

Assortative Matching and the Education Gap

Ximena Peña*

Georgetown University - Banco de la República

December 2006

Abstract

This paper attempts to explain the decrease and reversal of the education gap between males and females. Given a continuum of agents, the education decisions are modelled as an assignment game with endogenous types. In the first stage agents choose their education level and in the second they participate in the labor and marriage markets. Competition among potential matches ensures that the efficient education levels can always be sustained in equilibrium, but there may be inefficient equilibria. Combining asymmetries intrinsic to the modelled markets the model reproduces the observed education gap.

Key Words: Assortative matching, efficiency, gender, education.

JEL Classification Numbers: C78, D13, D61.

*I would like to thank my advisors James Albrecht, Axel Anderson and Susan Vroman for their inspired guidance. Also, many thanks to Luca Anderlini, Orazio Attanasio, Renee Bowen, Julián Castillo, Roger Lagunoff, Zaki Zahran, and seminar participants at Los Andes, Banco de la República, Fedesarrollo, Javeriana, Nacional, Rosario Universities, Econometric Society Meetings, Midwest and Midwest Theory Conferences for helpful comments. The usual disclaimer applies. xp@georgetown.edu, <http://www12.georgetown.edu/students/xp>

1 Introduction

In countries like the U.S., Colombia and Brazil, women have closed the education gap in college education and even surpassed men in college attainment, reversing the historical attainment advantage enjoyed by the latter. This paper proposes an explanation for the decrease and reversal of the education gap. Education has returns in the two main markets young adults participate in: labor and marriage market. In the labor market education enhances productivity and thus generates higher returns via the wage rate. In the marriage market, if spousal attributes are complementary, people marry with a like education partner and thus education is a vehicle to match with a better type. This paper highlights the importance of the search for a spouse in the education decisions of both sides of the market¹.

The education decisions are modelled as a continuous assignment game with endogenous types that can be described in two stages. The first stage is noncooperative: agents observe their education costs and simultaneously choose education investments. The returns to players are affected by the (equilibrium) investment choices of other agents. In the cooperative second stage agents match, produce ‘household good’ and work. Assuming transferable utility, spouses bargain over the fraction of the household surplus each appropriates. Higher investments in education generate an improved set of potential matches and a higher share of the surplus. The result of the model is a set of matched agents, a split of the surplus and a distribution of education across agents.

¹To keep the model tractable, we assume away the effect of education and marriage decisions on social outcomes such as fertility.

The decrease and eventual reversal of the education gap is explained by introducing asymmetries intrinsic to the modelled markets that affect the marginal benefit of education. Namely, the observed evolution of the gender (marginal) wage gap in the labor market and the relative abundance of females in the marriage market. Hence, the model generates asymmetric education decisions from sets of agents with identical cost distributions.

Two very recent papers are very close in spirit to this one. Chiappori, Iyigun and Weiss (2006) propose an alternative and complementary explanation to the reversal of the education gap. They develop a model with pre-marital schooling investment, marriage and labor markets where agents allocate their time between home and market production. They suggest that women may overtake men in schooling as a means to escape labor market discrimination because it is lower at high education levels. In addition, since females' household time obligations have declined over time, their marriage-market returns have also increased. The main differences between the Chiappori et al. (2006) and this paper are that while they have two 2 possible education levels we have a continuum. In addition, despite combining both labor and marriage market incentives, the proposed explanations are different and complementary.

Iyigun and Walsh (2006), embed pre-marital investments and spousal matching into a collective household model. In a purely marriage market setting, a corollary to their results is that when men are in short supply, women choose higher pre-marital investments. The main difference with this paper is that they do not model the labor market. Therefore, even though they suggest the same channel of incentives for females, by not including labor market incentives, cannot address the observed

decrease of the education gap.

The classical paper in the marriage market literature belongs to Becker (1973). Several models endogenize investment levels in a matching environment. Cole, Mailath and Postlewaite (CMP 2001a, 2001b) solve the hold-up problem by endogenizing investment specificity and introducing competition among agents for complementary investments. The model predicts the existence of multiple equilibria, and the ex-ante efficient levels of investment can always be attained given the optimal bargaining rule². Peters and Siow (2002) develop a model where parents invest in their children's wealth and spousal wealth is a public good in marriage. The authors find that when the marriage market is large the competitive equilibrium is efficient.

Other authors have attempted to explain some aspect of the education attainment ratio (see for example, Ríos-Rull and Sánchez, 2002 and Echevarría and Merlo, 1999). These papers use the household decision, that is, have parents determine investment levels in their children's attributes. We model the problem from the individual's perspective. Since young adults *own* the college education decision so long as education costs are appropriately modelled, this seems a better approach.

The contribution of the paper is to introduce asymmetries natural to the modelled markets and reproduce behavior of education attainment. The additional incentives for women to educate come from their relative abundance and its implications for marriage market prospects. There are alternative and complementary sources for increased incentives for women to educate, that to the best of my knowledge have

²This paper is similar to the continuum version of CMP for the special case of continuous best response functions. CMP deal with potentially discontinuous best response functions, since lack of coordinations can yield attribute distributions that imply such best replies. However, equilibria with continuous best response functions are a subset of the latter.

not been developed in the literature. For example, given a world where divorce rates have increased significantly and the probability of remarrying is greater for men, if women are more risk averse -there is evidence for this from the experimental economics literature (for a survey see Croson and Gneezy, 2004)- then they insure more than men against the risk of divorce by educating further.

The rest of the paper is organized as follows. The next section presents some stylized facts and a motivating example is worked out in Section 3. Section 4 develops the model, characterizes the equilibria, its efficiency and comparative statics results. Section 5 concludes. We appendicize all technical proofs and provide intuitive arguments in-text.

2 Stylized Facts

Women have been consistently investing more in education than men, to the extent that they have reversed the historical schooling attainment advantage enjoyed by the latter. Until the late 1970's the ratio of college attainment of men to women, that is, the ratio of the number of men to women with completed University education, was around 1.6 in the U.S. (Ríos-Rull and Sánchez, 2002). However, the education ratio experienced a dramatic decline and females overtook men in college education: Goldin et al. (2006) show that women have attained higher college graduation rates starting in the 1970 cohort. This trend is not specific to developed countries since a similar situation is observed in several Latin American countries, including Brazil and Colombia. To motivate the discussion show facts for Colombia will be shown

both because the changes in education gender attainment are qualitatively similar to what has been observed elsewhere, and to study the phenomenon in a developing country³. Figure (1) shows the education attainment ratio for individuals between 25 and 40 years of age with at least a college degree in Colombia⁴. The dashed line shows the ratio of males to females in the group, which consistently declined from 2.6 in the late 1970's to 0.79 in 2004. Starting in the early 90's women have surpassed men in their education attainment.

Despite a decreasing trend since the 1950's gender wage differentials persist in all industrialized nations, where *unconditional* gender wage differential (wage gap in what follows)⁵ is defined as the ratio between average male wage over average female wage. In the U.S. it has decreased from 1.59 in 1979 to 1.35 in 1989 and 1.25 in 1998 for full-time workers (Blau and Khan, 2004). A similar behavior is observed in developing countries. The thick solid line in Figure 1 shows the decrease in the wage gap in Colombia for agents with education of college or more: from over 2 in the early 80's to under 1.3 in 2005. Males are still paid higher level wages than females at any education level. Define the marginal wage rate as the difference in mean wage between agents with less than college and those with completed college or more, which measures the additional earnings associated to completing college

³An exercise that could be done is to repeat the stylized facts for other developed and developing countries, to get an idea of how the labor and marriage market incentives interact, and how much heterogeneity there is.

⁴Colombian calculations based on the September shift of the National Household Survey (NHS), significant for the 7 main cities. Calculations were also made for the percentage of males over percentage of females with university education and the qualitative results are the same.

⁵Calculations were also made for the *conditional* wage gap, that is, controlling for variables such as occupation. Despite the controls, qualitatively the results are very similar.

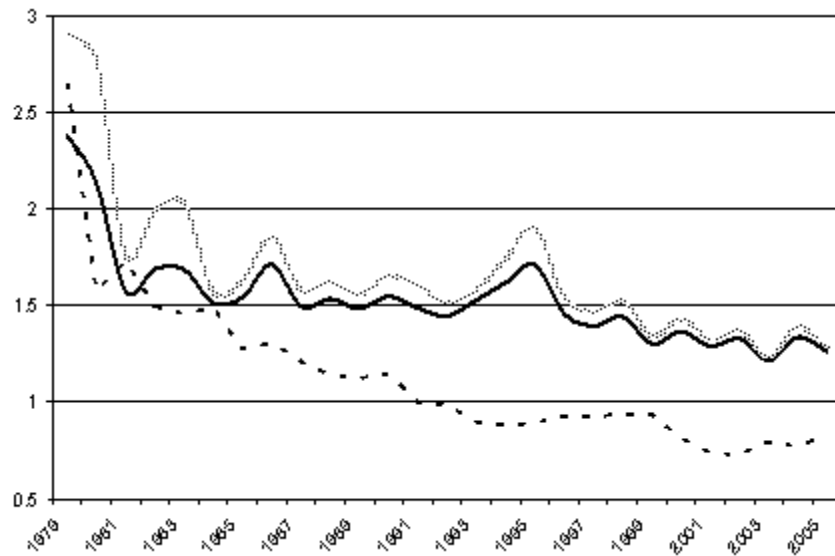


Figure 1: Education Attainment Ratio, Wage Gap and Marginal Wage Gap (Male/Female, College+, 25-40 years of age). The education attainment ratio (dashed line) decreases from over 2 in the late 70's to 0.8 in 2005. The wage gap (solid line) passes from over 2 at the beginning of the period to 1.3 in 2005. The marginal wage gap (solid grey line) follows very closely the behavior of the wage gap.

education. The grey line in Figure 1 is the marginal wage gap, i.e. the ratio between male and female college premiums. Marginal incentives move very closely to the observed change in the wage gap. Not only do women face lower wage *levels*, but also lower *marginal* wages⁶. Labor market incentives suggest that men should be more educated since education appears to be less fruitful for women. This makes the relative female over-education even more surprising.

This paper includes the marriage market as an additional source of incentives to agents. How much does marriage matter in the agents college choices? Ge (2006) finds that marriage plays a significant role in a female's college choice in the US, for a sample of high school white females from the National Longitudinal Survey of Youth 1979. When excluding the benefits from marriage the model's predicted college graduation drops 6 percentage points, from 38% to 32%.

The inclusion of the marriage market analysis aims at capturing two main effects: assortative matching across education levels and the effect of the relative scarcity of men. People do not marry randomly. Fernández, Guner and Knowles (2001) use household surveys from 34 countries to calculate the degree of correlation of spouses' education (marital sorting). They find that the average Pearson correlation between spousal education for the sample is 0.610, with a standard deviation of 0.106. Table (1) describes the education attainment of married females by the education attainment of their spouses in Colombia in 2003. Assortative mating within education

⁶Dougherty (2005) suggests that the rate of return to schooling appears to be greater for females than for males in the U.S., in spite of the existence of both a conditional and unconditional wage gap. This contradicts the findings presented for Colombia of a wage gap increasing in education levels. Other countries where returns to post-secondary education are higher for men than for women, especially after the 25th percentile, include the Netherlands and Sweden (see for example Albrecht, Björklund, Vroman, 2003 and Albrecht, Van Vuuren and Vroman, 2006).

2003				
Male\Female	Less than High School	Less than College	College+	Total
Less than High School	38%	11%	1%	49%
Less than College	10%	24%	4%	39%
College+	1%	4%	7%	12%
Total	49%	39%	12%	100%

Table 1: Assortative Matching (Colombia, 25-40 years of age)

classes is large since we observe that a the majority of matches fall along the main diagonal; the correlation coefficient between spousal education is 0.63.

Females face tougher competition than males in the marriage market. Factors affecting genders in a (potentially) different way such as death rates, imprisonment rates, immigration patterns or sexual orientation generate a relative scarcity of males in the marriage market. To measure the relative abundance of females we calculate the ratio between the fraction of matched males over fraction of matched females. Colombia data shows that for people between 20 and 65 years of age, this ratio was on average 1.15 for the 1979-2005 period, with a standard deviation of 0.02 and for agents between 25 and 40 years of age the figure is 1.07 with the same standard deviation. That is, 15% and 7% more males are matched than females, respectively. This generates increased competition and additional incentives for women to educate.

Summarizing, decreasing wage differentials and marginal wage gap suggest that still today men have more incentives to educate. The marriage market is introduced to model the effects assortative matching and the latent additional incentive for women to educate given that men are in short supply. It is the interaction of labor and marriage market incentives is necessary to reproduce the puzzling decrease and

reversal of the College attainment ratio.

3 Motivating Example

Let us motivate the model by an example. Assume a large population on each side of the market: males and females. This implies that a change of an attribute by a single agent on either side does not affect the division of matches unaffected by the attribute change. Also, an agent who increases his or her education level can match with other agents in the economy. Let's solve a parametrized Social Planner's problem for the marriage market in isolation, which is always supportable as an equilibrium to illustrate the matching and education investment decisions.

Females and males are indexed by their education costs c_f and c_m , which are uniformly distributed on $[\underline{c}, \bar{c}]$ and capture both the individual's natural ability and budget constraint. Let $\chi(\cdot)$ be the education choice function of a female with cost c_f , mapping it into an education level x ; $\varphi(\cdot)$ is the education choice function of a c_m male, mapping costs into education level y . The cost for a female of acquiring education x is $c_f x$ and the cost for a male of acquiring education y is $c_m y$. The household production generated by an (x, y) couple is given by, $\pi(x, y) = x^{1/3}y^{1/3}$. Note that $\pi_x, \pi_y, \pi_{xy} > 0$. Aggregate surplus is maximized by the association of likes: i.e. lowest cost female matching with the lowest cost male and so on. This type of positive assortative matching (PAM) is a consequence of the complementarity of the couple's education levels ($\pi_{xy} > 0$).

What is the effect of a having more females than males? All potential matches

will be made, starting with the lowest cost females, until there are no more available males; some high cost females will be left unmatched. The relative abundance of females is measured as the ratio of matched males over matched females: $\alpha \equiv \frac{\bar{c} - \underline{c}}{\tilde{c}_f - \underline{c}}$ with \tilde{c}_f being the cost of the ‘last’ matched female. An increase in α is equivalent to an increase of the ratio of females to males.

A matching is a rule associating a female to her mate. Given uniform cost functions, the matching is a straight line:

$$c_m = \mu^*(c_f) = \alpha c_f \quad \text{with} \quad \alpha \equiv \frac{\bar{c} - \underline{c}}{\tilde{c}_f - \underline{c}}$$

Efficiency requires that for each c_f and c_m , education choices must solve:

$$\underset{x, \varphi}{Max} \quad x^{1/3} y^{1/3} - c_f x - c_m y \quad \text{st.} \quad c_m = \alpha c_f \quad (\text{b})$$

The second order condition is satisfied and we have an interior maximum:

$$\chi(c_f) = \frac{1}{27\alpha c_f^3} \quad \text{and} \quad \varphi(c_m) = \frac{1}{27\alpha^2 c_f^3}$$

The ratio of education attainment of men to women is $\frac{\varphi(c_m)}{\chi(c_f)} = \frac{1}{\alpha}$. If $\alpha = 1$, then $\frac{\varphi(c_f)}{\chi(c_f)} = 1$ and both sides of the market are equally educated. However, when $\alpha > 1$ for a given c_f , $\frac{\varphi(c_f)}{\chi(c_f)} < 1$ and females are more educated than their respective matches, even for the lowest-cost pair. Conversely, if $\alpha < 1$, $\frac{\varphi(c_f)}{\chi(c_f)} > 1$ and males are more educated. This simple example illustrates the subtle point that in the marriage market an asymmetric gender composition of the population generates, through the

matching, different levels of education investments between men and women.

4 Model

The education decisions are modelled as an assignment game with endogenous types. In the first stage agent's education investments are determined while in the second agents match and split the surplus.

Let C_f, C_m be the distributions of education costs that admit a density across women and men, respectively, with⁷ $C_f = C_m$. There is a continuum of size m of males and size f of females. At $t = 0$ agents sample a constant cost of education, which captures their natural ability and budget constraint. Let χ and φ be education choice functions mapping education costs $c_f, c_m \in [0, 1]$ into desired education levels $x \in [0, \bar{x}], y \in [0, \bar{y}]$ for women and men, respectively. After educating, agents participate in the labor and marriage markets where they match, produce and bargain over the surplus.

In the second stage agents are identified by their publicly observable education level. In the labor market they face wage schedules $w = \langle w_f(x), w_m(y) \rangle$, such that w is C^2 and increasing. Let us assume that in equilibrium all workers are employed and that there is a marginal wage gap, as observed in the data: $w'_m(y) \geq w'_f(x) \forall x = y$. This will be exploited in the comparative statics exercises (Proposition 9).

A married couple, female x and male y , generates divisible output $\pi(x, y)$. π is C^2 , symmetric, strictly increasing and strictly supermodular (SPM) $\frac{\partial^2 \pi(x, y)}{\partial x \partial y} > 0$.

⁷There could potentially be asymmetries in costs distributions across genders as well as in the relative size of males and females, but to better understand the effect of changes in the composition of the population we leave $C_f = C_m$.

We normalize: $\pi(x, 0) = \pi(0, y) = \pi(x, \emptyset) = \pi(\emptyset, y) = 0 \quad \forall x, y$, where \emptyset means no match. Home production captures the quantity and quality of children, and the enjoyment of each other's company.

The total surplus of an (x, y) couple is given by the home production and labor market outcomes of spouses: $s(x, y) = \pi(x, y) + w_m(y) + w_f(x)$. Therefore, for a single agent the total surplus equals the wage rate. Marriage is individually rational. Let π and w_f, w_m be such that s is strictly concave in x, y .

4.1 Second Stage

Being a two stage game, it is solved by backward induction. Assume that the first stage decision rules χ, φ are C^1 and strictly monotone decreasing (which we later prove in Proposition 5). This, together with education cost distributions that admit a density, imply that the resulting education cumulative distributions G_m, G_f of men and women, respectively, admit a density $\mathbf{g}_m, \mathbf{g}_f$. Stage 2 is an assignment game where the scarce resource, i.e. highly educated agents, are allocated to maximize total social surplus. An outcome is a set of matched pairs and a split of the surplus.

A *matching* μ is a function $\mu : [0, \bar{x}] \rightarrow [0, \bar{y}] \cup \{\emptyset\}$, that associates to an x woman the education level of her mate $y = \mu(x)$ where μ is one-to-one on $\mu^{-1}(M)$ and \emptyset is interpreted as no match. A couple who are not matched under μ , but who prefer each other to their assignments, can *block* the matching since by rematching and sharing the resulting surplus, they are strictly better off.

A matching μ in education levels is *feasible* if:

$$\int_A g_f(x) dx = \int_A g_m(\mu(x)) dx \quad \forall A \subset [0, \bar{x}] \quad (1)$$

Let $v_f(x) \geq 0$ be the endogenously determined share of match surplus plus wage rate (or return) of a type x woman and $v_m(y) \geq 0$ that of a y man. A *stable bargaining outcome* is feasible⁸, individually rational and satisfies $\forall x, y$:

$$\begin{aligned} v_f(x) + v_m(y) &= s(x, y) && \text{for matched couples} \\ v_f(x) + v_m(y) &> s(x, y) && \text{for unmatched couples} \end{aligned} \quad (2)$$

In a stable bargaining outcome, within a gender, agents with the same attributes receive equal (gross) payoffs: "equal treatment" in CMP. As a result, there are no blocking pairs. Subtract $v_m(y)$ from both expressions in (2). It becomes apparent that in a stable outcome payoffs $v_f(x), v_m(y)$ are the upper envelope of the shares generated by the potential partners⁹. Equation (2) can thus be restated as an individual maximization where spousal education maximizes an individual's share of household surplus:

$$v_f(x) = \underset{y \in [0, \bar{y}]}{\text{Max}} \{s(x, y) - v_m(y)\} \quad \text{and} \quad v_m(y) = \underset{x \in [0, \bar{x}]}{\text{Max}} \{s(x, y) - v_f(x)\} \quad (3)$$

⁸For the case of discontinuous attribute choices, see CMP's (2001a) definition of feasibility.

⁹Since marriage is individually rational, the lower bound of the payoffs are the agent's wage, and hence payoffs are always positive.

When choosing spousal education level agents balance two opposing effects take place in this maximization: a higher education spouse implies higher joint surplus but also appropriates a higher share of the surplus. The split of surplus reflects what agents are willing to offer for different spouses: given complementarity, $s_{xy} > 0$, a low type male will always be outbid by a high type one for a female with high education level. Thus, competition generates mutually acceptable matches and an association of likes (Proposition 3).

An equilibrium in the second stage, taking the education decisions as given, is a matching and a split of the surplus.

Definition 1 *For given first stage choices a **matching equilibrium** is a pair (μ, v) such that:*

1. *The matching μ is feasible (Condition 1).*
2. *The outcome is stable (Condition 2).*

Studying the Social Planner's problem (SP) is interesting since in this stage the first and second welfare theorems obtain (Proposition 2). Given χ, φ , the SP chooses the *efficient matching* μ to maximize social welfare and, due to input complementarity, the assignment is positively assortative. If females are relatively more abundant, some will be unmatched. Let \tilde{x} denote the education level of the 'last' matched female and as will be shown in Proposition 3 feasibility of the matching can be restated as shown in the constraint set.

$$S(\chi, \varphi) = \underset{\mu}{Max} \int_0^{\bar{x}} s(x, \mu(x)) dG_f(x) \quad s.t. \quad G_f(x) = G_m(\mu(x)) \quad \forall x \in [\tilde{x}, \bar{x}] \quad (4)$$

If the mass of females exceeds that of males, provided that all possible pairs are matched, low education women are left unmatched¹⁰.

Gretsky, Ostroy and Zame (1992) -GOZ- prove that the SP's problem can be decentralized and hence equilibrium matchings are efficient. The associated matching pattern is equivalent to the existence of a stable bargaining outcome.

Proposition 2 (GOZ, 1992) *Given G_f, G_m , the Social Planner's problem can be decentralized: the first and second welfare theorems obtain.*

Let us characterize the matching equilibrium. Given Proposition 2, as long as the education distributions admit a density, there is an increasing, continuously differentiable and unique efficient matching.

Proposition 3 *Given G_f, G_m :*

1. μ is increasing.
2. The matching is unique and C^1 :

$$\mu(x) = G_m^{-1}(G_f(x)) \quad \forall x \in [\tilde{x}, \bar{x}] \quad (5)$$

¹⁰This implication of the frictionless model does not fit well what we observe in the data: single women have different levels of education, and many of them are highly educated. If frictions were introduced in the marriage market, we would obtain a statistical version of this result, i.e. highly educated women are more likely to marry.

Existence of a stable split of the surplus between spouses is immediate from Proposition 2. Proposition 4 characterizes the stable bargaining outcome, and shows that it induces efficient education levels since agents internalize the returns to their investments.

Proposition 4 *For given G_f, G_m :*

1. *For any stable bargaining outcome, v is strictly increasing and C^2 :*

$$v'_f(x) = s_1(x, \mu(x)) \forall x \in [0, \bar{x}] \quad \text{and} \quad v'_m(\mu(x)) = s_2(x, \mu(x)) \forall y \in [0, \bar{y}] \quad (6)$$

2. *v_m is strictly concave. v_f is strictly concave if the following condition holds¹¹:*

$$-\frac{s_{11}(x, \mu(x))}{s_{12}(x, \mu(x))} > \frac{\mathbf{g}_f(x)}{\mathbf{g}_m(G_m^{-1}(G_f(x)))} \forall x \in [0, \bar{x}]$$

4.2 First Stage

Let us now turn to the determination of education choice functions χ and φ . Recall that agents sample an education cost from distributions C_f, C_m that admit a density and choose the education level to maximize the (net) payoff:

$$\chi(c_f) = \arg \max_{x \in [0, \bar{x}]} \{v_f(x) - c_f x\} \quad \forall c_f \in [0, 1]$$

¹¹Note that if s was assumed convex in x, y , then both v_f and v_m are immediately convex.

$$\varphi(c_m) = \arg \max_{y \in [0, \bar{y}]} \{v_m(y) - c_m y\} \quad \forall c_m \in [0, 1] \quad (7)$$

From Proposition 4 v_f, v_m are strictly increasing and differentiable and so we can characterize the solution to the agent's problem.

Proposition 5 *Given C_f, C_m :*

1. χ and φ are continuous, differentiable and strictly decreasing:

$$\chi(c_f) = s_1^{-1}(c_f, \mu(c_f)) \quad \forall c_f \in [0, 1]$$

$$\varphi(c_m) = s_2^{-1}(c_f, \mu(c_f)) \quad \forall c_m \in [0, 1] \quad (8)$$

2. V_f, V_m are continuous and strictly concave:

$$V_f(c_f) = \underset{x \in [0, 1]}{\text{Max}} \{v_f(x) - c_f x\} \quad \text{and} \quad V_m(c_m) = \underset{y \in [0, 1]}{\text{Max}} \{v_m(y) - c_m y\}$$

Results from Proposition 5 in turn ensure that the required conditions for Proposition 4 are met: given that χ, φ are strictly monotone and C^1 , v is C^2 .

An equilibrium in the game are education choice functions and a split of surplus such that a single player's education decision is a best response given the other players' choices. Education distributions, in turn, are consistent with a stable outcome and a feasible matching.

Definition 6 *Given C_f, C_m , an **education equilibrium** is an array (χ, φ, μ, v) such that:*

1. *Agents maximize the net payoff*¹² (Condition 7).

2. (μ, v) is a **matching equilibrium**.

The efficient education decisions can be characterized using the SP's problem, since the second welfare theorem holds for the whole game (Proposition 7). First, let's formally define a *matching in education costs* as a function $\mu^* : [0, 1] \rightarrow [0, 1] \cup \{\emptyset\}$ that associates to a c_f woman her mate's cost $\mu^*(c_f)$ and \emptyset implies she's unmatched. If men are in short supply, some women will be unmatched. Let \tilde{c}_f denote the education cost of the 'last' matched female. Let $\mathbf{c}_f, \mathbf{c}_m$ be the cost densities. A feasible matching in education costs is given by:

$$\int_A \mathbf{c}_f(c_f) dc_f = \int_A \mathbf{c}_m(\mu^*(c_f)) dc_f \quad \forall A \subset [0, 1] \quad (9)$$

As will be apparent from Proposition 7, this implies:

$$\mu^*(c_f) = C_m^{-1}(C_f(c_f)) \quad \forall c_f \in [0, \tilde{c}_f] \quad (10)$$

The SP's problem is then:

¹²An alternative way to express this condition which evidences the equivalence with Nash Equilibrium is:

$\forall c_f, \forall \tilde{\chi} : V_f(\chi(c_f)) \geq V_f(\tilde{\chi}(c_f))$ and $\forall c_m, \forall \tilde{\varphi} : V_m(\varphi(c_m)) \geq V_m(\tilde{\varphi}(c_m))$. Note that an education equilibrium is a Subgame Perfect Nash Equilibrium, since it survives backward induction.

$$\begin{aligned}
S(C_f, C_m) &= \underset{\chi^*, \varphi^*, \mu^*}{Max} \int_0^1 [s(x, y) - c_f x - \mu^*(c_f) y] dC_f \\
s.t. \mu^*(c_f) &= C_m^{-1}(C_f(c_f)) \quad \forall c_f \in [0, \tilde{c}_f]
\end{aligned} \tag{11}$$

For given C_f, C_m that admit a density, efficiency requires that the efficient matching μ^* and education choice functions χ^*, φ^* maximize net social welfare S . Again, given SPM of S , the efficient matching implies PAM¹³ (Proposition 7).

We find that a one-to-one frictionless matching market with an ex-ante investment stage generates efficient investments despite the (potential) hold-up problem and investment externalities due to the complementarity assumptions. However, as in CMP, inefficient equilibria may exist as well due to a coordination failure: the absence in the other side of the market of agents with attributes that would induce the efficient education levels.

Proposition 7 *Given C_f, C_m :*

1. *There exists a unique solution to the Social Planner's problem $(\chi^*, \varphi^*, \mu^*)$. The efficient matching μ^* implies PAM.*
2. *An efficient education equilibrium exists.*
3. *If $s_i(1, 1) > 1$, $i = 1, 2$, there may be an inefficient over-investment equilibria where all agents choose the maximum level of education.*

¹³Kremer and Maskin (1996), assuming a discrete number of agents, find that given an asymmetric surplus function cross-matching around the median is more efficient than PAM. As the set of agents tend to infinity the *measure* of agents cross-matching tends to zero. Therefore, despite asymmetries in the marginal incentives, PAM is always efficient in this setting.

There is a unique solution to the SP's problem and the Theorem of the Maximum under convexity assumptions characterizes the solutions. The first order conditions for both the Social Planner's and the individual's problem are shown to be equivalent¹⁴; the SP's solutions are a subset of the decentralized equilibria and hence the set of decentralized equilibria is non-empty.

We now turn to comparative statics results. We introduce two asymmetries, decreasing marginal wage gap and relative abundance of females in the marriage market, and study their effect on the marginal return of education. The interaction between the two asymmetries theoretically replicates the decrease and reversal of the education gap.

First, let's formalize the way in which the marginal wage gap affects the education decisions of agents. The stylized facts suggest that the marginal wage for males is higher than for females. In the model, a higher marginal wage implies higher marginal return to an education level and therefore a man would choose higher education levels than a woman of the same education cost (Lemma 8). In the absence of other asymmetries labor market incentives, captured through a higher marginal wage for males, translates directly into an education gap favoring men.

Lemma 8 (*Marginal Incentives*) *For given C_f, C_m and an equal proportion of men and women, marginal labor incentives determine the education gap: if $w'_m(y) \geq w'_f(x) \forall x = y$ then $\frac{\varphi(c_f)}{\chi(c_f)} \geq 1$.*

A more challenging exercise has to do with changes in the relative abundance of agents and the effect on payoffs. The motivating example shows in a simple setting

¹⁴Existence of an equilibrium can be established directly using a fixed point argument.

that as the ratio of females to males increases, so do incentives for females to educate through the matching. When thinking about the problem in terms of the decentralized solution, an initial intuition suggests that an increase in the relative abundance of women translates into a set of matched females described by a truncated female education distribution G_f . Given PAM, the highest education woman matches with the top education man, and so on, until there are no more men to match with. The truncation would happen at the point where the mass of matched males equals the mass of females. However, the actual effect is more complex because education decisions, and hence education distributions, are equilibrium objects and thus changes in the relative abundance of females affect them. Education decisions χ, φ determine the matching μ ; once matched, agents bargain over the split of surplus v which in turn determines the education decisions χ, φ .

To disentangle the feedback effects described above and simplify the comparative statics let the relative abundance of women be a proportional increase in the number of females: there are $\alpha \in \mathbb{R}$ females per male for each education cost, with $\alpha \geq 1$. This definition is convenient because the education cost distributions remain equal between genders. Thus, results regarding asymmetric education decisions are not driven by differences in cost distributions.

From the second part of Proposition 7 we know that the unique efficient matching in education levels μ and the unique efficient matching in education costs μ^* are equivalent and describe the same assignment: they both associate the same ‘types’ together. The former describes the set of matched pairs in terms of the education costs while the latter does it using education levels. Therefore using μ^* the matching

is pinned down in terms of primitives of the model and we can identify the effect of changes in α on the education decisions, side-tracking the feedbacks between matching and education decisions.

In the absence of asymmetries and given identical education cost distributions, education decisions are the same for both sides of the market. This implies that men and women are equally educated. Introducing each asymmetry in isolation yields a clear cut result. On one hand, a marginal wage gap favoring men generates more educated males because it translates directly into higher relative marginal payoffs. On the other, men in short supply yields more educated women. The intuition behind this result is as follows: an increase in α is equivalent to having women distributed on a space that is more *dense*, generating increased competition for females vis-à-vis males. Tougher competition, captured through changes in the matching and hence through the allocated spouse, translates into higher education levels for women.

Proposition 9 *For given C_f, C_m , let $(\chi, \varphi, \mu^*, v)$ be an education equilibrium.*

1. *In the absence of asymmetries ($w'_m(y) = w'_f(x) \forall x = y$ and $\alpha = 1$), $\frac{\varphi(\mu^*(c_f))}{\chi(c_f)} = 1$ and there is no education gap.*
2. *If males face higher marginal wages for all education levels, ($w'_m(y) > w'_f(x) \forall x = y$ and $\alpha = 1$), there exists an education equilibrium $(\bar{\chi}, \bar{\varphi}, \mu^*, \bar{v})$ such that $\frac{\bar{\varphi}_m(\mu^*(c_f))}{\bar{\chi}_f(c_f)} > 1$ and males are more educated.*
3. *If there is a proportional increase in the number of females per male ($w'_m(y) = w'_f(x) \forall x = y$ and $\alpha > 1$), there exists an education equilibrium $(\bar{\chi}, \bar{\varphi}, \mu^*, \bar{v})$ such that $\frac{\bar{\varphi}_m(\mu^*(c_f))}{\bar{\chi}_f(c_f)} < 1$ and females are more educated.*

The two asymmetries work in opposite directions: a marginal wage gap favoring men implies they should be more educated while the relative abundance of women provides additional incentives for them. Therefore, the interaction of labor and marriage market incentives is required to explain the decrease and reversal of the education gap, which is a direct consequence of the previous Proposition.

Corollary 10 *Balancing out the labor and marriage market incentives the model theoretically reproduces the closing and reversal of the education gap: if $w'_m(\varphi(\mu^*(c_f))) \gtrsim w'_f(\chi(c_f))$ then $\frac{\varphi(\mu^*(c_f))}{\chi(c_f)} \gtrsim 1$.*

The Corollary states the conditions under which labor or marriage market incentives dominate. In this way this static model is able to reproduce the dynamics of the education gap. The relative abundance of females has been stable throughout the time period, generating a latent incentive for women to educate more. However labor market incentives, captured through the marginal wage gap, were so high that they outbalanced the marriage market incentives and generated more educated males. As the marginal wage gap decreased over time, incentives for males receded and marriage market incentives outweighed the labor market ones. During the final part of the period women faced higher overall incentives and overcame men in schooling attainment.

5 Conclusion and Extensions

The interaction between the fact that men are in short supply in the marriage market and higher returns to education for men in the labor market can theoretically repli-

cate the decrease and reversal of the education gap. Through comparative statics exercises this one-shot game is able to reproduce the described dynamics in the labor and marriage market. However, since dynamics are not modelled, features such as match dissolution cannot be addressed.

Since Chiappori et al. (2006) propose a complementary source for incentives, it would be interesting to perform a structural estimation of the models to determine the empirical relevance of the proposed explanations, including differences in risk aversion, for example for the different states in the U.S. The empirical difficulty of separately identifying the effect of changes in education costs and benefits will remain.

An possible extension is to introduce frictions in the labor and marriage markets. In doing so, the wage differential will be endogenized rather than assumed since the continuation value in the labor market depends on the marriage market prospects, which are better for men than women given their relative scarcity. In addition, by introducing frictions, the model will no longer predict unmatched low-education females when the ratio of females to males increases, but rather we would have unmatched females of all education levels.

6 Appendix: Omitted Proofs

Proposition 3

Proof.

1. Lorentz (1953) proved that SPM s is sufficient to obtain PAM for the continuous

case.

2. In a feasible matching $\mu(x)$ the mass of males equals the mass of females at every education level. Note since the matching is increasing, the feasibility condition (1) implies:

$$G_f(x) = G_m(\mu(x)) \quad \forall x \in [\tilde{x}, \bar{x}]$$

If G_m admits a density then it is strictly monotone and hence invertible. There exists a unique matching:

$$\mu(x) = G_m^{-1}(G_f(x)) \quad \forall x \in [\tilde{x}, \bar{x}]$$

Using the Implicit Function Theorem the derivative is:

$$\mu'(x) = \frac{g_f(x)}{g_m(G_m^{-1}(G_f(x)))} > 0$$

■

Proposition 4

Proof.

1. Expression (2) for a matched couple is: $v_m(\mu(x)) = s(x, \mu(x)) - v_f(x)$. Since s is differentiable then v_m is as well with derivative wrt $\mu(x)$:

$$v'_m(\mu(x)) = s_2(x, \mu(x)) > 0$$

Likewise, given differentiability properties of s , v_x is differentiable with derivative wrt x :

$$v'_f(x) = s_1(x, \mu(x)) + [s_2(x, \mu(x)) - v'_m(\mu(x))] \mu'(x)$$

Substituting the previous result we get $v'_f(x) = s_1(x, \mu(x)) > 0$.

2. Since s is C^2 we can take a derivative of $v'_m(\mu(x))$ wrt $\mu(x)$ from (6) to get:
 $v''_m(\mu(x)) = s_{22}(x, \mu(x)) < 0$. v_m is strictly concave and C^2 .

Now, taking a derivative of $v'_f(x)$ wrt x from (6) we have

$$v''_f(x) = s_{11}(x, \mu(x)) + s_{12}(x, \mu(x)) \mu'(x)$$

v_f has a continuous second derivative, and it is C^2 as long as G_f, G_m admit a density. v_f is strictly concave if $v''_f < 0$ which is equivalent to $-\frac{s_{11}(x, \mu(x))}{s_{12}(x, \mu(x))} > \mu'(x)$. Substituting $\mu'(x) = \frac{\mathbf{g}_f(x)}{\mathbf{g}_m(G_m^{-1}(G_f(x)))}$ the expression becomes $-\frac{s_{11}(x, \mu(x))}{s_{12}(x, \mu(x))} > \frac{\mathbf{g}_f(x)}{\mathbf{g}_m(G_m^{-1}(G_f(x)))} \forall x$.

■

Proposition 5

Proof.

1. To show continuity of χ and φ we apply the Theorem of the Maximum under convexity assumptions (Sundaram, 1999).

It is easy to see from the maximization problem that a higher education cost implies a lower education decision: χ, φ are strictly decreasing. Taking first order condition of (7) for a woman we have $v'_f(\chi(c_f)) = c_f$ which implies: $\chi(c_f) = v'^{-1}_f(c_f)$. For men the expression is: $\varphi(c_m) = v'^{-1}_m(c_m)$.

Given condition (6), these expressions can be stated in terms of the exogenous function s as $\chi(c_f) = s^{-1}_1(c_f, \mu(c_f))$, and $\varphi(c_m) = s^{-1}_2(c_f, \mu(c_f))$. Since s is C^2 then the education choice functions are C^1 with derivatives:

$$\chi'(c_f) = s^{-1}_{11}(c_f, \mu(c_f)), \quad \varphi'(c_m) = s^{-1}_{22}(c_f, \mu(c_f))$$

Since s is strictly concave the education choice functions are strictly decreasing.

2. Applying the Theorem of the Maximum under Convexity assumptions, V_f and V_m are continuous and strictly concave.

■

Proposition 7

Proof.

1. Recall the Planner's problem (11). Given that SMP of the surplus function together with strictly increasing education choice functions imply that the matching in costs is also increasing, and by arguments similar to Proposition 3, condition (9) can be rewritten as $1 - C_f(c_f) = 1 - C_m(\mu^*(c_f))$. The definition

of feasible matching in education costs, and given that C_m admits a density and is thus invertible, imply that there exists a unique matching in costs given by expression (10):

$$\mu^*(c_f) = C_m^{-1}(C_f(c_f)) \quad \forall c_f \in [0, \tilde{c}_f]$$

Recall that c_f, c_m are the cost densities. The matching is strictly increasing and the derivative is given by:

$$\mu^{*'}(c_f) = \frac{c_f'(c_f)}{c_m(\mu^*(c_f))} > 0$$

Given that s is continuous and strictly concave (both properties are preserved under integration) and that the constraint is a linear equality (the constraint set is trivially convex), the FOCs are necessary and sufficient for a global optima. There is a unique solution to the Social Planner's problem $(\mu^*, \chi^*, \varphi^*)$. Moreover, since the conditions for the Theorem of The Maximum under concavity (Sundaram, 1999) are met: the objective function is continuous and strictly concave and the constraint set is compact and continuous, $(\mu^*, \chi^*, \varphi^*)$ are continuous functions. In addition, $S(C_f.C_m)$ is continuous and strictly concave.

2. From part 1 a solution to the SP's problem always exists. To show that an efficient equilibrium always exists, we will show that the FOCs of the SP's and decentralized problems are equivalent.

For the decentralized solution let us focus on the female education decision, given the constraints she faces in terms of payoffs and matching

$$\begin{aligned}
& \underset{x \in [0, \bar{x}]}{Max} \{v_f(x) - c_f x\} \\
& st \ v_f(x) = s(x, \mu(x)) - v_m(\mu(x)) \quad \forall x \\
& \mu(x) = G_m^{-1}(G_f(x)) \quad \forall x
\end{aligned}$$

Note that by Proposition 5 the objective function is continuous and differentiable. To determine the characteristics of the constraint set, combine the two constraints to get one linear constraint. Hence, the constraint set is trivially compact, convex and continuous. Therefore, the problem is well defined.

Substitute the definition of feasible bargaining into the objective function and let λ_x be the multiplier associated to a feasible matching in education levels (5) for each education level. Again, \mathbf{g}_m and \mathbf{g}_f are the densities of the education distributions.

$$\underset{x \in [0, \bar{x}]}{Max} \{s(x, \mu(x)) - v_m(\mu(x)) - c_f x\} \quad st \ G_f(x) = G_m(\mu(x)) \quad \forall x$$

Taking FOC wrt x :

$$s_1(x, \mu(x)) + s_2(x, \mu(x)) \mu'(x) - v'_m(\mu(x)) \mu'(x) - c_f + \lambda_x [\mathbf{g}_f(x) - \mathbf{g}_m(\mu(x)) \mu'(x)] = 0 \tag{12}$$

From (6) we know that $v'_m(\mu(x)) = s_2(x, \mu(x))$. In addition, since the matching needs to be feasible, the constraint binds and the term in brackets drops out. There-

fore, the FOC simplifies to:

$$s_1(x, \mu(x)) = c_f$$

Now let's turn to a point-wise optimization of the Planner's problem (11). Let γ_x be the multiplier associated with the feasible matching in education costs (10). Optimality implies that for the Planner's female education choice the following holds:

$$s_1 - c_f + \gamma_x [\mathbf{c}_f(c_f) - \mathbf{c}_m(\mu^*(c_f)) \mu'] = 0 \quad (13)$$

Again, the constraint is binding and the term in brackets drops out. Therefore (12) and (13) are equivalent: $\chi(c_f) = \chi^*(c_f) \forall c_f$ and individual maximization yields the same education choices as the the Social Planner's. The same analysis can be done for men yielding $\varphi(c_m) = \varphi^*(c_m) \forall c_m$.

From Proposition 3 there exists a unique efficient matching in education levels, μ , determined by G_m and G_f . There is also a unique efficient matching in education costs, μ^* , pinned down by C_m and C_f (Proposition 7 Part 1). Since the strictly monotone functions χ and φ associate education and cost distributions (Proposition 5), μ and μ^* are equivalent and describe the same assignment: the centralized and decentralized matchings coincide. An efficient equilibrium always exists. ■

3.

Proof. Let us derive conditions under which the trivial over-education equilibrium can be sustained: G_f and G_m are degenerate at \bar{x}, \bar{y} , respectively, and all agents get the maximum education. For an agent to get the maximum education, from (7) the

following condition must hold:

$$v'_f(x) > c_f \quad \forall c_f \quad \text{and} \quad v'_m(y) > c_m \quad \forall c_m$$

Therefore, rewriting this expression using equation (6) we get $s_1(x, \mu(x)) > c_f$ and $s_2(x, \mu(x)) > c_m$. For all agents to fully educate, given $c_i \in [0, 1], i = m, f$, we need the following conditions to hold

$$s_i(\bar{x}, \bar{y}) > 1 \quad i = 1, 2$$

The benefits of getting full education need to surpass the costs for every possible education cost, in particular, the highest one. Hence, for some parameter configurations, there are multiple equilibria¹⁵.

Note that the previous condition, given the education distribution of agents, is equivalent to the SP's condition for optimality: in some cases it might be an efficient equilibrium for all the agents to fully educate. However, all other individuals fully educating might be the result of lack of coordination, rather than efficient decision-making.

Note that all agents getting no education can never be an equilibrium, since the relevant condition would be $s_i(0, 0) < 0 \quad i = 1, 2$, and s is strictly increasing by assumption. ■

¹⁵As long as the previous condition holds for every agent, over-education is an equilibrium, regardless of the relative abundance of females, that is, regardless of whether the deviating agents is matched or not since $v_f(\bar{x})$ and $v_m(\bar{y})$ are defined both for married and unmatched agents.

Lemma 8

Proof. Since $\alpha = 1$, $\mu^*(c_f) = c_f$. If $x = y$, $w'_m(y) \geq w'_f(x)$ given strict monotonicity of the wage function is equivalent to $w'^{-1}_m(c_f) \geq w'^{-1}_f(c_f)$. Given symmetric π , asymmetries in the marginal wage directly affect the marginal surplus $\frac{s_2^{-1}(c_f, c_f)}{s_1^{-1}(c_f, c_f)} \geq 1$. Condition (??) implies that the wage gap can be rewritten as $\frac{\varphi(c_f)}{\chi(c_f)} = \frac{s_2^{-1}(c_f, c_f)}{s_1^{-1}(c_f, c_f)} \geq 1$. It follows that marginal labor incentives determine the education gap: if $x = y$, $w'_m(y) \geq w'_f(x)$ then $\frac{\varphi(c_f)}{\chi(c_f)} \geq 1$. ■

Proposition 9

Proof.

1. Since $\alpha = 1$, $\mu^*(c_f) = c_f$. If for all $x = y$ and $w'_m(y) = w'_f(x)$, following a similar argument as in Lemma 8, by this implies that $\frac{\varphi(c_f)}{\chi(c_f)} = 1$: men and women are equally educated.
2. Since $\alpha = 1$, $\mu^*(c_f) = c_f$. If for all education levels $w'_m(y) > w'_f(x)$, by Lemma 8 $\frac{\varphi(c_f)}{\chi(c_f)} > 1$: men are more educated.
3. Since μ^* and μ are equivalent, let us rewrite the education gap by combining the individual's education choice functions (??) and the matching in costs (10): $\frac{\varphi(\mu^*(c_f))}{\chi(c_f)}$. By Proposition 5 φ is decreasing. Thus, if $\alpha > 1$ then $\mu^*(c_f) > c_f$ and $\frac{\varphi(\mu^*(c_f))}{\chi(c_f)} < 1$: women are more educated.

■

Corollary 10

Proof. For given $\bar{\alpha} > 1$, then $\mu^*(c_f) > c_f$. If labor market incentives dominate we have $w'_m(\varphi(\mu^*(c_f))) > w'_f(\chi(c_f))$ and following a similar logic to Lemma 1 males are more educated: $\frac{\varphi(\mu^*(c_f))}{\chi(c_f)} > 1$. For fixed marriage market incentives, as the marginal wage gap decreases and the condition reverses, $w'_m(\varphi(\mu^*(c_f))) < w'_f(\chi(c_f))$, the marriage market incentive dominates and women are more educated: $\frac{\varphi(\mu^*(c_f))}{\chi(c_f)} < 1$.

■

7 Bibliography

Albrecht, J., A. Björklund, and S. Vroman “Is There a Glass Ceiling in Sweden?” *Journal of Labor Economics*, volume 21 (January 2003), pp.145–177.

Albrecht, J. A. Van Vuuren and S. Vroman “Counterfactual Distributions with Sample Selection Adjustments: Econometric Theory and an Application to the Netherlands”, mimeo, Georgetown University, January 2006.

Anderson, A and L. Smith “Assortative Matching and Reputation”, mimeo, University of Michigan, September 2006.

Becker, G. “A Theory of Marriage. Part I” *Journal of Political Economy*, Vol. 81, No. 4. (Jul. - Aug., 1973), pp.813-846.

Blau, F. and L. Kahn, “The US Gender Pay Gap in the 1990s: Slowing Convergence” NBER Working Paper #10853, October 2004.

Chiappori, P., M. Iyigun and Y. Weiss “Investment In Schooling and the Marriage Market” IZA Discussion Paper No. 2454, November 2006.

Crosen, R. and U. Gneezy “Gender Differences in Preferences” mimeo, Harvard

University, September 2004.

Cole, H., G. Mailath and A. Postlewaite “Efficient Non-Contractible Investments in Large Economies” *Journal of Economic Theory*, 101, (December 2001), pp.333-373.

Dougherty, C. “Why is the Return to Schooling Higher for Woman Than For Men?” *Journal of Human Resources*, 40, (Summer 2005) pp.969-988.

Echevarría, C. and A. Merlo “Gender Differences in Education in a Dynamic Household Bargaining Model” *International Economic Review*, 40(2) (May 1999), pp.265-86.

Fernández, R., N. Guner and J. Knowles “Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality”, NBER Working Paper #8580, November 2001.

Ge, S. “Women’s College Choice: How much Does Marriage Matter?” Working Paper, University of Minnesota, January 2006.

Goldin, C., L. Katz and I Kuziemko “The Homecoming of American College Women: A Reversal of the College Gender Gap”, NBER Working Paper #12131, March 2006.

Gretsky, N. E. , J. Ostroy, and W. R. Zame, “The Nonatomic Assignment Model”, *Economic Theory* 1, (March 1992), pp.103-127.

Iyigun M. and R. Walsh “Building the Family Nest: Pre-Marital Investments, Marriage Markets and Spousal Allocations”, *Review of Economic Studies* (forthcoming).

Kremer, M. and E. Maskin “Wage Inequality and Segregation by Skill” *Quarterly*

Journal of Economics, forthcoming.

Lorentz, G. “An Inequality for Rearrangements” *American Mathematical Monthly*, 60 (1953), pp.176-179.

Peters, M. and A. Siow “Competing Pre-Marital Investments” *Journal of Political Economy*, vol. 110(3) (June 2002), pp.592-608.

Ríos-Rull, J. and V. Sánchez Marcos “College Attainment of Women” *Review of Economic Dynamics*, 5, 4 (October 2002) , pp.965-998.

Sundaram, R. K. *A First Course in Optimization Theory*, second printing (Cambridge University Press: Cambridge and New York, 1999).

Topkis, D.M. *Supermodularity and Complementarity*, (Princeton, N.J.:Princeton University Press, 1998).