)	0.437	(25.62)
σ_{lw}	2.751	(12.64)	4.786	(11.60)	2.839	(8.17)
σ_{le}	0.000	0.00	0.000		0.000	
σ_{we}	-0.028	(-0.41)	-0.050	(-0.87)	0.026	(0.63)
LR test against Model A	not rejected		rejected		rejected	

t-values are in the parentheses.

Table 1A. Key parameter estimates (Women, continued)

Model C, restricted model, $\sigma_{le} = \sigma_{we} = 0$						
γ^l	1.502	(1.76)	7.470	(12.58)	0.977	(0.37)
γ^w	0.125	(9.85)	0.186	(10.98)	0.230	(2.52)
$\sigma 2_l$	272.018	(30.38)	368.190	(18.70)	252.541	(23.41)
$\sigma 2_w$	0.301	(40.99)	0.284	(22.43)	0.366	(23.54)
$\sigma 2_e$	5.160	(19.64)	1.158	(16.61)	0.436	(25.83)
σ_{lw}	2.252	(10.82)	4.497	(10.91)	2.983	(8.76)
σ_{le}	0.000		0.000		0.000	
σ_{we}	0.000		0.000		0.000	
LR test against Model A	rejected		rejected		rejected	
Model D, restricted model,all correlations are zero						
γ^l	1.322	(1.38)	6.142	(12.21)	1.253	(0.49)
γ^w	0.133	(9.71)	0.165	(8.41)	0.267	(2.53)
$\sigma 2_l$	256.120	(29.56)	327.242	(17.92)	240.818	(22.30)
$\sigma 2_w$	0.320	(41.28)	0.333	(22.38)	0.398	(24.15)
$\sigma 2_e$	5.254	(19.32)	1.138	(17.23)	0.442	(24.97)
σ_{lw}	0.000		0.000		0.000	
σ_{le}	0.000		0.000		0.000	
σ_{we}	0.000		0.000		0.000	
LR test against Model A	rejected		rejected		rejected	

t-values are in the parentheses.

Table 1B. Key parameter estimates (Men)

	All		low edu.		high edu.	
Model A, unrestricted model						
γ^l	0.119	(0.27)	6.943	(3.30)	-1.256	(-0.56)
γ^w	0.193	(10.71)	0.390	(4.96)	0.466	(4.84)
$\sigma 2_l$	294.771	(29.27)	390.80	(11.27)	267.68	(23.50)
$\sigma 2_w$	0.335	(17.74)	0.372	(7.12)	0.324	(19.09)
$\sigma 2_e$	4.872	(18.17)	1.206	(13.23)	0.399	(18.65)
σ_{lw}	2.352	(7.38)	5.429	(4.72)	3.801	(8.69)
σ_{le}	3.214	(1.47)	-6.149	(-2.37)	-0.023	(-0.03)
σ_{we}	-0.394	(-4.45)	-0.291	(-3.06)	-0.106	(-2.72)
Model B, restricted model, $\sigma_{le} = 0$						
γ^l	0.970	(5.03)	5.509	(7.06)	-2.219	(-3.02)
γ^w	0.176	(10.76)	0.341	(5.86)	0.499	(5.06)
$\sigma 2_l$	5.708	(173.9)	6.021	(100.9)	5.544	(133.7)
$\sigma 2_w$	-1.180	(-25.39)	-1.158	(-13.56)	-1.013	(-11.73)
$\sigma 2_e$	1.616	(28.37)	0.223	(2.78)	-0.921	(-17.39)
σ_{lw}	3.128	(12.44)	4.571	(9.15)	3.476	(9.96)
σ_{le}						
σ_{we}	-0.298	(-3.61)	-0.179	(-2.50)	-0.125	(-3.12)
LR test against Model A	rejected		rejected		rejected	

t-values are in the parentheses.

Table 1B. Key parameter estimates(Men, continued)

Model C, restricted model, $\sigma_{le} = \sigma_{we} = 0$						
γ^l	0.049	(0.25)	6.922	(8.59)	-3.373	(-4.35)
γ^w	0.118	(17.50)	0.224	(10.26)	0.195	(7.00)
$\sigma 2_l$	5.715	(165.77)	6.068	(99.40)	5.590	(131.36)
$\sigma 2_w$	-1.212	(-35.72)	-1.251	(-22.00)	-1.127	(-21.52)
$\sigma 2_e$	1.625	(28.32)	0.243	(2.91)	-0.919	(-17.14)
σ_{lw}	2.714	(10.49)	4.726	(9.35)	3.811	(10.37)
σ_{le}	0.000		0.000		0.000	
σ_{we}	0.000		0.000		0.000	
LR test against Model A	rejected		rejected		rejected	
Model D, restricted model,all correlations are zero						
γ^l	0.481	(2.99)	6.035	(9.79)	0.715	(-1.10)
γ^w	0.135	(18.83)	0.190	(8.06)	0.375	(12.18)
$\sigma 2_l$	5.648	(157.92)	5.931	(89.15)	5.511	(129.07)
$\sigma 2_w$	-1.163	(-32.42)	-1.097	(-16.07)	1.110	(-23.31)
$\sigma 2_e$	1.676	(27.49)	0.230	(2.84)	0.883	(-15.05)
σ_{lw}	0.000		0.000		0.000	
σ_{le}	0.000		0.000		0.000	
σ_{we}	0.000		0.000		0.000	
LR test against Model A	rejected		rejected		rejected	

t-values are in the parentheses.

Table 2. Hours, wage, and education by education level

	Men		Women	
	Low edu	High edu	Low edu	High edu
e	11.30	15.47	11.20	15.44
l	44.23	43.73	29.88	33.57
$\hat{w}$	7.68	10.76	6.15	8.70
elasticity	1.21	-0.31	1.53	0.13

9 Appendix A

Proof of Claim 2.

The definition of b_{FT} and CRRA implies

$$b_{FT} = \alpha(\alpha)^{-\sigma}/x'(0).$$

Differentiating with respect to α yields

$$\frac{\partial b_{FT}}{\partial \alpha} = \frac{[(1-\sigma)\alpha]}{x'(0)(\alpha)^{\sigma+1}}$$

and so

$$\frac{\partial b_{FT}}{\partial \alpha} \gtreqless 0 \text{ as } (1-\sigma)\alpha \gtreqless 0.$$

For $b \in (b_{FT}, b_{PT})$, (3) implies $\partial l^*/\partial\alpha$ is given by

$$\frac{\partial l^*}{\partial \alpha} = \frac{1}{-bx'' - \alpha^2 u''}[u'(y_{PT}) + \alpha l^* u''(y_{PT})]$$

where $y_{PT} = \alpha l^*$. Concavity of x and u implies $\partial l^*/\partial\alpha > 0$ if and only if $u'(y_{PT}) + \alpha l^* u''(y_{PT}) > 0$. CRRA now implies

$$\frac{\partial l^*}{\partial \alpha} \gtreqless 0 \text{ as } (1-\sigma)\alpha l^* \gtreqless 0, \tag{16}$$

where $b \in (b_{FT}, b_{PT})$ implies $l^* \in (0,1)$.

The statement of the Claim follows from these facts and that l^* is strictly decreasing in b for $b \in (b_{FT}, b_{PT})$ with $l^* = 1$ at $b = b_{FT}$ and $l^* = 0$ at $b = b_{PT}$. b^c is defined where $(1-\sigma)\alpha l^* = 0$.

Proof of Theorem 1.

For any (α, b), (4) implies MR has slope:

$$\frac{\partial[MR]}{\partial \alpha} = wE_b\left[\frac{\partial l^*}{\partial \alpha}.[u' + \alpha w l^* u''] + w l^{*2} u''\right] \tag{17}$$

where the utility functions are evaluated at $C = \alpha w l^*$. We prove parts (a) and (b) separately.

(a) Note that $\frac{\partial l^*}{\partial \alpha} = 0$ in the full time and non-participant region and so $\frac{\partial[MR]}{\partial \alpha} \leq 0$; i.e. strictly increasing returns can only occur in the part-time region. In the part-time region, (3) implies

$$\frac{\partial l^*}{\partial \alpha} = w\{\frac{u' + \alpha w l^* u''}{-[bx'' + \alpha^2 w^2 u'']}\} \tag{18}$$

The bracketed term in (17) is then equivalent to

$$-\frac{\left[bx''(1-l^*) + \alpha^2 w^2 u''\right]}{w}\left[\frac{\partial l^*}{\partial \alpha}\right]^2 + wl^{*2}u''.$$

As $x'' < 0$ by assumption, (5) guarantees this term is strictly positive and so MR is strictly increasing in α which establishes (a).

(b) As pointed out in (a), $\frac{\partial l^*}{\partial \alpha} = 0$ in the full time and non-participant region and $\frac{\partial [MR]}{\partial \alpha} \leq 0$; i.e. there are decreasing returns to education. Consider now the part-time region. Using the above equation for $\frac{\partial l^*}{\partial \alpha}$, the bracketed term in (17) is

$$\begin{aligned} &\left[w\frac{[u' + \alpha w l^* u'']^2}{-[bx'' + \alpha^2 w^2 u'']} + wl^{*2}u''\right] \\ \leq \quad & w\left[\frac{[u' + \alpha w l^* u'']^2}{-\alpha^2 w^2 u''} + l^{*2}u''\right] \end{aligned}$$

as $-x'' > 0$. Rearranging this latter expression is equivalent to:

$$\frac{wu'[u' + 2\alpha w l^* u'']}{-\alpha^2 w^2 u''}.$$

But note that $\frac{\partial l^*}{\partial \alpha} \leq 0$ in this region implies implies $u' + \alpha w l^* u'' \leq 0$ and $u'' < 0$ implies this term is strictly negative; i.e. there are decreasing returns to education, which completes the proof of Theorem 1.

Proof of Proposition 2. V is defined by

$$V(a,b) = \int_a^{\alpha^*} (MR(\alpha, b) - c_a)\, d\alpha$$

and a^c is defined as the solution to $V(a,b) = 0$. Differentiating with respect to b yields

$$\frac{\partial V}{\partial b} = \int_a^{\alpha^*} \frac{\partial\, [MR(\alpha, b)]}{\partial b} d\alpha.$$

Now (7) implies $\partial[MR]/\partial b = 0$ outside of the part-time region. In the part-time region $\alpha \in (\alpha_{PT}, \alpha_{FT})$, (3) in Claim 1 implies that l^* is strictly decreasing in b. Further CRRA with $\sigma \leq 1$ implies

$$\frac{\partial}{dl^*}[l^* u'(\alpha w l^*)] = \frac{(1-\sigma)\alpha w l^*}{[\alpha w l^*]^{\sigma+1}} > 0.$$

(7) and $\sigma < 1$ now imply $\partial[MR]/\partial b < 0$ in the part-time region. Hence it follows that $\partial V/\partial b < 0$ for all $a < \alpha_{FT}$. As we have already shown that $\partial V/\partial a > 0$ at $a = a^c$, the definition of a^c and the Implicit Function Theorem now imply a^c increases with b_0.

10 Appendix B

Table B. Sample statistics

		Women		Men	
Variables	Definition	mean	Std. dev.		
marr	Married	0.775		0.802	
healthgood	Good health	0.694		0.730	
healthave	Average health	0.202		0.182	
lote	Language other than Eng.	0.007		0.005	
white	Ethnicity:white people	0.977		0.978	
age	Age	39.873	8.50	40.173	8.49
chlhh02	No. of children 0–2	0.094	0.30	0.092	0.30
chlhh34	No. of children 3–4	0.114	0.33	0.112	0.33
chlhh511	No. of children 5–11	0.468	0.75	0.371	0.68
chlhh1215	No. of children 12–15	0.288	0.57	0.235	0.52
chlhh1618	No. of children 16–18	0.076	0.28	0.059	0.25
htown	home owner	0.141		0.129	
htownmort	paying off mortgage	0.630		0.668	
regrse	Dummy Rest of South-east	0.104		0.108	
regsthwt	Dummy, South-west	0.048		0.054	
regeana	Dummy, East Anglia	0.024		0.022	
regemid	Dummy, East Midlands	0.049		0.061	
regnw	Dummy, North-west	0.105		0.112	
regyorks	Dummy, York	0.068		0.057	
regrn	Dummy, Rest of North	0.024		0.029	
regwales	Dummy, Wales	0.158		0.158	
regscot	Dummy, Scotland	0.181		0.185	

Table B (continued). Sample statistics

		Women		Men	
Variables	Definition	mean	Std. dev.		
prntmiss	Not lived with both bio. parents	0.101		0.061	
fathmiss	father information missing	0.076		0.032	
mothmiss	Father information missing	0.028		1.181	
siblings	No. of siblings	2.532	2.12	2.357	1.99
birthorder	Birth order	2.236	1.69	0.318	0.47
pacert	Father's edu.: certificate	0.147		2.110	1.62
pafcert	Father's edu.:post-school qual.	0.254		0.192	
pauni	Father's edu.: university degree	0.062		0.242	
macert	mother's edu.: certificate	0.243		0.059	
mafcert	mother's edu.:post-school qual.	0.153		0.291	
mauni	mother's edu.: university degree	0.038		0.137	
bkssome	Some books in house (age 14)	0.343		0.037	
bksvfew	Few books in house (age 14)	0.248		0.390	
kidsubu	In a suburban area as a child	0.225		0.320	
kidtown	In a small town as a child	0.295		0.219	
kidrural	In a rural area as a child	0.131		0.294	
firstborn	First born	0.305		0.113	
yrschool	Years of schooling	13.088	2.37	13.479	2.33
lothinc	(ln) non-labor income	6.391	1.38	5.680	1.64
Obs.	3839		2665		

11 Appendix C

Table C1A. ML estimates—Hours equation (Women)

	All		low edu.		high edu.	
constant	38.676	(3.88)	27.080	(-1.01)	38.785	(0.97)
age	98.202	(1.95)	34.884	(-0.52)	98.219	(1.40)
age^2	-162.093	(-2.65)	42.223	(0.52)	-162.082	(-1.85)
$eng2$	-8.153	(-1.83)	-4.043	(-0.79)	-8.187	(-0.89)
marr	7.546	(10.48)	4.497	(4.22)	7.561	(7.26)
$chlhh02$	-8.139	(-7.42)	-8.975	(-5.19)	-8.244	(-5.92)
$chlhh34$	-7.964	(-7.81)	10.791	(-6.60)	-8.243	(-6.29)
$chlhh511$	-4.831	(-11.00)	-4.069	(-5.91)	-4.859	(-8.15)
$chlhh1215$	-1.639	(-2.97)	-1.719	(-2.18)	-1.264	(-1.61)
$chlhh1618$	-0.800	(-0.72)	-1.759	(-0.99)	-0.594	(-0.40)
lothinc	-8.536	(-56.68)	-7.835	(-36.00)	-7.504	(-24.41)
htown	8.654	(8.410	13.201	(8.87)	8.506	(5.54)
htownmort	5.675	(4.93)	10.772	(6.43)	5.748	(3.36)
healthgood	3.130	(3.23)	5.592	(3.88)	2.961	(2.07)
healthaverage	4.910	(6.88)	8.215	(7.87)	4.779	(4.50)
γ^l	0.787	(1.11)	7.435	(4.04)	0.493	(0.18)
estimates of tHE Covariance matrix						
$\sigma 2_l$	289.714	(27.57)	361.405	(13.24)	274.144	(17.94)
$\sigma 2_w$	0.298	(34.42)	0.319	(11.90)	0.363	(19.52)
$\sigma 2_e$	5.124	(19.74)	1.098	(18.29)	0.443	(25.06)
σ_{lw}	3.580	(14.24)	5.473	(7.58)	4.062	(8.91)
σ_{le}	6.039	(3.00)	-4.669	(-2.26)	1.835	(1.75)
σ_{we}	0.144	(2.03)	-0.167	(-2.37)	0.045	(1.08)

t-values are in the parentheses.

Table C2A. ML estimates—Wage equation (Women)

	All		low edu.		high edu.	
constant	-0.634	(-1.74)	-2.204	(-2.33)	-1.664	(-1.08)
age	5.073	(3.71)	2.479	(1.11)	4.889	(2.16)
age^2	-6.032	(-3.58)	-1.868	(-0.69)	-6.117	(-2.18)
$eng2$	-0.138	(-0.77)	-0.044	(-0.16)	-0.238	(-0.40)
white	0.230	(1.63)	0.039	(0.31)	0.030	(0.14)
regrse	-0.030	(-0.87)	-0.010	(-0.18)	-0.004	(-0.07)
regsthwt	-0.139	(-2.82)	-0.057	(-0.80)	-0.130	(-1.60)
regeana	-0.079	(-1.08)	-0.022	(-0.23)	-0.128	(-0.91)
regemid	-0.083	(-1.62)	-0.049	(-0.70)	-0.057	(-0.66)
regnw	-0.062	(-1.74)	-0.030	(-0.56)	-0.049	(-0.84)
regyorks	-0.043	(-0.92)	-0.042	(-0.60)	-0.003	(-0.04)
regrn	-0.071	(-1.05)	0.057	(0.53)	-0.142	(-1.22)
regwales	-0.103	(-2.94)	-0.083	(-1.82)	-0.067	(-1.12)
regscot	-0.065	(-1.94)	-0.104	(-2.20)	0.025	(0.46)
married	0.023	(0.74)	0.045	(0.54)	0.013	(0.89)
γ^w	0.091	(7.04)	0.278	(4.44)	0.173	(1.88)

t-values are in the parentheses.

Table C3A. ML estimates—Education equation (Women)

	All		low edu.		high edu.	
Education equation						
constant	13.258	(23.85)	12.561	(52.80)	15.554	(62.44)
age	-2.038	(-3.93)	-2.898	(-9.07)	0.166	(0.78)
white	0.811	(1.73)	0.126	(0.74)	-0.232	(-1.09)
fathmiss	0.040	(0.25)	0.063	(0.70)	-0.177	(-2.07)
mothmiss	0.294	(1.33)	-0.008	(-0.06)	0.019	(0.17)
prntmiss	-0.115	(-1.28)	-0.180	(-2.57)	-0.003	(-0.06)
siblings	-0.149	(-5.21)	-0.064	(-3.51)	-0.001	(-0.10)
firstborn	0.124	(1.24)	0.021	(0.31)	0.085	(2.07)
birthorder	0.057	(1.45)	0.008	(0.32)	-0.007	(-0.38)
pacert	0.497	(4.15)	0.213	(2.71)	0.122	(2.49)
pafcert	0.330	(3.34)	0.145	(2.31)	0.052	(1.20)
pauni	1.398	(6.25)	0.465	(2.47)	0.366	(5.84)
macert	0.602	(5.77)	0.150	(2.23)	-0.009	(-0.21)
mafcert	0.758	(6.02)	0.132	(1.52)	0.024	(0.49)
mauni	1.244	(4.44)	0.345	(1.75)	0.220	(2.95)
bkssome	-0.274	(-3.08)	-0.159	(-2.73)	-0.101	(-2.75)
bksvfew	-0.742	(-7.13)	-0.294	(-4.46)	-0.134	(-2.60)
kidsubu	0.061	(0.59)	0.046	(0.69)	0.057	(1.34)
kidtown	-0.330	(-3.47)	-0.125	(-2.10)	-0.022	(-0.50)
kidrural	-0.052	(-0.42)	-0.058	(-0.71)	-0.056	(-1.09)
function	-10883		-4596		-4481.	

t-values are in the parentheses.

Table C1B. ML estimates—Hours equation (Men)

	All		low edu.		high edu.	
constant	32.993	(3.00)	-26.550	(-0.88)	65.630	(1.74)
age	82.741	(1.86)	-24.445	(-0.32)	5.369	(0.09)
age^2	-150.746	(-2.75)	64.289	(0.69)	-59.975	(-0.79)
$eng2$	-0.804	(-0.17)	-0.431	(-0.05)	-0.359	(-0.05)
marr	10.246	(10.34)	5.516	(3.59)	8.467	(6.79)
$chlhh02$	-3.172	(-2.26)	0.239	(0.10)	-8.479	(-5.36)
$chlhh34$	0.621	(0.54)	-0.286	(-0.16)	2.241	(1.47)
$chlhh511$	-1.616	(-2.80)	-0.387	(-0.42)	-1.006	(-1.35)
$chlhh1215$	-0.811	(-1.15)	0.704	(0.62)	-1.171	(-1.30)
$chlhh1618$	1.947	(1.32)	1.484	(0.62)	-0.127	(-0.08)
lothinc	-5.556	(-51.77)	-7.847	(-30.31)	-3.652	(-34.70)
htown	9.145	(7.81)	16.457	(8.64)	10.827	(7.03)
htownmort	5.281	(4.10)	9.603	(4.65)	12.436	(7.06)
healthgood	10.999	(9.60)	4.539	(2.50)	4.693	(3.04)
healthaverage	11.274	(12.94)	10.237	(7.31)	4.241	(3.68)
γ^l	0.119	(0.27)	6.943	(3.30)	-1.256	(-0.56)
estimates of the Covariance matrix						
$\sigma 2_l$	294.771	(29.27)	294.771	(29.27)	294.771	(29.27)
$\sigma 2_w$	0.335	(17.74)	0.335	(17.74)	0.335	(17.74)
$\sigma 2_e$	4.872	(18.17)	4.872	(18.17)	4.872	(18.17)
σ_{lw}	2.352	(7.38)	5.429	(4.72)	3.801	(8.69)
σ_{le}	3.214	(1.47)	-6.149	(-2.37)	-0.023	(-0.03)
σ_{we}	-0.394	(-4.45)	-0.291	(-3.06)	-0.106	(-2.72)

t-values are in the parentheses.

Table C2B. ML estimates—Wage equation (Men)

	All		low edu.		high edu.	
constant	-2.858	(-6.28)	-4.407	(-3.77)	-6.372	(-3.96)
age	11.409	(6.91)	9.006	(3.40)	6.324	(2.85)
age^2	-13.557	(-6.72)	-9.717	(-3.08)	-7.488	(-2.74)
lote	-0.438	(-2.58)	-0.024	(-0.02)	-0.754	(-2.47)
white	0.059	(0.61)	-0.097	(-0.43)	0.148	(1.30)
regrse	-0.027	(-0.59)	0.109	(1.46)	-0.037	(-0.66)
regsthwt	-0.073	(-1.30)	0.097	(1.13)	-0.159	(-2.14)
regeana	0.044	(0.40)	-0.033	(-0.23)	0.064	(0.46)
regemid	-0.102	(-1.64)	-0.060	(-0.70)	-0.093	(-1.07)
regnw	-0.133	(-2.84)	-0.074	(-1.06)	-0.084	(-1.40)
regyorks	-0.078	(-1.21)	-0.021	(-0.22)	-0.080	(-0.97)
regrn	-0.145	(-1.77)	-0.129	(-1.00)	-0.146	(-1.47)
regwales	-0.129	(-2.93)	-0.066	(-1.05)	-0.138	(-2.38)
regscot	-0.170	(-4.17)	-0.123	(-2.07)	-0.134	(-2.51)
married	0.134	(5.39)	0.204	(3.64)	0.082	(2.99)
γ^w	0.193	(10.71)	0.390	(4.96)	0.466	(4.84)

t-values are in the parentheses.

Table C3B. ML estimates—Education equation (Men)

	All		low edu.		high edu.	
constant	14.353	(32.67)	11.749	(29.79)	15.617	(97.63)
age	0.507	(0.87)	-1.836	(-4.19)	0.729	(3.28)
white	-0.980	(-3.30)	0.435	(1.41)	-0.325	(-3.15)
fathmiss	-0.239	(-1.25)	-0.089	(-0.57)	0.047	(0.55)
mothmiss	-0.309	(-1.20)	-0.089	(-0.51)	0.008	(0.07)
prntmiss	-0.232	(-1.85)	-0.217	(-2.38)	-0.124	(-2.25)
siblings	-0.219	(-6.34)	-0.073	(-2.76)	-0.047	(-2.78)
firstborn	0.222	(1.96)	0.009	(0.11)	0.070	(1.60)
birthorder	0.079	(1.70)	-0.001	(-0.04)	0.009	(0.40)
pacert	0.225	(1.73)	0.059	(0.61)	-0.024	(-0.50)
pafcert	0.328	(2.69)	0.192	(2.11)	-0.083	(-1.74)
pauni	1.367	(5.05)	0.298	(1.25)	0.259	(3.63)
macert	0.586	(4.94)	0.341	(3.88)	0.130	(3.04)
mafcert	0.670	(4.36)	0.309	(2.55)	0.203	(3.73)
mauni	1.037	(3.20)	0.627	(2.02)	0.315	(3.68)
bkssome	-0.207	(-1.87)	0.051	(0.59)	-0.097	(-2.46)
bksvfew	-0.404	(-3.41)	0.015	(0.17)	-0.095	(-2.05)
kidsubu	0.404	(3.39)	0.033	(0.36)	0.142	(3.16)
kidtown	0.272	(2.50)	0.019	(0.24)	0.052	(1.21)
kidrural	-0.062	(-0.41)	0.054	(0.52)	0.004	(0.06)
function	-8444		-3281		-3822	

t-values are in the parentheses.