University *of* Leicester

The Sex and the Uni: Educational Assortative Matching and Over-Education*

Alessandro Tampieri, University of Leicester, UK

The Sex and the Uni: Educational Assortative Matching and Over-Education[*]

Alessandro Tampieri[†]

University of Leicester

January 19, 2010

Abstract

Educational assortative matching encourages individuals to acquire education so as to increase the probability of marrying a high-income partner. But since everyone is more educated, the chances of a good match do not change. Hence over-education emerges, as in absence of educational assortative matching individuals could reach their optimal level of education by exploiting less educational resources. Over-education is stronger the higher the probability of educational assortative matching, the larger the relative importance of the partner's income in determining utility and ability levels, and the lower the cost of education. Government intervention can reach a socially efficient level of education through either a tax on education or income.

JEL Numbers: I21, J12

Keywords: educational assortative matching, over-education.

[Very Preliminary]

1 Introduction

In the last decades, the educational attainments of labor force in US and Europe have increased more rapidly than the skill requirements of available jobs (Vaisey, 2006, Moro Egido and Budria, 2007). The phenomenon by which individuals perform jobs for which they are overqualified is called "over-education". Although a large empirical literature concerned in measuring over-education[1] have been developing, there have been few contributions in order to understand the reasons behind it.

This paper proposes an explanation for the existence of over-education, based on the idea that acquiring education has two main effects. First, it increases income directly. Second, it increases the chance to marry someone met at school.

The school is one of the first places where people create their own social networks. At school they meet people, make friendships and social connections, and spend the most part of their youth. Youth is also the period where they are likely to fall in love and to find their own partner, thus there is a certain probability to meet the partner in the school environment. Partners met at school share similar education levels. We refer to the positive correlation in education levels between partners as "educational assortative matching"[2]. Several studies have shown strong evidence of increases in the educational resemblance of spouses since at least the 1940s (Kalmijn 1991a, 1991b; Mare 1991; Pencavel 1998; Quian, 1998, Qian and Preston 1993; Smits et al. 2000, Schwartz and Mare, 2005). Educational assortative matching reflects similarities between partners in their innate ability, since ability is similar in individuals who share the same school experience. Our idea is that the presence of educational assortative matching is the key point to understand the existence of over-education.

We build up a model of education and marriage market where the partner's income positively affects the individual's utility. Individuals maximise their expected utility with respect to education taking into account the expected

---

[1] For discussions, see Hartog, 2000 and McGuinness, 2006.

[2] The expression "assortative matching" has been coined by Gary Becker (1973), and it alludes to a relationship (either positive or negative) between characteristics of partners. As we refer to the similarity in level of education between partners, we specify that assortative matching is "educational".

matching in the marriage market, which can be random or assortative. The probability of educational assortative matching is exogenous and depends on the customs and the culture of the society considered. In the presence of educational assortative matching, acquiring education both increases income and the probability to marry a high-income partner. In fact, the more an individual acquires education, the more likely his or her partner will have high education and hence high income. This gives an incentive to increase the quantity of education acquired. Over-education is defined as the difference between the actual level of education and the level of education without educational assortative matching.

Our results suggests that educational assortative matching increases the quantity of education acquired to an inefficient level from a social point of view. If educational assortative matching were not present, individuals would reach their optimal level of education by using fewer educational resources. Over-education is stronger the larger the probability of educational assortative matching, the higher the relative importance of the partner's income in determining utility, the higher the ability levels of individuals, and the lower the rate cost of education.

What determines these results? The probability to be well-matched in the marriage market is affected by relative income, since individuals compare the potential partners' incomes between them. At the same time, the presence of educational assortative matching leads *everyone* to acquire more education. Hence, even if the absolute levels of education increase, the relative levels of education between individuals remain the same, and the probability to be matched with a high-income partner does not change. This approach is in the flavour of Akerlof (1976), where workers signal their ability through their fastness at work. In order to look abler, workers of a given ability work faster than they would. In our model, individuals observe the partner's education level as a signal of ability, and in order to look abler they acquire more education than they would.

Then we consider whether public intervention make individuals to reach the socially efficient level of education. We illustrates alternatives interventions through either a tax on education or a proportional tax of income. Both interventions can correct over-education, even though to apply the optimal tax

on education would require to know the individual's ability, while this is not necessary for the optimal tax on income.

To our knowledge, over-education has not been largely developed from a theoretical perspective, with few notable exceptions. Frank (1978) investigates the differentials in wages between men and women as a consequence of female overqualification. This is not caused by employers' discrimination against women but by family location decisions, since a family is more likely to move close to better jobs for husbands, sacrificing the wife's opportunities. Hence the role differences between men and women are essential for their results, and over-education is generated by a job search process. Compared to this paper, we abstract away from both gender differences and job search.

In Lommerud (1989), over-education occurs as individuals care about social status, represented by the relative income. Lommerud uses a tax on income in order to tax away differences in status, and on the other hand consider a subsidy on education slightly lower the tax on income in order to prevent the tendency to over-invest in education. In our model, we cannot correct the cause of over-education directly. Instead, through a tax either on education or income we affect the return to education: assortative matching is unaltered, but individuals acquires a first best quantity of education.

Konrad and Lommerud (2000) explain over-education through a household bargaining model where the young choose individually their level of education and, once married, they sacrifice their returns in education in favour of an optimal level of family public goods (i.e., to spend time with children, partner, and so on). Over-education emerges because the educational decisions affects the threat point of spouses. To over-invest in education is inefficient in order to optimise the quantity of the family public good, but leads to increase the threat point so as to be in an advantaged position in the household bargaining. In our model individuals choose separately their educational levels but there is not a cooperative game during the marriage life, and we do not consider the existence of family public goods.

This paper also shares in common with Peters and Siow (2001), Baker and Jacobsen (2005), Iyigun and Walsh (2005), Chiappori *et al.*(2006) and Nosaka (2007) the link between education, marriage and assortative matching. However, our key questions are whether the quantity of education acquired is

socially efficient and whether over-education emerges. Thus, we focus on how education is affected by the other aspects.

The remainder of the paper is organized as follows: Section 2 describes the model, the marriage market and the problem of educational choice. Section 3 shows the results. Section 4 illustrates government intervention. Section 5 concludes.

2   The model

We consider a continuum of individuals, with measure normalised to 1. Individuals differ in ability, denoted by θ and distributed according to density $f(\theta)$ with cumulative distribution function $F(\theta)$ and positive support in an interval $[\underline{\theta}, \overline{\theta}] \in \Re^+$. We refer to ability as every innate characteristic that contributes to the income potential. Let θ_i be the ability level of the individual $i \in [\underline{\theta}, \overline{\theta}]$.

There are two periods. In the first period, individuals attend school and decide the quantity of education to acquire. Education increases income in the second (working) period. Education is a variable $e \geq 0$. For every $i \in [\underline{\theta}, \overline{\theta}]$, we denote as e_i the quantity of education acquired by i. Education is costly for individuals. The utility cost of education depends on the quantity of education acquired. For every $i \in [\underline{\theta}, \overline{\theta}]$, we denote the utility cost of education as $C(e_i) = c\frac{e_i^2}{2}$, where $c > 0$.

In the second period, individuals work and are matched in the marriage market. The individual i's income is denoted[3] by $y_i(e_i, \theta_i) = e_i\theta_i$. In the marriage market, every individual is matched with another individual. For every $i \in [\underline{\theta}, \overline{\theta}]$, we denote by $p \in [\underline{\theta}, \overline{\theta}], p \neq i$ the partner of i. The partner's income is represented by $y_p(e_p, \theta_p) = e_p\theta_p$, where e_p and θ_p denote the partner's education and ability, respectively.

We assume that an individual's utility is determined by the own income, the partner's income and net of the cost of education. Thus, the utility of i is given by:

$$U_i(e_i, \theta_i, e_p, \theta_p) = e_i\theta_i + \alpha e_p\theta_p - c\frac{e_i^2}{2}, \tag{1}$$

_____

[3]Note that this formulation implies that education increases earnings more for abler individuals, as $\frac{\partial y_i(e_i, \theta_i)}{\partial \theta_i \partial e_i} > 0$.

where $\alpha \in [0,1]$ represents the relative importance of the partner's income in determining the individual's utility. We analyse the first and the second period in reverse order.

2.1 Marriage market

We start by describing the matching in the marriage market faced by individuals. According to equation (1), individuals prefer to be matched with a high-income partner, as this would increase the individual's benefit. The matching with a counterpart can be of two types: assortative[4] or random. Assortative matching occurs when an individual meets the partner at school. In this case partners share the same education levels. Every matching that occurs out of school is random. Here, the choice of education is not linked to the kind of matching.

Let $\beta \in [0,1]$ denote the exogenous probability that the matching is assortative. The size of β depends on the customs and the culture of the society we are considering. For instance, let us consider the differences between attending university in a campus or through a "distance-learning" program. In the first case, students spend time together, attend the same courses, share the same issues, and live in the same student accommodation. These individuals are more likely to incur in assortative matching. In the case of the distance learning program, individuals do not meet school friends, do not move into the campus, thus they are more likely to incur in random matching. Hence, if the campus system is more common in a given population, β is high, while where the distance-learning system is more common, β is low.

2.2 Educational choice

During the first period, individuals decide the quantity of education to acquire.

The future matching in the marriage market affects educational decisions today: in fact, in the case of assortative matching, an individual obtains informations about education and ability of the partner during the first period,

---

[4]Along the model, we simplify the terminology by referring to "assortative matching" meaning "educational assortative matching".

while in the case of random matching the individual can only guess the partner's characteristics.

In the case of assortative matching, every individual is matched with a partner of same education level, thus $e_p = e_i$. Consequently, by increasing the quantity of education acquired in the first period it is possible to increase the probability to be matched with a partner with high education (who will obtain a high income in the second period). Hence individuals may want to acquire more education in order to improve the probability of being married with a high-income partner.

In the case of random matching, an individual have no information about the partner's ability, thus the partner's expected ability is determined by θ_{av}, the average ability of the population: $\theta_p = \theta_{av} \equiv \int_{[\underline{\theta},\overline{\theta}] \in \Re^+} \theta f(\theta) d\theta$.

With assortative matching, during the first period individuals can infer exactly the partner's level of ability through his of her education level. For showing it, let us suppose that $F(\theta_p)$ is the education level of a partner with ability θ_p, and also that $F'(\theta_p) > 0$. We suppose that individual i acquires a generic level of education $\hat{e}$. Since we are in assortative matching, this individual will be matched with a partner with ability with education $F(\hat{\theta})$. Hence individual i can infer the partner's ability $\hat{\theta}$ as the inverse image of $F(\hat{\theta})$, so $\hat{\theta} = F^{-1}(\hat{e})$. If this holds, we can rewrite equation (1) as:

$$\hat{e}\theta_i + \alpha \left((1-\beta)(e_p\theta_{av}) + \beta\left(\hat{e}F^{-1}(\hat{e})\right)\right) - c\frac{\hat{e}^2}{2}. \tag{2}$$

In the equilibrium we consider, all type i individuals make identical choices, and so (2) is the expected utility in each type i individuals. The first part of equation (2) is the total benefit given by the individual's income, the second part is the total benefit given by the partner's income, and finally the third part is the total cost of education. The second part of (2) can be in turn decomposed into two parts: (i) $\alpha(1-\beta)(e_p\theta_{av})$, which represents the partner's expected income with random matching, and (ii) $\alpha\beta\left(\hat{e}F^{-1}(\hat{e})\right)$, which represents the partner's expected income with assortative matching. The decision problem is:

$$\max_{\hat{e}} \hat{e}\theta_i + \alpha \left((1-\beta)(e_p\theta_{av}) + \beta\left(\hat{e}_iF^{-1}(\hat{e})\right)\right) - c\frac{\hat{e}^2}{2}. \tag{3}$$

The first order condition of problem (3) is:

$$\theta_i + \alpha\beta \left(F^{-1}(\hat{e}) + \hat{e}\frac{d}{d\hat{e}}F^{-1}(\hat{e}) \right) - c\hat{e} = 0. \tag{4}$$

For simplicity, let us call $F^{-1}(\hat{e}) = g(\hat{e})$. In equilibrium, $\hat{e} = e_i$, the first order condition is:

$$\theta_i + \alpha\beta \left(g(e_i) + e_i\frac{d}{de_i}g(e_i) \right) - ce_i = 0. \tag{5}$$

The following Lemma shows the solution of equation (5).

Lemma 1 *For every $i \in [\underline{\theta}, \overline{\theta}]$, the level of education in equilibrium is $\overline{e}_i = \frac{(1+2\alpha\beta)\theta_i}{c}$*

Proof. See Appendix. ■

Note that $F'(\theta_i) > 0$ holds only if $\overline{e}_i$ is an increasing function. Differentiating $\overline{e}_i$ with respect to θ_i yields:

$$\frac{\partial}{\partial\theta_i}\frac{(1+2\alpha\beta)\theta_i}{c} = \frac{(1+2\alpha\beta)}{c} > 0,$$

which confirms the initial assumption.

2.3 Definition of over-education

We showed that individuals optimally choose a level of education $\overline{e}_i$ that allows them to maximise, according to the cost of education, both their future incomes and the probability to be matched with a high-income partner.

Now consider $\beta = 0$. This can be imagined as a situation where education is individually acquired (e.g. distance-learning course for everybody), so individuals cannot meet their partners at school. Now the first order condition of (3) is:

$$\theta_i = ce_i, \tag{6}$$

and the solution of (6) is $e_i^* = \frac{\theta_i}{c}$. As $\beta \in [0,1]$, e^* is lower than $\overline{e}$. This is intuitive. Without assortative matching, potential partners do not observe

8

in the first period the individuals' education level, therefore individuals have no reason to try to look abler to them. The increased quantity of education they used to acquire was optimal only because *everyone* were acquiring more education. We refer to the equilibrium with no assortative matching as *first best equilibrium*, since agents would reach their optimal quantity of education by employing less educational resources. Over-education is defined as the difference in quantity of education between the level of education in equilibrium $\bar{e}$ and the level of education with no assortative matching e^*.

Definition 1 *Let $\Delta e = \bar{e} - e^*$. Δe is defined as the level of over-education.*
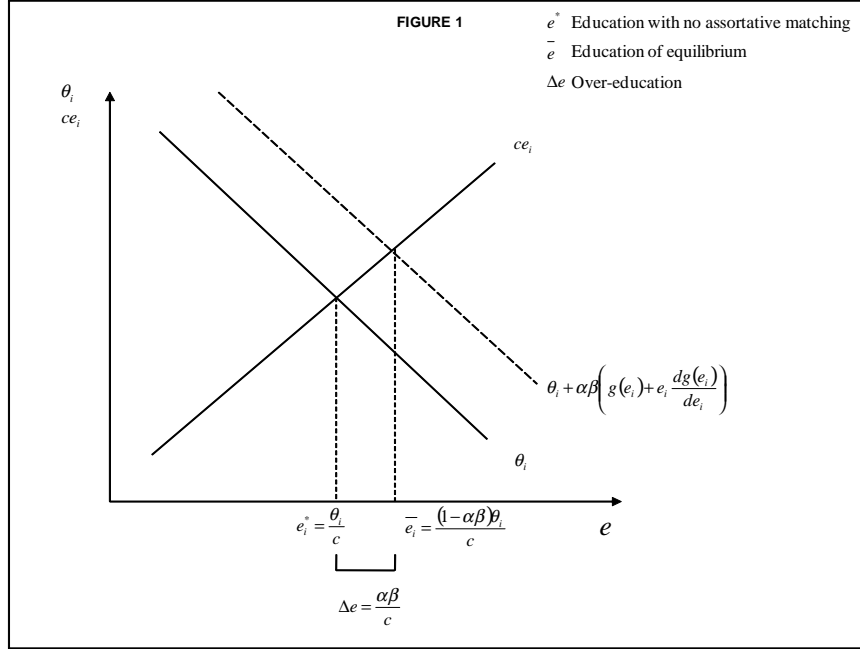
3 Results

We start this section by showing the level of over-education in this economy.

Proposition 1 *Let Δe be the level of over-education. This is given by $\Delta e = \frac{2\alpha\beta\theta_i}{c}$.*

Proof. It follows by the difference between $\bar{e}_i$ and e_i^*: $\Delta e = \bar{e} - e^* = \frac{(1+2\alpha\beta)\theta_i}{c} - \frac{\theta_i}{c} = \frac{2\alpha\beta\theta_i}{c}$. ∎

Figure 1 shows the level of education in equilibrium, with no assortative matching and the level of over-education. The optimal quantity of education, given by the point where the marginal benefit equates the marginal cost, is reached to a higher level of education compared to the case with no assortative matching.

The intuition of why over-education emerges with assortative matching is the following. Assortative matching increases the chance to marry a partner with same level of education, then *everyone* studies more in order to increase the probability to be married with a high-income partner. With no assortative matching, everyone studies *less* and so the probability to be married with a high income partner *does not change*. Hence, without assortative matching it would be possible to obtain the same result in terms optimal choice, but employing less educational resources, thus improving social welfare.

FIGURE 1

e^* Education with no assortative matching
$\overline{e}$ Education of equilibrium
Δe Over-education

In the figure:

ce_i

$\theta_i + \alpha\beta\left(g(e_i) + e_i\frac{dg(e_i)}{de_i}\right)$

θ_i

$e_i^* = \frac{\theta_i}{c}$ $\overline{e}_i = \frac{(1-\alpha\beta)\theta_i}{c}$ e

$\Delta e = \frac{\alpha\beta}{c}$

3.1 Comparative static

We now illustrate what happens to the equilibrium after a variation in the probability of assortative matching, in relative importance of the partner's income in determining utility, in ability levels and in the rate cost of education. The following proposition summarises these comparative static properties.

Proposition 2 *An increase either in assortative matching or in the relative importance of the partner's income in determining utility leads to an increase in over-education. Also, the more an individual has high ability level, the more he or she will be affected by over-education. Finally, as the rate cost of education increases, the level of over-education diminishes.*

Proof. Differentiation of Δe with respect to β, α, θ and c yields $\frac{\partial \Delta e}{\partial \beta} = \frac{2\alpha\theta_i}{c} > 0$, $\frac{\partial \Delta e}{\partial \alpha} = \frac{2\beta\theta_i}{c} > 0$, $\frac{\partial \Delta e}{\partial \theta_i} = \frac{2\alpha\beta}{c} > 0$, $\frac{\partial \Delta e}{\partial c} = -\frac{2\alpha\beta}{c^2} < 0$, respectively. ∎

The first part of Proposition 2 establish that, as assortative matching and importance of partner's income in determining utility increase, so over-education grows stronger. Both effects are intuitive: in a society where individuals are more likely to meet their partner from their school friends (β high), their optimal level of education will be higher. Moreover, as the relative importance of

10

the partner's income in determining utility increases, individuals invest more in education since to stay with a partner with high ability level is more valuable: this aspect leads to further over-education.

The second part of the proposition reveals that, the higher the ability level of an individual is, the higher the level of over-education. In order to compete for a better match between people with high ability levels, the quantity of education to acquire needs to be higher than between individuals with low ability levels.

Finally, the last part Proposition 2 says that the rate cost of education reduces the level of over-education.

4   Government intervention

In this section, we assume that there is a government whose objective is to reach the socially efficient level of education. To reduce over-education, the government evaluates to levy a tax.

Consider first a tax based on the quantity of education acquired. For every $i \in [\underline{\theta}, \overline{\theta}]$, the tax on education is $T = te_i$, where t is the tax rate. Thus, for every $i \in [\underline{\theta}, \overline{\theta}]$, the problem becomes:

$$\arg\max_{e_i} \; e_i\theta_i + \alpha\left((1-\beta)\left(e_p\theta_{av}\right) + \beta\left(\alpha e_i g(e_i)\right)\right) - c\frac{e_i^2}{2} - te_i. \qquad (7)$$

The first order condition of (7) is:

$$\theta_i + \alpha\beta\left(g(e_i) + e_i\frac{d}{de_i}g(e_i)\right) = ce_i + t. \qquad (8)$$

The following proposition show the solution of equation (8).

Lemma 2 *For every $i \in [\underline{\theta}, \overline{\theta}]$, the level of education in equilibrium with a tax on education is $e_i^t = \frac{(\theta_i - t)(1 + 2\alpha\beta)}{c}$.*

Proof. See Appendix. ∎

In order to reach the first best level of education, the quantity of education acquired with taxation needs to be equal to e^*, thus $\frac{(\theta_i - t)(1 + 2\alpha\beta)}{c} = \frac{\theta_i}{c}$. By explicating t we find the optimal tax rate.

11

Proposition 3 *The optimal tax on education is $T = et^*$, where $t^* = \frac{2\alpha\beta\theta_i}{2\alpha\beta+1}$.*

An alternative solution can be to levy a proportional tax on income. We denote the proportional tax rate as τ. For every $i \in [\underline{\theta}, \overline{\theta}]$, the problem is:

$$\underset{e_i}{\arg\max} \; e_i\theta_i\left(1-\tau\right) + \alpha\left((1-\beta)\left(e_p\theta_{av}\right) + \beta\left(\alpha e_i g(e_i)\right)\right) - c\frac{e_i^2}{2}. \qquad (9)$$

The first order condition of (10) is:

$$\theta_i + \beta\alpha\left(g(e_i) + e_i\frac{d}{de_i}g(e_i)\right) = ce_i + \tau\theta_i, \qquad (10)$$

and the level of education is determined by the following Lemma

Lemma 3 *For every $i \in [\underline{\theta}, \overline{\theta}]$, the level of education in equilibrium with a proportional tax on income is $e_i^\tau = \frac{\theta_i(1-\tau+\beta\alpha)}{c}$.*
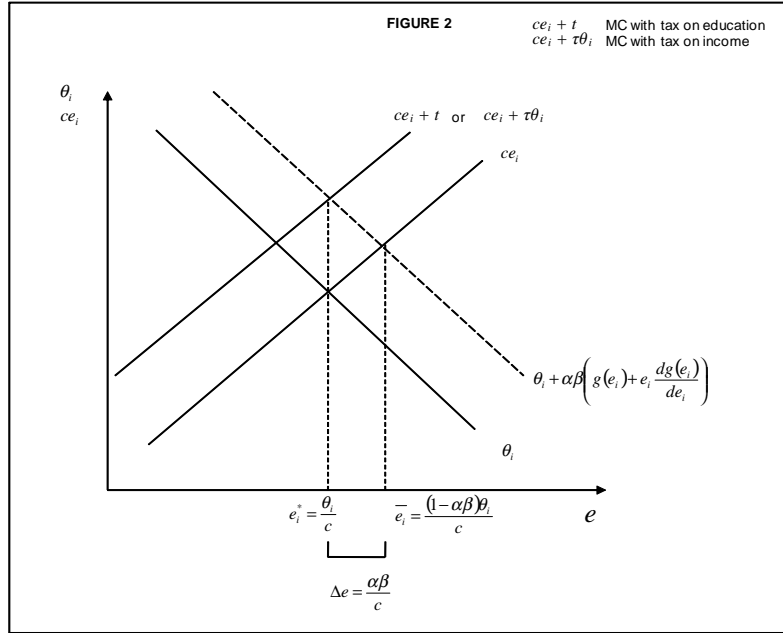
Proof. See Appendix. ∎

As before, we equate this level of education with the socially efficient level, $e_i^\tau = \frac{\theta_i(1-\tau+\beta\alpha)}{c} = \frac{\theta_i}{c}$. By explicating τ we find the optimal taxation on income.

Proposition 4 *The optimal proportional tax on income is $\tau^* = \beta\alpha$.*

Figure 2 shows the equilibrium where either $e_i t^*$ or τ^* are levied. In both cases, the tax corrects the distortion given by the presence of assortative matching by affecting the individual return to education. The presence of assortative matching still makes individuals to acquire education to improve the probability to a wealthier marriage, but now they acquire less education as their future income is diminished.

The implementation of the tax on education presents problems. To establish the level of t^* it is necessary to know the probability of assortative matching and the ability of individuals. While to measure the probability of assortative matching is possible, even if this can give debatable results (it can be controversial the way to measure it), the government cannot obtain informations about the individuals' ability. Choosing a level of ability as a benchmark, such as the average ability, may create more distortions. It is not going to solve

FIGURE 2

$ce_i + t$ — MC with tax on education
$ce_i + \tau\theta_i$ — MC with tax on income

θ_i
ce_i

$ce_i + t$ or $ce_i + \tau\theta_i$

ce_i

$\theta_i + \alpha\beta\left(g(e_i) + e_i \dfrac{dg(e_i)}{de_i} \right)$

θ_i

$e_i^* = \dfrac{\theta_i}{c}$ $\overline{e_i} = \dfrac{(1-\alpha\beta)\theta_i}{c}$ e

$\Delta e = \dfrac{\alpha\beta}{c}$

over-education to abler individuals, who are the more affected by the problem, and gives an incentive to less able individuals to study less.

On the other hand, the proportional tax on education is totally based on the probability of assortative matching and the importance of the partner's income in determining utility, thus it presents less problems in its implementation.

5 Concluding remarks

The paper derives the role of assortative matching in educational decisions. Individuals choose their level of education for increasing both their own income and the probability to marry a high-income partner. The pursuit of the latter occurs by the existence of educational assortative matching, and leads individuals to acquire a greater level of education. Nevertheless, since in the marriage market the relative and not the absolute level of education matters, the increased investment in education does not improve the chance in the marriage market. Hence the optimal level of education acquired would be socially efficient without the effects of assortative matching, as less resources would have been employed for obtaining the optimal quantity of education. We define the

difference between the optimal level of education with and without educational assortative matching as over-education.

Over-education is increasing as assortative matching, the importance of the partner's income in determining utility and the level of ability increase, while it is decreasing as the rate cost of education increases.

Finally, we consider whether a public intervention can correct over-education. We illustrates alternatives interventions through a tax on education or a proportional tax on income. Both interventions can reach the socially efficient level of education, although the tax on education presents more difficulties in the implementation.

An interesting extension of this paper may be to consider also assortative matching in terms of social class between partners Although educational and social class assortative matching are positively correlated, individuals with different social background may acquire the same level of education. Introducing assortative matching by social class may have different effects according to the social group we regard. On the one hand, the opportunity cost to acquire more education is generally higher for advantaged individuals since, for instance, they may have better job opportunities through the parental network. On the other hand, this can strengthen the effect on over-education for disadvantaged people, as assortative matching by social class is a further barrier in the attempt to improve the marriage matching through education. The introduction in our framework of assortative matching by social class is left for future work.

References

Akerlof, G.A. 1976 The Economics of Caste and of the Rat Race and Other Woeful Tales. Quarterly Journal of Economics, 90, pp. 599-617.

Becker, G. S. 1973. A theory of marriage: Part I. The Journal of Political Economy 81, (4) (Jul. - Aug.): 813-46.

Burdett, K., and M. G. Coles. 1999. Long-term partnership formation: Marriage and employment. Economic Journal 109, (456) (June): F307-34.

Burdett, K., and M. G. Coles. 1997. Marriage and class. The Quarterly Journal of Economics 112, (1) (Feb.): 141-68.

Chiappori, P. A., M.F. Iyigun, and Y. Weiss. 2006. Investment in schooling and the marriage market. SSRN eLibrary.

Frank, R. H. 1978. Why women earn less: The theory and estimation of differential overqualification. The American Economic Review 68, (3) (Jun.): 360-73.

Hartog, J. 2000. Over-education and earnings: where are we, where should we go? Economics of education review 19 (2000): 131-147.

Iyigun, M. F., and R. P. Walsh. 2007. Building the family nest: Premarital investments, marriage markets, and spousal allocations. Review of Economic Studies, Vol.74, (2) (April): 507-535.

Kalmijn, M. 1991a. Shifting Boundaries: Trends in Religious and Educational Homogamy. American Sociological Review 56: 786–800.

Kalmijn, M. 1991b. Status Homogamy in the United States. American Journal of Sociology 97: 496–523.

Konrad, K. A., and K. E. Lommerud. 2000. The bargaining family revisited. The Canadian Journal of Economics / Revue Canadienne d'Economique 33, (2) (May): 471-87.

Lommerud, K. E. 1989. Educational Subsidies When Relative Income Matters. Oxford Economic Papers, vol. 41(3), pages 640-52.

Mare, R.D. 1991. Five Decades of Educational Assortative Mating. American Sociological Review 56: 15–32.

McGuinness, S. 2006. Overeducation in the Labour Market, Journal of Economic Surveys 20(3), 387-418.

Moro Egido, A. I., and S. Budria. 2007. Overeducation and wages in europe: Evidence from quantile regression. FEDEA.

Pencavel, J. 1998. Assortative mating by schooling and the work behavior of wives and husbands. The American Economic Review 88: 326-9.

Peters, M. and A. Siow. 2002. Competing premarital investments. The Journal of Political Economy 110, (3) (Jun.): 592-608.

Qian, Z. and Preston, S.H. 1993. Changes in American Marriage, 1972 to 1987: Availability and Forces of Attraction by Age and Education. American Sociological Review 58: 482–95.

Qian, Z. 1998. Changes in Assortative Mating: The Impact of Age and Education, 1970–1990. Demography 35:279–92.

Schwartz, C., Mare, R. 2005. Trends in educational assortative marriage from1940 to 2003. Demography 42: 621–646.

Shimer, R., and L. Smith. 2000. Assortative matching and search. Econometrica 68, (2) (Mar.): 343-69.

Smits, J., W. Ultee, and J. Lammers. 2000. More or Less Educational Homogamy? A Test of Different Versions of Modernization Theory Using Cross-Temporal Evidence for 60 Countries. American Sociological Review 65: 781-88.

Vaisey, S. 2006 . Education and its Discontents: Overqualification in America, 1972-2002.

Appendix: proofs

Proof of Lemma 1

We start from equation (5):

$$\theta_i + \alpha\beta\left(g(e_i) + e_i\frac{d}{de_i}g(e_i)\right) - ce_i = 0.$$

Since $\theta_i = g(e_i)$ we substitute θ_i:

$$g(e_i) + \alpha\beta\left(g(e_i) + e_i\frac{d}{de_i}g(e_i)\right) = ce_i.$$

This is a linear differential equation. First, we explicit $\frac{d}{de_i}g(e_i)$:

$$\frac{\partial}{\partial e_i}\left(g\left(e_i\right)\right) = \frac{ce_i - \left(1 + \alpha\beta\right)g\left(e_i\right)}{\alpha\beta e_i},$$

then, we bring $\frac{(1+\alpha\beta)}{\alpha\beta e_i}g\left(e_i\right)$ in the LHS:

$$\frac{\left(1 + \alpha\beta\right)}{\alpha\beta e_i}g\left(e_i\right) + \frac{\partial}{\partial e_i}\left(g\left(e_i\right)\right) = \frac{ce_i}{\alpha\beta e_i}.$$

If we multiply both sides for $\exp\{\frac{(1+\alpha\beta)}{\alpha\beta}\ln e_i\}$, we have, in the LHS, the derivative of $g\left(e_i\right)e_i^{\frac{(1+\alpha\beta)}{\alpha\beta}}$ with respect to e_i:

$$\exp\{\frac{\left(1 + \alpha\beta\right)}{\alpha\beta}\ln e_i\}\left(\frac{\left(1 + \alpha\beta\right)}{\alpha\beta e_i}g\left(e_i\right) + \frac{\partial}{\partial e_i}\left(g\left(e_i\right)\right)\right) = \exp\{\frac{\left(1 + \alpha\beta\right)}{\alpha\beta}\ln e_i\}\left(\frac{ce_i}{\alpha\beta e_i}\right),$$

$$\int \exp\{\frac{\left(1 + \alpha\beta\right)}{\alpha\beta}\ln e_i\}\left(\frac{\left(1 + \alpha\beta\right)}{\alpha\beta e_i}g\left(e_i\right) + \frac{\partial}{\partial e_i}\left(g\left(e_i\right)\right)\right) = \int \exp\{\frac{\left(1 + \alpha\beta\right)}{\alpha\beta}\ln e_i\}\left(\frac{ce_i}{\alpha\beta e_i}\right),$$

$$g\left(e_i\right)\exp\{\ln e_i^{\frac{(1+\alpha\beta)}{\alpha\beta}}\} = \int\left(\exp\{\ln e_i^{\frac{(1+\alpha\beta)}{\alpha\beta}}\}\frac{ce_i}{\alpha\beta e_i}\right)$$

$$g\left(e_i\right)e_i^{\frac{(1+\alpha\beta)}{\alpha\beta}} = \frac{ce_i^{\frac{(1+2\alpha\beta)}{\alpha\beta}}}{1 + 2\alpha\beta}.$$

By explicating e_i, we obtain $e_i = \frac{(1+2\alpha\beta)g(e_i)}{c}$. Note that $g\left(e_i\right) = \theta_i$, hence we can rewrite $e_i = \frac{(1+2\alpha\beta)\theta_i}{c}$. ∎

Proof of Lemma 2

The first order condition of (7) is:

$$\theta_i + \alpha\beta\left(g(e_i) + e_i\frac{d}{de_i}g(e_i)\right) = ce_i + t.$$

Before to start, since $\theta_i = g(e_i)$ we substitute θ_i:

$$g(e_i) + \alpha\beta\left(g(e_i) + e_i\frac{d}{de_i}g(e_i)\right) = ce_i + t.$$

17

This is a linear differential equation. First, we explicit $\frac{d}{de_i}g(e_i)$:

$$\frac{d}{de_i}\left(g\left(e_i\right)\right) = \frac{ce_i + t - \left(1 + \alpha\beta\right)g\left(e_i\right)}{\alpha\beta e_i},$$

then, we bring $\frac{(1+\alpha\beta)}{\alpha\beta e_i}g\left(e_i\right)$ in the LHS:

$$\frac{\left(1 + \alpha\beta\right)}{\alpha\beta e_i}g\left(e_i\right) + \frac{d}{de_i}\left(g\left(e_i\right)\right) = \frac{ce_i + t}{\alpha\beta e_i}.$$

If we multiply both sides for $\exp\{\frac{(1+\alpha\beta)}{\alpha\beta}\ln e_i\}$, we have, in the LHS, the derivative of $g\left(e_i\right)e_i^{\frac{(1+\alpha\beta)}{\alpha\beta}}$ with respect to e_i:

$$\exp\{\frac{\left(1 + \alpha\beta\right)}{\alpha\beta}\ln e_i\}\left(\frac{\left(1 + \alpha\beta\right)}{\alpha\beta e_i}g\left(e_i\right) + \frac{d}{de_i}\left(g\left(e_i\right)\right)\right) = \exp\{\frac{\left(1 + \alpha\beta\right)}{\alpha\beta}\ln e_i\}\left(\frac{ce_i + t}{\alpha\beta e_i}\right),$$

$$\int \exp\{\frac{\left(1 + \alpha\beta\right)}{\alpha\beta}\ln e_i\}\left(\frac{\left(1 + \alpha\beta\right)}{\alpha\beta e_i}g\left(e_i\right) + \frac{d}{de_i}\left(g\left(e_i\right)\right)\right) = \int \exp\{\frac{\left(1 + \alpha\beta\right)}{\alpha\beta}\ln e_i\}\left(\frac{ce_i + t}{\alpha\beta e_i}\right),$$

$$g\left(e_i\right)\exp\{\ln e_i^{\frac{(1+\alpha\beta)}{\alpha\beta}}\} = \int\left(\exp\{\ln e_i^{\frac{(1+\alpha\beta)}{\alpha\beta}}\}\frac{ce_i}{\alpha\beta e_i}\right),$$

$$g\left(e_i\right)e_i^{\frac{(1+\alpha\beta)}{\alpha\beta}} = \frac{ce_i^{\frac{(1+2\alpha\beta)}{\alpha\beta}}}{1 + 2\alpha\beta} + te_i^{\frac{1+\alpha\beta}{\alpha\beta}}.$$

By explicating e_i and noting that $g\left(e_i\right) = \theta_i$, we obtain $e_i = \frac{(\theta_i - t)(1 + 2\alpha\beta)}{c}$.∎

Proof of Lemma 3

The first order condition of (9) is:

$$\theta_i\left(1 - \tau\right) + \beta\alpha\left(g(e_i) + e_i\frac{d}{de_i}g(e_i)\right) - ce_i = 0.$$

Before to start, since $\theta_i = g(e_i)$ we substitute θ_i:

$$g(e_i)\left(1 - \tau+\right) + \beta\alpha g(e_i) + \beta\alpha e_i\frac{d}{de_i}g(e_i) - ce_i = 0.$$

This is a linear differential equation. First, we explicit $\frac{d}{de_i}g(e_i)$:

$$\frac{d}{de_i}g(e_i) = \frac{c}{\beta\alpha} - \frac{g(e_i)\left(1-\tau+\beta\alpha\right)}{\beta\alpha e_i},$$

then, we bring $\frac{(1+\alpha\beta)}{\alpha\beta e_i}g\left(e_i\right)$ in the LHS:

$$\frac{\left(1-\tau+\beta\alpha\right)}{\beta\alpha e_i}g(e_i) + \frac{d}{de_i}g(e_i) = \frac{c}{\beta\alpha}.$$

If we multiply both sides for $\exp\{\frac{(1-\tau+\beta\alpha)}{\beta\alpha}\ln(e)\}$, we have, in the LHS, the derivative of $g\left(e_i\right)e_i^{\frac{(1+\alpha\beta)}{\alpha\beta}}$ with respect to e_i:

$$\exp\{\frac{\left(1-\tau+\beta\alpha\right)}{\beta\alpha}\ln(e)\}\left(\frac{\left(1-\tau+\beta\alpha\right)}{\beta\alpha e_i}g(e_i) + \frac{d}{de_i}g(e_i)\right) = \exp\{\frac{\left(1-\tau+\beta\alpha\right)}{\beta\alpha}\ln(e)\}\frac{c}{\beta\alpha},$$

$$\int \exp\{\frac{\left(1-\tau+\beta\alpha\right)}{\beta\alpha}\ln(e)\}\frac{\left(1-\tau+\beta\alpha\right)}{\beta\alpha e_i}g(e_i) + \exp\{\frac{\left(1-\tau+\beta\alpha\right)}{\beta\alpha}\ln(e)\}\frac{d}{de_i}g(e_i) =$$

$$\int \exp\{\frac{\left(1-\tau+\beta\alpha\right)}{\beta\alpha}\ln(e)\}\frac{c}{\beta\alpha},$$

$$\exp\{\ln(e^{\frac{(1-\tau+\beta\alpha)}{\beta\alpha}})\}g(e_i) = c\int \frac{\exp\{\ln(e^{\frac{(1-\tau+\beta\alpha)}{\beta\alpha}})}{\beta\alpha},$$

$$g\left(e_i\right)e_i^{\frac{(1+\alpha\beta)}{\alpha\beta}} = c\frac{e^{\frac{(1-\tau+\beta\alpha)}{\beta\alpha}+1}}{\left(1-\tau+\beta\alpha\right)}.$$

By explicating e_i and noting that $g\left(e_i\right) = \theta_i$, we obtain $e_i = \frac{\theta_i(1-\tau+\beta\alpha)}{c}$. $\blacksquare$