Yale University

EliScholar – A Digital Platform for Scholarly Publishing at Yale

Cowles Foundation Discussion Papers | Cowles Foundation

3-1-2015

The Marriage Market, Labor Supply and Education Choice

Pierre-André Chiappori

Monica Costa Dias

Costas Meghir

Follow this and additional works at: https://elischolar.library.yale.edu/cowles-discussion-paper-series

 Part of the Economics Commons

Recommended Citation

Chiappori, Pierre-André; Dias, Monica Costa; and Meghir, Costas, "The Marriage Market, Labor Supply and Education Choice" (2015). *Cowles Foundation Discussion Papers*. 2425.
https://elischolar.library.yale.edu/cowles-discussion-paper-series/2425

This Discussion Paper is brought to you for free and open access by the Cowles Foundation at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Cowles Foundation Discussion Papers by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.

THE MARRIAGE MARKET, LABOR SUPPLY AND EDUCATION CHOICE

By

Pierre-Andre Chiappori, Monica Costa Dias, and Costas Meghir

March 2015

COWLES FOUNDATION DISCUSSION PAPER NO. 1994



COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

The Marriage Market, Labor Supply and Education Choice

Pierre-Andre Chiappori*, Monica Costa Dias†and Costas Meghir‡

March 15, 2015

Abstract

We develop an equilibrium lifecycle model of education, marriage and labor supply and consumption in a transferable utility context. Individuals start by choosing their investments in education anticipating returns in the marriage market and the labor market. They then match based on the economic value of marriage and on preferences. Equilibrium in the marriage market determines intrahousehold allocation of resources. Following marriage households (married or single) save, supply labor and consume private and public under uncertainty. Marriage thus has the dual role of providing public goods and offering risk sharing. The model is estimated using the British HPS.

Acknowledgements: This research has greatly benefited from discussions with Jim Heckman, Ana Reynoso and Harald Uhlig as well as participants in a seminar at the University of Chicago and the Becker Friedman Institute conference in honor of Gary Becker. Monica Dias thanks the ESRC for funding. Costas Meghir thanks the Cowles foundation and the ISPS for funding. Pierre André Chiappori thanks the NSF (award 1124277) for funding. The usual disclaimer applies.

*Columbia University
†IFS
‡Yale University, NBER and IFS

1 Introduction

1.1 Matching on human capital

The present paper explores the intersection of two fundamental Beckerian insights: human capital and matching. We are now used to considering education as an investment, whereby agents give up present consumption for higher income and consumption tomorrow. Similarly, we routinely think of marriage in terms of a matching game, in which couples create a surplus that is distributed between spouses, according to some endogenous rule that reflects equilibrium constraints. Still, the interaction between these notions remains largely unexplored. In particular, whether individuals, on the marriage market, can be expected to match assortatively on human capital is largely an open question. For instance, in the presence of domestic production, one may in some cases expect negative assortative matching, a point stressed by Becker himself in his seminal 1973 contribution.

Even if household production is disregarded, the analysis of matching on human capital raises challenging questions. Recent work on the dynamics of wages and labor supply has emphasized the importance of productivity shocks, which typically take a multiplicative form. It follows that higher human capital comes with higher expected wages, but also possibly with more wage volatility. In such a context, whether an educated individual, receiving a large but highly uncertain income, will match with a similar spouse or will trade lower spousal expected income for a lower volatility is not clear. While any individual probably prefers a wealthier spouse, even at the cost of higher volatility, how this preference varies with the individual's own income process - the crucial determinant of assortativeness when intra-couple transfers are allowed, which is our case - is far from obvious.

We believe that the interaction between educational choices and matching patterns is of cru-

cial importance for analyzing the long-run effects of a given policy. When considering the consequences of, say, a tax reform, standard labor supply models, whether unitary or collective, typically take education and family composition as given. While such assumptions make perfect sense from a short-term perspective, they may severely bias our understanding of the reform's long-term outcome. Taxation and welfare programs have a double impact on incentives to invest in human capital. On the one hand, they directly affect the returns from the investment perceived on the labor market. On the other hand, they also influence matching patterns, hence the additional returns reaped on the marriage market - the so-called 'marital college premium', whose importance for human capital investment has been emphasized by several recent contributions (Chiappori, Iyigun and Weiss, 2009, from now on CIW; Chiappori, Salanié and Weiss, 2014, from now on CSW). Added to that is the effect that taxes and welfare have on insurance, which can also affect both marital patterns and investment in human capital. In the long run, these aspects may be of major importance.

The main motivation of the present paper is precisely to provide an explicit framework in which these effects can be conceptually analyzed and empirically quantified. Our model has several, original features. Following a Beckerian tradition, we model marriage as a frictionless, matching game in a Transferable Utility (TU) framework with risk averse agents. Individual utilities have an economic and a non economic component. The economic gain from marriage is twofold: spouses share a public good, and also insure each other against productivity shocks. In addition, marriage provides idiosyncratic, non-monetary benefits, which are additively separable and education-specific, as in Choo and Siow (2006, henceforth CS) and CSW. The TU property implies that, once married, households behave as a single decision-maker (unitary household). Despite its obvious shortcomings, this property considerably simplifies the analysis of the couple's dynamics of consumption and labor supply.

We consider a three-stage model, and assume Pareto efficiency and full commitment. We abstract from issues relating to divorce and our full commitment assumption precludes renegotiation; these are important questions we wish to address as this research agenda develops. Agents first independently invest in human capital; their decision is driven by their idiosyncratic ability, their idiosyncratic cost of investment (which may for instance reflect borrowing constraints), and the expected returns on investment - which is itself determined by the equilibrium prevailing on the relevant markets. In the second stage, individuals match on the marriage market, based on their human capital and their idiosyncratic preferences for marriage. Finally, the last period is divided into T subperiods, during which couples or singles consume private and public goods, save and supply labor subject to permanent and transitory wage shocks, very much like standard lifecycle models.

As is usual, such a game can be solved backwards, starting with the third stage. Due to the TU assumption, the analysis of the dynamic labor supply model exactly characterizes the total surplus generated by marriage, while it is compatible with any intra-couple distribution of surplus. The matching game in the second stage is defined by the distribution of human capital among men and women, as resulting from investment during the first stage, and the expected surplus generated in the third stage. Crucially, equilibrium conditions on the marriage market fully determine the intra-household allocation of the surplus for all possible levels of human capital. In particular, these conditions allow the characterization for each individual of the consequences, in expected utility terms, of the various levels of human capital they may choose to acquire. This 'education premium', in turn, determines education decisions in the first stage. In essence, therefore, investments in the first stage are modeled under a rational expectations logic: agents anticipate a given vector of returns to education, and the resulting decisions lead to an equilibrium in the marriage market which is compatible with these expectations.

In this context, the impact of any given policy reform can be considered along several dimensions. Coming back to the tax reform example, the short term impact can be analyzed from the dynamic labor supply model of the third stage: existing couples (and singles) respond to changes in income tax by adjusting their labor supply and their public and private consumptions. From a longer-term perspective, however, matching on the marriage market will also be affected; typically, the respective importance of economic and non economic factors will vary, resulting in changes in the level of assortativeness on human capital, therefore ultimately in inter- and intra-household inequality. Finally, the changes affect the returns on investment in human capital both directly (through their impact on after tax income) and indirectly (by their consequences on the marriage market); they can therefore be expected to propagate to human capital investments. Imperfect as it may be, our approach is the first to consider all these aspects in a unified and theoretically consistent framework.

1.2 Existing literature

Our paper is a direct extension of the collective models of Chiappori (1988,1992) and Blundell, Chiappori and Meghir (2005) amongst others. In these models the intra-household allocations are Pareto efficient and there is no time/dynamic dimension. Both these restrictions are relaxed here. Thus the framework we use is directly related to intertemporal models of labor supply and savings over the life-cycle, such as Mazzocco (2007), who uses a collective framework, and Attanasio, Low and Sanchez-Marcos (2008) and Low, Meghir and Pistaferri (2010) who focus respectively on female and male labor supply. Similarly in a recent paper Blundell et al. (2015) consider female labor supply over the lifecycle in a context where household composition is changing over the lifecycle but exogenously. More closely to this paper Low et al. (2015) allow for endogenous marriage decisions with limited commitment in a partial equilibrium context

with frictions but treating education as exogenous. Jacquemet and Robin (2011) and Goussé (2013) specify an equilibrium model of marriage with frictions and labor supply with frictions. Their model draws from Shimer and Smith (2000) and the complementarity arises from the production of public goods that depends on the wages of both spouses. Their model does not include savings and the only source of uncertainty is exogenous divorce. Moreover it does not allow for endogenous education choices. Finally, precursors of this paper are CIW (2008), which specifies a theoretical model of education decisions, the marriage market and time at home, and CSW, which provides an empirical estimation; however, both papers adopt a reduced form specification in which marital gains are recovered from matching patterns without analyzing actual behavior.

Our model is also related to recent developments on matching models under transferable utility (see Chiappori and Salanié 2015 for a recent survey). In particular, the stochastic structure representing idiosyncratic preferences for marriage is directly borrowed from CS and CSW. Our framework, however, introduces several innovations. First, agents match on human capital - unlike CS, where they match on age, and CSW, where they match on education. Human capital, in our framework, depends on education but also on innate ability. In principle, the latter is not observed by the econometrician. However, observing agents' wage and labor supply dynamics (during the third stage) allows us to recover the joint distribution of education and ability, therefore of human capital. A second difference is that both CS and CSW identify the structural model under consideration from the sole observation of matching patterns. As a result, CS is exactly identified under strong, parametric assumptions, whereas identification in CSW comes from the observation of multiple cohorts together with parametric restrictions on how surplus may change across cohorts. In our case, on the contrary, our structural model of household labor supply allows to identify preferences, therefore the surplus function. The matching model, therefore, is over identified, and allows to recover the intra-couple alloca-

tion of surplus while generating additional, testable restrictions. Lastly, this identification, together with the knowledge of the joint distribution of ability and education, enable us to explicitly model the process of educational choice. As a consequence, we can evaluate the long term impact of a given policy reform on human capital formation. While the link between intra-household allocation and investment in human capital has already been analyzed from a theoretical perspective,[1] our approach is, to the best of our knowledge, the first to explore it empirically through a full-fledged structural model.

2 The model

2.1 Time structure

We model the life-cycle of a cohort of women $f \in \mathcal{F}$ and men $m \in \mathcal{M}$, so time and age will be used interchangeably and commonly represented by t. The individual's life cycle is split into three stages, indexed 1 to 3. In stage 1, individuals invest in human capital by choosing an education level; this investment depends on their innate ability and their cost of education, as well as on the perceived benefits of this investment. The ability of agent i, denoted θ_i, belongs to a finite set of classes, $\Theta = \left(\theta^1, ..., \theta^N\right)$. Education costs are continuously distributed, and the agent can choose between a finite number of education levels, $\mathcal{S} = \left(S^1, ..., S^J\right)$. At the end of period 1, each agent is thus characterized by human capital (or productivity type) $H(s, \theta)$, which is a summary measure of education and innate ability. The distribution of human capital has a finite support $\mathcal{H}$ of cardinality (at most) $J \times N$. So at this stage the agent belongs to a finite set of classes $\mathcal{H} = \left(H^1, ..., H^{J \times N}\right)$ that fully characterise his/her prospects in the marriage and labour markets.

[1] See for instance CIW.

In stage 2, individuals draw a vector of marital preferences and enter the marriage market; the latter is modelled as a frictionless matching process based on one observable characteristic, the level of human capital, and on unobservable marital preferences. At the end of stage 2, some individuals are married whereas the others remain single forever.

Stage 3 (the 'working life' stage) is divided into T periods; in each period, individuals, whether single or married, observe their (potential) wage and non labor income, and decide on consumptions and labor supplies. Credit markets are assumed complete, so that agents can, during their active life, borrow or save at the same interest rate. Following a collective logic (Chiappori 1988, 1992), decisions made by married couples are assumed Pareto-efficient. Moreover, the intra-household allocation of private consumption (therefore of welfare) is endogenous, and determined by commitments made at the matching stage. In particular, we do not consider divorce or separation in this model.

2.2 Economic utilities

The lifetime utility of agent i is the sum of three components. The first is the expected, discounted sum of economic utilities generated during the periods t of i's third stage of life by consumptions and labour supply; the second other is the subjective utility of marriage (or singlehood) generated by the agents' marital preferences; and the third is the utility cost of education attendance. In what follows, we consider the following economic utilities at date t of stage 3:

$$u_{it}\left(Q_t, C_{it}, L_{it}\right) = \ln\left(C_{it}Q_t + \alpha_{it}L_{it}Q_t\right) \tag{1}$$

where L is time off paid work and C and Q are private and public consumptions, respectively. We take labor supply choices to be discrete: agents choose either to participate to the labor market ($L = 0$) or not to ($L = 1$).

The choices of consumptions, labour supply and savings are driven by time-varying preferences and income. First, wages at age t are determined by the person's age and human capital, itself a function of education s_i and ability θ_i, and also by an idiosyncratic productivity shock that may have a transitory and a permanent component. Formally:

$$w_{it} = W_G\left(H_i, t\right) e_{it} \tag{2}$$

where w_{it} denotes i's earnings at age t, $G = M, F$ indexes i's gender group, W_G is the aggregate, gender-specific price of human capital class H_i at age t, $H_i = H\left(s_i, \theta_i\right)$ is i's human capital, and e_{it} is an idiosyncratic shock. Second, preferences may vary; in practice, the α_{it} are random variables.

Two remarks can be made on these utilities. From an ordinal viewpoint, they belong to Bergstrom and Cornes' Generalized Quasi Linear (GQL) family. As a consequence, at any period and for any realization of family income, they satisfy the Transferable Utility (TU) property. For a given couple (m, f), any conditional (on employment and savings) Pareto efficient choice of consumption and public goods maximizes the sum of the spouse's exponential of utilities:[2]

$$\exp u_i\left(Q_t, C_{it}, L_{it}\right) + \exp u_j\left(Q_t, C_{jt}, L_{jt}\right) = \left(C_{it} + C_{jt} + \alpha_{it} L_{it} + \alpha_{jt} L_{jt}\right) Q_t \tag{3}$$

[2]In the static model, one can use $\exp u_i$ as a particular cardinalization of i's preferences. Then any (ex post) efficient allocation maximize some weighted sum of utilities of the form $\exp u_i\left(Q_t, C_{it}, L_{it}\right) + \mu \exp u_j\left(Q_t, C_{jt}, L_{jt}\right) \geq \bar{u}_j$ under a budget constraint. Here, the maximand is equal to

$$\left(C_{it} + \mu C_{jt} + \alpha_{it} L_{it} + \mu \alpha_{jt} L_{jt}\right) Q_t$$

and the first order conditions with respect to private consumptions (assuming the latter are positive) give:

$$Q_t = \lambda_t = \mu Q_t$$

where λ_t is the Lagrange multiplier of the budget constraint. It follows that $\mu = 1$, implying that any Pareto efficient solution with positive private consumptions must maximize the sum of $\exp u_i$.

Solving this program gives the optimal choice of private and public consumptions at each period, conditional on labor supplies and savings. The latter are then determined from a dynamic perspective, by maximizing the expected value of the discounted sum (over periods t to the end of life) of utilities.

The second remark adopts a cardinal viewpoint. The Von Neumann-Morgenstern utilities defined by (1) belong to the ISHARA class, defined by Mazzocco (2007). By a result due to Schulhofer-Wohl (2006), this implies that the TU property also obtains *ex-ante*, in expectations. In particular, there exists a specific cardinalization of each agent's lifetime economic utility such that any household maximizes the sum of lifetime utilities of its members, under an intertemporal household budget constraint. Specifically, we show below the following result. Take a couple (m, f) with respective human capital H_m and H_f, and let V_m, V_f denote their respective, lifetime expected utility. Then there exists a function $\Upsilon(H_m, H_f)$ such that the set of Pareto efficient allocations is characterized by:

$$\exp\left\{\frac{1-\delta}{1-\delta^T}V_m\right\} + \exp\left\{\frac{1-\delta}{1-\delta^T}V_f\right\} = \exp\left\{\frac{1-\delta}{1-\delta^T}\Upsilon(H)\right\}$$

The crucial remark, now, is that the expression

$$\bar{U}_i = \exp\left(\frac{1-\delta}{1-\delta^T}V_i\right) \tag{4}$$

is an increasing function of V_i; therefore, in the stage 2 matching game, it is a specific (and convenient) representation of i's utility. If we define

$$g(H_m, H_f) = \exp\left\{\frac{1-\delta}{1-\delta^T}\Upsilon(H)\right\}$$

the previous relationship becomes:

$$\bar{U}_m + \bar{U}_f = g\left(H_m, H_f\right)$$

which shows that we are in a TU context even *ex-ante*, since the Pareto frontier is an unweighted sum of these utility indices. The function $g\left(H_m, H_f\right)$, when evaluated at the point of marriage, is the economic value generated by marriage. An important consequence is that, throughout the third stage (their working life), couples behave as a single decision maker maximizing the function g (or equivalently Υ). In particular, a standard, unitary model of dynamic labor supply can be used at that stage.

Alternatively, agents may choose to remain single; then they maximize the discounted sum of expected utility under an intertemporal, individual budget constraint. We denote $V^S\left(H_m\right)$ and $V^S\left(H_f\right)$ the respective lifetime economic utility of a single male (female) with human capital H_m $\left(H_f\right)$. Note these expressions, again, are expectations taken over future realizations of the preferences and wages shocks; they are contingent on the information known at the date of marriage, namely each person's ability and education, as summarized by the person's human capital. In line with the previous notations, we then define:

$$\bar{U}_i^S = \exp\left\{\frac{1-\delta}{1-\delta^T} V^S\left(H_i\right)\right\} \tag{5}$$

Finally, for any man m with human capital H_m and any woman f with human capital H_f, the difference between the economic value that would be generated by their marriage, $g\left(H_m, H_f\right)$, and the sum of m's and f's respective expected utility as singles is the *economic* surplus generated by the marriage. Again, it depends only on both spouses' productivity and education,

and is denoted

$$\Sigma\left(H_m, H_f\right) = g\left(H_m, H_f\right) - \bar{U}_m^S - \bar{U}_f^S. \tag{6}$$

Note that all these expressions refer to the same cardinalization of lifetime expected utilities, given by (4).

2.3 Marital preferences

Our representation of marital preferences follow that of CS and CSW. Before entering the marriage market, agent i draws a vector $\beta_i = (\beta_i^0, \beta_i^H, H \in \mathcal{H})$, where β_i^H represents i's subjective satisfaction of being married to a spouse with human capital H and β_i^0 denotes his/her subjective satisfaction of remaining single. We assume that the total gain generated by the marriage of man m with human capital H_m and woman f with human capital H_f is the sum of the economic gain $g\left(H_m, H_f\right)$ defined above and the idiosyncratic preference shocks β:

$$g_{mf} = g\left(H_m, H_f\right) + \beta_m^{Hf} + \beta_f^{H_m} \tag{7}$$

and the resulting surplus is:

$$\Sigma_{mf} = \Sigma\left(H_m, H_f\right) + \left(\beta_m^{Hf} - \beta_m^0\right) + \left(\beta_f^{Hm} - \beta_f^0\right) \tag{8}$$

Again, the function $\Sigma\left(H_m, H_f\right)$ is defined as the expected *economic* lifetime surplus for a couple with human capital composition (H_m, H_f), over and above what they would each obtain as singles. The remaining part of the expression relates to the non-economic benefits of marriage.[3]

[3]It should be stressed that our interpretation of β_i^H as i's subjective utility of being married to a spouse with human capital H is by no means the only possible. Alternatively, β_i^H could be some unobserved characteristic of i that is identically valued by all spouses with human capital H. The crucial property is that this term enhances total surplus in a way that does not depend on the spouse's identity, but only on her/his human

Importantly, it is a restriction of this model that the idiosyncratic preferences of m, as described by the random vector β_m, only depend on the education of m's spouse, not on her identity. In other words, non-pecuniary preferences are over people with different levels of human capital, not over specific persons. This assumption is crucial, because it allows to fully characterize the stochastic distribution of individual utilities at the stable match (see CSW and Chiappori and Salanié 2015).[4]

2.4 Second stage matching game

At the end of the first stage, agents are each characterised by their human capital H, a function of their innate ability θ and education s. The male and female populations are therefore distributed over the space $\mathcal{H}$, which consists of $N \times J$ classes. Moreover, agents draw their marital preferences at the beginning of stage 2, which we assume independent of their human capital. They then enter a matching game under TU, in which the surplus function for any potential match is given by (8). As usual, a matching is defined by a measure on the product space of male and female characteristics (i.e., $\mathcal{H} \times \mathcal{H}$) and two sets of individual utility levels, (U_m) and (U_f), such that for any pair (m, f) *on the support of the measure* - that is, for any couple that matches with positive probability:

$$U_m + U_f = g_{mf}$$

Intuitively, the pair (U_m, U_f) describes how the total gain g_{mf} generated by the possible marriage of m and f would be divided between the spouses.

capital.

[4]Moroever, the introduction, in the marital gain generated by the couple (m, f), of match-specific terms of the form ε_{mf} would raise specific difficulties in our frictionless framework. For instance, if the ε are assumed i.i.d., then when the number of individuals becomes large the fraction of singles goes to zero (and their conditional utility tends to infinity). See Chiappori, Nguyen and Salanié 2015 for a precise discussion.

The matching is stable if (i) no married person would rather be single, and (ii) no two individuals would strictly prefer being married to each other to remaining in their current situation. A direct consequence is that for *any* pair (m, f), it must be the case that:[5]

$$U_m + U_f \geq g_{mf}$$

Now, a crucial result by Chiappori, Salanié and Weiss is the following:

Theorem 1. ***(Chiappori, Salanié and Weiss 2015)*** *If the surplus is given by (8), then there exist* $2(NJ)^2$ *numbers -* $\bar{U}_M(H_m, H_f)$ *and* $\bar{U}_F(H_m, H_f)$ *for* $(H_m, H_f) \in \mathcal{H}^2$ *- such that:*

1. *For any* (H_m, H_f)

$$\bar{U}_M(H_m, H_f) + \bar{U}_F(H_m, H_f) = g(H_m, H_f) \tag{9}$$

2. *For any* m *with human capital* H_m *married to* f *with human capital* H_f*,*

$$\begin{aligned} U_m &= \bar{U}_M(H_m, H_f) + \beta_m^{H_f} \text{ and} \\ U_f &= \bar{U}_F(H_m, H_f) + \beta_f^{H_m} \end{aligned} \tag{10}$$

Proof. See Chiappori, Salanié and Weiss (2015). □

In words, the utility of any man m at the stable matching is the sum of a deterministic component, which only depends on his and his spouse's human capital, and of m's idiosyncratic net preference for marrying a spouse with that human capital; the same type of result obtains for women. For notational consistency, if i remains single we consider the class of his spouse

[5] If this inequality was violated for some couple (m, f), one could conclude that m and f are not matched (then an equality would obtain) but should be matched (since each of them could be made better off than their current situation), a violation of stability.

to be 0, and we define

$$\bar{U}_M(H,0) = \bar{U}_F(0,H) = 0 \text{ for all } H$$

Note that the characterization of utilities provided by (10) refers to a specific cardinalization of individual utilities, defined by (U_m, U_f); technically, this is the particular cardinalization that exhibits the TU property. Obviously, it can equivalently be translated into the initial cardinalization; in that case, the total, expected utility of person i is:

$$V_i = \frac{1-\delta^T}{1-\delta}\ln(U_i) = \frac{1-\delta^T}{1-\delta}\ln\left(\bar{U}_{G(i)}(H_m, H_f) + \beta_i^{H_j}\right) \tag{11}$$

where $G(i)$ is the gender of i and H_j denotes the human capital of i's spouse.

An immediate corollary is the following:

Corollary 1. *1. For any man m with human capital H_m, m's spouse at the stable matching has human capital H_f if and only if the following inequalities hold for all $H \in \mathcal{H} \cup \{0\}$:*

$$\bar{U}_M(H_m, H_f) + \beta_m^{H_f} \geq \bar{U}_M(H_m, H) + \beta_m^H$$

Similarly, m is single if and only if:

$$\bar{U}_M(H_m, H_f) + \beta_m^{H_f} \leq \beta_m^0 \quad \text{for all } H_f \in \mathcal{H}$$

2. For any woman f with human capital H_f, f's spouse at the stable matching has human capital H_m if and only if the following inequalities hold for all $H \in \mathcal{H} \cup \{0\}$:

$$\bar{U}_F(H_m, H_f) + \beta_f^{H_m} \geq \bar{U}_F(H, H_f) + \beta_f^H$$

Similarly, m is single if and only if:

$$\bar{U}_F\left(H_m, H_f\right) + \beta_f^{H_m} \leq \beta_f^0 \quad \textit{for all } H_m \in \mathcal{H}$$

3. *The ex-ante expected utility of a man m with human capital H_m is:*

$$A_M(H_m) = \mathbb{E}\left[\max_{H_f \in \mathcal{H} \cup \{0\}} \left(\bar{U}_M\left(H_m, H_f\right) + \beta_m^{H_f}\right)\right] \tag{12}$$

and the ex-ante *expected utility of a female agent f with human capital H_f is:*

$$A_F(H_f) = \mathbb{E}\left[\max_{H_m \in \mathcal{H} \cup \{0\}} \left(\bar{U}_F\left(H_m, H_f\right) + \beta_f^{H_m}\right)\right] \tag{13}$$

where the expectation is over the realization of unobserved preferences for spouse's types, β_m and β_f for men and women respectively.

The main implication of this result is that marital choices in stage 2 can be modeled as individual, discrete choice problems, in which the thresholds $\bar{U}_M\left(H_m, H_f\right)$ and $\bar{U}_F\left(H_m, H_f\right)$ can be identified using standard techniques. Note, however, that these parameters are not independent, since they have to satisfy the restrictions (9); we will return to this point later on. Also, note that these ex-ante expected utilities only depend on the individual's stock of human capital.

2.5 First stage: the education choice

In the first stage of life, individuals decide upon the level of educational investment. We assume there are three choices, corresponding to three classes in $\mathcal{S}$: statutory schooling, high school and college. Each level of education s is associated with a cost $c_s(X, \upsilon_s)$ where X are

observable characteristics and υ_s is an unobservable cost.

Defining human capital as a function of schooling and ability $H(s, \theta)$, education choice is defined by

$$\text{for man } m \quad : \quad s_m = \arg\max_{s \in \mathcal{S}} \{A_M(H(s, \theta_m)) - c_s(X_m, \upsilon_{sm})\} \tag{14}$$

$$\text{for woman } f \quad : \quad s_f = \arg\max_{s \in \mathcal{S}} \{A_F(H(s, \theta_f)) - c_s(X_f, \upsilon_{sf})\} \tag{15}$$

where $\mathbb{E}U$ and $\mathbb{E}V$ are defined in equations 12 and 13 for males and females, respectively, and where the subscript s indexes schooling level s. Individuals are assumed to know their ability at that point, but this may not be observable by the econometrician. Education choice takes into account both the returns in the labor market and the returns in the marriage market, which are embedded in the value functions for each choice.

3 Solving the model

It is instructive to outline the solution of the problem. As is standard in dynamic models of the lifecycle, the model is solved working backwards from the end of life. We therefore start with the last period of the third stage. As mentioned before, the TU property implies that any married couple behaves as a single decision maker maximizing the sum of the spouses' (exponential of) utilities: the Pareto weights associated with our original logarithmic cardinalization of utilities, which determine the intrahousehold allocation of welfare, do not affect aggregate household consumption, savings and individual labor supply decisions. Singles maximize their own utility. Both maximizations are subject to an intertemporal budget constraint.

3.1 Employment, consumption and savings during the working life

We start with the labor supply and consumption decisions. The form of preferences allows to easily derive consumptions from savings and labor supply choices; savings are then chosen to satisfy the conditional (on labor supply) intertemporal optimality condition; optimal labor supply is then the solution to a discrete choice problem.

3.1.1 General solution to the couple's problem in period t

In Appendix A we derive the solution to the last period of life, T. Many of the properties of that last period, such as the separability of the Pareto weights in the individual value function, carry over to the general solution for any of the earlier periods. Here we show the form of the solution for an earlier period, $t < T$.

Consumptions Each period t sees the arrival of new information on each spouse's preferences for working and productivity, $\alpha_t = (\alpha_{mt}, \alpha_{ft})$ and $e_t = (e_{mt}, e_{ft})$. Choice is also conditional on the other circumstances faced by the couple, namely savings carried over from the previous period, K_{t-1}, and the spouses' human capital, $H = (H_m, H_f)$. Given the information set $(\alpha_t, e_t, K_{t-1}, H)$, we first consider the couple's consumption decisions conditional on savings and employment, K_t and $L_t = (L_{mt}, L_{ft})$. For the within period problem of resource allocation to private consumption (C) and public good (Q), we can use the exponential cardinalization of individual preferences. The couple thus solves:

$$\max_{Q_t, C_t} \quad Q_t \left(C_t + \alpha_{mt} L_{mt} + \alpha_{ft} L_{ft} \right)$$

$$\text{under the budget constraint} \quad w_{mt} + w_{ft} + y_t^C + RK_{t-1} = K_t + C_t + w_{mt} L_{mt} + w_{ft} L_{ft} + pQ_t$$

Here $w_{mt} + w_{ft}$ is the couple's total ('potential') labor income in period t, and y_t^C is the couple's non labor income. Note that the latter may depend on individual labor supplies and earnings, which allows for means tested benefits and taxes as well as benefits that depend on participation, such as unemployment insurance or earned income tax credits. Wages are as defined in equation (2) and considered net of income taxes. Finally, R is the risk-free interest rate at which savings accumulate over periods, $C_t = C_{mt} + C_{ft}$ is total expenditure in the private consumption of spouses, and pQ_t is total expenditure in the public good.

Conditional on savings and labour supply, the solutions for public and private consumptions are

$$\begin{aligned} Q_t(K_t, L_t) &= \frac{y_t^C + RK_{t-1} - K_t + w_{mt}(1 - L_{mt}) + w_{ft}(1 - L_{ft}) + (\alpha_{mt}L_{mt} + \alpha_{ft}L_{ft})}{2p} \\ C_t(K_t, L_t) &= y_t^C + RK_{t-1} - K_t + w_{mt}(1 - L_{mt}) + w_{ft}(1 - L_{ft}) - pQ_t(K_t, L_t) \\ &= pQ_t(K_t, L_t) - (\alpha_{mt}L_{mt} + \alpha_{ft}L_{ft}) \end{aligned}$$

where consumptions are written as functions of (K_t, L_t) to highlight the fact that they are conditional solutions.

Efficient risk sharing conditional on savings and employment We now consider the intra-household allocation of resources during period t from an *ex-ante* perspective - that is, before the realization of the shocks. Here, efficiency is relative to sharing the (wages and preferences) risks. In this context, it requires the maximization of a weighted sum of expected utilities, obviously using the *initial*, logarithmic cardinalization. If μ denotes the wife's Pareto weight corresponding to that cardinalization, the standard efficiency condition imposes that the ratio of marginal utilities of private consumption be constant (and equal to the Pareto

weight) for all periods and all realizations of the random shocks:

$$\frac{\partial u_{mt}\left(Q_t, C_{mt}, L_{mt}\right)}{\partial C_{mt}} = \mu \frac{\partial u_{ft}\left(Q_t, C_{ft}, L_{ft}\right)}{\partial C_{ft}}$$

Note that the Pareto weight μ is a price endogenously determined in the marriage market. Thus, it only depends on the information available then, namely the human capital of both spouses (H_m, H_f). Moreover, it remains constant over the couple's working life - a direct implication of efficiency under full commitment. Efficient risk sharing then yields private consumptions as follows:

$$\begin{aligned} C_{mt} &= \frac{1}{1+\mu} pQ_t - \alpha_{mt} L_{mt} \\ C_{ft} &= \frac{\mu}{1+\mu} pQ_t - \alpha_{ft} L_{ft}. \end{aligned}$$

Therefore, the conditional (on employment and savings) instantaneous indirect utilities are

$$v_{mt} = 2 \ln Q_t\left(K_t, L_t\right) + \ln p + \ln \frac{1}{1+\mu} \tag{16}$$

$$v_{ft} = 2 \ln Q_t\left(K_t, L_t\right) + \ln p + \ln \frac{\mu}{1+\mu} \tag{17}$$

Note that Q_t is also a function of the entire state space, including the wage and preference shocks, savings and human capital, $(e_t, \alpha_t, K_{T-1}, H)$. We therefore write $v_{it}\left(K_t, L_t; e_t, \alpha_t, K_{t-1}, H, \mu\right)$.

Expected value functions Appendix A shows that, for period T:

$$\begin{aligned}
E_{T|T-1}V_{mT}\left(e_T,\alpha_T,K_{T-1},H,\mu\right) &= I_T\left(e_{T-1},\alpha_{T-1},K_{T-1},H\right)+\ln\frac{1}{1+\mu}\\
E_{T|T-1}V_{fT}\left(e_T,\alpha_T,K_{T-1},H,\mu\right) &= I_T\left(e_{T-1},\alpha_{T-1},K_{T-1},H\right)+\ln\frac{\mu}{1+\mu}\\
\text{where } I_T\left(e_{T-1},\alpha_{T-1},K_{T-1},H\right) &= E_{T|T-1}\max_{L_T}\left[\max_{K_T}\left\{2\ln Q_T\left(L_T,K_T\right)+\ln p\right\}\mid e_{T-1},\alpha_{T-1}\right]
\end{aligned}$$

where expectations are taken over the (education-specific) distribution of (e_t,α_t) conditional on their realization at $t-1$. Note that here $K_T=0$ since bequests are not being considered.

Given the conditional instantaneous indirect utilities in (16)-(17), it is easy to show by recursion that the additive separability of the Pareto weight carries over to earlier periods:

$$\begin{aligned}
E_{t|t-1}V_{mt}\left(e_t,\alpha_t,K_{t-1},H,\mu\right) &= I_t\left(e_{t-1},\alpha_{t-1},K_{t-1},H\right)+\ln\left(\frac{1}{1+\mu}\right)\sum_{\tau=t}^{T}\delta^{\tau-t}\\
E_{t|t-1}V_{ft}\left(e_t,\alpha_t,K_{t-1},H,\mu\right) &= I_t\left(e_{t-1},\alpha_{t-1},K_{t-1},H\right)+\ln\left(\frac{\mu}{1+\mu}\right)\sum_{\tau=t}^{T}\delta^{\tau-t}
\end{aligned}$$

where δ is the discount factor. The common term in the individual value functions, I_t, is defined recursively by

$$I_t\left(e_{t-1},\alpha_{t-1},K_{t-1},H\right)=\max_{L_t}E_{t|t-1}\left[\max_{K_t}\left\{2\ln Q_t\left(L_t,K_t\right)+\ln p+\delta I\left(e_t,\alpha_t,K_t,H\right)\right\}\mid e_{t-1},\alpha_{t-1}\right]$$

where expectations are taken over the (education-specific) distribution of (e_t,α_t) conditional on (e_{t-1},α_{t-1}). A crucial feature of the above expressions is that the Pareto weight μ affects individual welfare but drops out of the aggregate value function I, reflecting the TU property. This then implies that the intertemporal optimality condition for savings (Euler equation) is the same for both spouses. For any choice of labor supplies (including the optimal one),

conditional optimal savings ($K_t^*(L_t)$) satisfy:

$$2\frac{\partial \ln Q_t\left(K_t, L_t\right)}{\partial K_t} + \delta \frac{\partial I_{t+1}\left(e_t, \alpha_t, K_t, H\right)}{\partial K_t} = 0$$

Finally, the optimal choice of labor supplies are defined by

$$\left(L_{mt}^*, L_{ft}^*\right) = \underset{L_t \in \{0,1\}^2}{\arg\max} \left\{2 \ln Q_{it}\left(K_t^*(L_t), L_t\right) + \ln p + \delta I_{t+1}\left(e_t, \alpha_t, K_t^*(L_t), H\right)\right\}.$$

The single's problem is a close replica of the couple's problem, just simpler, and its solution can be derived using the same approach as briefly discussed in Appendix B.

3.1.2 The first period after marriage

The Markov processes for (e_t, α_t) start at date $t = 1$, and initial savings are set to zero. So the functions I_1 and I_1^S do not depend on past values of the shock or on past investment, but only on human capital; we denote them respectively by $\Upsilon\left(H\right)$ and $\Upsilon^S\left(H_i\right)$. It follows that the expected *economic* utility, at marriage, of each spouse is given by:

$$V_m\left(H, \mu\right) = \Upsilon\left(H\right) + \left(\sum_{\tau=0}^{T-t} \delta^\tau\right) \ln\left(\frac{1}{1+\mu}\right) \tag{18}$$

$$\text{and } V_f\left(H, \mu\right) = \Upsilon\left(H\right) + \left(\sum_{\tau=0}^{T-t} \delta^\tau\right) \ln\left(\frac{\mu}{1+\mu}\right) \tag{19}$$

which depends on the spouses' respective levels of human capital and on the Pareto weight μ that results from the matching game in the earlier lifecycle stage 2. For singles, expected lifetime utility is simply:

$$V^S\left(H_i\right) = \Upsilon^S\left(H_i\right)$$

3.2 Matching

We now move to the second stage, i.e. the matching game. Remember that marriage decisions are made before preferences and productivity shocks (α, e) are realized, and that we assume full commitment. We first compute the expected utility of each spouse, conditional on the Pareto weight μ. We then show that the model can be reinterpreted as a matching model under TU; finally, we compute the equilibrium match and the corresponding Pareto weights.

3.2.1 Formal derivation

Consider a match between man with human capital H_m and woman with human capital H_f. The spouses' expected, economic lifetime utilities are given by (18)-(19). However, an alternative cardinalization, already introduced in (4), turns out to be more convenient here. Specifically, define $\bar{U}_i$ by:

$$\bar{U}_i = \exp\left(\frac{1-\delta}{1-\delta^T} V_i\right) \tag{20}$$

then if $H = (H_m, H_f)$:

$$\bar{U}_m \exp\left\{-\Upsilon(H)\frac{1-\delta}{1-\delta^T}\right\} = \frac{1}{1+\mu}, \bar{U}_f \exp\left\{-\Upsilon(H)\frac{1-\delta}{1-\delta^T}\right\} = \frac{\mu}{1+\mu}$$

and finally:

$$\bar{U}_m + \bar{U}_f = \exp\left\{\frac{1-\delta}{1-\delta^T}\Upsilon(H)\right\} = g(H)$$

which expresses that the sum of individual, *economic* utilities add up to the marital gain $g(H)$. Lastly, we can add the idiosyncratic shocks to both sides of this equation; we finally have that,

for any married couple (H_m, H_f):

$$\bar{U}_m + \beta_m^{H_f} + \bar{U}_f + \beta_f^{H_m} = g(H) + \beta_m^{Hf} + \beta_f^{H_m} = g_{mf}$$

The matching game, therefore, has a *transferable utility* structure: if the utility of person i is represented by the particular cardinal representation $\left(\bar{U}_m, \bar{U}_f\right)$, then the Pareto frontier is a straight line with slope -1.

In particular, whether matching will be assortative on human capital or not, depends on the supermodularity of function g. One can easily check that the sign of the second derivative $\partial^2 g / \partial H_m \partial H_f$ is indeterminate (and can be either positive or negative depending on the parameters); so this needs to be investigated empirically.[6]

Clearly, one can equivalently use any of the two cardinalizations described before; remember, though, that the Pareto weight μ refers to the initial cardinalization (V_m, V_f). This Pareto weight μ is match-specific; as such, it might in principle depend on the spouses' stocks of human capital, but also on their marital preferences. However, the following result, which is a direct corollary of Theorem 1, states that this cannot be the case:

Corollary 2. *At the stable match, consider two couples* (m, f) *and* (m', f') *such that* $H_m = H_{m'}$ *and* $H_f = H_{f'}$. *Then the Pareto weight is the same in both couples*

[6]Since matching involves a random component, if the population is large enough all combinations of human capital will match with positive probability. Still, a result due to Graham ($$$) states that, for the stochastic structure under consideration here, for any two levels H and $\bar{H}$ of human capital, the total number of 'assortative couples' (i.e., $H - H$ or $\bar{H} - \bar{H}$) will exceed what would be expected under purely random matching if and only if the deterministic function g is supermodular for H and $\bar{H}$ - i.e.:

$$g(H, H) + g\left(\bar{H}, \bar{H}\right) \geq g\left(\bar{H}, H\right) + g\left(H, \bar{H}\right)$$

That is what is meant by assortative matching.

Proof. From (10) in Theorem, we have that:

$$\begin{aligned} U_m &= \bar{U}_m + \beta_m^{H_f} = \bar{U}_M\left(H_m, H_f\right) + \beta_m^{H_f} \text{ and} \\ U_f &= \bar{U}_f + \beta_f^{H_m} = \bar{U}_F\left(H_m, H_f\right) + \beta_f^{H_m} \end{aligned}$$

It follows that

$$\bar{U}_m = \bar{U}_M\left(H_m, H_f\right) \text{ and } \bar{U}_f = \bar{U}_F\left(H_m, H_f\right)$$

Since

$$\bar{U}_i = \exp\left(\frac{1-\delta}{1-\delta^T} V_i\left(H, \mu\right)\right)$$

we conclude that μ only depends on (H_m, H_f). □

3.3 The first stage: Education Choice

The solution to the matching problem allows us to construct the expected value of marriage for males and females, conditional on each of the three education levels. At this point the stochastic structure is provided by the shock to the costs of education. Moreover, exogenous shifters of education choice can be included as elements of the cost function of education. Given this, the education choice is described in equation (14)-(15). Empirically this will be the basis of a logistic regression. However, while the individual is assumed to know their ability, this needs to be integrated out of the education choice model.

4 Identification

The model as presented now requires a distributional assumption for identification of the Pareto weights. However, this can be relaxed if we are willing to allow preferences for marriage to depend on exogenous variables that do not affect the surplus from marriage.

To do this we still assume that marriage generates a surplus, which is the sum of an 'economic' component, reflecting the gains arising when marriage from both risk sharing and the presence of a public good, and a non monetary term reflecting individual, idiosyncratic preferences for marriage. The economic part is, as before, a deterministic function of the spouses' respective levels of human capital; its distribution between husband and wife is endogenous and determined by the equilibrium conditions on the marriage market. Regarding the non monetary part, however, we assume that the non monetary benefit of agent i $(= m, f)$ is the sum of a systematic effect, which depends on some of i's observable characteristics (but not on his spouse's), and of an idiosyncratic term; as before, we assume that the idiosyncratic term, modeled as a random shock, only depends on the human capital of i's spouse. Equation (8) is thus replaced with:

$$\Sigma_{mf} = \Sigma\left(H_m, H_f\right) + \left(X_m a^{H_m, H_f} + \beta_m^{Hf} - \beta_m^0\right) + \left(X_f b^{H_m, H_f} + \beta_f^{Hm} - \beta_f^0\right) \tag{21}$$

where X_i is a vector of observable characteristics of agent i. For instance, X_i may include the education levels of i's parents; a possible interpretation being that an individual's preferences for the spouses human capital is directly affected by the individual's family background. Many alternative interpretations are possible; the crucial assumption, here, is simply that the surplus depends on both X_m and X_f but not on their interaction. Also, note that the coefficients a and b may depend on both spouse's human capital.

In such a setting, one can, under standard, full support assumptions, identify the vectors of parameters a^{H_m,H_f}, b^{H_m,H_f} and the distribution of $\beta_m^{Hf} - \beta_m^0$ and $\beta_f^{Hm} - \beta_f^0$ (up to the standard normalizations). To see why, note that Theorem 1 and Corollary 1 can be extended in the following way:

Theorem 2. *If the surplus is given by (21), then there exist* $2(NJ)^2$ *numbers -* $\bar{U}_M(H_m, H_f)$ *and* $\bar{U}_F(H_m, H_f)$ *for* $(H_m, H_f) \in \mathcal{H}^2$ *- such that:*

1. *For any* (H_m, H_f)

$$\bar{U}_M(H_m, H_f) + \bar{U}_F(H_m, H_f) = g(H_m, H_f)$$

2. *For any* m *with human capital* H_m *married to* f *with human capital* H_f*,*

$$\begin{aligned} U_m &= \bar{U}_M(H_m, H_f) + X_m a^{H_m,H_f} + \beta_m^{H_f} \text{ and} \\ U_f &= \bar{U}_F(H_m, H_f) + X_f b^{H_m,H_f} + \beta_f^{H_m} \end{aligned}$$

with the normalization $a^{H_m,0} = b^{0,H_f} = 0$.

Proof. Assume that m and m' have the same human capital H_m, and their respective partners f and f' have the same human capital H_f. Stability requires that:

$$U_m + U_f = g(H_m, H_f) + X_m a^{H_m,H_f} + \beta_m^{H_f} + X_f b^{H_m,H_f} + \beta_f^{H_m} \tag{22}$$

$$U_m + U_{f'} \geq g(H_m, H_f) + X_m a^{H_m,H_f} + \beta_m^{H_f} + X_{f'} b^{H_m,H_f} + \beta_{f'}^{H_m} \tag{23}$$

$$U_{m'} + U_{f'} = g(H_m, H_f) + X_{m'} a^{H_m H_f} + \beta_{m'}^{H_f} + X_{f'} b^{H_m H_f} + \beta_{f'}^{H_m} \tag{24}$$

$$U_{m'} + U_f \geq g(H_m, H_f) + X_{m'} a^{H_m H_f} + \beta_{m'}^{H_f} + X_f b^{H_m H_f} + \beta_f^{H_m} \tag{25}$$

Subtracting (22) from (23) and (25) from (24) gives

$$U_{f'} - U_f \geq (X_{f'} - X_f)\, b^{H_m, H_f} + \beta_{f'}^{H_m} - \beta_f^{H_m} \geq U_{f'} - U_f \tag{26}$$

hence

$$U_{f'} - U_f = (X_{f'} - X_f)\, b^{H_m, H_f} + \beta_{f'}^{H_m} - \beta_f^{H_m}$$

It follows that the difference $U_f - X_f b^{H_m, H_f} - \beta_f^{H_m}$ does not depend on f, i.e.:

$$U_f - X_f b^{H_m, H_f} - \beta_f^{H_m} = \bar{U}_F\left(H_m, H_f\right)$$

The proof for m is identical. □

As before, an immediate consequence is the following:

Corollary 3. *1. For any man m with human capital H_m, m's spouse at the stable matching has human capital H_f if and only if the following inequalities hold for all $H \in \mathcal{H} \cup \{0\}$:*

$$\bar{U}_M\left(H_m, H_f\right) + X_m a^{H_m, H_f} + \beta_m^{H_f} \geq \bar{U}_M\left(H_m, H\right) + X_m a^{H_m, H} + \beta_m^{H}$$

Similarly, m is single if and only if:

$$\bar{U}_M\left(H_m, H_f\right) + X_m a^{H_m, H_f} + \beta_m^{H_f} \leq \beta_m^0 \quad \text{for all } H_f \in \mathcal{H}$$

2. For any woman f with human capital H_f, f's spouse at the stable matching has human capital H_m if and only if the following inequalities hold for all $H \in \mathcal{H} \cup \{0\}$:

$$\bar{U}_F\left(H_m, H_f\right) + X_f b^{H_m, H_f} + \beta_f^{H_m} \geq \bar{U}_F\left(H, H_f\right) + X_f b^{H, H_f} + \beta_f^{H}$$

Similarly, m is single if and only if:

$$\bar{U}_F\left(H_m, H_f\right) + X_f b^{H_m, H_f} + \beta_f^{H_m} \leq \beta_f^0 \quad \textit{for all } H_m \in \mathcal{H}$$

3. *The ex-ante expected utility of a man m with human capital H_m is:*

$$A_M(H_m) = \mathbb{E}\left[\max_{H_f \in \mathcal{H} \cup \{0\}} \left(\bar{U}_M\left(H_m, H_f\right) + X_m a^{H_m, H_f} + \beta_m^{H_f}\right)\right] \tag{27}$$

and the ex-ante *expected utility of a female agent f with human capital H_f is:*

$$A_F(H_f) = \mathbb{E}\left[\max_{H_m \in \mathcal{H} \cup \{0\}} \left(\bar{U}_F\left(H_m, H_f\right) + X_f b^{H_m, H_f} + \beta_f^{H_m}\right)\right] \tag{28}$$

where the expectation is over the distribution of unobserved preferences for spouse's types, β_m and β_f for men and women respectively.

It follows that the marital choice of any male m (female f) with human capital H_m (H_m) boils down to a standard, multinomial choice discrete model; the standard identification results apply. However, in the version of this paper we rely on an extreme value distribution for individual utilities and not on covariates.

Beyond this, there are other important aspects of identification because both education and marriage are endogenous in our model. A key identifying assumption is that marriage does not cause changes in wages. In other words any correlation of wages and marital status is attributed to composition effects. However, education does cause changes in wages and it is likely that the ability composition of the various education groups differ: labor market ability is known when educational choices are made in our model. To control for the endogeneity of education we allow the costs of education to depend on parental background and parental

income, all of which are assumed to be excluded from wages and employment. In other words we assume that the entire influence of the parents works through educational attainment. This is clearly a strong assumption that can be relaxed particularly if we are in a position to exploit policy reforms that affect educational attainment but not wages directly.

5 Data

Estimation uses the 18 annual waves (1991 to 2008) of the British Household Panel Survey (BHPS). In this panel, apart from those who are lost through attrition, all families in the original 1991 sample and subsequent booster samples remain in the panel from then onwards. Other individuals have been added to the sample in subsequent periods — sometimes temporarily — as they formed families with original interviewees or were born into them. All members of the household aged 16 and above are interviewed, with a large set of information being collected on demographic characteristics, educational achievement, employment and earned income. Crucial to our analysis, the family relationship between members of the household can be determined.

We use the longitudinal information on all individuals in the original and booster samples during their prime working years, between the ages of of 23 and 50. To this we add information on the spouses they marry to during the observation window. Marital status is assessed for those ever observed aged 30 or above. Amongst them, singles are those who are never observed as married or co-habiting. All others are classified as being in couples and for them we keep all observations over the duration of their first marriage. In total, the final dataset contains information on education, employment and earnings for 4,317 couples, 1131 single women and 937 single men. Of these, 60% are observed for at least 5 years. In total, the sample size is

just short of 43,000 observations.

In the empirical analysis, employment is defined as working at least 5 hours per week. Earnings are measured on a weekly basis. We use the central 90% of the distribution of pre-tax real earnings for employees only. Since our model does not deal with macroeconomic growth and fluctuations, we net out aggregate earnings growth from earnings. Finally, we consider 3 education levels, corresponding to secondary education (leaving school at 16), high school and university degree.

6 Empirical specification and estimation

6.1 Earnings process

Individual earnings vary by gender, education, idiosyncratic ability and age. We assume they follow a cubic polynomial in age, with all coefficients being education- and gender-specific. We estimate the following earnings equation by gender (g and education (s):

$$\ln w_{it} = \ln W(\theta_i) + \delta_1 t + \delta_2 t^2 + \delta_3 t^3 + e_{it} + \epsilon_{it} \quad (29)$$

$$e_{it} = \rho e_{it-1} + \xi_{it} \quad (30)$$

where i indexes individual and t is age, w_{it} are the earnings of individual i when aged t, e is the productivity shock, assumed to follow an AR(1) process with normal innovations, and ϵ is a transitory shock that we interpret as measurement error (classical). W is the market wage faced by individual i of ability type (θ). We consider a discrete distribution of ability, with a 2-points support.

6.2 Preferences

Conditional on gender, education and marital status, we model labour supply as a function of age and unobserved preferences for working. This is a simple specification meant to capture the variation in employment over the lifetime without controlling for the presence of children and other important factors in driving labour supply decisions. The following formalises the preferences coefficient represented by α_{it} in the theoretical model:

$$\alpha_{it} = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3 + \eta_i + u_{it} \tag{31}$$

where the parameters ($\alpha_0,\ \alpha_1\ \alpha_2$) are specific to gender, education and marital status. The variable η represents unobserved heterogeneity in preferences for working, accounting for persistent differences in labour supply across individuals that are not fully explained by differences in earnings capacity. It is assumed to follow a discrete 2-point distribution independent of ability but possibly related with education. Since preferences for working are revealed only after the marriage stage, its distribution is also independent of the spouse's characteristics. Finally, u is a transitory preference shock, drawn from a normal distribution.

6.3 Estimation

This model can be estimated based on the method of moments. The process involves solving the life-cycle model and the resulting equilibrium in the marriage market for a particular parameter vector. Once this has been done we can construct moments from simulated data. Estimation then involves choosing the parameters that best match the equivalent data moments. Note that in solving the model we need to account for the fact that any change in educational attainment has implications for the marriage market and of course any change in the marriage market and

the implied intrahousehold allocations has implications for educational choice. This implies that the fixed point in the matching game needs to be computed in each iteration of the estimation process, whereby individuals educational choice is taken under some beliefs about the distribution of human capital in the marriage market that is actually realised. Hence, the full estimation can be very time consuming.

The estimates presented here are based on a stepwise procedure and not on the implementation that requires the full solution mentioned above.[7] Here we start by estimating earnings equations in levels (as in 29) and the stochastic process of wages controlling for the endogeneity of education using a control function approach. This regression provides a full description of the lifecycle wage profiles, excluding the effect of ability, and of the returns to education.

We next take these parameters as fixed and estimate the parameters of the ability distribution (θ) and of preferences (31), including the unobserved preferences for work. In total there are 45 parameters that need to be estimated to recover the conditional distributions of ability: earnings levels by gender and ability (12 parameters in total), the probability weights for the spouses joint distribution of abilities by education (another 27 parameters), and the probability weights for the distribution of ability by education and gender among single individuals (the final 6 parameters). In addition, there are also 66 preference-related parameters to be estimated, including the coefficients on age, the parameters in the distribution of unobserved preferences and the variances of the transitory preference shocks. In matching the model moments to the data we take into account the endogenous selection into employment, as implied by the structure of the model.

Estimation relies on a total of 328 moments describing the distribution of log earnings net of age effects and employment choices. This list includes the means, variances and several quantiles of

[7]We will study the feasibility of the full solution approach to estimation in future versions.

the earnings distribution, the regression coefficients of employment on a quadratic polynomial in age and moments describing the individual-level persistency of employment, measured by the proportion of years working amongst those observed for at least 5 years, all by education, gender and marital status. For couples, it also includes quantiles of the joint distribution of earnings. The choice of moments is driven by the nature of the parameters we need to identify. Specifically, the distribution of unobserved heterogeneity in earnings is identified by the remaining unexplained cross sectional distribution of earnings. In doing this, selection into employment is taken into account by the model itself. The distribution of unobserved heterogeneity in preferences is identified by the distribution of the proportion of time working over each individual observation window. A full list of data and simulated moments together with the diagnostics of the quality of fit can be found in appendix D. Appendix C presents the estimated parameters.

The solution to the matching game yields estimates of the Pareto weights. It relies on Theorem 1 and its Corollary 1. First note that there are six possible levels of human capital for both men and women, corresponding to the interaction between two levels of ability and three levels of education (Statutory or Secondary, High School and College). Consider man m with human capital H_m. The utility he gets from marrying a wife with human capital H is $\bar{U}(H_m, H) + \beta_m^H$. Among men of human capital H_m, the probability of marrying woman H is given by:

$$\begin{aligned} \Pr(H \mid H_m) &= \Pr\left(\bar{U}(H_m, H) + \beta_m^H \geq \bar{U}(H_m, H') + \beta_m^{H'}\right) \\ &= \Pr\left(\beta_m^H - \beta_m^{H'} \geq \bar{U}(H_m, H') - \bar{U}(H_m, H)\right) \end{aligned}$$

for all $H' \in \mathcal{H} \cup \{0\}$. If we assume that the β_m^H are extreme value distributed, as we do here, we obtain a multinomial logit structure, the estimation of which gives the expected utilities $\bar{U}(H_m, H)$ for all H (with the normalization $\bar{U}(H_m, 0) = 0$).

Similarly, m is single if and only if:

$$\bar{U}\left(H_m, H\right) + \beta_m^H \leq \beta_m^0 \quad \text{for all } H \in \mathcal{H}$$

Recall that $\bar{U_m}$ is

$$\begin{aligned} \bar{U}\left(H_m, H\right) &= \exp \frac{1-\delta}{1-\delta^T} V_m(H_m, H, \mu) \\ \text{where } V_m(H_m, H, \mu) &= \Upsilon(H_m, H) + \exp \frac{1-\delta^T}{1-\delta} \ln(\mu(H_m, H)) \end{aligned}$$

Note that these functions are known up to a transformation of the Pareto weights. This then implies that the Pareto weights can be estimated if one observes the probabilities $\Pr\left(H \mid H_m\right)$. Moreover, the observation of similar quantities for women results in over-identification, with two full sets of identification conditions.

Although matching here is on a trait that is only partly observed in the data, we can use the lifetime of earnings and employment histories in couples and singles to estimate the conditional (on education) distribution of ability in couples and singles, respectively $\Pr\left(\theta_m, \theta_f \mid s_m, s_f\right)$ and $\Pr\left(\theta_i \mid s_i\right)$. The marriage market outcomes - specifically $\Pr(H \mid H_m)$ and $\Pr(H \mid H_f)$ for all $H \in \mathcal{H} \cup \{0\}$ and each $H_m \in \mathcal{H}$ and $H_f \in \mathcal{H}$ - can be recovered by applying a simple conditional probability rule:

$$\begin{aligned} \Pr\left(H \mid H_i\right) &= \Pr\left(S, \theta \mid S_i, \theta_i\right) \\ &= \frac{\Pr\left(S, S_i, \theta, \theta_i\right)}{\Pr\left(S_i, \theta_i\right)} \\ &= \frac{\Pr\left(\theta, \theta_i \mid S, S_i\right) \Pr\left(S, S_i\right)}{\sum_{s \in \mathcal{S} \cup \{0\}} \Pr\left(\theta_i \mid s, S_i\right) \Pr\left(s, S_i\right)} \end{aligned}$$

for $i = m, f$. Note that all the quantities after the third equality are either directly ob-

served in the data ($\Pr(s, S_i)$ for all $s \in \mathcal{S} \cup \{0\}$) or estimated from the third stage problem ($\Pr(\theta, \theta_i \mid S, S_i)$ and $\Pr(\theta_i \mid s, S_i)$).

Finally, the function describing the utility cost of education is estimated within the structural model by the method of moments, conditional on the expected returns to education in the marriage and labor markets. The excluded variables are parental income when the respondent was aged 16 and family background as described by the first 2 principal components of a set of variables describing the environment in the respondent's parental home.[8] In estimating the cost of education, we match the regression coefficients of education attainment on these variables.

6.4 Education Choice

There are three levels of educational attainment possible: Secondary school (Statutory schooling), High school, corresponding to A-levels or equivalent and University, corresponding to 3-year degrees or above.

Estimating education choice is important for being able to solve for new equilibria and for carrying out full structural estimation of the model as described above. However, given the simplified estimation approach we have followed, the education choice component is not strictly necessary other than for constructing counterfactual equilibria. This is a central motivation of this paper, particularly because we are interested in analysing specific policies. However, we leave the simulation of counterfactuals to the the next version of the paper.

Estimation of the returns to education is undertaken by a multinomial logit mixed by the ability distribution, which enters the expected lifecycle value for each education choice. Preferences

[8]These include mother's and father's education, number of siblings and siblings' order, whether away from any of the parents, books at home, ethnic background.

for work or marriage are not known by the individual at this point, so are integrated out of the expected utilities.

7 Results

The estimates on the earnings equation and the distribution of ability as well as the preference parameters are presented in Appendix C, since they are not of a central interest in themselves. Perhaps the most pertinent fact to surface from these estimates is that the probability of being low ability for single man is between 0.75 and 0.5 (the latter for university graduates). However, the probability of being low ability for single women is between 0.11 (university graduates) and 0.44 (statutory schooling) (see Table 7 in Appendix C). Thus single men tend to be low ability, while single women tend to be drawn from the high ability part of the distribution.

7.1 The Surplus

In Table 1 we present the economic surplus of marriage for various couples. We start by ranking individuals by their human capital as shown in the table. This is implied by the present value of earnings for different educational ability combinations. For men the lower ability individuals have lower human capital than the higher ability ones, whatever the level of education; and of course human capital increases with education, given ability. For women the ranking is not as straightforward: for example women with statutory education of higher ability have more human capital than lower ability high school graduates.

The Table then presents the economic surplus for all possible 36 combinations of human capital for couples. There are two important conclusions from this. First, the gradient of the surplus is much steeper with respect to female human capital that it is for male. This is because the

impact of education on female earnings (conditional on employment) and employment itself is much higher for women than it is for men. We show this in Figures 1 and 2. Hence a large part of the variation in the surplus is explained by the human capital of the woman. Second, the surplus is super modular and implying positive assortative matching if it were not for preferences for marriage as implied by the random preferences β_i^H. This can be seen by noticing that for any 2×2 submatrix, the sum of diagonal terms exceeds the sum of off-diagonal ones.

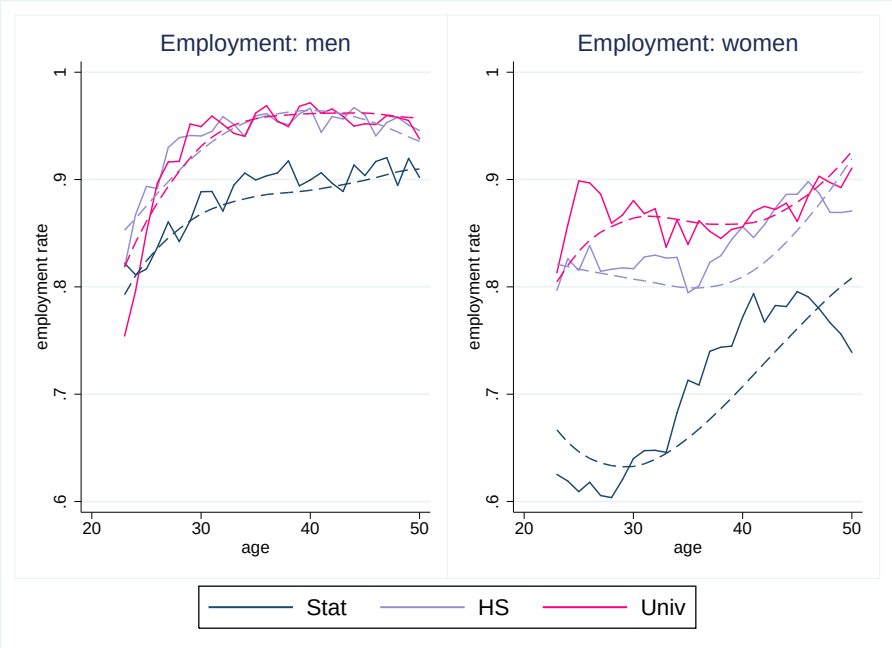
Table 1: Economic surplus from marriage by human capital of both spouses

	Women					
	Lowest HC $(s,\theta)=(1,1)$	2 $(2,1)$	3 $(1,2)$	4 $(2,2)$	5 $(3,1)$	Highest HC $(3,2)$
Men						
Lowest H: $(s,\theta)=(1,1)$	30.89	61.47	121.00	156.16	232.56	289.75
2: $(2,1)$	47.85	84.77	151.83	193.45	279.40	343.65
3: $(3,1)$	53.73	95.01	171.75	218.71	312.37	384.23
4: $(1,2)$	59.73	107.34	193.96	248.71	354.94	437.64
5: $(2,2)$	90.53	146.94	248.90	314.23	434.91	530.23
Highest H: $(3,2)$	89.56	150.56	262.62	333.59	462.07	565.35

Notes: Education levels 1, 2, and 3 correspond to statutory education, high school and university, respectively. Ability types 1 and 2 stand for low and high productivity.

This supermodularity result is quite surprising per se. Moreover, it generates an interesting prediction regarding matching pattern. Indeed, a result due to Graham (2013) states the following. In a model of the type being considered here, take any two levels of human capital for men, $H_m < H'_m$, and any two levels of human capital for women, $H_f < H'_f$. Consider the subpopulation of couples in which all husbands have either H_m or H'_m and all wives have either H_f or H'_f. If the corresponding, deterministic submatrix is supermodular - that is, if

Figure 1: BHPS data and model predictions: employment of men and women over the lifecycle



Notes: Full lines are for BHPS data and the dashed lines are for model simulations.

the double difference:

$$g\left(H_m, H_f\right) + g\left(H'_m, H'_f\right) - g\left(H'_m, H_f\right) - g\left(H_m, H'_f\right)$$

is positive, then it should be the case that in the corresponding subpopulation, the number of couples matched assortatively (i.e., (H_m, H_f) and $\left(H'_m, H'_f\right)$) should be larger than what would be expected under random matching. While we did not try a systematic test, this prediction seems remarkably well satisfied by the data. Of the 125 submatrices that can be constructed in that way, only 7 give a difference between actual and random smaller than -.01, and none smaller than -.015, whereas 30 are larger than .01, with a maximum of .04. In other words, the model predicts a high level of assortative matching that is actually found in the data. This is

Figure 2: BHPS data and model predictions: log annual earnings for men and women over the lifecycle



Notes: Full lines are for BHPS data and the dashed lines are for model simulations. Annual earnings in real terms (£1,000, 2008 prices).

all the more interesting that the prediction comes from the computation of the surplus, which does *not* involve actual matching patterns (but only labor supply behavior).

7.2 The Pareto Weights

The estimation approach we follow allows us to back out the implied Pareto weights. These describe the allocation of resources within the household in the context of the equilibrium observed in the data. Note, however, that if we wished to compute counterfactuals we would need to solve the model under a new policy regime. This would imply changes in all components of the model including new Pareto weights. In other words as we change the environment

and affect the economic surplus of marriage we will also affect educational choices, matching patterns and intrahousehold allocations, implied by the Pareto weights.

Table 2 shows the estimates of the Pareto weights for women.[9] These should be compared to a male Pareto weight of one. Since we can estimate the Pareto weights using either the matching probabilities of women to men or vice versa the model is overidentified; the weights we present have imposed equality using minimum distance.

In principle, the relationship between a person's human capital and Pareto weight needs not be strictly monotonic; Pareto weights also reflect relative scarcity of spouses at each level of human capital, and therefore depend on the entire distribution. Still, we do observe that the wife's Pareto weight is monotonically increasing in her human capital and mostly decreasing in the husband's. Moreover, Pareto weights are always increasing in education.

Among couples of college graduates with higher ability, the Pareto weights are basically equal for men and women. However, if a low skill man marries a highest skill women (a very rare combination) her Pareto weight is 2.6 times his. If a low skill woman marries the highest skill man her weight is only about 5% of his. It is also interesting to note that in most cases along the diagonal (except for the top human capital) the intra-household resource allocations favour the man, with her Pareto weight varying between 20%-65% of his. Women only manage to achieve intra-marital equality or better when they obtain a college degree.

A limitation of Pareto weights is that they only give a partial picture of intra-household allocation. They depend on the specific cardinalization of utility chosen. Since, in our setting, women and men are not found to have identical preferences (particularly regarding labor supply), the interpretation of the *value* of the wife's Pareto weights should not be taken too literally. A

[9]Standard errors for the Pareto weights were bootstrapped using 250 draws from the asymptotic distribution of the parameters driving decisions in the third stage of live. For each replication, the Pareto weights were calculated using men's and women's conditional matching probabilities and the estimated variance is that of the resulting sample of Pareto weights.

Table 2: Women's Pareto weights by spouses' human capital

	Women					
	Lowest HC $(s,\theta)=(1,1)$	2 $(2,1)$	3 $(1,2)$	4 $(2,2)$	5 $(3,1)$	Highest HC $(3,2)$
Men						
Lowest H: $(s,\theta)=(1,1)$	0.192	0.275	0.455	0.678	1.621	2.606
	(.08)	(.12)	(.10)	(.16)	(.43)	(.28)
2: $(2,1)$	0.135	0.192	0.314	0.459	1.813	1.491
	(.06)	(.08)	(.07)	(.11)	(.46)	(.37)
3: $(3,1)$	0.106	0.166	0.290	0.441	1.395	2.570
	(.04)	(.06)	(.05)	(.08)	(.66)	(.23)
4: $(1,2)$	0.096	0.147	0.244	0.378	2.041	1.956
	(.03)	(.05)	(.04)	(.07)	(.58)	(.45)
5: $(2,2)$	0.059	0.087	0.152	0.226	0.649	1.315
	(.02)	(.03)	(.03)	(.05)	(.43)	(.37)
Highest H: $(3,2)$	0.051	0.085	0.153	0.153	0.635	1.049
	(.01)	(.02)	(.02)	(.02)	(.25)	(.38)

Notes: SE in parenthesis under estimates. Education levels 1, 2, and 3 correspond to statutory education, high school and university, respectively. Ability types 1 and 2 stand for low and high productivity.

more telling measure is the allocation of welfare and private consumption - something our structural model allows us to reconstruct. This is done in Table 3. We see, in particular, that the allocation of welfare is much less unequal than what the sole consideration of Pareto weights might suggest. Moreover, the distribution of private consumption is much more unequal than that of welfare; actually, in the cells corresponding to high skilled husbands and low skilled wives, all private consumption goes to the husband. Indeed, in our model people get utility from private and public consumption, time off paid work, and their idiosyncratic preferences. Time off work and private consumption are substitutes, while both are complements of public consumption. Therefore, our results suggest that in couples where the husband is much more skilled than the wife, most of her utility comes from time off work (she is indeed less likely to work), public consumption, and her marital preference.[10] Note that this finding is totally

[10]Note that, given the strong super-modularity of the economic component, a marriage between spouses of

consistent with another feature of the data, mentioned earlier - namely, that unskilled women have more time off off paid work than men of any skill level (and particularly skilled ones).

Table 3: Sorting patterns and the sharing of consumption and welfare

	Women's HC					
	lowest	2	3	4	5	highest
Men's HC: lowest						
% matches in cell	0.070	0.014	0.040	0.018	0.003	0.002
man's share in welfare	0.60	0.57	0.54	0.51	0.47	0.45
man's share in consumption	0.93	0.90	0.61	0.53	0.38	0.23
Men's HC: 2						
% matches in cell	0.041	0.020	0.040	0.040	0.012	0.002
man's share in welfare	0.62	0.59	0.55	0.53	0.47	0.48
man's share in consumption	0.96	0.98	0.72	0.66	0.30	0.31
Men's HC: 3						
% matches in cell	0.005	0.004	0.002	0.008	0.008	0.006
man's share in welfare	0.63	0.60	0.56	0.53	0.48	0.45
man's share in consumption	1.01	0.99	0.74	0.68	0.41	0.23
Men's HC: 4						
% matches in cell	0.139	0.030	0.075	0.058	0.005	0.008
man's share in welfare	0.63	0.60	0.56	0.54	0.46	0.47
man's share in consumption	0.99	0.99	0.79	0.73	0.33	0.33
Men's HC: 5						
% matches in cell	0.052	0.024	0.059	0.069	0.003	0.027
man's share in welfare	0.66	0.63	0.59	0.56	0.51	0.48
man's share in consumption	1.02	1.03	0.86	0.83	0.67	0.44
Men's HC: highest						
% matches in cell	0.006	0.008	0.011	0.032	0.009	0.050
man's share in welfare	0.66	0.62	0.58	0.58	0.51	0.49
man's share in consumption	1.02	1.02	0.86	0.89	0.69	0.52

Lastly, it is crucial to keep in mind that the Pareto weights, and more generally the patterns of intra-household distribution of resources and welfare, are not structural parameters but endogenous entities reflecting the conditions in the marriage market. The present estimations

very different skills signals very large values of the corresponding marital preference. These are rare events, as shown by the size of the cells where spouses have very different skills (see first row for each level of the man's human capital in table 3).

reflect the patterns we see in the data. Changes in taxes and benefits, or indeed of other aspects of the economic environment, will affect both the economic surplus of marriage and (in the long run) individual decisions regarding marriage and human capital investments. Obviously, marital patterns, including intra-household allocations, will be impacted. For instance, a policy reform that subsidizes education will alter the supply of workers in the various education groups; this will affect both the degree of assortative matching and the equilibrium allocation of surplus.

8 Education choice

Finally, Table 4 presents the estimation of the results for education choice. These originate from a multinomial logit mixed by the ability distribution, which enters the expected lifecycle value for each education choice. Preferences for work or marriage are not known by the individual at this point, so are integrated out of the expected utilities.

Table 4: Cost of education

	Men		Women	
	HS	Univ	HS	Univ
constant	0.702	2.656	0.543	2.639
	(.02)	(.03)	(.01)	(.01)
background factor 1	0.073	-0.285	-0.000	-0.186
	(.01)	(.01)	(.01)	(.01)
background factor 2	-0.014	0.002	0.123	-0.109
	(.01)	(.02)	(.02)	(.02)
log parental income	-0.084	-0.212	-0.008	-0.180
	(.01)	(.03)	(.01)	(.01)

The costs of education are a function of two family background factors and of parental income when the individual was 16. The two factors are the first two principal components from a set

of variables characterizing the family.[11]

The results clearly indicate that both family background and family income affect education choice. Specifically family income reduces the costs of high school and university education for men. For women, they only affect the costs of attending university, but not the costs of attending high school.

This model of education choice can be used both for analyzing the lifecycle impact of education policy that shifts the costs of education and for simulating the equilibrium effects of policies that change education returns or indeed other policies that change the economic value of marriage.

9 Concluding Remarks

In this paper we have presented an equilibrium model of education choice, marriage and lifecycle labor supply, savings and public goods in a world with uncertainty in the labor market. Our framework relies on a transferable utility setting, which allows us to potentially simulate policies that change the economic environment at any stage of the lifecycle. Matching in the marriage market is stochastic and trades off the economic value of marriage with random preferences for type of mate (defined by their human capital). On the economic side, the final structure of matching is driven both from the demand for public goods and from a risk sharing motive.

We find that the surplus from marriage is indeed super-modular, pushing towards positive assortative matching, with any departures from perfect sorting being driven by random prefer-

[11]The family background variables include education of both parents (5 levels each), dummy for no siblings, dummy for 3 or more siblings, dummy for whether subject is the first child, books in childhood home (3 levels) and dummy for whether lived with both parents when aged 16.

ences for mates. We also find that the human capital of women is a very strong determinant of marital surplus, more so than the human capital of men. Finally, we show that generally most resources flow to men: only women with college degrees achieve equality (or better) within the household. However, the apparently large level of inequality in private consumption is partly compensated by differences in labor supply; all in all, the allocation of welfare, although rarely uniform, is much less unequal than suggested by the sole consideration of private consumptions.

This paper is a first step towards a rich research agenda analyzing the interactions of marriage, labor markets and educational choices. Important generalizations will include allowing for imperfectly transferable utility, generalizing the model to allow for divorce and finally allowing for limited commitment. These are important issues that will lead to better understanding of marriage markets and intra-household inequality. However they are also challenging. Our framework here shows, however that such equilibrium models can be rich in implications and valuable for the understanding of the longer term effects of policies.

Finally, the framework developed in this model, complicated as it may be, relies on two simple but extremely powerful insights. One is that marital sorting patterns – who marries whom – have an important, economic component, which can be analyzed in terms of 'complementarity' or 'substitutability' (in modern terms, super- or sub-modularity) of the surplus created within marriage; the second, that the intra household allocation of resources (therefore of welfare) is related to the equilibrium conditions prevailing on the 'marriage market', and should therefore be analyzed using the 'theory of optimal assignments' (aka matching models). Both insights are explicitly present in Becker's 1973 JPE masterpiece. That, more than forty years later, we can still find much to learn in exploiting these insights is an obvious tribute to the importance of Becker's contribution.

Appendix A: The solution of the household problem in the last period of life

Consumptions Take a man $m \in \mathcal{M}$ with human capital H_m married to a woman $f \in \mathcal{F}$ with with human capital H_f. All the results below are conditional on the time-invariant human capital of the spouses, $H = (H_m, H_f)$, and we omit them for simplicity.

The problem of this couple at time T is

$$\begin{array}{ll} \max\limits_{Q_T, C_T, L_T} & Q_T \left(C_T + \alpha_{mT} L_{mT} + \alpha_{fT} L_{fT} \right) \\ \text{s.t.} & \text{budget constraint: } y_T^C + w_{mT} + w_{fT} + RK_{T-1} = C_T + w_{mT} L_{mT} + w_{fT} L_{fT} + pQ_T \\ & \text{wage equation (2)} \end{array}$$

Here, K_{T-1} denotes savings accumulated at the end of period $T-1$ and transferred to period t at the risk-free interest factor R; and $Y_T =$ is the couple's total ('potential') income in period t. $y_T^C + w_{mT} + w_{fT}$ is the sum of the maximum possible labor income, $w_{mt} + w_{ft}$ (where total possible working time has been normalized to 1 for each individual), and the couple's non labor income, y_T^C. Note that the latter may depend on individual labor supplies and earnings.

Since T is the last period of life and bequests are not being considered in this problem, the optimal savings is $K_T = 0$ and the problem is static. We can thus derive total household consumptions as functions of labor supplies:

$$\begin{aligned} Q_T &= \frac{y_T^C + w_{mT}\left(1 - L_{mT}\right) + w_{fT}\left(1 - L_{fT}\right) + RK_{T-1} + \left(\alpha_{mT} L_{mT} + \alpha_{fT} L_{fT}\right)}{2p} \text{ and} \\ C_T &= \frac{y_T^C + w_{mT}\left(1 - L_{mT}\right) + w_{fT}\left(1 - L_{fT}\right) + RK_{T-1} - \left(\alpha_{mT} L_{mT} + \alpha_{fT} L_{fT}\right)}{2} \end{aligned}$$

and the sum of utilities becomes:

$$\frac{1}{4}\left(y_T^{L_{mT},L_{fT}} + w_{mT}\left(1-L_{mT}\right) + w_{fT}\left(1-L_{fT}\right) + RK_{T-1} + \left(\alpha_{mT}L_{mT} + \alpha_{fT}L_{fT}\right)\right)^2$$

Labour supplies The pair (L_{mT}, L_{fT}) can take four values - namely $(0,0), (1,0), (0,1)$ and $(1,1)$; and efficient labor supplies solve the program:

$$\max_{(L_{mT},L_{fT})\in\{0,1\}^2} \frac{1}{4}\left(y_T^C\left(L_{mT},L_{fT}\right) + w_{mT}\left(1-L_{mT}\right) + w_{fT}\left(1-L_{fT}\right) + RK_{T-1} + \left(\alpha_{mT}L_{mT} + \alpha_{fT}L_{fT}\right)\right)^2$$

Therefore labor supply patterns depend on the realization of the preference shoks α_{mT} and α_{fT}. Specifically:

- conditional on the woman's labor supply, L_{fT}, the man does not work ($L_{mT} = 1$) if $w_{mT} + y_T^C\left(0, L_{fT}\right) \geq \alpha_{mT} + y_T^C\left(1, L_{fT}\right)$, and will work otherwise
- similarly, conditional on L_{mT}, the woman does not work ($L_{fT} = 1$) if $w_{fT} + y_T^C\left(L_{mT}, 0\right) \geq \alpha_{fT} + y_T^C\left(L_{mT}, 1\right)$, and will work otherwise

Note that (generically on the realization of the shocks) *all* Pareto-efficient allocations correspond to the same labor supply pattern; this is a direct consequence of the (ordinal) TU property. The various efficient allocations differ only by the allocation of private consumption.

Efficient risk sharing We now consider the allocation of private consumption during the last subperiod from an *ex-ante* perspective - that is, before the realization of the shocks. Efficiency, here, is relative to sharing the (wages and preferences) risks. Efficiency, in this context, requires the maximization of a weighted sum of expected utilities, obviously using the

initial, logarithmic cardinalizations. If μ denotes the wife's Pareto weight corresponding to that cardinalization, the standard efficiency condition becomes:

$$\frac{\partial u_m\left(Q_T, C_{mT}, L_{mT}\right)}{\partial C} = \mu \frac{\partial u_f\left(Q_T, C_{fT}, L_{fT}\right)}{\partial C}$$

This gives:

$$\begin{aligned} C_{mT} &= \frac{1}{1+\mu} pQ_T - \alpha_{mT} L_{mT} \\ C_{fT} &= \frac{\mu}{1+\mu} pQ_T - \alpha_{fT} L_{fT} \end{aligned}$$

and finally indirect utilities:

$$\begin{aligned} v_{mT} &= 2\ln Q_T + \ln p + \ln \frac{1}{1+\mu} \\ v_{fT} &= 2\ln Q_T + \ln p + \ln \frac{\mu}{1+\mu} \end{aligned}$$

Note that Q_T depends on the realization of the wage and preferences shocks, e_T and α_T, as well as savings, non labor income and the spouses' respective stocks of human capital (through their impact on wages); we therefore write $Q_T\left(e_T, \alpha_T, K_{T-1}, H\right)$ and $v_{iT}\left(e_T, \alpha_T, K_{T-1}, H, \mu\right)$, where $H = (H_m, H_f)$.

Expected value functions We assume that the unobserved productivity shocks and preferences for time off paid work, (e, α), follow a first-order Markov process. Then, the expected

value functions are

$$\begin{aligned} V_{mT}\left(e_{T-1},\alpha_{T-1},K_{T-1},H,\mu\right) &= \mathrm{E}_{T|T-1}\left[v_{mT}\left(e_T,\alpha_T,K_{T-1},H,\mu\right)\mid e_{T-1},\alpha_{T-1}\right] \\ &= I_T\left(e_{T-1},\alpha_{T-1},K_{T-1},H\right)+\ln\left(\frac{1}{1+\mu}\right) \quad \text{and} \\ V_{fT}\left(e_{T-1},\alpha_{T-1},K_{T-1},H,\mu\right) &= I_T\left(e_{T-1},\alpha_{T-1},K_{T-1},H\right)+\ln\left(\frac{\mu}{1+\mu}\right) \end{aligned}$$

where

$$\begin{aligned} I_T\left(e_{T-1},\alpha_{T-1},K_{T-1},H\right) &= \mathrm{E}_{T|T-1}\left[2\ln Q_T\left(e_T,\alpha_T,K_{T-1},H\right)+\ln p\mid e_{T-1},\alpha_{T-1}\right] \\ &= 2\int \ln Q_T\left(e_T,\alpha_T,K_{T-1},H\right)\,dF\left(e_T,\alpha_T\mid e_{T-1},\alpha_{T-1}\right)+\ln p \end{aligned}$$

Appendix B: Employment, consumption and savings for singles

At time t, a single individual i chooses $(L_{it},Q_{it},C_{it},K_{it})$ to maximise lifetime utility:

$$\left(C_{it}Q_{it}+\alpha_{it}L_{it}\right)+\delta I^S_{i,t+1}\left(e_{it},\alpha_{it},K_{it},H_i\right)$$

subject to the budget constraint

$$w_{it}\left(1-L_{it}\right)+y^S_{it}+RK_{i,t-1}=K_{it}+C_{it}+pQ_{it}$$

where $w_{it}\left(1-L_{it}\right)$ is the individual's labor income and y^S_t is non labor income, itself possibly a function of employment and labor income.

Conditionally on labour supply and savings, the consumptions are

$$\begin{aligned} Q_{it}\left(K_{it}, L_{it}\right) &= \frac{y_{it}^{S} + RK_{i,t-1} - K_{it} + w_{it}\left(1 - L_{it}\right) + \alpha_{it} L_{it}}{2p} \\ C_{it}\left(K_{it}, L_{it}\right) &= pQ_{it}\left(K_{it}, L_{it}\right) - \alpha_{it} L_{it} \end{aligned}$$

and the choice of (K_{it}, L_{it}) solves the maximization problem

$$V_{it}^{S}\left(e_{it}, \alpha_{it}, K_{i,t-1}, H_i\right) = \max_{L_{it}, K_{it}} \left\{2 \ln Q_{it}\left(K_{it}, L_{it}\right) + \ln p + \delta I_{i,t+1}^{S}\left(e_{it}, \alpha_{it}, K_{it}, H_i\right)\right\}$$

where

$$\begin{aligned} & I_{it}^{S}\left(e_{i,t-1}, \alpha_{i,t-1}, K_{i,t-1}, H_i\right) \\ = \; & \mathbb{E}_{t|t-1} \max_{L_{it} \in \{0,1\}} \left[\max_{K_{it}} \left\{2 \ln Q_{it}\left(K_{it}, L_{it}\right) + \delta I_{i,t+1}^{S}\left(e_{it}, \alpha_{it}, K_{it}, H_i\right)\right\} \mid e_{i,t-1}, \alpha_{i,t-1}\right] \end{aligned}$$

Then, conditionally on employment, optimal savings, K_{it}^{S}, solves the intertemporal optimality condition:

$$2\frac{\partial \ln Q_t}{\partial K_t} + \delta \frac{\partial I_{t+1}^{S}}{\partial K_t} = 0$$

Finally, employment is

$$L_{it}^{S} = \arg\max_{L_t \in \{0,1\}} \left\{2 \ln Q_{it}\left(K_{it}(L_t), L_t\right) + \ln p + \delta I_{i,t+1}\left(e_{it}, \alpha_{it}, K_{it}^{S}, H\right)\right\}.$$

Appendix C: Estimates of model parameters

Table 5 contains estimates of the parameters in the stochastic wage process. Most of these parameters were estimated in the first stage reduced form model of education choice and wages.

The exception are the wage levels (row 1 in the table) and the ability premium (row 3), which are estimated within the structural model taking employment choice into account. In here, ability types 1 and 2 stand for low and high productivity, respectively. What is interesting to notice here is that the returns to education are more important for women than men, a finding illustrated in figure 2 by the narrowing of the gender wage gap with education. The high market premium of education for women can be partly driven by the short working hours that women with statutory education do (see Blundell et al., 2015). Moreover, education narrows the ability wage gap among women, with a premium that is much more modest for university graduates than other groups.

Table 5: Earnings process by gender and education

		Men			Women		
		Stat	HS	Univ	Stat	HS	Univ
(1)	log earnings (ab 1, stat ed, age 23)	2.52			1.95		
		(.01)			(.01)		
(2)	education premium		0.260	0.381		0.434	0.717
			(.04)	(.07)		(.07)	(.10)
(3)	ability premium (type 2)	0.52	0.51	0.50	1.00	0.95	0.28
		(.01)	(.01)	(.03)	(.01)	(.01)	(.03)
(4)	age (δ_1)	0.475	0.606	0.923	-0.232	0.144	0.738
		(.05)	(.04)	(.07)	(.08)	(.07)	(.08)
(5)	age squared (δ_2)	-0.252	-0.302	-0.524	0.125	-0.172	-0.620
		(.04)	(.03)	(.06)	(.07)	(.06)	(.08)
(6)	age cubic (δ_3)	0.042	0.050	0.094	-0.017	0.052	0.153
		(.01)	(.01)	(.01)	(.02)	(.02)	(.02)
(7)	Autocorr coeff (ρ)	0.502	0.594	0.416	0.811	0.820	0.886
		(.115)	(.131)	(.226)	(.104)	(.067)	(.122)
(8)	Var innov in prod (σ^2_ξ)	0.024	0.012	0.026	0.030	0.035	0.022
		(.005)	(.005)	(.049)	(.006)	(.004)	(.006)
(9)	Var ME (σ^2_ϵ)	0.004	0.007	0.001	0.012	0.005	0.011
		(.004)	(.005)	(.049)	(.004)	(.002)	(.006)
	N	9,116	11,990	4,291	8,432	7,469	3,962

Notes: SE in brackets under estimates. Earnings are in logs of £1,000 per year, 2008 prices.

Tables 6 and 7 show the probability weights in the distribution of ability in couples and for singles, respectively. Estimates for couples are conditional on the education of both spouses and each square displays the mass in all points in the conditional distribution, thus adding up to 1. The table discloses some interesting regularities, with ability type 2 (the more productive type) being relatively more frequent amongst more educated couples. Among singles, ability type 1 (low productivity) is more prevalent for those with basic education only, and single men are comparatively more likely to be of this ability type then single women.

Table 6: Probability weights for the joint distribution of ability in couples by spouses' education

		Men					
		Stat Ed		HS		Univ	
		ability 1	ability 2	ability 1	ability 2	ability 1	ability 2
Women	Stat Ed						
	ability 1	0.216	0.429	0.213	0.270	0.200	0.242
		(.01)	(.02)	(.02)	(.02)	(.07)	(.07)
	ability 2	0.124	0.231	0.209	0.308	0.103	0.456
		(.02)		(.02)		(.05)	
	HS						
	ability 1	0.117	0.247	0.132	0.156	0.078	0.155
		(.01)	(.02)	(.01)	(.02)	(.04)	(.04)
	ability 2	0.153	0.483	0.262	0.450	0.154	0.613
		(.02)		(.01)		(.04)	
	Univ						
	ability 1	0.148	0.287	0.276	0.071	0.112	0.118
		(.05)	(.40)	(.08)	(.06)	(.04)	(.10)
	ability 2	0.093	0.472	0.051	0.603	0.078	0.692
		(.08)		(.09)		(.03)	

Finally, estimates of the preference parameters are presented in table 8.

Table 7: Proportion of ability type 1 among singles by gender and education

	secondary	high school	university
men	0.753	0.705	0.505
	(.03)	(.02)	(.03)
women	0.439	0.171	0.109
	(.04)	(.06)	(.08)

Table 8: Preference parameters and distribution of unobserved heterogeneity in preferences for employment

	Men			Women		
	Stat	HS	Univ	Stat	HS	Univ
Couples						
intercept (α_0)	-0.012	0.229	0.649	0.574	0.312	1.412
	(1.36)	(0.58)	(0.26)	(0.18)	(0.19)	(0.13)
age (α_1)	-0.151	-0.056	0.009	0.061	0.125	0.033
	(0.17)	(0.19)	(0.39)	(0.13)	(0.17)	(0.16)
age squared (α_2)	0.087	-0.132	0.111	-0.294	0.023	0.074
	(0.14)	(0.10)	(0.19)	(0.13)	(0.15)	(0.13)
age cubic (α_3)	-0.034	0.062	-0.041	0.051	-0.064	-0.028
	(0.06)	(0.05)	(0.08)	(0.05)	(0.05)	(0.05)
Singles						
intercept (α_0)	0.102	0.766	1.464	1.018	0.810	1.463
	(1.36)	(0.43)	(0.26)	(0.07)	(0.10)	(0.23)
age (α_1)	0.195	0.198	0.073	0.225	-0.080	-0.002
	(0.42)	(0.54)	(0.12)	(0.31)	(0.30)	(0.38)
age squared (α_2)	0.111	0.118	0.100	-0.245	0.141	-0.046
	(0.26)	(0.93)	(0.15)	(0.37)	(0.29)	(0.30)
age cubic (α_3)	-0.035	-0.078	-0.109	0.045	-0.055	-0.015
	(0.06)	(0.32)	(0.06)	(0.11)	(0.09)	(0.09)
Unobserved preferences						
low utility from work ($\eta = 2$)	2.325	1.758	1.050	1.127	0.955	0.495
	(1.40)	(0.49)	(0.36)	(0.11)	(0.19)	(0.17)
probability utility type 2	0.606	0.628	0.448	0.551	0.480	0.559
	(0.04)	(0.05)	(0.16)	(0.07)	(0.09)	(0.09)
Var transitory pref shock (u)	0.921	0.957	1.040	0.130	1.000	0.928
	(0.27)	(0.34)	(0.26)	(0.05)	(0.07)	(0.14)

Notes: SE in brackets under estimates.

Appendix D: Fit

This appendix contains tables showing all data moments used in estimation and their simulated counterparts, together with the ratio of the discrepancy between the two moments and the standard error of the data estimate.

Table 9: Distr log earnings net of age effects: single men

Moment	Data	Simulated	SE data	No. SE diff
secondary education				
mean	2.643	2.645	0.031	0.052
var	0.143	0.125	0.016	1.126
P(earnings<Q10)	0.100	0.096	0.021	0.213
P(earnings<Q25)	0.250	0.272	0.029	0.762
P(earnings<Q50)	0.500	0.517	0.038	0.440
P(earnings<Q75)	0.750	0.744	0.032	0.172
P(earnings<Q90)	0.900	0.897	0.025	0.116
high school				
mean	2.792	2.907	0.031	3.702
var	0.137	0.106	0.015	2.147
P(earnings<Q10)	0.100	0.034	0.019	3.396
P(earnings<Q25)	0.250	0.128	0.031	3.989
P(earnings<Q50)	0.500	0.383	0.038	3.073
P(earnings<Q75)	0.750	0.666	0.032	2.658
P(earnings<Q90)	0.900	0.867	0.023	1.412
university education				
mean	2.623	2.860	0.042	5.660
var	0.175	0.114	0.025	2.435
P(earnings<Q10)	0.100	0.013	0.026	3.306
P(earnings<Q25)	0.250	0.082	0.040	4.186
P(earnings<Q50)	0.500	0.288	0.046	4.560
P(earnings<Q75)	0.750	0.561	0.040	4.771
P(earnings<Q90)	0.900	0.767	0.025	5.323

Table 10: Distr log earnings net of age effects: single women

Moment	Data	Simulated	SE data	No. SE diff
secondary education				
mean	2.418	2.412	0.042	0.130
var	0.233	0.297	0.025	2.536
P(earnings<Q10)	0.100	0.108	0.020	0.377
P(earnings<Q25)	0.250	0.312	0.032	1.948
P(earnings<Q50)	0.500	0.530	0.045	0.682
P(earnings<Q75)	0.750	0.729	0.040	0.536
P(earnings<Q90)	0.900	0.871	0.026	1.153
high school				
mean	2.631	2.646	0.035	0.434
var	0.218	0.210	0.020	0.407
P(earnings<Q10)	0.100	0.075	0.017	1.484
P(earnings<Q25)	0.250	0.277	0.031	0.845
P(earnings<Q50)	0.500	0.526	0.038	0.676
P(earnings<Q75)	0.750	0.749	0.034	0.039
P(earnings<Q90)	0.900	0.887	0.023	0.581
university education				
mean	2.739	2.755	0.031	0.495
var	0.158	0.166	0.020	0.414
P(earnings<Q10)	0.100	0.080	0.020	0.999
P(earnings<Q25)	0.250	0.278	0.032	0.869
P(earnings<Q50)	0.500	0.575	0.045	1.686
P(earnings<Q75)	0.750	0.737	0.034	0.389
P(earnings<Q90)	0.900	0.842	0.025	2.339

Table 11: Distr log earnings net of age effects: men in couples

Moment	Data	Simulated	SE data	No. SE diff
secondary education				
mean	2.794	2.801	0.012	0.561
var	0.127	0.116	0.006	1.777
P(earnings<Q10)	0.100	0.106	0.008	0.718
P(earnings<Q25)	0.250	0.246	0.013	0.340
P(earnings<Q50)	0.500	0.475	0.016	1.512
P(earnings<Q75)	0.750	0.755	0.015	0.366
P(earnings<Q90)	0.900	0.915	0.010	1.514
high school				
mean	2.913	2.924	0.011	1.033
var	0.120	0.113	0.006	1.181
P(earnings<Q10)	0.100	0.082	0.008	2.276
P(earnings<Q25)	0.250	0.257	0.012	0.563
P(earnings<Q50)	0.500	0.522	0.015	1.487
P(earnings<Q75)	0.750	0.769	0.012	1.514
P(earnings<Q90)	0.900	0.891	0.008	1.069
university education				
mean	2.840	2.860	0.018	1.100
var	0.105	0.110	0.010	0.492
P(earnings<Q10)	0.100	0.094	0.014	0.414
P(earnings<Q25)	0.250	0.289	0.023	1.663
P(earnings<Q50)	0.500	0.540	0.027	1.455
P(earnings<Q75)	0.750	0.722	0.022	1.262
P(earnings<Q90)	0.900	0.850	0.013	3.908

Table 12: Distr log earnings net of age effects: women in couples

Moment	Data	Simulated	SE data	No. SE diff
secondary education				
mean	2.205	2.226	0.021	1.042
var	0.277	0.255	0.010	2.174
P(earnings<Q10)	0.100	0.047	0.008	6.386
P(earnings<Q25)	0.250	0.191	0.013	4.482
P(earnings<Q50)	0.500	0.547	0.018	2.557
P(earnings<Q75)	0.750	0.775	0.015	1.649
P(earnings<Q90)	0.900	0.887	0.012	1.073
high school				
mean	2.504	2.557	0.016	3.274
var	0.283	0.236	0.010	4.527
P(earnings<Q10)	0.100	0.030	0.007	9.707
P(earnings<Q25)	0.250	0.233	0.012	1.375
P(earnings<Q50)	0.500	0.534	0.014	2.333
P(earnings<Q75)	0.750	0.748	0.013	0.170
P(earnings<Q90)	0.900	0.874	0.008	3.083
university education				
mean	2.690	2.720	0.021	1.434
var	0.227	0.181	0.013	3.551
P(earnings<Q10)	0.100	0.043	0.010	5.510
P(earnings<Q25)	0.250	0.273	0.017	1.343
P(earnings<Q50)	0.500	0.580	0.022	3.716
P(earnings<Q75)	0.750	0.762	0.018	0.641
P(earnings<Q90)	0.900	0.863	0.012	3.028

Table 13: Distribuition log earnings net of age effects: men in couple by spouses's education

Moment	Data	Simulated	SE data	No. SE diff
secondary education				
mean: spouse second educ	2.801	2.776	0.019	1.306
mean: spouse high school	2.901	2.841	0.024	2.530
mean: spouse univ educ	2.924	2.839	0.055	1.563
var: spouse second educ	0.123	0.121	0.009	0.119
var: spouse high school	0.112	0.104	0.011	0.694
var: spouse univ educ	0.109	0.114	0.053	0.100
high school				
mean: spouse second educ	2.818	2.921	0.021	4.919
mean: spouse high school	2.935	2.904	0.019	1.641
mean: spouse univ educ	2.838	2.994	0.029	5.466
var: spouse second educ	0.097	0.115	0.010	1.846
var: spouse high school	0.108	0.111	0.007	0.362
var: spouse univ educ	0.074	0.112	0.011	3.498
university education				
mean: spouse second educ	2.905	2.901	0.080	0.057
mean: spouse high school	2.986	2.792	0.035	5.517
mean: spouse univ educ	2.873	2.892	0.023	0.842
var: spouse second educ	0.109	0.109	0.019	0.017
var: spouse high school	0.109	0.110	0.016	0.043
var: spouse univ educ	0.098	0.106	0.012	0.658

Table 14: Distribuition log earnings net of age effects: women in couple by spouses's education

Moment	Data	Simulated	SE data	No. SE diff
secondary education				
mean: spouse second educ	2.182	2.178	0.031	0.134
mean: spouse high school	2.320	2.272	0.040	1.197
mean: spouse univ educ	2.453	2.498	0.079	0.561
var: spouse second educ	0.243	0.233	0.014	0.737
var: spouse high school	0.275	0.274	0.024	0.051
var: spouse univ educ	0.249	0.271	0.043	0.529
high school				
mean: spouse second educ	2.471	2.481	0.033	0.295
mean: spouse high school	2.527	2.551	0.025	0.984
mean: spouse univ educ	2.721	2.738	0.038	0.455
var: spouse second educ	0.254	0.238	0.020	0.836
var: spouse high school	0.271	0.237	0.017	2.062
var: spouse univ educ	0.172	0.185	0.028	0.449
university education				
mean: spouse second educ	2.653	2.689	0.074	0.475
mean: spouse high school	2.690	2.593	0.043	2.261
mean: spouse univ educ	2.781	2.798	0.035	0.480
var: spouse second educ	0.172	0.195	0.045	0.492
var: spouse high school	0.206	0.180	0.025	1.024
var: spouse univ educ	0.191	0.164	0.022	1.193

Table 15: Male employment: statutory education

Moment	Data	Simulated	SE data	No. SE diff
regression for couples: intercept	1.06860	0.87293	0.02121	9.23
regression for couples: age	0.03887	0.01420	0.00428	5.76
regression for couples: age sq	-0.00049	-0.00076	0.00006	4.47
regression for singles: intercept	0.88852	0.84386	0.07393	0.60
regression for singles: age	0.01244	0.01985	0.01283	0.58
regression for singles: age sq	-0.00019	-0.00174	0.00018	8.61
couples: % time employed	0.85368	0.87108	0.00735	2.37
couples: % time employed < 0.2	0.03745	0.00214	0.00608	5.81
couples: % time employed < 0.4	0.06894	0.01334	0.00782	7.11
couples: % time employed < 0.6	0.12000	0.06755	0.01018	5.15
couples: % time employed < 0.8	0.22638	0.27611	0.01233	4.03
singles: % time employed	0.69148	0.71585	0.04262	0.57
singles: % time employed < 0.2	0.15982	0.04248	0.03287	3.57
singles: % time employed < 0.4	0.23744	0.12655	0.04356	2.55
singles: % time employed < 0.6	0.32877	0.28053	0.05033	0.96
singles: % time employed < 0.8	0.41553	0.53451	0.04873	2.44

Table 16: Male employment: high school education

Moment	Data	Simulated	SE data	No. SE diff
regression for couples: intercept	1.01065	0.92178	0.01423	6.25
regression for couples: age	0.00990	0.00559	0.00291	1.48
regression for couples: age sq	-0.00013	-0.00032	0.00004	4.77
regression for singles: intercept	0.93754	0.90353	0.05288	0.64
regression for singles: age	0.02716	0.01282	0.01022	1.40
regression for singles: age sq	-0.00032	-0.00089	0.00014	4.04
couples: % time employed	0.90267	0.91491	0.00571	2.14
couples: % time employed < 0.2	0.01560	0.00000	0.00373	4.18
couples: % time employed < 0.4	0.03853	0.00258	0.00573	6.27
couples: % time employed < 0.6	0.06972	0.01876	0.00743	6.86
couples: % time employed < 0.8	0.15505	0.15525	0.01113	0.02
singles: % time employed	0.77206	0.87081	0.03885	2.54
singles: % time employed < 0.2	0.09730	0.00404	0.03226	2.89
singles: % time employed < 0.4	0.14054	0.01515	0.03905	3.21
singles: % time employed < 0.6	0.21081	0.06465	0.04191	3.49
singles: % time employed < 0.8	0.35135	0.28788	0.04686	1.35

Table 17: Male employment: university education

Moment	Data	Simulated	SE data	No. SE diff
regression for couples: intercept	1.00585	0.92263	0.01731	4.81
regression for couples: age	0.01140	0.00837	0.00403	0.75
regression for couples: age sq	-0.00014	-0.00032	0.00005	3.65
regression for singles: intercept	1.07402	0.85947	0.07998	2.68
regression for singles: age	0.05276	0.02384	0.01515	1.91
regression for singles: age sq	-0.00065	-0.00125	0.00020	2.98
couples: % time employed	0.91593	0.96071	0.00774	5.79
couples: % time employed < 0.2	0.01099	0.00000	0.00508	2.16
couples: % time employed < 0.4	0.02857	0.00085	0.00734	3.78
couples: % time employed < 0.6	0.05275	0.00593	0.00984	4.76
couples: % time employed < 0.8	0.13187	0.05466	0.01541	5.01
singles: % time employed	0.83326	0.88915	0.02499	2.24
singles: % time employed < 0.2	0.03125	0.00588	0.01964	1.29
singles: % time employed < 0.4	0.05208	0.00588	0.02594	1.78
singles: % time employed < 0.6	0.12500	0.04706	0.03745	2.08
singles: % time employed < 0.8	0.30208	0.26078	0.04236	0.97

Table 18: Female employment: statutory education

Moment	Data	Simulated	SE data	No. SE diff
regression for couples: intercept	0.76869	0.90939	0.02226	6.32
regression for couples: age	0.04346	0.00012	0.00430	10.08
regression for couples: age sq	-0.00046	-0.00048	0.00006	0.40
regression for singles: intercept	0.49194	0.90581	0.06901	6.00
regression for singles: age	-0.01154	-0.00120	0.01303	0.79
regression for singles: age sq	0.00020	-0.00078	0.00017	5.77
couples: % time employed	0.73492	0.75502	0.01019	1.97
couples: % time employed < 0.2	0.10941	0.10332	0.00900	0.68
couples: % time employed < 0.4	0.18308	0.19053	0.01193	0.62
couples: % time employed < 0.6	0.27061	0.29117	0.01325	1.55
couples: % time employed < 0.8	0.38731	0.40240	0.01302	1.16
singles: % time employed	0.63499	0.66898	0.02357	1.44
singles: % time employed < 0.2	0.22517	0.19458	0.02466	1.24
singles: % time employed < 0.4	0.28808	0.28133	0.02688	0.25
singles: % time employed < 0.6	0.37417	0.37470	0.02889	0.02
singles: % time employed < 0.8	0.47682	0.49096	0.02942	0.48

Table 19: Female employment: high school education

Moment	Data	Simulated	SE data	No. SE diff
regression for couples: intercept	0.72012	0.84802	0.02175	5.88
regression for couples: age	-0.01167	0.00798	0.00404	4.86
regression for couples: age sq	0.00020	-0.00031	0.00005	10.11
regression for singles: intercept	0.76698	0.88375	0.05174	2.26
regression for singles: age	-0.00908	0.01084	0.00946	2.11
regression for singles: age sq	0.00013	-0.00077	0.00013	6.93
couples: % time employed	0.80985	0.88491	0.01145	6.56
couples: % time employed < 0.2	0.05478	0.00242	0.00679	7.71
couples: % time employed < 0.4	0.10741	0.01413	0.01090	8.56
couples: % time employed < 0.6	0.18260	0.06519	0.01398	8.40
couples: % time employed < 0.8	0.29538	0.24440	0.01985	2.57
singles: % time employed	0.80274	0.83946	0.01912	1.92
singles: % time employed < 0.2	0.06349	0.01256	0.01507	3.38
singles: % time employed < 0.4	0.11640	0.04348	0.02333	3.13
singles: % time employed < 0.6	0.19577	0.12271	0.02811	2.60
singles: % time employed < 0.8	0.31746	0.34106	0.03239	0.73

Table 20: Female employment: university education

Moment	Data	Simulated	SE data	No. SE diff
regression for couples: intercept	0.74616	0.93515	0.03096	6.10
regression for couples: age	-0.02034	0.00702	0.00576	4.75
regression for couples: age sq	0.00028	-0.00018	0.00008	5.75
regression for singles: intercept	0.95145	0.90494	0.06031	0.77
regression for singles: age	0.01467	0.01228	0.01198	0.20
regression for singles: age sq	-0.00017	-0.00060	0.00016	2.70
couples: % time employed	0.82513	0.99241	0.01459	11.47
couples: % time employed < 0.2	0.05502	0.00000	0.01174	4.69
couples: % time employed < 0.4	0.09091	0.00000	0.01396	6.51
couples: % time employed < 0.6	0.16507	0.00181	0.01913	8.53
couples: % time employed < 0.8	0.23923	0.00724	0.02375	9.77
singles: % time employed	0.88439	0.92647	0.02286	1.84
singles: % time employed < 0.2	0.01111	0.00000	0.01076	1.03
singles: % time employed < 0.4	0.03333	0.00851	0.01858	1.34
singles: % time employed < 0.6	0.11111	0.03617	0.03342	2.24
singles: % time employed < 0.8	0.16667	0.16170	0.04080	0.12

Table 21: Joint distribution of log earnings net of age effects for men with secondary education, by women's education

Moment	Data	Simulated	SE data	No. SE diff
Women's education: secondary				
$\ln W_m \in Q_1$, $\ln W_f \in Q_1$	0.184	0.210	0.020	1.282
$\ln W_m \in Q_1$, $\ln W_f \in Q_2$	0.110	0.094	0.015	1.131
$\ln W_m \in Q_1$, $\ln W_f \in Q_3$	0.023	0.035	0.006	1.888
$\ln W_m \in Q_2$, $\ln W_f \in Q_1$	0.178	0.206	0.020	1.429
$\ln W_m \in Q_2$, $\ln W_f \in Q_2$	0.135	0.111	0.017	1.397
$\ln W_m \in Q_2$, $\ln W_f \in Q_3$	0.062	0.064	0.011	0.167
$\ln W_m \in Q_3$, $\ln W_f \in Q_1$	0.126	0.147	0.017	1.233
$\ln W_m \in Q_3$, $\ln W_f \in Q_2$	0.111	0.082	0.016	1.793
$\ln W_m \in Q_3$, $\ln W_f \in Q_3$	0.070	0.053	0.013	1.351
Women's education: high school				
$\ln W_m \in Q_1$, $\ln W_f \in Q_1$	0.126	0.159	0.018	1.801
$\ln W_m \in Q_1$, $\ln W_f \in Q_2$	0.126	0.083	0.019	2.225
$\ln W_m \in Q_1$, $\ln W_f \in Q_3$	0.038	0.060	0.009	2.535
$\ln W_m \in Q_2$, $\ln W_f \in Q_1$	0.163	0.199	0.017	2.071
$\ln W_m \in Q_2$, $\ln W_f \in Q_2$	0.148	0.118	0.017	1.850
$\ln W_m \in Q_2$, $\ln W_f \in Q_3$	0.114	0.093	0.018	1.189
$\ln W_m \in Q_3$, $\ln W_f \in Q_1$	0.103	0.139	0.018	1.977
$\ln W_m \in Q_3$, $\ln W_f \in Q_2$	0.088	0.080	0.016	0.545
$\ln W_m \in Q_3$, $\ln W_f \in Q_3$	0.094	0.071	0.022	1.045
Women's education: university				
$\ln W_m \in Q_1$, $\ln W_f \in Q_1$	0.036	0.122	0.028	3.060
$\ln W_m \in Q_1$, $\ln W_f \in Q_2$	0.083	0.121	0.048	0.790
$\ln W_m \in Q_1$, $\ln W_f \in Q_3$	0.071	0.124	0.045	1.155
$\ln W_m \in Q_2$, $\ln W_f \in Q_1$	0.250	0.104	0.064	2.271
$\ln W_m \in Q_2$, $\ln W_f \in Q_2$	0.155	0.140	0.054	0.281
$\ln W_m \in Q_2$, $\ln W_f \in Q_3$	0.083	0.153	0.033	2.131
$\ln W_m \in Q_3$, $\ln W_f \in Q_1$	0.036	0.050	0.025	0.572
$\ln W_m \in Q_3$, $\ln W_f \in Q_2$	0.107	0.089	0.045	0.392
$\ln W_m \in Q_3$, $\ln W_f \in Q_3$	0.179	0.097	0.057	1.436

Notes: (Q_1, Q_2, Q_3) denote the 3 thirds of the distribution of earnings. They are measured separately by education for men and women.

Table 22: Joint distribution of log earnings net of age effects for men who completed high school, by women's education

Moment	Data	Simulated	SE data	No. SE diff
Women's education: secondary				
$\ln W_m \in Q_1$, $\ln W_f \in Q_1$	0.133	0.143	0.020	0.514
$\ln W_m \in Q_1$, $\ln W_f \in Q_2$	0.095	0.066	0.017	1.776
$\ln W_m \in Q_1$, $\ln W_f \in Q_3$	0.044	0.046	0.014	0.139
$\ln W_m \in Q_2$, $\ln W_f \in Q_1$	0.198	0.218	0.023	0.851
$\ln W_m \in Q_2$, $\ln W_f \in Q_2$	0.175	0.114	0.023	2.703
$\ln W_m \in Q_2$, $\ln W_f \in Q_3$	0.065	0.088	0.014	1.604
$\ln W_m \in Q_3$, $\ln W_f \in Q_1$	0.113	0.175	0.014	4.548
$\ln W_m \in Q_3$, $\ln W_f \in Q_2$	0.096	0.082	0.015	0.918
$\ln W_m \in Q_3$, $\ln W_f \in Q_3$	0.080	0.068	0.016	0.716
Women's education: high school				
$\ln W_m \in Q_1$, $\ln W_f \in Q_1$	0.109	0.154	0.014	3.357
$\ln W_m \in Q_1$, $\ln W_f \in Q_2$	0.088	0.091	0.012	0.224
$\ln W_m \in Q_1$, $\ln W_f \in Q_3$	0.057	0.069	0.010	1.160
$\ln W_m \in Q_2$, $\ln W_f \in Q_1$	0.161	0.207	0.015	3.083
$\ln W_m \in Q_2$, $\ln W_f \in Q_2$	0.143	0.122	0.014	1.438
$\ln W_m \in Q_2$, $\ln W_f \in Q_3$	0.106	0.099	0.015	0.503
$\ln W_m \in Q_3$, $\ln W_f \in Q_1$	0.126	0.100	0.016	1.668
$\ln W_m \in Q_3$, $\ln W_f \in Q_2$	0.108	0.080	0.013	2.211
$\ln W_m \in Q_3$, $\ln W_f \in Q_3$	0.101	0.077	0.016	1.551
Women's education: university				
$\ln W_m \in Q_1$, $\ln W_f \in Q_1$	0.117	0.148	0.027	1.168
$\ln W_m \in Q_1$, $\ln W_f \in Q_2$	0.128	0.179	0.034	1.536
$\ln W_m \in Q_1$, $\ln W_f \in Q_3$	0.117	0.177	0.029	2.101
$\ln W_m \in Q_2$, $\ln W_f \in Q_1$	0.069	0.088	0.014	1.291
$\ln W_m \in Q_2$, $\ln W_f \in Q_2$	0.231	0.119	0.041	2.734
$\ln W_m \in Q_2$, $\ln W_f \in Q_3$	0.153	0.135	0.030	0.585
$\ln W_m \in Q_3$, $\ln W_f \in Q_1$	0.036	0.053	0.012	1.471
$\ln W_m \in Q_3$, $\ln W_f \in Q_2$	0.078	0.052	0.024	1.109
$\ln W_m \in Q_3$, $\ln W_f \in Q_3$	0.072	0.048	0.018	1.299

Notes: (Q_1, Q_2, Q_3) denote the 3 thirds of the distribution of earnings. They are measured separately by education for men and women.

Table 23: Joint distribution of log earnings net of age effects for men with university education, by women's education

Moment	Data	Simulated	SE data	No. SE diff
Women's education: secondary				
$\ln W_m \in Q_1$, $\ln W_f \in Q_1$	0.144	0.158	0.057	0.241
$\ln W_m \in Q_1$, $\ln W_f \in Q_2$	0.122	0.063	0.058	1.022
$\ln W_m \in Q_1$, $\ln W_f \in Q_3$	0.022	0.058	0.018	2.022
$\ln W_m \in Q_2$, $\ln W_f \in Q_1$	0.089	0.238	0.036	4.113
$\ln W_m \in Q_2$, $\ln W_f \in Q_2$	0.189	0.072	0.056	2.081
$\ln W_m \in Q_2$, $\ln W_f \in Q_3$	0.044	0.090	0.037	1.231
$\ln W_m \in Q_3$, $\ln W_f \in Q_1$	0.222	0.221	0.074	0.016
$\ln W_m \in Q_3$, $\ln W_f \in Q_2$	0.100	0.045	0.036	1.523
$\ln W_m \in Q_3$, $\ln W_f \in Q_3$	0.067	0.055	0.035	0.329
Women's education: high school				
$\ln W_m \in Q_1$, $\ln W_f \in Q_1$	0.145	0.156	0.035	0.301
$\ln W_m \in Q_1$, $\ln W_f \in Q_2$	0.046	0.050	0.015	0.268
$\ln W_m \in Q_1$, $\ln W_f \in Q_3$	0.008	0.023	0.004	3.530
$\ln W_m \in Q_2$, $\ln W_f \in Q_1$	0.176	0.280	0.024	4.356
$\ln W_m \in Q_2$, $\ln W_f \in Q_2$	0.145	0.086	0.025	2.367
$\ln W_m \in Q_2$, $\ln W_f \in Q_3$	0.074	0.059	0.019	0.790
$\ln W_m \in Q_3$, $\ln W_f \in Q_1$	0.156	0.231	0.029	2.609
$\ln W_m \in Q_3$, $\ln W_f \in Q_2$	0.145	0.069	0.026	2.937
$\ln W_m \in Q_3$, $\ln W_f \in Q_3$	0.105	0.046	0.026	2.249
Women's education: university				
$\ln W_m \in Q_1$, $\ln W_f \in Q_1$	0.137	0.176	0.021	1.820
$\ln W_m \in Q_1$, $\ln W_f \in Q_2$	0.111	0.110	0.020	0.064
$\ln W_m \in Q_1$, $\ln W_f \in Q_3$	0.052	0.089	0.015	2.550
$\ln W_m \in Q_2$, $\ln W_f \in Q_1$	0.166	0.192	0.025	1.031
$\ln W_m \in Q_2$, $\ln W_f \in Q_2$	0.193	0.119	0.026	2.801
$\ln W_m \in Q_2$, $\ln W_f \in Q_3$	0.108	0.092	0.021	0.791
$\ln W_m \in Q_3$, $\ln W_f \in Q_1$	0.061	0.104	0.016	2.716
$\ln W_m \in Q_3$, $\ln W_f \in Q_2$	0.085	0.070	0.015	0.945
$\ln W_m \in Q_3$, $\ln W_f \in Q_3$	0.086	0.049	0.019	1.944

Notes: (Q_1, Q_2, Q_3) denote the 3 thirds of the distribution of earnings. They are measured separately by education for men and women.

References

[1] Attanasio, Orazio, Hamish Low and Virginia Snchez-Marcos (2008) Explaining Changes in Female Labor Supply in a Life-Cycle Model, The American Economic Review, Vol. 98, No. 4 (Sep., 2008), pp. 1517-1552

[2] Becker, Gary S. (1974) A Theory of Social Interactions Journal of Political Economy, Vol. 82, No. 6 (Nov. - Dec., 1974), pp. 1063-1093

[3] Becker, Gary S. (1991) A Treatise on the Family, Harvard University Press

[4] Blundell, R, Monica Costa Dias, Costas Meghir Jonathan Shaw (2015) Marriage, Social Insurance and Labor Supply, NBER Working Paper.

[5] Blundell, Richard W., Pierre-André Chiappori, Thierry Magnac and Costas Meghir (2007) Collective Labour Supply: Heterogeneity and Non-Participation The Review of Economic Studies, Vol. 74, No. 2 (Apr., 2007), pp. 417-445

[6] Blundell, Richard W., Pierre André Chiappori, and Costas Meghir (2005) Collective Labor Supply with Children Journal of Political Economy, Vol. 113, No. 6 (December 2005), pp. 1277-1306

[7] Browning, Martin and P. A. Chiappori (1998) Efficient Intra-Household Allocations: A General Characterization and Empirical Tests Econometrica, Vol. 66, No. 6 (Nov., 1998), pp. 1241-1278

[8] Browning, Martin, P.A. Chiappori, and Y. Weiss, 2013, *Family Economics,* Cambridge University Press.

[9] Chiappori, Pierre-André (1988) Rational Household Labor Supply Econometrica, Vol. 56, No. 1 (Jan., 1988), pp. 63-90

[10] Chiappori, Pierre-André (1991) Nash-Bargained Household Decisions: A Rejoinder International Economic Review, Vol. 32, No. 3 (Aug., 1991), pp. 761-762

[11] Chiappori, Pierre-André (1992) Collective Labor Supply and Welfare Journal of Political Economy, Vol. 100, No. 3 (Jun., 1992), pp. 437-467

[12] Chiappori, Pierre-Andr, Murat Iyigun, and Yoram Weiss (2009) Investment in Schooling and the Marriage Market, American Economic Review 2009, 99:5, 1689?1713

[13] Chiappori, Pierre-André and B. Salanié (2015) "The Econometrics of Matching Models", *Journal of Economic Literature*

[14] Choo, Eugene and Aloysius Siow (2006) Who Marries Whom and Why, Journal of Political Economy, Vol. 114, No. 1 (February 2006), pp. 175-201

[15] Goussé, M. (2013) "Marriage Market and Intra-Household Allocation : Evolution of preferences and transfers in the UK from 1991 to 2008", mimeo Sciences Po

[16] Jacquemet, Nicolas and Jean-Marc Robin (2011) Marriage with Labor Supply, CES Working Papers 2011.50

[17] Low, Hamish, Costas Meghir, and Luigi Pistaferri (2010) Wage Risk and Employment Risk over the Life Cycle, American Economic Review 100 (September 2010): 1432?1467

[18] Low, Hamish, Costas Meghir, Luigi Pistaferri, Alessandra Voena (2015), Marriage, Social Insurance and Labor Supply, mimeo.

[19] Mazzocco, M. (2007) "Household Inter Temporal Behaviour: A Collective Characterization and a Test of Commitment" *The Review of Economic Studies* Vol. 74, No. 3 (Jul., 2007) , pp. 857-895

[20] Shimer Robert and Lones Smith (2000) Assortative Matching and Search Econometrica, Vol. 68, No. 2 (Mar., 2000), pp. 343-369

[21] Schulhofer-Wohl, Sam (2006), "Negative Assortative Matching of Risk-Averse Agents with Transferable Expected Utility," *Economics Letters* 92(3), 383–88.

[22] Voena, A. (2013) "Yours, Mine and Ours: Do Divorce Laws Affect the Intertemporal Behavior of Married Couples?", mimeo University of Chicago