Tax Policy and Work Specialisation*

A. L. Booth
Economics Department,
University of Essex
Wivenhoe Park CO4 3SQ, UK

M.G. Coles
Economics Department
University of Essex
Wivenhoe Park CO4 3SQ, UK

December 4, 2007

Abstract

This paper considers how asymmetric tax treatment, where labour market earnings are taxed but household production is untaxed, affects educational choice and labour supply. We show that taxes on labour market earnings can generate a large (non-marginal) switch to home production and the ensuing deadweight losses are large. Using a cross-country panel, we find that gender differences in labour supply responses to tax policy can play an important role in explaining differences in aggregate labour supply across countries.

Keywords: Increasing returns, tax policy, gender, labour supply, education

JEL Classification: H24, H3, J22, J24, J31.

---

*We are grateful to the ARC for financial support under Discovery Project Grant No. DP0556740. We also thank Florence Jaumotte for providing access to her OECD tax database, Hiau Joo Kee for excellent data assistance, and Patricia Apps, Bas Jacobs and Ian Walker for helpful comments.

1 Introduction

This paper considers how asymmetric tax treatment, where labour market earnings are taxed but household production is not taxed, affects educational choice and labour supply. A key insight of the model is that individuals have an incentive to specialise; either to focus on home production, or to invest in general human capital and work mainly in the labour market. For reasons that will become clear, we show why women, who typically have greater labour supply elasticities than men, might face increasing returns to education. We further show that a tax on labour market earnings can generate a large (non-marginal) switch to home production and that the ensuing deadweight loss is not a small Harberger triangle.

There is a large literature which analyses optimal education choice and dynamic labour supply within a lifecycle framework (see Trostel and Walker (2006) and the references contained therein). As checking second order conditions is complicated in such frameworks, the typical approach is to assume an interior solution and characterise a solution to the first order conditions. But there is good reason to believe the second order conditions might fail. For example consider a one-period textbook case where the agent chooses education e, labour supply l and consumption c to solve the utility maximisation problem

$$\max_{e,c,l} u(c, 1-l) \text{ s.t. } pc \leq M + [wH(e)]\, l - \gamma e,$$

where $H(e)$ describes the worker's general human capital given education e, w is the market wage rate for skills, and γ is the cost of acquiring education. As labour market earnings $wlH(e)$ exhibit increasing returns to scale in education and labour supply, this problem is not a concave programming problem. Thus second order conditions are likely to fail and corner solutions apply. For example, it is an empirical fact that many exit education at compulsory school leaving age. It is also an empirical regularity that some do not participate in the workplace and instead focus on home production. The same second order condition problem is faced by more complicated dynamic models in which earnings are also of the form wHl.[1]

In this paper we extend the above simple optimisation problem to allow for a two-period

[1] For example see the influential paper Trostel (1993) among many others.

model of home production and individual heterogeneity in home and workplace productivity. As education and labour supply are complements in the above earnings function, $wH(e)l$, and in the model to be developed below, educational choice and labour supply will be positively correlated across individuals. By increasing earned wages in the workplace, more education tends to increase individual labour supply. But it is not difficult to see that the return to education is affected crucially by the anticipated utilization of education; i.e., by expectations of future labour supply. If one does not anticipate being in the workplace for long, there is little sense in making a costly educational investment that will bring only a small market return. Thus more education and greater labour supply are mutually reinforcing choices. Using cross-sectional data across a huge array of countries, Trostel and Walker (2006) show there is a universal strong positive correlation between individual education choice and labour supply. Their insights are also consistent with the trend increase in female education and participation rates in virtually all OECD countries (see Jaumotte (2003)).

But here we go a step further and argue these re-inforcing effects may generate increasing marginal returns to education. Specifically, we will show that the expected marginal return to education is proportional to $l^*(e;\theta)wH'(e)$. This term is composed of two effects:

(i) $wH'(e)$ is the Mincerian return to education - it describes the increase in the market wage rate through an increase in education;

(ii) $l^*(e;\theta)$ is the optimal labour supply choice of an individual with education e and characteristics θ, and thus describes the utilisation rate of human capital in the workplace.

When utility is linear in consumption, we will show there are increasing marginal returns to education if $l^*(e;\theta)wH'(e)$ is increasing in education, which in turn requires that labour supply is sufficiently elastic. As Trostel and Walker (2006) find the elasticity of $l^*(e;\theta)$ with respect to education is, on average, around four times larger for women than for men, then women are much more likely to face increasing returns to education. Indeed there are necessarily increasing returns for individuals at the non-participant margin, as the marginal return to education is zero when $l^* = 0$.

Of course in a competitive environment with no taxation, the phenomenon of private increasing returns to education does not, by itself, yield a market failure. But in the analysis to be developed below, we show that - once taxes are imposed on labour market earnings while home production remains untaxed - private increasing returns lead to large switches in behaviour. A worker who otherwise might invest in education and participate in the labour market (paying income tax to the government), instead switches to non-participation and pure home production (paying no tax). An important contribution of this paper is to show that it is typically women who experience the correspondingly large deadweight losses.

[**Insert Figure 1 near here**]

Our paper also uses an international panel data set to describe how different tax policies have affected average working-age male and female participation rates. As Figure 1 clearly demonstrates, male participation rates are typically high and closely clustered, in contrast to the much lower and more heterogeneous female participation rates. Using fixed-effects estimation and controlling for demographics, we find that average tax rates, taxes on second earners and child benefits have significant detrimental effects on female participation rates. As an illustration of the magnitude of the policy effects, we calculate what US participation would look like if the US had the fiscal policy values of a typical Scandinavian country, Sweden, which has one of the highest female participation rates in OECD countries.

1.1 Related Literature

Time-use studies show that non-participating women of working age are typically engaged in home-production rather than leisure (see for example Apps and Rees, 1996; Apps, 2003). Indeed Burda et al. (2007) establish that female and male leisure hours are roughly equal. Instead it is the allocation of work hours between the workplace and domestic production which differs significantly between the genders. The central theme of this paper is to consider how tax policy distorts the allocation of work hours between the workplace and domestic production, and how education choice is also affected.[2]

Perhaps the closest paper to ours is Bovenberg and Jacobs (2005) (but also see Jacobs (2005) and Jacobs and Bovenberg (2007)). That paper considers optimal tax policy where the government taxes labour income but, as workers also underinvest in education, it in addition offers

[2] In a series of important papers, Apps and Rees (1996; 1999) consider how tax policy affects labour supply and household production but they do not consider educational investments, which are our main focus.

education subsidies. As their framework is closely related to the one to be developed in our paper, it is at first sight surprising they do not need to consider increasing returns. However, the critical difference between the frameworks is they assume marginal home productivity is zero at the non-participation margin; i.e. where $l = 0$.[For the referees: in Appendix B (not intended for publication), we illustrate how their arguments are affected when marginal home productivity is strictly positive.] Whenever marginal home productivity is sufficiently large, non-participation may become a binding constraint and increasing returns are then a robust phenomenon. Unfortunately increasing returns imply first order conditions are no longer sufficient to describe optimal behaviour.[3] Furthermore a marginal tax analysis is no longer valid, as we shall show that small changes in tax rates can lead to discontinuous jumps in educational investment and labour supply. It is not surprising then that the theoretical literature typically avoids this non-concavity issue. But the optimisation theory, when properly done, is interesting since the switch to home production can lead to (discontinuously) large deadweight losses.

In a neglected paper, Rosen (1983) provides the appropriate intuition for the results identified in our model below. He considers a labour market model with two skills where in the first period the worker invests in either or both skills, and in the second period then allocates time across each skill. He shows agents tend to specialise - to invest mainly in one skill and allocate time to utilising that skill. In our context, individuals specialise and become either work specialists or home specialists. Work specialists have high participation rates in the labour market and so enjoy a higher market return to investments in general human capital. Work specialists thus tend to invest in high levels of education. Conversely home specialists have low participation rates (or perhaps work part-time) and, on assumption that post-compulsory schooling improves domestic productivity less than market-sector productivity, invest less in education. An individual's optimal choice of specialism depends on comparative advantage arguments - how productive is that individual in the home relative to the workplace. But we shall show that the partition between the two specialisations depends on taxes: higher income taxes imply more individuals become home specialists. A small marginal increase in income tax leads some to switch to home specialisation. The large drop in tax paid (a tax payer is

[3] If marginal home productivity is very small, the Bovenberg-Jacobs approach may still apply, as increasing returns only occur over a small region and a solution to the first order conditions will describe optimality.

switching to domestic production and pays no tax) yields a correspondingly large deadweight loss.

Rios-Rull (1993) considers optimal skills acquistion in a competitive economy with home and market production, while Booth and Coles (2007a) consider that decision within an imperfectly competitive labour market. Those papers do not consider how income taxes distort education and labour supply. Booth and Coles (2007a), however, argue the government might offer publicly provided childcare benefits. By paying such benefits conditional on participation in the labour market, this tax instrument encourages greater female participation rates and so targets the large deadweight losses identified here. Finally Schindler and Weigert (2007) neatly finesse the second order condition problem highlighted here by assuming education increases the probability of earning a high wage w^H in the second period, and there are only two wage outcomes $w = w^L, w^H$. The first period problem then reduces to

$$\max_e p(e)V^H + (1 - p(e))V^L - \gamma e,$$

where $p(e)$ is the probability of the high wage outcome and V^i is the second period payoff which depends only on the wage outcome $i = L, H$ (i.e. V^i does not otherwise depend on e). Assuming p concave then guarantees a concave programming problem.

1.2 Outline of the Paper

The next section decribes the model and Section 3 derives the optimal education and labour supply choices of individuals for a given tax environment. Section 4 considers how changes in the tax environment distort those decisions and establishes there are large deadweight losses when there are increasing (private) returns to education. It argues that women are most likely to bear those costs. Using an international panel data set, Section 5 then describes how tax policy has affected average male and female participation rates across several OECD countries.

2 The Model

A representative individual's (pre-tax) earnings are wl, where w is the (competitive) market wage rate for that individual's skills, and l is labour supplied to the market. Following Mincer (1958), the wage rate is $w = w(a, e)$ where a describes the worker's ability and e is education. Our insights are driven entirely by the fact that earnings wl then exhibit increasing returns to scale in e, l. Unfortunately increasing returns significantly complicate the analysis as we must formally consider the failure of second order conditions. To keep the analysis manageable, many aspects of the model are kept deliberately simple. Generalization in several ways is straightforward but would unnecessarily complicate the presentation and obscure the relevant insights.

Young people typically make their human capital investments prior to meeting their future partners and before raising families. We therefore consider educational choice in a two period framework in which educational choice is made in the first period given expectations of second period home productivity. This timing is not critical to the results - there are (joint) increasing returns in earnings regardless of whether education e and labour supply l are chosen simultaneously or sequentially.

Thus consider a representative worker who 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. In the second period, home productivity b is realised. The worker then has a unit time endowment where time $l \in [0, 1]$ is spent working in the labour 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 carrying out other domestic activities.[4]

To keep the algebra under control, we assume the market wage increases linearly with human capital; i.e. $w = w_0(a + e)$ with $w_0 > 0$. This assumption is not critical to the results - it simplifies the algebra as w_0 then describes the Mincerian rate of return to education and is the same for all. The cost of attaining education level e is also linear, γe, where $\gamma > 0$ and is also the same for all. It is straightforward to show that the results below also hold when higher

[4] For people with no children, h might be pure leisure, although time-use studies show that, even in partnered households without children, considerable time is spent on home-related activities such as cooking and cleaning.

ability types have lower education costs.[5] It is also useful to abstract from income effects by assuming γ is a disutility cost (i.e. acquiring education is costly only in that it requires passing exams) and that all have zero initial wealth.

The government's tax program is described by a pair (S, τ) with $S > 0$, $\tau \in [0, 1]$. Given an individual's gross labour market earnings $y_G = wl$, after-tax income is

$$y = S + (1 - \tau)wl.$$

Thus S/τ defines a break even level of income: workers with pre-tax earnings $y_G < S/\tau$ receive a net transfer $S - \tau y_G$ from the government, while those earning $y_G > S/\tau$ pay net tax $\tau y_G - S$. We refer to S as the level of social insurance or lump-sum transfers (e.g. a mother might receive child benefit payments) and τ as the marginal tax rate on additional earnings. The worker's second period budget constraint is then

$$pc \leq S + (1 - \tau)wl = S + (1 - \tau)w_0(a + e)l$$

where c is consumption and p is the price of the consumption good which we normalise $p = 1$. Note the critical non-concavity: (after-tax) earnings have increasing returns in e and l.

Again for simplicity assume second period utility is additively separable in consumption and home production $h = 1 - l$; i.e.

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

where u, x are strictly increasing, strictly concave and twice differentiable functions. Note this specification implies workers with higher b have a higher marginal return to home production.

We assume education increases general human capital in the workplace but does not improve domestic productivity. Of course this need not be the case; e.g. Rosen (1983). But in our context, it seems unlikely that university degree schemes much improve domestic skills. Indeed

[5] A more general specification might instead assume $w(.)$ 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 which depends on type. 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 $y_G = lw(a, \widehat{c_a}^{-1}(k))$ continue to imply joint increasing returns to l and k. Assuming $w(.)$ is linear is a useful simplification which implies w_0 can be interpreted as the Mincer 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 never optimal.

if education were to increase home productivity as much as it increases workplace productivity, we would not observe the strong positive association between education and hours of work found in Trostel and Walker (2006). For simplicity, then, we assume (post-compulsory) education has a negligible impact on domestic productivity; i.e. the worker takes domestic productivity b as given.[6]

As utility is strictly increasing in c, the budget constraint always binds and so consumption $c = S + (1-\tau)y_G$. As the time constraint implies $h = 1 - l$, the worker's second period optimisation problem is equivalent to choosing $l \in [0,1]$ to solve

$$\max_{l\in[0,1]} u(S + (1-\tau)w_0\alpha l) + bx(1-l), \tag{1}$$

where productivity $\alpha = a + e$ is given in the second period. This objective function is strictly concave in l and so standard first order conditions fully describe the maximum. Claim 1 below describes those conditions. Given optimal labour supply and tax parameters (S, τ), second period utility is

$$U_2^*(\alpha, b; S, \tau) = \left[\max_{l\in[0,1]} u(S + (1-\tau)w_0\alpha l) + bx(1-l)\right].$$

The worker in the first period chooses education to solve

$$V_1(a, b; S, \tau) = \max_{\alpha \geq a} \left[U_2^*(\alpha, .) - \gamma[\alpha - a]\right]. \tag{2}$$

We shall show below that $U_2^*(\alpha, .)$ is not concave in α. Although the optimal education rule, denoted $\alpha^*(a, b)$, is (generically) unique, there may be several solutions to the first order conditions for optimality. Thus simply solving the first order conditions is not sufficient to identify optimal education choice. The following not only fully determines optimal $\alpha^*(.)$, it also describes how varying the tax program (S, τ) affects that decision and ex-post labour supply.

[6] In less developed countries, maternal education is found to have a significant effect on child quality. Our model is not intended to capture this effect and our empirical work relates to developed countries, as will be seen.

3 Optimal Education and Labour Supply

In this section we take the tax parameters (S, τ) as given and solve for the worker's optimal first period education choice and second period labour supply. The section that follows then considers how changing tax parameters (S, τ) affects those choices.

3.1 Second period labour supply

Given second period productivity parameters (α, b), the worker's optimal second period labour supply choice, properly denoted $l^*(\alpha, b; S, \tau)$, solves (1). As (S, τ) is held fixed in this section, however, we simplify notation here by subsuming reference to S, τ.

As there may be corner solutions, define the following functions

$$\begin{aligned} b_{PT}(\alpha) &= (1-\tau)\alpha w_0 u'(S)/x'(1), \\ b_{FT}(\alpha) &= (1-\tau)\alpha w_0 u'(S + (1-\tau)w_0\alpha)/x'(0). \end{aligned}$$

Note that b_{PT} is linear and increasing in α, but b_{FT} is non-linear and may be a decreasing function of α (see Figures 2a and 2b below). For now note that concavity of u and x implies $b_{PT} \geq b_{FT}$ with strict inequality if either u or x is strictly concave. Figures 2a and 2b plot these functions when $u(.)$ exhibits constant relative risk aversion (CRRA).

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

Claim 1. Optimal Second Period Labour 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

$$bx'(1 - l^*) = \alpha w_0 (1-\tau) u'(S + (1-\tau)\alpha w_0 l^*). \tag{3}$$

Claim 1 describes the Kuhn-Tucker conditions implied by (1). People with very high home productivity, $b > b_{PT}$, do not participate in the labour 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 labour supply implies $l^* \in (0, 1)$ and (3) describes the optimal trade-off between home production and employment in the market sector. Although one might interpret l^* as the worker's average participation rate over a working lifetime, 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 participation region while $b \geq b_{PT}$ is the non-participant region.

Given this description of optimal labour supply in the second period, the next step is to determine the optimal education choice e in the first period. In this first period problem, we show there are increasing marginal returns to education for some education levels.

The optimal education choice e depends on how second period labour supply l^* varies with productivity. Standard comparative statics establish that labour supply l^* is strictly decreasing in home productivity in the part-time region (of course l^* is constant in the constrained regions). If utility is linear in consumption, labour supply l^* is unambigously increasing with α. 'Risk aversion' is more complicated as there are 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.[7] Claim 2 describes how l^* varies with α in this case.

Claim 2. Optimal Labour 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 $\alpha < S/[(\sigma-1)(1-\tau)w_0]$, l^* and b_{FT} are both increasing in α;

(b) for $\alpha > S/[(\sigma-1)(1-\tau)w_0]$, 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 2a and 2b depict these two cases. Figure 2a describes the thresholds b_{PT} and b_{FT}

[7] In our context, with no uncertainty, CRRA refers of course to the degree of concavity of the utility-of-consumption function.

for low levels of risk aversion, $\sigma < 1$. Claim 2 implies labour 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 2b, is strictly increasing in α.

Figures 2a, 2b here.

3.2 First period education

Given the characterization of l^* above, we now consider the optimal education choice in the first period. Note that a worker who invests to productivity level $\alpha \geq a$ in the first period obtains expected utility

$$U_1(\alpha, .) \equiv [u(S + (1-\tau)\alpha w_0 l^*) + bx(1-l^*)] - \gamma[\alpha - a],$$

with l^* as defined in Claim 1. A most important object for what follows is

$$MR = (1-\tau) w_0 l^* u'(c), \tag{4}$$

where $c = S + (1-\tau)\alpha w_0 l^*$. Totally differentiating U_1 with respect to α, noting that l^* is chosen optimally, the Envelope Theorem implies

$$\frac{dU_1}{d\alpha} = MR - \gamma.$$

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 $MR = (1-\tau)w_0 l^*$. Thus the marginal return to education is the Mincer rate of return (net of tax) multiplied by expected labour supply. Since Claim 2 with $\sigma = 0$ 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 marginal returns to education. The reason is simple - very low α workers who do not participate in the labour 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 labour supply, and hence the utilisation rate of human capital, 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. We now simplify by assuming a CRRA utility function with $\sigma \leq 1$.[8]

To describe the optimal education choice, we need to describe MR as a function of α. To do this, first define $\alpha_{PT}(b)$ as the inverse function of $b = b_{PT}(\alpha)$; i.e. $\alpha_{PT} = (b_{PT})^{-1}(b)$. This implies

$$\alpha_{PT} = bx'(1)/[(1-\tau)w_0 u'(S)].$$

Note that $\alpha = \alpha_{PT}(b)$ simply relabels the locus labelled $b = b_{PT}$ in Figure 2a.

We also need to define the inverse function of $b_{FT}(\alpha)$. Note for $\sigma < 1$ that Claim 2 implies $b = b_{FT}(\alpha)$ is a strictly increasing function. Hence its inverse function is also well-defined and so define $\alpha_{FT} = (b_{FT})^{-1}(b)$ and $\alpha_{FT}(.)$ is also an increasing function. $\alpha_{FT}(b)$ corresponds to the locus labelled b_{FT} in Figure 2a. Figure 2a and (4) now imply $MR = MR(\alpha, b)$ where

$$\begin{aligned} MR &= 0 \text{ if } \alpha \leq \alpha_{PT}(b) \\ &= (1-\tau)w_0 l^* u'(S + (1-\tau)\alpha w_0 l^*) \text{ if } \alpha \in (\alpha_{PT}(b), \alpha_{FT}(b)) \\ &= (1-\tau)w_0 u'(S + (1-\tau)\alpha w_0) \text{ if } \alpha \geq \alpha_{FT}(b). \end{aligned} \tag{5}$$

Figure 3 below graphs MR by productivity α, given $\sigma \leq 1$ and b fixed, and on the assumption MR is monotonic over the part-time region. For productivities $\alpha \leq \alpha_{PT}(b)$, the worker does not participate in the labour market and so $MR = 0$. For productivities $\alpha \geq \alpha_{FT}(b)$,the worker chooses $l^* = 1$ and MR is then decreasing in α as the marginal utility of consumption decreases

[8] The results are qualitatively identical with $\sigma > 1$ but the exposition is more complicated as Figure 1b implies the full participation 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 labour supply is then decreasing with productivity - see Figure 1b) but the full particpation region may not exist.

with after tax earnings. There are necessarily 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 and so it is not necessarily the case that MR is increasing over the entire part-time region. For ease of exposition, we shall assume MR is single peaked in this region. Although MR is continuous in α (as labour 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 of the part-time region.[9]

Figure 3 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

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

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

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

or an interior optimum

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

Assuming MR is single-peaked as drawn in Figure 3, there are two candidate optima. A local maximum occurs where $MR(\alpha, b) = \gamma$ on the decreasing portion of the marginal revenue curve and we let $\alpha^*(b)$ denote that solution (where MR single-peaked implies α^* is unique). The other candidate maximum is that the worker chooses zero education where such a choice is optimal only if $MR(a, b) \leq \gamma$.

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

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

which describes the surplus to educating up to α^*. If $V > 0$ the optimal education choice implies $\alpha = \alpha^*(b)$ is optimal as it generates positive value relative to no education. The converse is

[9] See the Appendix which describes the slope of MR.

implied by $V < 0$; the worker is better off choosing no education $\alpha = a$. The optimal education choice therefore depends on the sign of V.

Figure 3 depicts the critical ability a^c where $V(a^c, b) = 0$; i.e. the two shaded areas are equal. A worker with ability $a = a^c$ is indifferent between no education and education to α^*. As a^c must lie on the increasing portion of MR, it follows that $a^c < \alpha_{FT}$. Proposition 1 now establishes that lower ability workers, those with $a < a^c$ choose zero education, while higher ability workers invest to $\alpha^* \gg a^c$. The large discontinuity arises as there are increasing marginal returns to education.

Proposition 1. For given b, suppose MR is single-peaked and suppose that peak occurs at ability $\widehat{a}$. Then for any $\gamma \in (0, MR(\widehat{a}, .))$, an ability $a^c < \widehat{a}$ exists where:

(i) workers with ability $a < a^c$ choose $\alpha = a$ (no education) and ex-post choose low labour supply;

(ii) workers with ability $a \in [a^c, \alpha^*(a)]$ choose $\alpha = \alpha^*(a) \gg a$ and ex-post choose much higher labour supply.

Proof. For any $\gamma < MR(\widehat{a}, .)$, continuity and singlepeakedness of MR implies an $a^c < \widehat{a}$ exists where $V(a^c, b) = 0$ (though a^c may be negative). As

$$\frac{\partial V}{\partial a} = \gamma - MR(a, b)$$

it follows immediately that $V(a, b) < 0$ for $a < a^c$. Thus workers with ability $a < a^c$ choose no education. It also follows straightforwardly that $V(a, b) > 0$ for all $a \in [a^c, \alpha^*)$ (as $V(\alpha^*, b) = 0$) and so such types invest to α^*. This completes the proof of Proposition 1.

Increasing returns to education implies discontinuous education choice. Low ability types with $a < a^c$ choose no education and, as $a^c < \widehat{a} \leq \alpha_{FT}$, these workers either do not participate in the labour 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 3, participate in full time employment in the second period. Of course it is the switch to full time employment which makes the first period education decision worthwhile.

We refer to workers with abilities $a \leq a^c(b)$ as home specialists: such workers do not invest in workplace human capital and have relatively low labour supplies $l^* < 1$. An unrealistic implication of Proposition 1, however, is that very high ability workers, those with abilities $a \geq \alpha^*(b)$ also choose no education. This feature occurs as we have assumed risk averse workers, a wage function $w(a, e)$ which is additive in a and e and education costs which are the same for all. This feature disappears if we instead assume workers are risk neutral, a wage function $w(a, e)$ where ability and education are complementary inputs (so that higher ability workers have a greater Mincerian return to education) and/or education costs c_a which decrease with ability a. Higher ability types will then invest in more education. We do not consider such extensions since the increasing returns to education issue, which is of central interest here, is clearly robust to such variations.

Proposition 2 now shows how home specialisation 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 against 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.

4 Policy and Welfare

The previous section characterized the optimal education investments and ex-post labour supply choices of individuals given tax policy parameters (S, τ). The central feature is that the market dichotomises into *home specialists*, those with abilities (a, b) satisfying $a < a^c(b)$ who choose no education and have low market sector participation rates, and *work specialists*, those with abilities $a > a^c(b)$ who invest significantly in education and have high participation rates. We now consider how changes in tax policy affect those choices and describe the corresponding deadweight losses.

As the optimal choices depend on the underlying tax policy (S, τ), we now extend the notation. Specifically, optimal labour supply is now properly denoted $l = l^*(\alpha, b; S, \tau)$, the marginal return to education is $MR(\alpha, b; S, \tau)$, and the marginal home specialist is $a = a^c(b; S, \tau)$. For ease of exposition we maintain a CRRA utility function with $\sigma \leq 1$.

Proposition 3. Tax Policy and Home Specialists.

$a^c(b; S, \tau)$ is strictly increasing in S and τ.

Proof is in the Appendix.

With increasing returns to education, the marginal home specialist compares no education - which implies ex-post productivity $\alpha = a$ (resulting in low ex-post labour supply) - with investing to productivity $\alpha = \alpha^* \gg a$ (resulting in high ex-post labour supply). As an increase in the income tax rate reduces the return to education, this implies a^c increases with τ - more workers become home specialists.

The impact of sociual security S on education incentives is more subtle. The insight is that home specialists have low earnings in the second period (their labour market productivity is low and they choose low labour supply). As their marginal utility of consumption is relatively high, an increase in S raises their marginal payoff more relative to being educated and working full-time with relatively high earnings. Lump-sum transfers lower the value V of a switch to a higher education level (and higher consumption), and so increases a^c. Note this disincentive disappears if $u(.)$ is linear.

4.1 Deadweight Losses

Given the labour market is competitive and there are no externalities by assumption, the marginal social return to investment is simply the private marginal return when $S = \tau = 0$. Hence define the marginal social return to education:

$$SR(\alpha, b) = MR(\alpha, b; 0, 0).$$

A useful insight is the marginal social return to education is simply a special case of the previous analysis and so also exhibits increasing returns. Let $a^P(b) \equiv a^c(b; 0, 0)$ denote the socially efficient marginal home specialist and $\alpha^P(b) \equiv \alpha^*(b;0,0)$ denote the socially efficient investment level (for higher ability types). Note for any $S, \tau > 0$, Proposition 3 implies $a^P(b) < a^c(b;S,\tau)$; i.e. too many workers become home specialists.

Figure 4 plots SR and MR for given $S, \tau > 0$. The proof of Proposition 3 implies MR must lie below SR. It can also be shown that α_{PT}, α_{FT} lie to the right compared to their values when $S = \tau = 0$.

Figure 4 here.

Figure 4 depicts the deadweight losses implied by the tax program for the marginal home specialist $a = a^c$. As $a^P < a^c$, the socially optimal outcome is that the worker invests to α^P where $SR = \gamma$. If the marginal home specialist invests to α^*, the deadweight loss due to the tax program is the light-shaded Harberger triangle labelled DWL_2. For workers with higher abilities, $a > a^c$, the deadweight loss implied by the tax program always corresponds to such (small) Harberger triangles.

But suppose instead the marginal home specialist $a = a^c$ takes the no education option, $\alpha = a^c$. As the worker is indifferent between $\alpha = a$ and α^*, the additional deadweight loss due to this no education choice is the area between SR and MR over productivies $\alpha \in [a^c, \alpha^*]$. This additional area is dark-shaded and labelled DWL_1 in Figure 4. The large substitution effect induced by increasing marginal returns to education implies the deadweight loss is not a small Harberger triangle; instead the loss can be very large. That loss reflects the total loss in tax as the worker switches to home specialisation. Of course workers with abilities $a \in [a^S, a^c)$ strictly prefer the no education choice while the socially optimal decision is that they invest to α^S. The corresponding deadweight losses are large.

5 Tax Policy and Participation Rates: Some Evidence

Although the model in the previous section is highly stylised, it shows clearly why increasing returns to education in the earnings function, $E = w(a, e)l$, leads naturally to task specialisation. Depending on productivity parameters (a, b), an individual chooses between work or home specialisation. Proposition 1 identifies the partition $a^c(b; S,\tau)$ where workers with ability $a < a^c$ prefer to become home specialists. Proposition 2 establishes home specialists are characterised by relatively high home productivity and these types choose low education and have low participation rates in the labour market. Assuming for cultural or biological reasons, that women tend to be more productive than men in the home, then the model implies women are more likely to become home specialists than men. This assumption is also consistent with the fact that male participation rates tend to be very high. Proposition 3 establishes the home specialist partition a^c is strictly increasing in income tax rates τ and lump-sum transfers S. Thus tax policy has a potentially large impact on female education and participation rates with correspondingly large deadweight losses.

The aim of this section is to describe how different tax policies across various OECD countries have affected average male and female participation rates in the labour market. The data are an unbalanced panel of 20 OECD countries over the period 1980 to 2001.[10] The data were kindly provided by Florence Jaumotte (see Jaumotte (2004) for a full explanation of the construction of the variables).

Figure 1 in the Introduction plots average female participation rates against male participation rates for each country over this time period. There is considerable heterogeneity in female participation rates. Spain, Italy, Ireland and Korea have low female participation rates. The highest female participation rate is in Sweden, closely followed by Iceland, Finland and Denmark. In contrast, there is much less variation in male participation rates. In terms of the model, it may be helpful to think of a similar distribution of women in each OECD country reacting to tax policies that differ across countries. For instance, the tax program in Italy induces a larger fraction of the women to switch into home specialisation than does the tax program in Sweden. In contrast, men are not affected by such differences in policies, because their values of b are closer to zero.

Our empirical analysis uses several tax measures, described below.

(i) The *average tax rate*, calculated as the average tax rate for a single childless person at 67% of the average production wage (APW).[11]

[10] The countries for which we have data are Australia, Austria, Belgium, Canada, Czech Republic, Germany, Denmark, Finland, France, Great Britain, Ireland, Italy, Korea, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the USA.

[11] See Jaumotte (2004) for more details about the variables. In the regressions, our choice of functional form

Note this tax variable is derived from the relevant country's tax code, it is not average tax paid. As the model makes clear, the marginal home specialist calculates when young the expected tax payable if s/he were to invest in education and become a work specialist, and she then compares the resulting payoff against that obtained by being a home specialist with no taxes on home production. Thus, unlike the standard tax-literature approach, it is the average tax which drives the home specialisation decision, and not the marginal tax rate. The above tax variable is a measure of that average tax. However we are unable with our data to distinguish young female cohorts who are making their educational choices from the older women who made their decisions some time ago and for whom such investments are sunk costs. For the younger women, we would expect the current average tax rate to have a larger effect than the older women, but we are unable to test for this here.

(ii) The *tax wedge 2nd earner*. This is calculated as the ratio of 'tax second earner' and the average tax rate of a single individual earning the same gross income of the 2nd earner.[12]

Given that men typically have higher education rates, and assuming that women have higher home productivities, the second earner is likely to be the female partner. High taxes on second earners are then roughly equivalent to raising τ for women, and so are likely to increase the number of female home specialists and have a negative effect on female participation rates.

(iii) *Child benefits*. This variable is defined as the percentage increase in household disposable income from child benefits for two children at a gross earnings level of 133% of the APW (of which 33% is earned by the wife).[13]

Child benefit payments are an example of a non-taxable lump-sum transfer. Proposition 3 predicts that such benefits will increase home specialisation; i.e. lower female participation rates.

(iv) *Public spending on child care*. This measures state spending on childcare (including formal daycare and preschool expenditures) as a percentage of GDP.

---

for these variables was determined after appropriate specification checks.

[12] The tax second earner is defined as {1- [(Household net income when the wife's earnings are 67% of the APW)- (Household net income when the wife does not work)]/[[(Household gross income when the wife's earnings are 67% of the APW)- (Household gross income when the wife does not work)]}. The husband is assumed to earn 100% of APW and the household is assumed to have two children. The difference between gross and net income includes income taxes, employee social security contributions and the like.

[13] Thus Child Benefits={[Difference in household income when the household earns 133% of APW and has two children and that same household-type without any children]/[Household income without any children]}.100.

We interpret childcare subsidies as a subsidy on labour market participation for families with children. Booth and Coles (2007a) provide an extended discussion of this policy in the context of an imperfectly competitive labour market.

(v) To control for changing demographics, we also condition on the variable, the *number of kids aged 0-14 per woman aged 15-64.*

Table 1 gives the means of these variables.

Table 1: Means of Variables (Pooled OECD Data)

Variable	Mean
Average tax rate	24.41
Tax wedge 2nd earner	1.36
Index of child benefits including tax allowances	7.38
Public spending child care as % GDP	0.72
Number of kids aged 0-14 per woman aged 15-64	0.60
% Female workforce participation, age group 25-54	70.03
% Male workforce participation, age group 25-54	93.15

Panels A and B of Table 2 reports fixed-effects (FE) estimates of the reduced form female and male participation equations (with t-statistics in parentheses). Column [1] reports the specifications *without* a time trend, while column [2] reports the results *with* a time trend.

First consider the impact of the average tax rate for a single person. From panel A, we see that female participation is declining in the average tax rate for a single person, although the effect is quite small in both specifications (with and without a time trend). The FE estimates of male participation are reported in Panel B of Table 2. Average tax rates are also associated with a decline in employment for men, but in absolute terms this is far smaller than was found for women..

Next consider the impact of the tax wedge second earner. This is is associated with significantly lower female participation, as expected. The magnitude of this decline is relatively large, espcially so for the specification without the time trend. In contrast, higher tax wedges are associated with higher male participation, although the magnitude is very small. As the tax wedge increases, women participate less and so their men participate more, although this effect for men is imprecisely estimated once the time trend is included.

The third tax policy variable is the variable proxying lump-sum transfers S, namely child benefits including tax allowances. In the theory section we showed that more generous lump-sum transfers reduce the return to working in the labour market and so lead to lower education and female participation rates. This is borne out by the statistically significant negative association, the magnitude of which is very similar across specifications. For men, however, family cash transfers have a significant positive effect in both specifications. Our theory suggested a positive effect only if utility is linear in consumption.

Table 2: Fixed Effects Estimates of Female and Male Participation Rates

A. Female Participation	**[1]**	**[2]**
Ln average tax rate single	-0.049	-0.059
	(1.42)	(1.98)
Ln tax wedge 2nd earner	-0.208	-0.132
	(3.76)	(2.69)
Ln child benefits & tax allowances	-0.063	-0.057
	(2.73)	(2.86)
Ln public spending child care	0.039	-0.029
	(2.37)	(1.66)
Ln number of kids	-0.816	-0.820
	(12.53)	(14.60)
Time trend		0.006
		(6.76)
R^2 - within	0.733	0.802
R^2 - between	0.003	0.070
R^2 - overall	0.053	0.029
B. Male Participation		
Ln average tax rate single	-0.016	-0.012
	(1.65)	(1.68)
Ln tax wedge 2nd earner	0.031	0.003
	(1.99)	(0.23)
Ln child benefits & tax allowances	0.014	0.012
	(2.21)	(2.51)
Ln public spending child care	-0.030	-0.005
	(6.48)	(1.12)
Ln number of kids	-0.057	-0.055
	(3.14)	(4.11)
Time trend		-0.002
		(10.39)
R^2 - within	0.313	0.623
R^2 - between	0.107	0.134
R^2 - overall	0.064	0.398
No. of observations	158	158
No. of countries	20	20

The fourth tax policy variable is *public spending child care.* This is associated with a statistically significant increase in female participation (see Panel A). However, once the time trend is included, this variable becomes statistically insignificant, probably because there is a trend increase in public expenditure on child-care over the period. From Panel B, we see that public spending on child-care is associated with a statistically significant decline in male participation, in contrast to the positive impact it has on female participation. This is likely to reflect the fact that, as childcare expenditure increases, women can participate more and this allows their men to supply less labour. But for men this variable becomes insignificantly different from zero with the inclusion of the time trend.

Finally, an increase in the number of dependent children per woman can be interpreted as a rise in b and this is associated with a significant reduction in female participation.[14] For men, the child dependency variable has a small negative effect in contrast to the more pronounced negative effect on female participation. This is robust across all specifications.

5.1 Discussion

Figure 1 showed that there are big differences across countries in female participation rates. These differences are explained to a large degree by our family-related tax policy variables and the proxy for demographics. As expected, the controls were less able to explain male participation rates .

To illustrate the magnitude of the policy effects, we calculate what US participation would look like if the US had the tax policy values of a typical Scandinavian country, Sweden. To carry out this counter-factual exercise, we use the estimates reported in column [1] of Table 2 to show predicted labour force participation for the USA.[15] The first row of Table 3 gives these predictions for US women (74% participation) and men (93% participation). The other rows of Table 3 shows how these predictions change as policies are altered to Swedish values.[16]

[14] From Proposition 2, we know that $a^c(b)$ is increasing in b. A fall in the number of children might be interpreted as a decline in b, and consequently we would expect female education and participation to increase. Our estimates show that the impact of a decline in the number of children aged 0-14 per woman is indeed associated with an increase in female participation.

[15] For these calculations we use the specifications without the time trend.

[16] For completeness, if the U.S. had Swedish child dependency rates then, ceteris paribus, female participation would increase to 83.2% and male participation increases to 93.6%.

Table 3: Comparative Statics for Predicted US Participation, Swedish Tax Policy

	Comparative static changes	Participation (%)	Participation (%)
		Women	Men
[1]	All US values	74.14	92.81
[2]	US with Swedish tax wedge 2nd earner	77.77	92.15
[3]	US, both Swedish tax policy variables	77.00	91.85
[4]	US, with Swedish child benefits & tax allowance	70.22	93.94
[5]	US with Swedish childcare expenditures	79.42	88.03

The second row shows how female participation would increase if the US were to lower its tax wedge for second earners to the Swedish value. Such a policy shift would increase female participation by 3.6 percentage points to 77.8%, a large effect, but would leave male participation little affected. Row [3] shows the combined effects of adopting Swedish tax policy as summarised by the two variables average tax rate single and tax wedge second earner.[17] Predicted female participation in the USA when Swedish tax policy is introduced is 77%. This is a net increase because the negative effect of shifting to a higher average tax rate is outweighed by the positive effect of shifting to a lower tax wedge on second earners. The male participation rate remains virtually unchanged.

Row [4] shows what happens when we restore tax policy to US values and instead change the index for family cash transfers from the average US value of 4.1 to the average Swedish value of 9.8. As shown, predicted US female participation declines to 70% while male participation increases to nearly 94%, perhaps reflecting substitution between family members in response to increasing cash transfers.

Row [5] shows what happens when US public childcare expenditure as a percentage of GDP is increased to match Swedish rates.[18] It can be seen that US female participation increases to almost 80%, an increase of over 5 percentage points. Male participation drops, possibly again reflecting substitution between partners within a household as more women become work specialists. This emphasises the importance of state childcare expenditures as a targeted subsidy.[19] In summary, although some have suggested it is a puzzle that Sweden is

[17] The US value for the average tax rate single is 22.16 while for Sweden it is much higher, at 29.6. However the tax wedge is higher in the US at 1.33, while it is much lower at 1.0 in Sweden.

[18] This involves a shift from just 0.474% to 1.8753%.

[19] Such expenditures might also have the additional effect of improving the human capital of children, as argued recently by the UK Government, but that is a separate issue not considered here.

characterized by both high average tax rates and high labour supply, there are other important tax policies affecting female labour supply, as our example has illustrated.

6 Conclusion

This paper considered optimal educational investment and labour supply and showed that there are increasing returns in the earnings function. Individual labour market responses to tax policy are shown to be sensitive to home productivity. Specifically, increasing returns implies that a tax on labour income can generate large, non-marginal substitution effects, driving those with a comparative advantage in home production out of the labour market. Assuming home productivity varies substantially by gender, the model predicts that individual responses to fiscal policy will vary significantly across men and women.

Consistent with the theory, our empirical results indicate that gender differences in labour supply responses to tax policy can play an important role in explaining differences in aggregate labour supply across countries. Our estimates show that female participation is declining in the average tax rate for a single person and the tax wedge for second earners, and is increasing in public expenditure on childcare. Female participation is also declining in lump-sum income transfers as proxied by child benefits. In summary, while high tax rates - especially on second earners - encourage women to switch from market to home production, these distortions can be partially offset by targeted employment subsidies such as state-funded childcare. Our analysis suggests that the co-existence in some countries, such as Sweden, of high average tax rates and high labour supply is in fact consistent with the observation that labour supply responses vary by gender in response to heterogeneity in family-related fiscal policies.

In future work we hope to develop our framework further to allow for endogenous fertility decisions. We have already made a start, in our companion paper Booth and Coles (2007b) in which we model match formation. While endogenising fertility decisions would undoubtedly greatly complicate the model, it will be interesting to see if such an approach might yield additional insights.

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), pp. 779-804.

[2] Apps, PF (2003). 'Gender, Time Use and Models of the Household'. IZA Discussion Paper No. 796, June.

[3] Apps Patricia F. and Ray Rees (1996). "Labour supply, household production and intra-family welfare distribution." *Journal of Public Economics*, 60 (2), May, 199-219.

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

[5] Booth, AL and M G Coles (2007a) "A Microfoundation for Increasing Returns in Human Capital Accumulation and the Under-Participation Trap", *European Economic Review,* 51 (7), October, 1661-1681.

[6] Booth, AL and M G Coles (2007b) "Education, Matching and the Allocative Value of Romance". Mimeo, University of Essex.

[7] Burda, M, D Hamermesh and P Weil (2007)."Total work, gender and social norms". IZA Discussion Paper 2705, March.

[8] Jacobs, B. (2005). "Optimal Income Taxation with Endogenous Human Capital". *Journal of Public Economic Theory*, 7(2), 295-315.

[9] Bovenberg, A. Lans, and Bas Jacobs (2005). "Redistribution and Education Subsidies are Siamese Twins", *Journal of Public Economics*, 89, 2005-2035.

[10] Jacobs, Bas and Lans Bovenberg (2007), "Generalized Results on Optimal Taxation and Education Policy and Conditions for Weak Efficiency in Human Capital Production", mimeo: University of Amsterdam/Tilburg.

[11] Jaumotte, Florence (2004) ”Labour Force Participation of Women: Empirical Evidence on the Role of Policy and Other Determinants in OECD Countries”. *OECD Economic Studies*, No. 37, pp51-108. Paris: OECD.

[12] Mincer, Jacob (1958) "Investment in Human Capital and Personal Income Distribution", *The Journal of Political Economy*, 66(4), Aug, 281-302.

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

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

[15] Schindler, Dirk and Benjamin Weigert (2007) "Endogenous Human Capital Risk and Public Policy." Paper presented at EALE Meetings, Oslo, September.

[16] Trostel, P.A. (1993). "The effect of taxation on human capital". Journal of Political Economy, 101(2), 327-350.

[17] Trostel P.A. and I. Walker (2006). "Education and Work", Education Economics,14 (4), 377-399.

7 Appendix A

Proof of Claim 2.

The definition of b_{FT} and CRRA implies

$$b_{FT} = \alpha(1-\tau)w_0(S+(1-\tau)\alpha)^{-\sigma}/x'(0).$$

Differentiating with respect to α yields

$$\frac{\partial b_{FT}}{\partial \alpha} = \frac{(1-\tau)w_0[S+(1-\sigma)(1-\tau)w_0\alpha]}{(S+(1-\tau)w_0\alpha)^{\sigma+1}x'(0)}$$

and so

$$\frac{\partial b_{FT}}{\partial \alpha} \gtrless 0 \text{ as } S + (1-\sigma)(1-\tau)w_0\alpha \gtrless 0.$$

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

$$\frac{\partial l^*}{\partial \alpha} = \frac{(1-\tau)w_0}{-bx'' - \alpha^2(1-\tau)^2 w_0^2 u''}[u'(y_{PT}) + \alpha(1-\tau)w_0 l^* u''(y_{PT})]$$

where $y_{PT} = S + (1-\tau)\alpha w_0 l^*$. Concavity of x and u implies $\partial l^*/\partial\alpha > 0$ if and only if $u'(y_{PT}) + \alpha(1-\tau)w_0 l^* u''(y_{PT}) > 0$. CRRA now implies

$$\frac{\partial l^*}{\partial \alpha} \gtrless 0 \text{ as } S + (1-\sigma)(1-\tau)\alpha w_0 l^* \gtrless 0, \tag{6}$$

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

The statement of the Claim follows from these facts and (a) l^* is strictly decreasing in b for $b \in (b_{FT}, b_{PT})$, (b) $l^* = 1$ at $b = b_{FT}$ and (c) $l^* = 0$ at $b = b_{PT}$. b^c is defined where $S + (1-\sigma)(1-\tau)w_0\alpha l^* = 0$.

Proof of Proposition 2. Recall V is defined by

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

and a^c is then defined by the implicit function $V(a^c, b) = 0$. Differentiating $V(.)$ with respect to b, noting that $MR(\alpha^*, b) = \gamma$, yields

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

Now (5) 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

$$\frac{\partial}{dl^*}[l^* u'(S + (1-\tau)\alpha w_0 l^*)] = \frac{S + (1-\sigma)(1-\tau)\alpha w_0 l^*}{[S + (1-\tau)\alpha w_0 l^*]^{\sigma+1}} > 0.$$

(5) and $\sigma < 1$ now imply $\partial[MR]/\partial b < 0$ in the part-time region. Hence we have $\partial V/\partial b < 0$. As

the proof of Proposition 1 implies $\partial V/\partial a > 0$ at $a = a^c$, the Implicit Function Theorem implies a^c increases with b.

Proof of Proposition 3. In the extended notation, V is defined by

$$V(a, b; S, \tau) = \int_a^{\alpha^*} (MR(\alpha, b; S, \tau) - \gamma)\, d\alpha.$$

and a^c is given by the implicit function $V(a^c, b; S, \tau) = 0$. Differentiating V wrt S, noting that $MR(\alpha, b; S, \tau) = \gamma$ at α^*, yields

$$\frac{\partial V}{\partial S} = \int_a^{\alpha^*} \frac{\partial\,(MR(\alpha, b; S, \tau.))}{\partial S} d\alpha.$$

Now (5) with $\sigma < 1$ implies MR does not change with S in the non-participant region (it is zero) and is strictly decreasing in S in the part-time[20] and full-participation regions. Hence $\partial V/\partial S < 0$. As $\partial V/\partial a > 0$ at $a = a^c$, the Implicit Function Theorem implies a^c increases with S.

Similarly

$$\frac{\partial V}{\partial \tau} = \int_a^{\alpha^*} \frac{\partial\,(MR(\alpha; b_0, S, \tau.))}{\partial \tau} d\alpha.$$

(5) with $\sigma < 1$ again implies MR does not change in the non-participant region (it is zero) and is strictly decreasing in τ in the part-time and full participation regions. Hence a^c increases with τ. This completes the proof of Proposition 3.

[20]For $\alpha \in (\alpha_{PT}, \alpha_{FT})$, (3) implies l^* decreases with S while total earnings, $S+(1-\tau)\alpha w_0 l^*$ increase. Together these imply that MR falls within the part-time region.

Figure 1: Female and Male Market-sector Participation

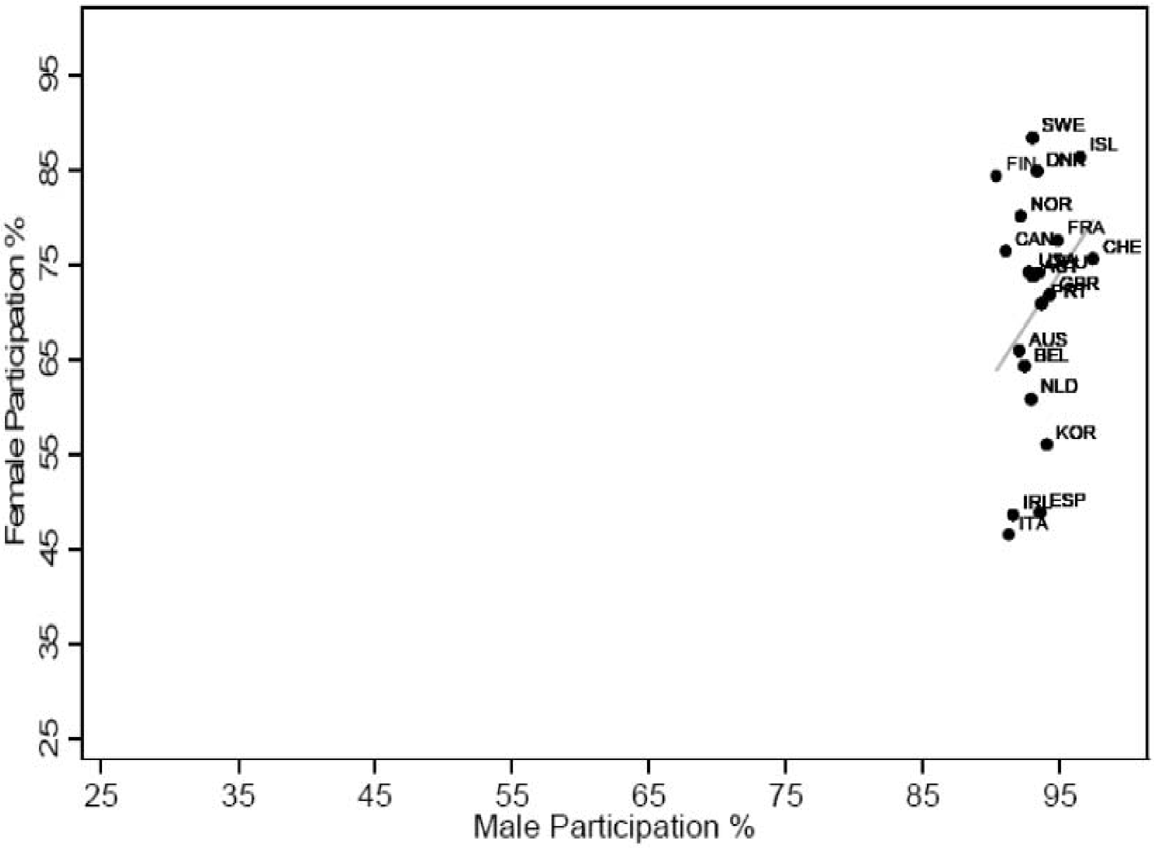


Figure 2(a): Labour Supply ($\sigma < 1$)

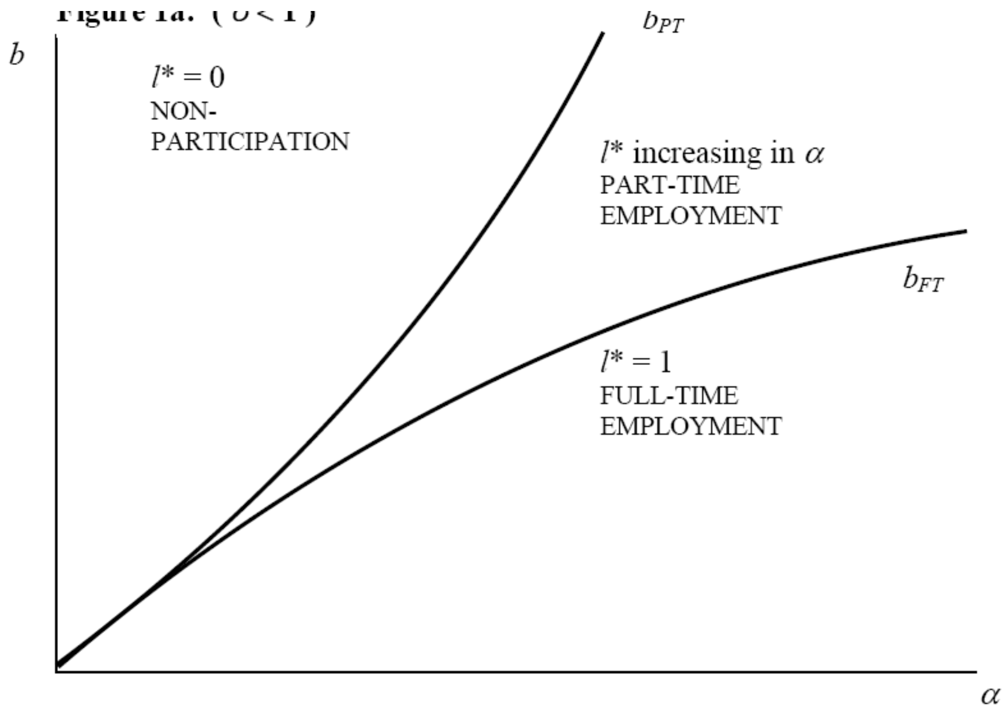


Figure 2(b): Labour Supply ($\sigma > 1$)

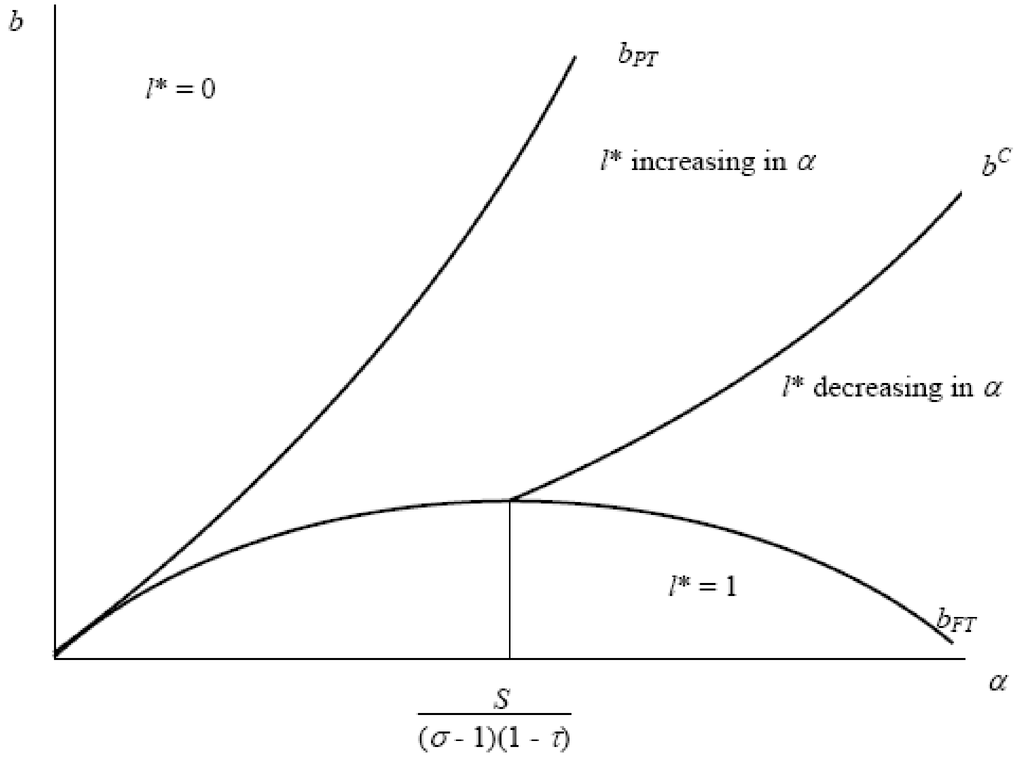


Figure 3: Optimal Education Choice (b_o given)

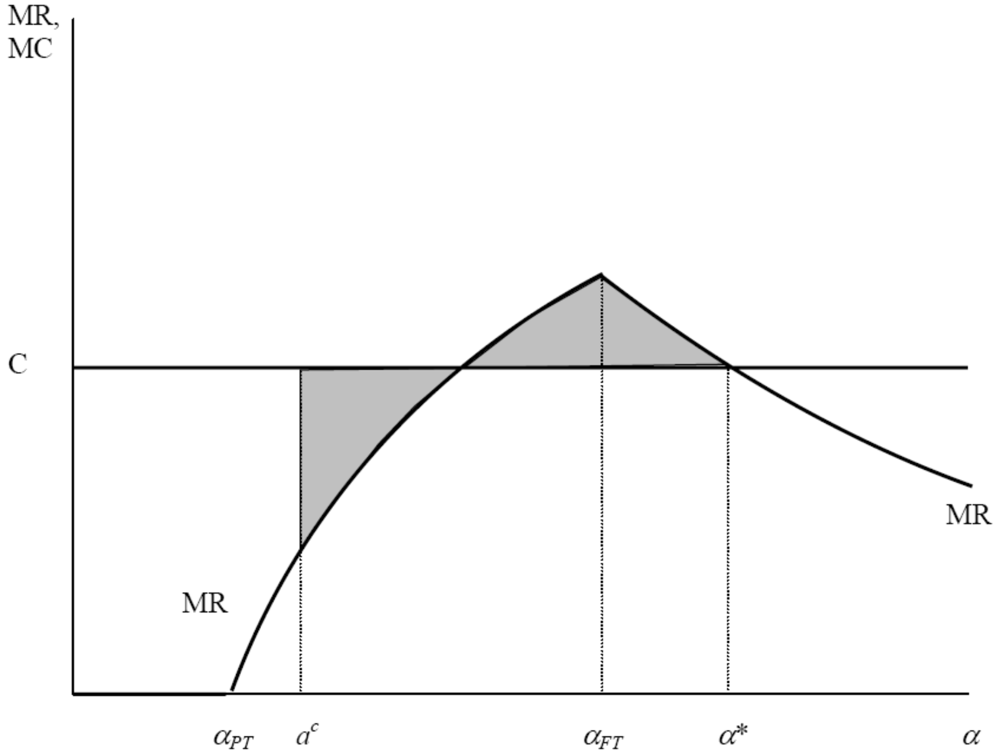
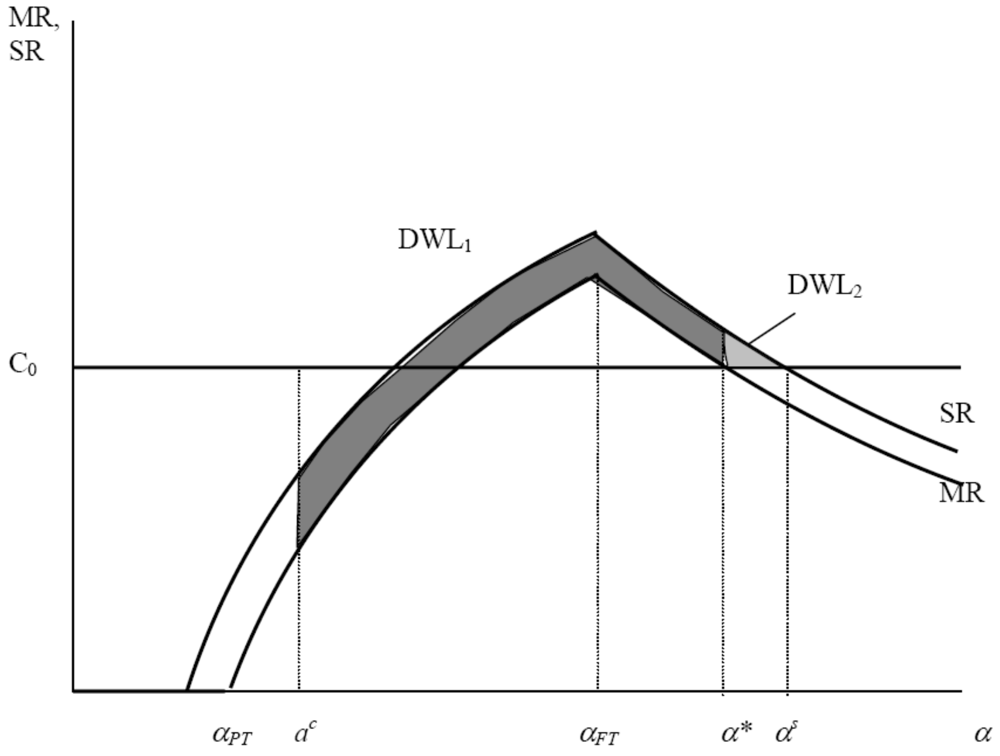


Figure 4: Deadweight Losses



0.1 Appendix B (for referees, not for publication)

0.2 Synopsis of Bovenberg-Jacobs (2005) structure

With no taxation, the representative worker in BJ chooses consumption c, labour supply l and education e to solve

$$\max u(c, l) = c - \frac{l^{1+1/\varepsilon}}{1 + 1/\varepsilon}$$

subject to the budget constraint

$$c = w_0 H(e) l - \gamma e$$

with $H(e) = e^\beta$. BJ correctly claim that second order conditions are satisfied if $\varepsilon > 0$ and $\beta(\varepsilon + 1) < 1$.

Substituting out c using the budget constraint, the optimal (second period) labour supply choice reduces to

$$l^*(e) = \arg\max_{l \geq 0} \left[w_0 e^\beta l - \gamma e - \frac{l^{1+1/\varepsilon}}{1 + 1/\varepsilon} \right].$$

As $\varepsilon > 0$ implies this problem is concave in l, the focs imply optimal

$$l^*(e) = w_0^\varepsilon e^{\varepsilon\beta}.$$

Thus second period maximised payoff (after some algebra) is

$$U_2^*(e) = \frac{[w_0 e^\beta]^{\varepsilon+1}}{\varepsilon + 1} - \gamma e.$$

(In the first period) the worker chooses e to maximise $U_2^*(e)$. This is a concave problem if $\beta(\varepsilon + 1) < 1$. The first order condition then describes the global maximum which is

$$\frac{dU_2^*}{de} = w_0^{\varepsilon+1} \beta [e^{\beta(\varepsilon+1)-1}] - \gamma = 0.$$

The first term is the marginal return to education, MR, and is decreasing in e if $\beta(\varepsilon + 1) < 1$ (as was assumed).

The Booth/Coles approach instead correctly shows $MR \equiv w_0 H'(e) l^*(e)$. In this BJ case, it is

$$MR \equiv w_0 [\beta e^{\beta-1}] w_0^\varepsilon e^{\varepsilon\beta}$$

which, not surprisingly, is the first term in the previous equation. Booth/Coles argue that when the participation constraint binds; i.e. at $l^* = 0$, then MR=0. Continuity then implies MR must be increasing in e for a range of e. Yet in the above case, MR is strictly decreasing in e. So is there a contradiction? The answer is no - the non-participation constraint never binds in BJ. The special property of BJ above is that the marginal disutility of labour is zero at $l = 0$. Consider then the following variation.

0.3 Extended Bovenberg-Jacobs.

Suppose we keep everything the same but change worker preferences:

$$u(c, l) = c - \frac{[l + b]^{1+1/\varepsilon}}{1 + 1/\varepsilon},$$

where $b > 0$. BJ is the special case $b = 0$. These extended preferences remain concave in labour supply, but with $b > 0$, $\partial u/\partial l$ is now strictly negative at $l = 0$. Repeating the BJ methodology, the optimal (second period) labour supply choice is

$$l^*(e) = \arg\max_l \left[w_0 e^\beta l - \gamma e - \frac{(l+b)^{1+1/\varepsilon}}{1+1/\varepsilon} \right].$$

As $\varepsilon > 0$ this problem is concave in l and optimality implies

$$\begin{aligned} l^*(e) &= 0 \text{ if } b \geq w_0^\varepsilon e^{\varepsilon\beta}, \\ l^*(e) &= w_0^\varepsilon e^{\varepsilon\beta} - b \text{ if } b < w_0^\varepsilon e^{\varepsilon\beta}; \end{aligned}$$

Substituting this into the worker's payoff function, the second period maximised payoff (after some algebra) is

$$\begin{aligned} U_2^*(e) &= -\gamma e - \frac{b^{1+1/\varepsilon}}{1+1/\varepsilon} \text{ if } w_0^\varepsilon e^{\varepsilon\beta} \leq b, \\ &= \frac{[w_0 e^\beta]^{\varepsilon+1}}{\varepsilon+1} - \gamma e - b w_0 e^\beta \text{ if } w_0^\varepsilon e^{\varepsilon\beta} > b; \end{aligned}$$

U_2^* is no longer concave in e. For e satisfying $w_0^\varepsilon e^{\varepsilon\beta} < b$, the worker chooses $l^* = 0$ and $MR = 0$.For e satisfying $w_0^\varepsilon e^{\varepsilon\beta} > b$, the marginal return to education is

$$MR = \beta w_0 e^{\beta-1} \left[w_0^\varepsilon e^{\beta\varepsilon} - b \right].$$

Inspection shows that MR has the structure identified in Booth/Coles vis-a-vis:

$$\begin{aligned} MR &= 0 \text{ if } e \leq e_1 \\ MR \text{ is increasing if } e &\in (e_1, e_2) \\ MR \text{ is decreasing if } e &\geq e_2, \end{aligned}$$

where e_1, e_2 depend on the model parameters.

Note then that MR is not a continuous function of b. Specifically MR has the above structure for any $b > 0$, and $MR = 0$ at $e = 0$. But $MR = \infty$ at $e = 0$ when $b = 0$ (in BJ with $b = 0$, $e \to 0$ implies $l^*(e) \to 0$ but $H'(e) \to \infty$. The condition $\beta(\varepsilon + 1) < 1$ ensures $l^*(e)H'(e) \to \infty$ in this limit].

Conclusion: The introduction of a strictly positive disutility of labour supply at the non-participant margin (i.e. at $l = 0$) yields the Booth/Coles structure.