THE MARRIAGE MARKET, LABOR SUPPLY AND EDUCATION CHOICE

By

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

March 2015
Revised July 2016

COWLES FOUNDATION DISCUSSION PAPER NO. 1994R

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[‡]

July 18, 2016

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 commodities 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. We thank six anonymous referees for their detailed and constructive comments, as well as participants in seminars at the University of Chicago, the Becker Friedman Institute conference in honor of Gary Becker and the conference in honor of Martin Browning. Monica Costa 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, IZA and IFS

1 Introduction

1.1 Matching on human capital

The present paper explores the intersection of two fundamental Beckerian concepts: 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 crucial 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. Education policy, 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 characteristics - ability, cost of investment (which may for instance reflect borrowing constraints) and preferences for marriage - 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 joint distribution of human capital and marital preferences 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. Similar arguments can be made for education policy or welfare reform. Imperfect as it may be, our approach is the first to consider all these aspects in a unified and theoretically consistent framework.

In the next subsection we discuss some of the existing literature. Then in Sections 2 and 3 we present the model and develop its solution. Section 4 discusses the identification of the distribution of marital preferences. In Section 5 we present the data and in Section 6 we present our estimation approach. In Sections 7 and 8 we discuss the empirical results and a counterfactual simulation. Section 9 concludes and discusses future avenues of research.

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 there is no time/dynamic dimension. This restriction is 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. (2016) 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. Goussé, Jacquemet and Robin (2016) specify an equilibrium model of marriage and labor supply based on search 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.[1]

Our model is also related to recent developments on the econometrics of 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

[1] Theoretical models with prematch investments include Cole, Mailath and Postlewaite (2001), Nöldeke and Samuelson (2015) and Peters and Siow (2002).

joint distribution of education and ability, therefore of human capital ; interestingly, these distributions are sufficient to fully characterize the equilibrium. 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 us to identify preferences, therefore the surplus function. The matching model, therefore, is over identified, and allows us to recover the intra-couple allocation 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,[2] 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 draw a vector of marital preferences and invest in human capital by choosing one of three educational levels. This investment depends

[2] See for instance CIW.

on their innate ability and their cost of education, as well as on the perceived benefits of this investment, including the benefits to be received on the marriage market; the latter, in turn, are directly influenced by marital preferences. 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) $H\left(s, \theta\right)$, 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 characterize his/her prospects in the marriage and labour markets.

In stage 2, individuals enter the marriage market; the latter is modeled as a frictionless matching process based on the level of human capital, and on 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 T periods of i's third stage of life by consumptions and labour supply; the second 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 non-market time 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:[3]

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

Solving this program gives the optimal choice of aggregate household 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.

From a cardinal perspective, 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 and V_f denote their respective, lifetime expected utility. There exists a function $\Upsilon(\mathbf{H})$, where $\mathbf{H} = (H_m, H_f)$, such that the set of

[3] 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 $\exp u_i (Q_t, C_{it}, L_{it}) + \mu \exp u_j (Q_t, C_{jt}, L_{jt}) \geq \bar{u}_j$ under a budget constraint. Here, the maximand is equal to

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

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 $\exp u_i$.

Pareto efficient allocations is characterized by:

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

This function is explicitly derived in Section 3. The crucial point, 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 (expected) utility. If we define

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

the previous relationship becomes:

$$\bar{U}_m+\bar{U}_f=\Gamma\left(H_m,H_f\right) \tag{5}$$

which shows that we are in a TU context even *ex-ante*, since the Pareto frontier is, for these utility indices, a straight line with slope -1 for all wages and incomes. The function $\Gamma\left(H_m,H_f\right)$, 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 Γ (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 $V_m^s\left(H_m\right)$ and $V_f^s\left(H_f\right)$ the respective lifetime economic utility of a single male (female) with human capital

H_m (H_f). 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:

$$\bar{U}_i^{\varsigma} = \exp\left\{\frac{1-\delta}{1-\delta^T} V_i^{\varsigma}\left(H_i\right)\right\} \tag{6}$$

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, $\Gamma\left(H_m, H_f\right)$, 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

$$\Sigma\left(H_m, H_f\right) = \Gamma\left(H_m, H_f\right) - \bar{U}_m^{\varsigma} - \bar{U}_f^{\varsigma}. \tag{7}$$

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 investing, agent i draws a vector $\beta_i = (\beta_i^0, \beta_i^H$ where $H \in \mathcal{H})$, 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

$\Gamma(H_m, H_f)$ defined above and the idiosyncratic preference shocks β:

$$\Gamma_{mf} = \Gamma(H_m, H_f) + \beta_m^{H_f} + \beta_f^{H_m} \tag{8}$$

and the resulting surplus is:

$$\Sigma_{mf} = \Sigma(H_m, H_f) + \left(\beta_m^{H_f} - \beta_m^0\right) + \left(\beta_f^{H_m} - \beta_f^0\right) \tag{9}$$

Again, the function $\Sigma(H_m, H_f)$ 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.[4] 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 us to fully characterize the stochastic distribution of individual utilities at the stable match (see CSW and Chiappori and Salanié 2015).[5]

[4]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 capital.

[5]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.

2.4 Second stage matching game

At the end of the first stage, agents are each characterised by their marital preferences, but also by their human capital H, a function of their innate ability θ and education s. From the second perspective, the male and female populations are therefore distributed over the space $\mathcal{H}$, which consists of $N \times J$ classes. They then enter a matching game under TU, in which the surplus function for any potential match is given by (9).

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:

$$U_m + U_f = \Gamma_{mf}$$

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. 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:[6]

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

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

Theorem 1. ***(Choo and Siow, 2006; Chiappori, Salanié and Weiss, 2015)*** *If the surplus is given by (9), then there exist* $2\left(NJ\right)^2$ *numbers -* $\bar{U}_M\left(H_m, H_f\right)$ *and* $\bar{U}_F\left(H_m, H_f\right)$ *for*

[6]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.

$(H_m, H_f) \in \mathcal{H}^2$ *- such that:*

1. *For any* (H_m, H_f)
$$\bar{U}_M(H_m, H_f) + \bar{U}_F(H_m, H_f) = \Gamma(H_m, H_f) \tag{10}$$

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

$$\begin{aligned} U_m &= \bar{U}_M(H_m, H_f) + \beta_m^{H_f} \text{ and} \\ U_f &= \bar{U}_F(H_m, H_f) + \beta_f^{H_m} \end{aligned} \tag{11}$$

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 to be 0.

Note that the characterization of utilities provided by (11) 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:

$$V_i = \frac{1-\delta^T}{1-\delta} \ln(U_i) = \frac{1-\delta^T}{1-\delta} \ln\left(\bar{U}_{g_i}(H_m, H_f) + \beta_i^{H_j}\right) \tag{12}$$

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 $H \in \mathcal{H} \cup \{0\}$:*

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

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 $H \in \mathcal{H} \cup \{0\}$:

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

3. The utility of a man m with human capital H_m and preferences β_m is:

$$A_M(H_m) = \max_{H_f \in \mathcal{H} \cup \{0\}} \left(\bar{U}_M \left(H_m, H_f\right) + \beta_m^{H_f}\right) \tag{13}$$

and the utility of a female agent f with human capital H_f and preferences β_m is:

$$A_F(H_f) = \max_{H_m \in \mathcal{H} \cup \{0\}} \left(\bar{U}_F \left(H_m, H_f\right) + \beta_f^{H_m}\right) \tag{14}$$

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 (10); 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

$$\text{for man } m \quad : \quad s_m = \arg\max_{s \in \mathcal{S}} \{A_M(H(s, \theta_m)) - c_s(X_m, \upsilon_{ms})\} \tag{15}$$

$$\text{for woman } f \quad : \quad s_f = \arg\max_{s \in \mathcal{S}} \{A_F(H(s, \theta_f)) - c_s(X_f, \upsilon_{fs})\} \tag{16}$$

where A_M and A_F are defined in equations 13 and 14 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.

Lastly, the structure of that stage is a simultaneous move game: agents each choose their education independently, but the payoffs they will receive depend on the human capital distribution on both sides of the market, which results from other players' investment. This, potentially, raises existence and uniqueness issues that will be discussed below.

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 us 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/age t sees the arrival of new information on each spouse's preferences for working and productivity, $\boldsymbol{\alpha}_t = (\alpha_{mt}, \alpha_{ft})$ and $\mathbf{e}_t = (e_{mt}, e_{ft})$. 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, $\mathbf{H} = (H_m, H_f)$. Given the information set $(\boldsymbol{\alpha}_t, \mathbf{e}_t, K_{t-1}, \mathbf{H})$, we first consider the couple's consumption decisions conditional on savings and employment, K_t and $\mathbf{L}_t = (L_{mt}, L_{ft})$. 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:

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

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

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

where consumptions are written as functions of $(K_t, \mathbf{L}_t)$ 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 relates to sharing the wage and preference risks. In this context, it requires the maximization of a weighted sum of expected utilities, using the *initial*, logarithmic cardinalization, which reflects preferences towards risk. 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:[7]

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

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:

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

[7] See for example Townsend (1994).

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

$$v_{mt} = 2\ln Q_t\left(K_t, \mathbf{L}_t\right) + \ln p + \ln\frac{1}{1+\mu} \tag{17}$$
$$v_{ft} = 2\ln Q_t\left(K_t, \mathbf{L}_t\right) + \ln p + \ln\frac{\mu}{1+\mu} \tag{18}$$

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

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

$$\begin{aligned}
E_{T|T-1}V_{mT}\left(\mathbf{e}_T, \boldsymbol{\alpha}_T, K_{T-1}, \mathbf{H}, \mu\right) &= I_T\left(\mathbf{e}_{T-1}, \boldsymbol{\alpha}_{T-1}, K_{T-1}, \mathbf{H}\right) + \ln\frac{1}{1+\mu} \\
E_{T|T-1}V_{fT}\left(\mathbf{e}_T, \boldsymbol{\alpha}_T, K_{T-1}, \mathbf{H}, \mu\right) &= I_T\left(\mathbf{e}_{T-1}, \boldsymbol{\alpha}_{T-1}, K_{T-1}, \mathbf{H}\right) + \ln\frac{\mu}{1+\mu} \\
\text{where } I_T\left(\mathbf{e}_{T-1}, \boldsymbol{\alpha}_{T-1}, K_{T-1}, \mathbf{H}\right) &= E_{T|T-1}\max_{L_T, K_T}\left[2\ln Q_T\left(\mathbf{L}_T, K_T\right) + \ln p \mid \mathbf{e}_{T-1}, \boldsymbol{\alpha}_{T-1}\right]
\end{aligned}$$

where expectations are taken over the (education-specific) distribution of $(\mathbf{e}_t, \boldsymbol{\alpha}_t)$ 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 (17)-(18), it is easy to show by recursion that the additive separability of the Pareto weight carries over to earlier periods:

$$\begin{aligned}
E_{t|t-1}V_{mt}\left(\mathbf{e}_t, \boldsymbol{\alpha}_t, K_{t-1}, \mathbf{H}, \mu\right) &= I_t\left(\mathbf{e}_{t-1}, \boldsymbol{\alpha}_{t-1}, K_{t-1}, \mathbf{H}\right) + \ln\left(\frac{1}{1+\mu}\right)\sum_{\tau=t}^{T}\delta^{\tau-t} \\
E_{t|t-1}V_{ft}\left(\mathbf{e}_t, \boldsymbol{\alpha}_t, K_{t-1}, \mathbf{H}, \mu\right) &= I_t\left(\mathbf{e}_{t-1}, \boldsymbol{\alpha}_{t-1}, K_{t-1}, \mathbf{H}\right) + \ln\left(\frac{\mu}{1+\mu}\right)\sum_{\tau=t}^{T}\delta^{\tau-t}
\end{aligned}$$

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

$$I_t\left(\mathbf{e}_{t-1},\boldsymbol{\alpha}_{t-1},K_{t-1},\mathbf{H}\right)=E_{t|t-1}\max_{L_t,K_T}\left[2\ln Q_t\left(\mathbf{L}_t,K_t\right)+\ln p+\delta I\left(\mathbf{e}_t,\boldsymbol{\alpha}_t,K_t,\mathbf{H}\right)\mid \mathbf{e}_{t-1},\boldsymbol{\alpha}_{t-1}\right]$$

where expectations are taken over the (education-specific) distribution of $(\mathbf{e}_t,\boldsymbol{\alpha}_t)$ conditional on $(\mathbf{e}_{t-1},\boldsymbol{\alpha}_{t-1})$. 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 $(K_t^*(\mathbf{L}_t))$ satisfy:

$$2\frac{\partial\ln Q_t\left(K_t,\mathbf{L}_t\right)}{\partial K_t}+\delta\frac{\partial I_{t+1}\left(\mathbf{e}_t,\boldsymbol{\alpha}_t,K_t,\mathbf{H}\right)}{\partial K_t}=0$$

Finally, the optimal choice of labor supplies are defined by

$$\left(L_{mt}^*,L_{ft}^*\right)=\underset{\mathbf{L}_t\in\{0,1\}^2}{\arg\max}\left\{2\ln Q_{it}\left(K_t^*(\mathbf{L}_t),\mathbf{L}_t\right)+\ln p+\delta I_{t+1}\left(\mathbf{e}_t,\boldsymbol{\alpha}_t,K_t^*(\mathbf{L}_t),\mathbf{H}\right)\right\}.$$

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 $(\mathbf{e}_t,\boldsymbol{\alpha}_t)$ 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 $\Upsilon\left(\mathbf{H}\right)$ and $\Upsilon^S\left(H_i\right)$. It follows that the expected *economic* utility, at marriage, of each spouse is given by:

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

$$\text{and} \quad V_f(\mathbf{H},\mu) = \Upsilon(\mathbf{H}) + \left(\sum_{\tau=0}^{T-t}\delta^\tau\right)\ln\left(\frac{\mu}{1+\mu}\right) \tag{20}$$

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:

$$V^\varsigma(H_i) = \Upsilon^\varsigma(H_i)$$

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 $(\boldsymbol{\alpha},\mathbf{e})$ 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 (19)-(20). However, an alternative cardinalization, already introduced in (4), turns out to be more convenient here. Specifically, define $\bar{U}_i$ by:

$$\bar{U}_i = \exp\left(\frac{1-\delta}{1-\delta^T}V_i\right) \tag{21}$$

then if $\mathbf{H} = (H_m, H_f)$:

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

and finally:

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

which expresses that the sum of individual, *economic* utilities add up to the marital gain $\Gamma(\mathbf{H})$.

Lastly, we can add the idiosyncratic shocks to both sides of this equation; we finally have that, for any married couple $\mathbf{H} = (H_m, H_f)$:

$$\bar{U}_m + \beta_m^{H_f} + \bar{U}_f + \beta_f^{H_m} = \Gamma(\mathbf{H}) + \beta_m^{H_f} + \beta_f^{H_m} = g_{mf} \tag{22}$$

The matching game, therefore, has a *transferable utility* structure: if the utility of person i is represented by the particular cardinal representation $\left(\bar{U}_i\right)$, 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 Γ, given iid shocks (β_m, β_f). One can easily check that the sign of the second derivative $\partial^2\Gamma/\partial H_m \partial H_f$ is indeterminate (and can be either positive or negative depending on the parameters); so this needs to be investigated empirically.[8]

[8] A result due to Graham (2011) states that, for the iid stochastic structure, for any two levels H and $\bar{H}$ of human capital, the total number of 'assortative couples' (i.e., $H$and H or $\bar{H}$and $\bar{H}$) will exceed what would be expected under purely random matching if and only if the deterministic function Γ is supermodular for H and $\bar{H}$ - i.e.:

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

That is what is meant by assortative matching.

Clearly, one can equivalently use any of the two cardinalizations described before to study marital sorting, because under TU matching patterns are driven by ordinal preferences; 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*

Proof. From (11) in Theorem 1, 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

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

we conclude that μ only depends on (H_m, H_f). $\square$

3.3 The first stage in the lifecycle: 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 realization of random marital preferences and the costs of education, which can include exogenous shifters. Given this, the education choice is described in equation (15)-(16).

As mentioned earlier, the first stage can be modeled as a normal form game, where each player's payoff depends on the other players' decisions. As such, existence has to be demonstrated; and neither uniqueness nor efficiency are guaranteed. We now discuss these issues.

The central idea, due to Cole, Mailath and Postlewaite (2001) and Nöldeke and Samuelson (2015), is to consider what we shall call an *auxilliary game*, defined as follows. Assume that stages 1 (investment) and 2 (matching), instead of taking place sequentially, are simultaneous. That is, consider the *two* stage game in which:

- At stage 1, agents match (based on their idiosyncratic charateristics, namely ability, education costs and marital preferences), and choose their education. In particular, matched pairs *jointly* (and efficiently) choose their respective investments in human capital.

- Stage 2 (the 'working life' stage) is identical to stage 3 of the initial game; i.e., it is divided into T periods, during which individuals, whether single or married, observe their (potential) wage and non labor income, and decide on consumptions and labor supplies.

Again, the auxiliary game can be solved by backwards induction. The behavior of a given couple is described as before; in particular, and using the same cardinalizations as in Subsection 3.2, the function Γ defined by (22) still characterizes the total surplus generated by a given match. Then stage 1 is a typical matching model under transferable utility. A key remark is that both utility when single and surplus when married are continuous functions of all characteristics. The main result is:

Proposition *Any stable match of the auxiliary game is a Nash equilibrium of the initial game.*

Proof. This is a direct consequence of Propositions 1 and 2 in Nöldeke and Samuelson (2015); an equilibrium in the initial game corresponds to Nöldeke and Samuelson's notion of an ex post equilibrium, whereas a stable match of the auxiliary game is an ex ante equilibrium. The only condition is continuity of the payoff functions, which is guaranteed in our case. □

This result, in turn, has several implications. The most obvious one regards existence. The Proposition implies that whenever there exists a stable match of the auxiliary game, then there exists a Nash equilibrium of the initial game. Moreover, the auxiliary game is a standard matching game under TU; stability, in this context, is equivalent to surplus maximization. In the end, the existence of a Nash equilibrium in the initial game boils down to the existence of a solution to an optimal transportation problem (i.e., finding a measure that maximizes total surplus subject to conditions on the marginals). It is well known[9] that such existence obtains under mild continuity and compactness conditions that are satisfied in our case.

Uniqueness is a more difficult issue. Note, first, that the stable match of the auxiliary game is 'generically' unique, in the sense that a maximization problem has 'in general' a unique solution: while it is always possible to construct situations in which the maximum is reached for different solutions, such cases are in general not robust to small perturbations.[10] This, however, does *not* imply uniqueness of the Nash equilibrium in the initial game. Indeed, Cole, Mailath and Postlewaite (2001) and Nöldeke and Samuelson (2015) provide examples of 'coordination failures', whereby an alternative, Nash equilibrium of the game involves all agents investing in a globally suboptimal, but individually rational way.[11] One intuition is that, because of supermodularity, an agent's optimal investment is typically an increasing function

[9] See for instance Chiappori, McCann and Nesheim (2010) and Chiappori, McCann and Pass (2015)

[10] A precise definition of the 'genericity' concept invoked in this - admitedly vague - statement would require transversality arguments in functional spaces, which would be well beyond the scope of this paper.

[11] Nöldeke and Samuelson (2015) provide a set of conditions that are sufficient for uniqueness of ex post equilibria. These conditions, however, are quite restrictive and cannot be expected to hold in our context.

of the other agents' education. In a context where the other agents underinvest, a person's best response is typically to underinvest as well, and these best responses may sometimes form a Nash equilibrium.

While the study of coordination failures is an interesting topic, we will not pursue it in the present context. From an empirical perspective, there is little evidence of global underinvestment in education. Theoretically, in the presence of (potential) multiple equilibria, a natural solution is to use an equilibrium refinement concept. In our case, a natural criterion is Pareto efficiency, because, unlike most games, one of the equilibria is always Pareto efficient. In what follows, we shall therefore concentrate on the ('generically' unique) stable matching of the auxilliary game as the relevant Nash equilibrium.

4 Identification of the distribution of marital preferences.

The model as presented now requires a distributional assumption on marital preferences 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 (9) is thus replaced with:

$$\Sigma_{mf} = \Sigma\left(H_m, H_f\right) + \left(X_m a^{H_m,H_f} + \beta_m^{H_f} - \beta_m^0\right) + \left(X_f b^{H_m,H_f} + \beta_f^{H_m} - \beta_f^0\right) \tag{23}$$

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 $\beta_m^{H_f} - \beta_m^0$ and $\beta_f^{H_m} - \beta_f^0$ (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 (23), then there exist* $2\left(NJ\right)^2$ *numbers -* $\bar{U}_M\left(H_m, H_f\right)$ *and* $\bar{U}_F\left(H_m, H_f\right)$ *for* $\left(H_m, H_f\right) \in \mathcal{H}^2$ *- such that:*

1. *For any* $\left(H_m, H_f\right)$

$$\bar{U}_M\left(H_m, H_f\right) + \bar{U}_F\left(H_m, H_f\right) = \Gamma\left(H_m, H_f\right)$$

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:

$$U_m + U_f = \Gamma(H_m, H_f) + X_m a^{H_m,H_f} + \beta_m^{H_f} + X_f b^{H_m,H_f} + \beta_f^{H_m} \tag{24}$$

$$U_m + U_{f'} \geq \Gamma(H_m, H_f) + X_m a^{H_m,H_f} + \beta_m^{H_f} + X_{f'} b^{H_m,H_f} + \beta_{f'}^{H_m} \tag{25}$$

$$U_{m'} + U_{f'} = \Gamma(H_m, H_f) + X_{m'} a^{H_m H_f} + \beta_{m'}^{H_f} + X_{f'} b^{H_m H_f} + \beta_{f'}^{H_m} \tag{26}$$

$$U_{m'} + U_f \geq \Gamma(H_m, H_f) + X_{m'} a^{H_m H_f} + \beta_{m'}^{H_f} + X_f b^{H_m H_f} + \beta_f^{H_m} \tag{27}$$

Subtracting (24) from (25) and (27) from (26) gives

$$U_{f'} - U_f \geq (X_{f'} - X_f) b^{H_m,H_f} + \beta_{f'}^{H_m} - \beta_f^{H_m} \geq U_{f'} - U_f \tag{28}$$

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.:

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

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 $H \in \mathcal{H} \cup \{0\}$:*

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

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 $H \in \mathcal{H} \cup \{0\}$:

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

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

$$A_M(H_m) = \mathbb{E}\left[\max_{H_f \in \mathcal{H}\cup\{0\}} \left(\bar{U}_M\left(H_m, H_f\right) + X_m a^{H_m,H_f} + \beta_m^{H_f}\right)\right] \tag{29}$$

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

$$A_F(H_f) = \mathbb{E}\left[\max_{H_m \in \mathcal{H}\cup\{0\}} \left(\bar{U}_F\left(H_m, H_f\right) + X_f b^{H_m,H_f} + \beta_f^{H_m}\right)\right] \tag{30}$$

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_f) 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 residual parental income, when the child was 16, after removing the effects of parental background (see below). The key idea is that children need to be at least in part financed by their parents and if the latter suffer an adverse liquidity shock this may inhibit educational attainment.

5 Data

Estimation uses the 18 annual waves (1991 to 2008) of the British Household Panel Survey (BHPS), which includes interviews with all household members over 16 and follows them even when they start their own household.

We select two sub-samples drawn from the original members of the panel and those added later. The main sample comprises longitudinal information for individuals born between 1951 and 1971 between the ages of 25 and 50.[12] To this sample we add information on the spouses they marry during the observation window. To avoid under-estimating marriage rates, those who are not observed past age 30 are dropped from the sample. Overall, the final dataset contains information on education, employment, earnings and family demographics for 4,295 families, 3,046 of which are couples and 629 and 620 are single women and men, respectively. Of these, over 60% are observed for at least 5 years. We exclude Northern Ireland as it is only

[12] For couples, we take the reference year of birth to be that of the wife.

surveyed in the booster samples, during the 2000s. In total, the sample size is just over of 41,000 observations.

In the resulting sample, singles are defined as individuals who are never observed married or cohabiting. For the rest, who are classified as married, we only use observations during their first observed marriage.

In estimating educational participation we use parental income observed when the child is 16.[13] This information is only available for individuals who are observed living with their parents at that age. So for this part of the model we use an additional smaller sample of individuals, born between 1973 and 1985, containing parental income information when the young respondent is aged between 16 and 18 and completed education by the age of 23. This sample includes 1,245 individuals, 636 of whom are women.

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 5%-98% of the distribution of pre-tax real earnings for employees only. Since our model does not deal with macroeconomic growth and fluctuations, we subtract aggregate earnings growth from earnings. Finally, we consider 3 education levels, corresponding to secondary education (leaving school at 16), high school and university (college) degree.

[13] We also need family background, since we actually use the residual parental income as we explain in the estimation section.

6 Empirical specification and estimation

6.1 Outline of estimation

We estimate the model in three steps. In a first step we estimate the age profiles of weekly earnings (interchangeably referred to as wages) by gender and education. This is done outside the model, based on the control function approach to allow for endogenous selection into work and for the endogeneity of education. For the former we use policy variation in out-of-work income as an instrument. For the latter we use the residual from a regression of parental income when the person was 16 on their family background characteristics.

The next two steps take these wage profiles as given and are performed within the model. Since we assume that preferences for work are drawn *after* the matching stage, we can separately estimate the lifecycle model post marriage, exploiting the TU structure (which implies that lifecycle labor supply, the public good and household consumption do not depend on the Pareto weights). Given estimates for preferences and the distribution of unobserved ability, we can then estimate the economic value of marriage for each type of match (by ability and education - which define human capital) and for singles. In a final stage, taking these values as given, we can estimate the preferences that drive marriage, the parameters driving the costs of education and the implied Pareto weights. We now provide details on this procedure as well as our specification.

6.2 Earnings process

The earnings w_{it} of Individual (i) vary by gender (g), education (s), ability (θ) and age (t). We thus estimate the following earnings equation:

$$\ln w_{it} = \ln W(\theta_i, s, g) + \delta_1^{gs} t_i + \delta_2^{gs} t_i^2 + \delta_3^{gs} t_i^3 + e_{it} + \epsilon_{it} \tag{31}$$

$$e_{it} = \rho^{gs} e_{it-1} + \xi_{it} \tag{32}$$

where e is the productivity shock, assumed to follow an AR(1) process with normal innovations ξ_{it} whose variance is $\sigma^2_{\xi,gs}$. ϵ_{it} is an independently and identically distributed shock with variance $\sigma^2_{\epsilon,gs}$ that we interpret as measurement error. $W(\theta_i, s, g)$ is the market wage faced by an individual of ability type (θ_i), gender g and schooling s. Ability is assumed to follow a distribution with two points of support. While this can be viewed as an approximation from the econometric point of view, it also simplifies the marital matching problem by defining a finite number of individual types.[14]

In a first step we estimate the education and gender specific age profiles using a control function approach as in Heckman (1979) to allow for endogenous selection into employment and for the endogeneity of education. For this purpose we use a reduced form binary choice model for employment driven by an index $z_1'\beta_E$. The education reduced form is taken to be an unordered discrete choice among three levels (Secondary, High School and University). This choice is driven by two separate indices $z_2'\beta_{HS}$ and $z_2'\beta_U$ as in a random utility model with three alternatives. z_1 are the instruments for employment and z_2 are the instruments for education choice. We then use the regression

$$\ln \tilde{w}_{it} = \delta_0^{gs} + \delta_1^{gs} t + \delta_2^{gs} t^2 + \delta_3^{gs} t^3 + \lambda_E(z_1'\beta_E) + \lambda_{ed}(z_2'\beta_{HS}, z_2'\beta_U) + v_{it} \tag{33}$$

where $\ln \tilde{w}_{it}$ are detrended wages and $\lambda_E(z_1'\beta_E)$ is a control function for employment and

[14]Three education groups and two ability types, giving us six classes of individuals for the matching game.

$\lambda_{ed}(z_2'\beta_{HS}, z_2'\beta_U)$ is a control function to account for the endogeneity of education.[15]

We use a probit for employment to estimate the index $z_1'\beta_E$ and we then approximate $\lambda_E(z_1'\beta_E)$ by a quadratic function in the Mills ratio evaluated at the estimated index. As an instrument we use the predicted residual out-of-work income. This is a residual from a regression of predicted out of work income (based on the welfare system at the relevant time) on household demographic composition and marital status (which are used in the calculation of welfare benefits). Since we have many years of data we are identifying the impact of out of work income on the basis of how it changes for different demographic groups over time. The probit also includes time and age dummies.[16]

For education we use a multinomial logit model for the three levels of education we consider (Secondary, High School and University) to estimate the two indices. We then approximate $\lambda_{ed}(z_2'\beta_{HS}, z_2'\beta_U)$ by a quadratic function in the probability of attending high school and the probability of attending college. As instruments for education we use residual parental income and its squared.[17] This is a residual of a regression of parental income when the person was 16, on a set of family background variables. What is left is assumed to reflect a liquidity shock at the time the individuals are making education choices.[18] If individuals are to an extent liquidity constrained when making education decisions this residual will affect educational outcomes.[19]

[15]We make the simplifying assumption that we can control for selection into employment by this simple control function without accounting for the dependence of the employment and the education reduced form. This is possible to relax if we estimate the model all in one step, but this is highly time consuming.

[16]Out of work income varies over time because of policy changes and over types of individual. It thus provides important exogenous variation for identifying selection into work. However, our structural model does not account for the tax and welfare system, something that we are intending to do in future. In estimating the age education profiles outside the model we are thus able to use policy induced information that we could not use if we estimated the entire model in one step.

[17]Ideally we would need two instruments, or assume that education choices are ordered. Our estimates are almost identical if we assume that choices are ordered.

[18]see Blundell et al. (2016).

[19]Family background includes the education of both parents (five levels each), number of siblings and sibling order (dummies for no siblings, three or more siblings, and whether respondent is the first child), books in childhood home (three levels) and whether lived with both parents when aged 16.

We assume that the liquidity shock does not affect either wages or preferences in later life, but only the costs of education.

Having estimated the education and gender specific age profiles we then take these as known and proceed to estimate the remaining components that determine post-marital behavior, namely the stochastic process of wages characterized by $\sigma^2_{\xi,gs}$ and ρ^{gs} and the variance of the measurement error σ^2_ϵ, as well as the preferences of leisure and the distributions of unobserved ability and unobserved preferences. This is described in the next section.

6.3 Estimating preferences and the distribution of wages

As already noted, and repeated here for convenience, the period utility function for the household, consisting of a man m and a woman f with education s_m and s_f respectively, can be written as

$$Q_t\left(C_t + \alpha_{mt}^{M s_m \gamma} L_{mt} + \alpha_{ft}^{F s_f \gamma} L_{ft}\right)$$

An important feature, which we exploit in estimating the model, is that household utility and therefore public good consumption and labor supply do not depend on the Pareto weights. To capture the way labor supply varies over the lifecycle without having to control explicitly for the presence of children and other important taste shifters over the lifecycle, we specify the α parameters to be a polynomial in age (t):

$$\alpha_{it}^{gs\gamma} = \alpha_0^{gs\gamma} + \alpha_1^{gs\gamma} t + \alpha_2^{gs\gamma} t^2 + \alpha_3^{gs\gamma} t^3 + \eta_i + u_{it} \tag{34}$$

where the parameters ($\alpha_\ell^{gs\gamma},\ \ell = 0, ..., 3$) are specific to gender (g), education (s) *and* marital status ($\gamma = 1$ for married and zero for single). In other words in our model, preferences for singles and married individuals can differ. 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. Individuals draw preferences for work after the matching stage from a distribution that depends on education and has two points of support (although this can easily be relaxed since individuals do not match on this). This assumption allows us to take marital sorting as exogenous for labor supply and to estimate the model for the post-marital choices separately. Finally, u is an independently and identically distributed normal shock, drawn each period.

In general identification of preferences requires some variables to affect labor supply only through wages. Various strategies are followed in the literature. For example, Blundell et al. (2016) identify their labor supply model by using tax reforms that affect the return to work but not preferences.[20] In this simpler model we do not use this source of exogenous variation, although in principle one could extend our model to allow for taxes and welfare benefits and thus exploit policy changes. This is beyond the scope of the paper, but it is certainly part of our future research program. Here the identification problem is resolved because of the very tight specification of the utility function.

We set the annual discount factor δ to 0.98 and the annual interest rate to 0.015, implying that agents have some degree of impatience.[21] All other parameters are estimated using the method of simulate moments.[22] In our model there are 36 possible types of marital matches and individuals may also be single.[23] For each possible match we simulate wages and labor supply for the entire lifecycle and construct several simulated moments that we then match to the equiv-

[20]See also Blundell, Duncan and Meghir (1998).
[21]Same discount rate and interest rate was used in Blundell et al. (2016).
[22]McFadden (1989) and Pakes and Pollard (1989).
[23]Each person can have one of two ability levels and one of three education levels.

alent data moments. This provides us with estimates for the joint distribution of unobserved ability in all couple types ($\Pr(\theta_m, \theta_f \mid s_m, s_f)$ for married couples and $\Pr(\theta_i \mid s_i)$, $i = m, f$ for singles), estimates of the α parameters, the distribution of preference heterogeneity (η), and the stochastic process of wages (given the pre-estimated age profiles for each of the three education groups). Critical to our strategy is the fact that we have estimates of the age education profiles for men and women, from the previous step. By constructing wage series that are censored whenever an individual decides not to work (based on the model) and matching the resulting moments to those observed we control for selection into work when estimating the stochastic process of wages.

From these estimates we can recover the marital sorting patterns by ability and education, as well as the unconditional distribution of ability for men and women. Given this, the marriage market outcomes - $\Pr(H \mid H_m)$ 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:

$$\Pr(H \mid H_i) \equiv \Pr(S, \theta \mid S_i, \theta_i) = \frac{\Pr(S, S_i, \theta, \theta_i)}{\Pr(S_i, \theta_i)} = \frac{\Pr(\theta, \theta_i \mid S, S_i) \Pr(S, S_i)}{\sum_{s \in \mathcal{S} \cup \{0\}} \Pr(\theta_i \mid s, S_i) \Pr(s, S_i)}$$

for $i = m, f$ and where $H = H(\theta, S)$. All the quantities in the fourth expression are either directly observed in the data ($\Pr(s, S_i)$ for all $s \in \mathcal{S} \cup \{0\}$) or estimated from this estimation stage ($\Pr(\theta, \theta_i \mid S, S_i)$ and $\Pr(\theta_i \mid s, S_i)$).

Heuristically, identification works as follows: the autocovariance structure of wage growth identifies the stochastic process of wages. The cross sectional dispersion of wages and their serial dependence that is not explainable by the stochastic process identifies the distribution of unobserved heterogeneity in earnings. The age profiles of participation (for each education and gender), given the already estimated age profiles of wages, identify the age effects on labor supply. Finally, since unobserved heterogeneity induces persistence in employment choices, the degrees of individual labour market attachment over 5 years or more, as well as changes with education, identifies the distribution of unobserved preference heterogeneity, given the functional forms we choose.[24]

6.4 Preferences for marital sorting and education

In our model individuals choose education at a first stage in life and then enter the marriage market. We allow for three levels of educational attainment: Secondary (statutory schooling), High School (corresponding to A-levels or equivalent) and University, corresponding to 3-year degrees or above. We interchangeably use the term college for this group. At the point where they make the education and the matching decisions, ability of all individuals and their preferences for partners are observable by all. Preferences for work are not known.

As discussed earlier this choice process can lead to many equilibria, one of which is efficient (see Nöldeke, G., and L. Samuelson, 2015). This equilibrium is equivalent to one where individuals choose education level and type of partner at the same time. We assume that the data is characterized by that equilibrium and we thus estimate preferences for type of partner and the determinants of education choice in one step. What follows is a discrete choice problem for men and women, respectively, where each chooses one option out of all possible combinations of education and types of spouse. Since we are assuming that the observed patterns correspond to the efficient equilibrium we can then back out the Pareto weights, which are the prices that decentralize this market.

[24] The 328 moments include the means, variances and several quantiles of 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. 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 value for a woman f with human capital $H_f = H_F(\theta_f, s_f)$ marrying a man m with human capital $H_m = H_M(\theta_m, s_m)$ is the sum of an economic component and a random preference component for a type of spouse (defined by education and ability). In the earlier steps we have estimated the parameters that allow us to compute the economic component for all possible matches defined by the ability and education of each member of the couple, up to the Pareto weight, which we will identify in this step. We can also compute the economic value of being single for all types.

This utility can also be interpreted as the value of choosing both the level of education and type of partner, given own ability, if we net out the costs of education. We define these costs for individual i to be

$$c_i^{gs} = \iota_0^{gs} + \iota_1^{gs} y_i^p + \kappa_i^{gs}$$

where the parameters ι^{gs} are gender and education specific. We include the residual parental income at 16, y_i^p, (described earlier) as a determinant of the costs of education.

The utility for female f with ability θ_f of choosing type of male partner $H = H_M(\theta, s)$ and of own education s_f is given by

$$\tilde{U}_F(\theta_f, s_f, H) =$$

$$\bar{U}_F\left(H, H_F(\theta_f, s_f)\right) + \varphi_F(H_F(\theta_f, s_f))\mathbf{1}(H = 0) + \tilde{\varphi}_F\left(\theta_f, |s - s_f|\right)\mathbf{1}(H_F(\theta_f, s_f) \neq 0) - c_f^{s_f F} + \beta_f^H$$

where $\bar{U}_F\left(H, H_F\right)$ corresponds to the economic value of marriage and β_f^H is a random preference component for a type of spouse H. To this we have added extra components of marital preferences. Specifically, $\varphi_F(H_F(\theta_f, s_f))$ is a set of six fixed coefficients (one for each type) measuring the non-economic utility component for remaining single; and $\tilde{\varphi}_F(\theta_f, |s - s_f|)$ is a set of coefficients capturing the (dis)taste for disparity in the educational attainment of spouses. To preempt, these coefficients proved to be important for fitting the sorting patterns in the data. Otherwise the simpler model predicted too little sorting.

The optimal schooling and partner choice is obtained by

$$(s_f^*, H^*) = argmax_{s_f,H}(\tilde{U}_F(\theta_f, s_f, H) \;\; \forall \;\; H, s_f)$$

Assuming that β_f^H follows an extreme value I distribution the probability of any observed choice *given* κ is given my the multinomial logit with 21 alternatives to choose from.[25] To obtain the probabilities that need to be matched with those observed in the data we integrate out κ, which is assumed to be normally distributed, thus relaxing the distributional assumption and in particular the independence of irrelevant alternatives.

Recall that $\bar{U}_F$ is

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

where

$$V_F(H, H_F, \mu(H, H_F)) = \Upsilon(H, H_F) + \frac{1-\delta^T}{1-\delta} \ln \frac{\mu(H, H_F)}{1+\mu(H, H_F)}$$

where the μ are such that each individual has a well defined utility value.[26] The key point is that $\Upsilon(H_M, H_F)$ can be computed on the basis of the estimates from the lifecycle estimation

[25] Three levels of own education, and a partner with one of two levels of ability and one of three levels of education or remain single: $3\times(3 \times 2 + 1)$.

[26] Recall that individual utility is in logs and hence the resulting argument after intrahousehold allocations has to be positive.

stage for each pair of (H_M, H_F). We treat the Pareto weights as unknown parameters, along with the various preference parameters, in the estimation problem.

The 36 Pareto weights, for each possible matched pair of (H_M, H_F), can be fully identified just by using the observed female choices (who they marry and what education level they choose). However, in equilibrium, we can also use the male choices to identify the same Pareto weights, which provides a set of overidentifying restrictions. This level of overidentification originates from the fact that we can estimate the economic value of marriage from the lifecycle problem, which generalizes the Choo and Siow (2006) approach. Using the male choices as well as the female ones is also necessary for estimating the preferences for being single and for marrying a different type of spouse than oneself for both genders.

We obtain the estimates by the method of moments estimator, using simultaneously the male and the female choices where we match the observed choice probabilities to the equivalent ones implied by the model. In doing so we also minimize the distance between the predicted marital patterns based on the male choices and those based on the female ones, thus finding the parameters that are most consistent with equilibrium. The extent to which the resulting predicted patterns differ from each other is a diagnostic for whether the model can rationalise the observed pattern as an equilibrium in the marriage market.

7 Results

All estimates relating to the earnings equations, the results on the distribution of ability as well as the preference parameters determining labor supply choices are presented in Appendix C, since they are not of a central interest in themselves. We also present details on the overall model fit.

7.1 The Parameters for marital sorting and education choice.

Table 1 presents the preference parameters governing marital sorting. Preferences for remaining single increase with education. Higher ability men have a lower preference for being single, while the contrary is true for higher ability women. The parameters on the panel below reflect the utility cost of educational disparity within a couple as perceived by each partner: the more disparate the educational levels, the higher is the utility cost, but this differs substantially by gender and ability. One exception to the preference for similarity are lower ability men who prefer a spouse that is one education level above or below them.

Table 1: Utility shifters - preferences for remaining single and for marrying similarly educated spouses

	Men by ability		Women by ability	
	low	high	low	high
Preference for remaining single, by education ($\varphi_F(H_F(\theta_f, s_f))$)				
Secondary	1.502	-1.373	-0.164	1.590
	(0.207)	(0.583)	(0.093)	(0.565)
High School	2.207	-0.912	1.142	3.630
	(0.575)	(0.457)	(0.164)	(0.862)
University	3.385	-0.420	3.896	5.552
	(0.827)	(0.106)	(1.545)	(1.753)
Preference for differently educated spouses ($\tilde{\varphi}_F(\theta_f, \|s - s_f\|)$)				
One educational level diff	0.344	-0.698	-0.104	-0.114
	(0.034)	(0.243)	(0.244)	(0.116)
Two educational levels diff	-2.732	-1.377	-4.287	-0.873
	(0.660)	(0.358)	(1.548)	(0.114)

Asymptotic standard errors in parentheses computed using the bootstrap.

Table 2 shows the costs of education implied by the estimates. A reduction in parental income at age 16 reduces both high school participation and college attendance (increases the cost of education). Although the coefficients for the effect on college are lower, this is no surprise

Table 2: Utility cost of education

	men		women	
	HS	Univ	HS	Univ
constant	1.073	3.912	1.864	6.294
	(0.337)	(1.321)	(0.503)	(2.341)
Parental income at 16 (residual)	-0.335	-0.160	-0.523	-0.075
	(0.155)	(0.095)	(0.135)	(0.037)

Asymptotic standard errors in parentheses computed using the bootstrap.

since college attendance takes place 2-3 years later and hence the effect of the shock may have been attenuated by that time.[27]

7.2 The marital surplus

We start by ranking individuals by their human capital as measured by their lifecycle earnings earnings capacity, which are a function of ability and education. For women earnings capacity increases with education and ability. However, being a university graduate implies higher earnings, whatever the level of ability. For men high ability high school graduates have a higher earnings capacity than lower ability university graduates. Table 3 reports the ranking of human capital by education and ability and the value of being single. The value of being single increases monotonically with individual human capital for both men and women. However, the increase is much steeper for women, which is part of the reason why single women tend to be drawn from the higher part of the human capital distribution.

Table 4 presents the economic surplus for all possible 36 combinations of human capital for couples. This is the economic value of marriage, over and above the value of remaining single.

[27] Remember that the parental income is a residual, where the effects of family background have been removed. Moreover, we observe family income at the time the child was 16 only for a (younger) subsample of the data. In a richer model it would be desirable to also control for family background, which would affect wages and preferences potentially. However this would increase the state space and the resulting possible matches beyond the capabilities of our data.

Table 3: Human Capital and the value of being single

			Females			
HC Rank	1	2	3	4	5	6
Education (ability)	Secondary (L)	High School (L)	Secondary (H)	High School (H)	University (L)	University (H)
Value of Single	33.4	61.8	88.6	88.6	191.5	293.0
			Males			
HC Rank	1	2	3	4	5	6
Education (ability)	Secondary (L)	High School (L)	Secondary (H)	University (L)	High School (H)	University (H)
Value of Single	114.9	150.1	220.4	276.7	300.2	441.1

HC: Human Capital; H: High ability. L: Low ability; higher rank corresponds to higher human capital.

There are two important conclusions from this. First, the gradient of the surplus is much steeper with respect to female human capital than it is with respect to male. This is because the impact of education on female earnings (conditional on employment) and on employment itself is much higher for women than it is for men. We show this in Figures 1 and 2.[28] Hence a large part of the variation in the surplus is explained by the human capital of the woman. Second, the surplus is generally super modular, particularly for higher levels of human capital. This will push towards 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 for most 2×2 submatrices, the sum of diagonal terms exceeds the sum of off-diagonal ones. In particular, all 2x2 matrices at the top of the human capital distribution (i.e. those including the top three levels for each gender) are supermodular. Similarly, all 2x2 submatrices involving human capital levels not immediately adjacent are positive, suggesting that violations of supermodularity, although possible, are mostly 'local'.

7.3 Marital patterns

The share of the surplus and the marital patterns drive the choice of partner. The Pareto weights implied by the choices of males are not restricted to be the same as those implied

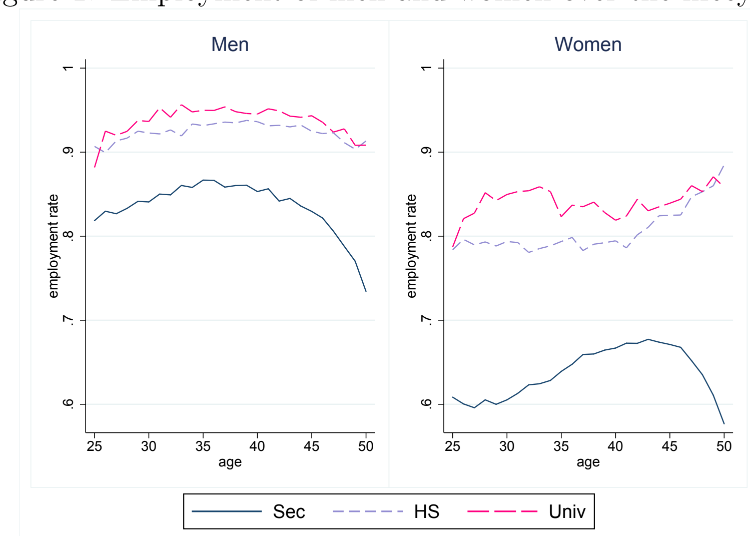
[28]These being earnings they include hours dimensions as well, which are not modeled here. In interpreting male and female differences it is important to note that many women work part time, at varying degrees over the lifecycle, while men nearly always work full time.

Table 4: Economic surplus from marriage

Men's educ and ability	women's ability and education					
	Sec (L)	HS (L)	Sec (H)	HS (H)	Univ (L)	Univ (H)
Sec (L)	85.06	148.88	189.26	189.10	197.17	245.39
HS (L)	82.61	144.33	189.53	185.97	199.87	249.21
Sec (H)	129.54	210.34	266.84	264.88	299.85	370.86
Univ (L)	101.45	176.79	241.15	232.27	268.43	338.90
HS (H)	139.01	220.91	288.21	281.00	326.74	405.43
Univ (H)	142.96	234.71	317.10	305.31	366.01	460.91

Rows and Columns ordered by male and female human capital respectively. L and H signify low and high ability respectively.

Figure 1: Employment of men and women over the lifecycle



by the observed choices of the females. In equilibrium they should be, but since the model is heavily overidentified this will in general not be the case in a finite sample even if the restrictions we impose are true in the population. In a final step of estimation we choose the Pareto weights that minimize the difference in implied marital patterns when comparing male and female choices. The matches are shown in Table 5 and are remarkably similar. For 50% of

Table 5: Marital Matching patterns

	Women's education					
Men's educ	Sec	HS	Univ	Sec	HS	Univ
	Simulated Proportions Men's choices			Data Proportions		
Sec	0.322	0.068	0.001	0.291	0.094	0.014
HS	0.159	0.124	0.026	0.156	0.126	0.032
Univ	0.007	0.047	0.048	0.019	0.044	0.053
	Simulated Proportions Women's choices			Data Proportions		
Sec	0.328	0.071	0.001	0.291	0.094	0.014
HS	0.158	0.125	0.029	0.156	0.126	0.032
Univ	0.007	0.050	0.050	0.019	0.044	0.053

The numbers represent cell proportions.

married couples both partners have the same level of education; however there are substantial amount of marriages that do not follow this rule. Hence along the educational dimension the sorting patterns are not perfectly assortative and the model is able to fit this.

Table 6 shows the composition of the singles sorted by their level of human capital.[29] Seventy three percent of single men are low ability,[30] compared to 30% in the population. By contrast among single women only twenty seven percent are low ability,[31] compared to 40% in the population. Once ranked by the value of human capital, as measured by potential earnings over the lifecycle, we still see that the majority of single women are high human capital and the majority of single men are at the lower end. Finally, the complete set of marital patterns conditional on being married, as implied by our model, are presented in Table 7. The matches that actually form will depend on the Pareto weights, and we turn to these now.

[29]Here and in what follows we rank individuals by the level of their potential lifecycle earnings, which depends on education and ability. The lowest level is denoted by 1 and increases up to 6.

[30]Human capital level is 1, 2 and 4.

[31]Human capital level is 1, 2 and 5.

Figure 2: Log annual potential earnings for men and women

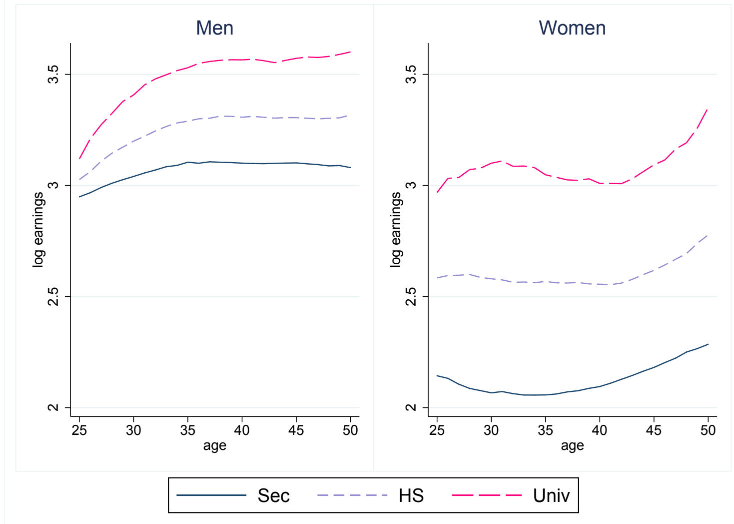


Table 6: Proportion of singles by level of human capital.

Level of Human Capital	1	2	3	4	5	6
Women	0.11	0.08	0.14	0.39	0.07	0.21
Men	0.22	0.31	0.07	0.20	0.16	0.04

Levels of human capital in increasing order: A. Men 1: Secondary - low ability, 2: High School Low ability, 3: Secondary High ability, 4: University Low ability, 5: High School - High ability, 6: University High ability; B. Women 1: Secondary - low ability, 2: High School- Low ability, 3: Secondary High ability, 4: High School - High ability, 5: University L- Low ability, 6: University High ability.

7.4 Sorting and the sharing rule

The estimated Pareto weights reveal the allocation of welfare within the household in the context of the equilibrium observed in the data. This takes into account the public good and the labor supply/leisure decision. Non-market time and private consumption are perfect substitutes, while both are complements of public consumption. Table 8 shows the men's share

Table 7: Sorting patterns conditional on marriage by eduction and ability

	women's ability and education					
	Sec (L)	HS (L)	Sec (H)	HS (H)	Univ (L)	Univ (H)
Sec (L)	0.041	0.025	0.044	0.015	0.000	0.000
HS (L)	0.031	0.011	0.034	0.006	0.004	0.004
Sec (H)	0.145	0.030	0.171	0.016	0.000	0.001
Univ (L)	0.000	0.010	0.001	0.007	0.003	0.004
HS (H)	0.053	0.084	0.078	0.053	0.012	0.014
Univ (H)	0.000	0.025	0.009	0.017	0.025	0.028

Rows and Columns ordered by male and female human capital respectively. L and H signify low and high ability respectively. Cell proportions reported.

in the gains from marriage that clear the market.[32]

Table 8: Sharing rule

	women's ability and education					
Men's educ and ability	Sec (L)	HS (L)	Sec (H)	HS (H)	Univ (L)	Univ (H)
Sec (L)	0.835	0.363	0.512	0.161	0.248	0.162
	(0.261)	(0.114)	(0.148)	(0.080)	(0.073)	(0.040)
HS(L)	0.938	0.606	0.596	0.376	0.038	0.019
	(0.335)	(0.254)	(0.212)	(0.152)	(0.024)	(0.021)
Sec (H)	0.613	0.449	0.448	0.294	0.070	0.089
	(0.225)	(0.172)	(0.155)	(0.127)	(0.047)	(0.052)
Univ (L)	0.937	0.857	0.936	0.661	0.429	0.353
	(0.330)	(0.343)	(0.335)	(0.231)	(0.165)	(0.110)
HS (H)	0.767	0.495	0.582	0.362	0.228	0.196
	(0.252)	(0.193)	(0.188)	(0.142)	(0.037)	(0.065)
Univ (H)	0.699	0.756	0.747	0.614	0.415	0.360
	(0.330)	(0.285)	(0.262)	(0.213)	(0.136)	(0.121)

Notes: Male Share of Surplus. Asymptotic standard errors in parentheses computed using the bootstrap. Ordering of cells by male and female human capital respectively. L and H signify low and high ability respectively

In principle, the relationship between a person's human capital and share of welfare need not be strictly monotonic; the share also reflects relative scarcity of spouses at each level of human capital, and therefore depends on the entire distribution. Still, we see that in most (but not all)

[32]Note that, given super-modularity of the economic component (when this is the case), a marriage between spouses of very different skills signals large values of the corresponding marital preference.

cases the male share declines in his wife's human capital. Among couples of college graduates with higher ability, the share favors women. Low skill men marrying the lowest skill women (1st column) benefit from very high shares. However, if a low skill man marries a highest skill woman (a very rare combination - 0.01% of the population) she gets 84% of welfare, reflecting a very high Pareto weight for her. Among 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.

The sorting we observe and the resulting Pareto weights are driven by the structure of the surplus. We have already seen how this varies as a function of human capital. The way it changes across groups is driven both by human capital at the time of matching and by its stochastic properties, since marriage allows risk sharing. In Figure 3 we show how the aggregate surplus[33] varies when we change the variance of earnings of men and women by the same factor. The figure shows that as uncertainty rises the economic value of marriage relative to being single increases because of risk sharing. Halving the variance reduces the aggregate surplus from marriage by 10%, and doubling it increases it by 15%.

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 reflect the patterns we see in the data. In what follows we carry out a counterfactual simulation, that will yield new Pareto weights and marital patterns.

One of the key points of our approach is that part of the returns to education can be accounted for by marriage and in particular by the sharing of the marital gains. Thus, ignoring the preferences for marital status, marital returns to high school account for 73% of the entire

[33]Weighted average across all matches, with weights the probability of a match with baseline risk and conditional on marriage.

Figure 3: The impact of risk on the marital surplus

return to HS for women, assuming optimal choice of partner. Marital returns account for 58% of the female college premium. Both these numbers demonstrate the importance of marriage in determining the returns to education for women. For men there is a similar impact, but it is smaller: the respective effects are 14% for high school and 21% for college.

8 Counterfactual Simulations

The model offers us a way of interpreting the data as well as the possibility of counterfactual analysis with an emphasis on longer run outcomes. Here we examine the impact of reducing the costs of university education by 5%. The key mechanism that can cause the realignments in the marriage market and indeed change the welfare of men and women is the increased supply of college graduates of both genders, which will affect the types of individuals that enter the marriage market. The changes in the implied Pareto weights will then feed back into the education choice.

The final equilibrium distribution of education shows an increase in the supply of both male and female college graduates, but more so for women (see Table 9). These supply changes are associated with both changes in the matching patterns and the welfare share for each type of match. The former are shown in Tables 10 and 11 while the latter in Table 12. Generally there is an increase in the proportion of college educated women marrying non-college educated men, as well as an increase in matches among college educated men and women. As shown in Table 11 the proportion of singles increases.

Table 9: Education distribution

	Men		women	
	baseline	low cost Univ	baseline	low cost Univ
Distribution of education				
Sec	0.452	0.429	0.542	0.517
HS	0.398	0.385	0.329	0.318
Univ	0.150	0.186	0.128	0.166

Table 10: Changes in the matching patterns

	Women's education and ability					
Men's educ and ability	Sec (L)	HS (L)	Sec (H)	HS (H)	Univ (L)	Univ (H)
Sec (L)	-0.07	-0.12	-0.18	-0.04	-0.02	-0.01
HS (L)	-0.22	-0.07	-0.13	-0.02	0.34	0.02
Sec (H)	-0.41	-0.10	-0.50	-0.06	-0.01	0.00
Univ (L)	0.00	0.08	-0.02	0.06	0.01	0.09
HS (H)	-0.11	-0.22	-0.27	-0.18	0.05	0.12
Univ (H)	0.00	0.20	0.06	0.12	0.49	0.55

NUmbers correspond to changes in the proportion of each cell. Ordering of cells by male and female human capital respectively. L and H signify low and high ability respectively

However Table 12 reveals very interesting changes in the way welfare is distributed within the household. A negative value means a decrease in the male share in favor of the female one. These results imply that subsidizing women to increase college attendance can increase their

Table 11: Marital patterns

	Men		women	
	baseline	low cost Univ	baseline	low cost Univ
proportion remaining single	0.193	0.196	0.190	0.197
proportion marrying equally educated spouse	0.502	0.497	0.500	0.494

share of welfare, particularly for the lower skill ones, as well as for those who marry high skill men. For example the welfare share of low ability male college graduates marrying low ability high school graduates declines by 2.6 percentage points. Female college graduates however do not necessarily gain themselves: the share of college graduates declines. These patterns are driven by the change in the relative scarcity of partners at each skill level.

Underlying these results is the convergence to a new long run equilibrium, with changes in educational attainment relative to the immediate effect induced by the subsidy, as the marital patterns change and the Pareto weights adjust. In future research it will be important to examine how changes in welfare benefits and their targeting affects marital patterns and life cycle work and consumption decisions.

Table 12: Changes in the sharing rule - percentage points

	Women's educ and ability					
Men's educ and ability	Sec (L)	HS (L)	Sec (H)	HS (H)	Univ (L)	Univ (H)
Sec (L)	-0.2	0.7	0.4	0.5	-0.2	-2.2
HS (L)	-1.3	0.2	0.2	0.2	5.5	2.8
Sec (H)	-0.1	0.3	0.0	0.3	0.2	0.9
Univ (L)	-0.5	-2.6	0.5	-2.4	1.0	0.7
HS (H)	-0.1	0.1	-0.2	0.1	2.2	1.6
Univ (H)	0.0	-2.2	-1.5	-1.9	0.1	0.2

Ordering of cells by male and female human capital respectively. L and H signify low and high ability respectively. A negative number corresponds to a decline in the male share.

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 simulate policies that change the economic environment at any stage of the lifecycle. Matching in the marriage market is stochastic but frictionless 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 super-modular nearly everywhere, pushing towards positive assortative matching, with any departures from perfect sorting being driven by random preferences for mates and by some local departure from supermodularity. 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. The model is able to replicate matching patterns very well, despite the fact we do not allow for frictions.

Generally high human capital women get more than half of the marital surplus, while men marrying low human capital women get most of the surplus. These shares reflect the existing equilibrium in the data. However, the share of welfare is endogenous and changes in the supply of men and women of different levels of human capital can change them. Thus in our counterfactual simulation, where we reduce the costs of education, inducing more to graduate from college we find that the share of *low* human capital women increases, while the share of low ability college graduate women declines.

Lastly, our model sheds light on the determinants of human capital investments. Two conclusions emerge. First, non-economic factors play an important role in both the decision to marry and the marital patterns conditional on marriage and therefore indirectly affect the return to education. This is by no means surprising. However, our approach allows us to quantify the magnitude of these effects; we find them to be quite large. Second, and more importantly, we can explicitly decompose the returns to education into those perceived on the labor market and those reaped in addition on the marriage market (the "marital college premium"). Our results are unambiguous: the benefits perceived through marriage (through risk sharing and the joint consumption of public goods) are dominant. Our analysis, therefore, confirms the notion, put forth by CIW, that any empirical analysis that omits marital gains and concentrates exclusively on the labor market may be severely biased. An important implication is that a policy (e.g., a tax reform) that directly affects the returns on human capital investment will also alter the respective importance of economic and non-economic factors for the determination of matching patterns, further influencing incentives to invest; these equilibrium effects will typically amplify the initial impact, resulting in potentially large long-term consequences that should not be ignored.

This paper is a first step towards a rich research agenda analyzing the interactions of marriage, labor markets and educational choices. Generalizations will include allowing for imperfectly transferable utility,[34] 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 –

[34]see Galichon et al. (2016) for a recent contribution along these lines.

have an important, economic component, which can be analyzed in terms of 'complementarity' or 'substitutability' (in technical 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.

References

[1] Attanasio, Orazio, Hamish Low and Virginia Sánchez-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, 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

[5] 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

[6] Blundell, R, Monica Costa Dias, Costas Meghir Jonathan Shaw (2016) Female Labor Supply, Human Capital and Welfare Reform, forthcoming Econometrica.

[7] Blundell, Richard, Duncan Alan, and Meghir Costas. "Estimating Labor Supply Responses Using Tax Reforms." Econometrica 66.4 (1998): 827-61.

[8] 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

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

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

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

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

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

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

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

[16] Cole, H. L.,G. J. Mailath and A. Postlewaite (2001): "Efficient Non-Contractible Investments in Large Economies," *Journal of Economic Theory*, 101 (2), 333–373.

[17] Galichon, Alfred and Kominers, Scott Duke and Weber, Simon (2016), Costly Concessions: An Empirical Framework for Matching with Imperfectly Transferable Utility (March 22, 2016). Available at SSRN: http://ssrn.com/abstract=2535394 or http://dx.doi.org/10.2139/ssrn.2535394

[18] Goussé, Marion, Nicolas Jacquemet and Jean-Marc Robin (2016): "Marriage, Labor Supply, and Home Production: A Longitudinal Microeconomic Analysis of Marriage, Intra-Household Bargaining and Time Use Using the BHPS, 1991-2008", CIRPEE Working Paper 16-01

[19] Graham, B. (2011), "Econometric Methods for the Analysis of Assignment Problems in the Presence of Complementarity and Social Spillovers", in J. Benhabib, M. Jackson and A. Bisin eds: Handbook of Social Economics, North-Holland, Vol. 1B, 965-1052.

[20] Heckman, J. (1979). Sample Selection Bias as a Specification Error. Econometrica, 47(1), 153-161.

[21] 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

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

[23] 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

[24] McFadden, D. (1989). "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," *Econometrica*, vol. 57(5): 995-1026.

[25] Nöldeke, G., and L. Samuelson (2015), "Investment and Competitive Matching", *Econometrica*, Vol. 83,No. 3, 835–896

[26] Pakes, A. and D. Pollard (1989). "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, Vol. 57(5): 1027-1057.

[27] Peters, Michael, and Siow Aloysius. "Competing Premarital Investments." Journal of Political Economy 110, no. 3 (2002): 592-608.

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

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

[30] Townsend, Robert M. "Risk and Insurance in Village India." Econometrica 62, no. 3 (1994): 539-91.

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

10 Online Appendix

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, $\mathbf{H} = (H_m, H_f)$, and we omit them for simplicity. Each period/age t sees the arrival of new information on each spouse's preferences for working and productivity, $\boldsymbol{\alpha}_t = (\alpha_{mt}, \alpha_{ft})$ and $\mathbf{e}_t = (e_{mt}, e_{ft})$.

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

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

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:

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

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, $\mathbf{e}_T$ and $\boldsymbol{\alpha}_T$, as well as savings, non labor income and the spouses' respective stocks of human capital (through their impact on wages); we therefore write $Q_T\left(\mathbf{e}_T, \boldsymbol{\alpha}_T, K_{T-1}, \mathbf{H}\right)$ and $v_{iT}\left(\mathbf{e}_T, \boldsymbol{\alpha}_T, K_{T-1}, \mathbf{H}, \mu\right)$, where $\mathbf{H} = \left(H_m, H_f\right)$.

Expected value functions We assume that the unobserved productivity shocks and preferences for time off paid work, $(\mathbf{e}, \boldsymbol{\alpha})$, follow a first-order Markov process. Then, the expected value functions are

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

where

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

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 maximize 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

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

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

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

where

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

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

$$2\frac{\partial\ln Q_{t}}{\partial K_{t}}+\delta\frac{\partial I_{t+1}^{\varsigma}}{\partial K_{t}}=0$$

Finally, employment is

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

Appendix C: Estimates of model parameters

Table 13 contains estimates of the parameters in the stochastic wage process. The education premium (row 2) and the education-specific age profiles (rows 4-6) were estimated in the first stage reduced form model of education choice, employment and wages. Estimates of the first stage regressions can be found in tables 14 and 15. The other parameters were 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 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 13: Earnings process by gender and education

		Men			Women		
		Sec	HS	Univ	Sec	HS	Univ
(1)	log earnings (ab 1, stat ed, age 23)	3.01			2.45		
		(.01)			(.03)		
(2)	education premium		0.065	0.096		0.246	0.497
			(0.029)	(0.036)		(0.055)	(0.075)
(3)	ability premium (type 2)	0.148	0.143	0.095	0.457	0.287	0.280
		(.03)	(.02)	(.04)	(.03)	(.06)	(.06)
(4)	age (δ_1)	0.339	0.622	1.002	-0.242	0.177	0.664
		(0.063)	(0.062)	(0.092)	(0.114)	(0.111)	(0.180)
(5)	age squared (δ_2)	-0.160	-0.318	-0.527	0.128	-0.205	-0.597
		(0.052)	(0.050)	(0.077)	(0.091)	(0.092)	(0.146)
(6)	age cubic (δ_3)	0.023	0.053	0.091	-0.007	0.063	0.156
		(0.013)	(0.012)	(0.019)	(0.021)	(0.022)	(0.035)
(7)	Autocorr coeff (ρ)	0.836	0.825	0.889	0.845	0.909	0.855
		(0.051)	(0.070)	(0.039)	(0.049)	(0.027)	(0.047)
(8)	Var innov in prod (σ^2_ξ)	0.27	0.023	0.021	0.046	0.072	0.074
		(0.005)	(0.004)	(0.006)	(0.008)	(0.012)	(0.013)
(9)	Var ME (σ^2_ϵ)	0.016	0.011	0.006	0.008	0.006	0.004
		(0.011)	(0.009)	(0.019)	(0.016)	(0.019)	(0.020)
	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.

Table 14: Education regressions - multinomial logit

	men	women
	High School	
parental income residual	0.477*	0.576**
	(0.244)	(0.266)
parental income residual squared	-0.068*	-0.181
	(0.037)	(0.325)
parental income missing	-0.707***	-1.202***
	(0.102)	(0.115)
intercept	0.457***	0.626***
	(0.095)	(0.109)
	University	
parental income: residual	1.389***	0.843***
	(0.342)	(0.311)
parental income: residual squared	-0.459**	-0.199
	(0.216)	(0.351)
parental income missing	-0.878***	-1.562***
	(0.131)	(0.133)
intercept	-0.329***	0.009
	(0.121)	(0.123)
N	4,207	4,610

Notes: Three alternatives are modeled: Secondary School, High School and University. Standard Errors in brackets under estimates. Parental income is net of the effects of long-term characteristics of the parental household including parental education, whether both parents at home, number of siblings, sibling order, ethnicity and region of residence.

Table 15: Employment regressions - probit

	men	women
parental income: residual	0.332***	0.578***
	(0.113)	(0.129)
parental income: residual squared	-0.122*	0.080
	(0.069)	(0.196)
Out of work income: residual	-0.685***	-0.253***
	(0.107)	(0.040)
intercept	0.765**	0.473***
	(0.301)	(0.067)
N	19,325	19,306

Notes: SE in brackets under estimates. Parental income is net of the effects of long-term characteristics of the parental household including parental education, whether both parents at home, number of siblings, sibling order, ethnicity and region of residence. Unearned income is net of the effects of family composition. Regressions also include year and age dummies.

Tables 16 and 17 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.

Finally, estimates of the preference parameters are presented in table 18.

Table 16: Probability weights for the joint distribution of ability in couples by spouses' education

		Men					
		Secondary		High School		University	
		ability 1	ability 2	ability 1	ability 2	ability 1	ability 2
Women	Secondary						
	ability 1	0.128	0.543	0.174	0.488	0.208	0.434
		(0.031)	(0.055)	(0.042)	(0.065)	(0.114)	(0.230)
	ability 2	0.071	0.258	0.248	0.090	0.094	0.263
		(0.039)		(0.039)		(0.118)	
	High School						
	ability 1	0.099	0.489	0.137	0.334	0.174	0.225
		(0.060)	(0.090)	(0.055)	(0.100)	(0.093)	(0.104)
	ability 2	0.079	0.333	0.118	0.412	0.077	0.525
		(0.082)		(0.063)		(0.116)	
	University						
	ability 1	0.092	0.345	0.081	0.099	0.070	0.099
		(0.199)	(0.236)	(0.064)	(0.128)	(0.052)	(0.102)
	ability 2	0.077	0.486	0.070	0.749	0.131	0.700
		(0.161)		(0.083)		(0.071)	

Table 17: Proportion of ability type 1 among singles by gender and education

	Secondary	High School	University
men	0.737	0.654	0.863
	(0.062)	(0.087)	(0.195)
women	0.417	0.125	0.155
	(0.155)	(0.092)	(0.281)

Table 18: Preference parameters and distribution of unobserved heterogeneity in preferences for employment

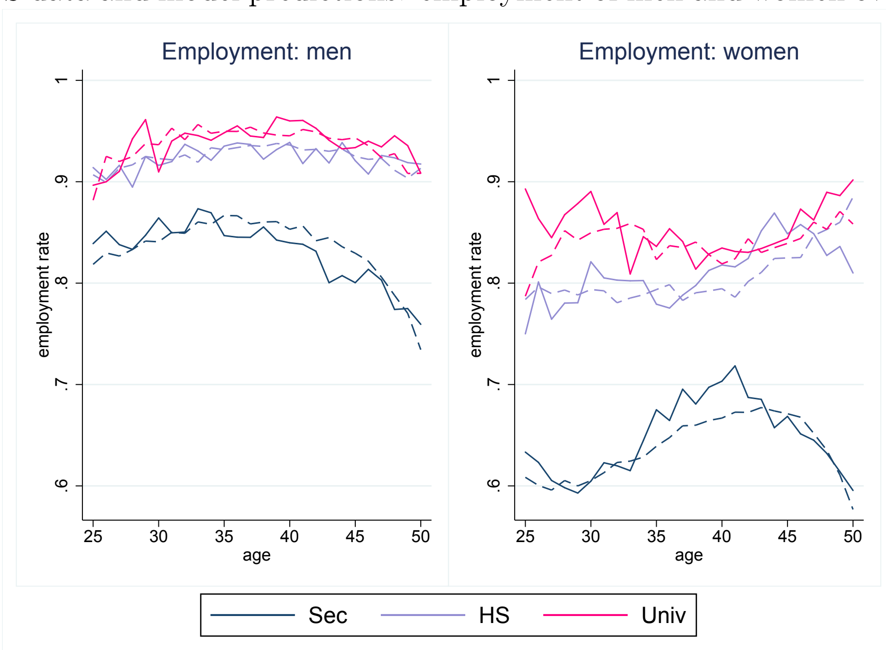
	Men			Women		
	Stat	HS	Univ	Stat	HS	Univ
Couples						
intercept (α_0)	0.259	0.244	0.607	0.700	0.702	1.366
	(0.331)	(0.990)	(1.099)	(0.313)	(0.481)	(0.413)
age (α_1)	-0.137	-0.023	-0.010	-0.092	-0.128	0.040
	(0.250)	(0.348)	(0.586)	(0.497)	(0.505)	(0.365)
age squared (α_2)	-0.172	-0.163	0.079	-0.561	0.143	0.065
	(0.252)	(0.252)	(0.380)	(0.407)	(0.276)	(0.224)
age cubic (α_3)	0.104	0.083	-0.005	0.229	-0.057	-0.023
	(0.100)	(0.069)	(0.188)	(0.099)	(0.114)	(0.110)
Singles						
intercept (α_0)	-0.300	1.011	1.295	0.606	1.070	2.267
	(0.444)	(0.886)	(1.260)	(0.398)	(0.434)	(0.537)
age (α_1)	1.460	0.200	-0.099	0.524	-0.096	-0.521
	(0.436)	(0.754)	(1.043)	(0.434)	(0.378)	(0.452)
age squared (α_2)	-0.424	0.129	-0.079	-0.204	-0.157	-0.569
	(0.478)	(0.776)	(0.703)	(0.402)	(0.696)	(0.369)
age cubic (α_3)	0.032	-0.075	0.068	0.064	0.078	0.275
	(0.254)	(0.208)	(0.205)	(0.200)	(0.350)	(0.064)
Unobserved preferences						
low utility from work ($\eta = 2$)	2.478	1.770	1.072	1.160	0.928	0.517
	(0.272)	(0.948)	(1.286)	(0.416)	(0.356)	(0.531)
probability utility type 2	0.617	0.634	0.453	0.542	0.483	0.557
	(0.059)	(0.112)	(0.279)	(0.194)	(0.302)	(0.268)
Var transitory pref shock (u)	0.564	1.007	1.074	0.370	0.676	0.942
	(0.226)	(0.490)	(0.963)	(0.233)	(0.191)	(0.496)

Notes: SE in brackets under estimates.

Appendix D: Fit

This appendix describes the fit of the model. We start by comparing the employment and earnings profiles in the data with those obtained from model simulations in Figures 4 and 5. The model is capable of capturing well these profiles.

Figure 4: 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 rest of this appendix contains tables showing all data moments used in estimation and their simulated counterparts, together with the ratio of the difference between the two moments and the standard error of the data estimate.

Table 19: Distr log earnings net of age effects: single men

Moment	Data	Simulated	SE data	No. SE diff
secondary education				
mean	2.731	2.710	0.041	0.53
var	0.153	0.155	0.022	0.10
P(earnings<Q10)	0.100	0.160	0.023	2.59
P(earnings<Q25)	0.250	0.306	0.037	1.52
P(earnings<Q50)	0.500	0.528	0.047	0.61
P(earnings<Q75)	0.750	0.759	0.042	0.22
P(earnings<Q90)	0.900	0.880	0.036	0.53
high school				
mean	2.798	2.799	0.049	0.02
var	0.197	0.130	0.029	2.32
P(earnings<Q10)	0.100	0.068	0.026	1.16
P(earnings<Q25)	0.250	0.220	0.041	0.71
P(earnings<Q50)	0.500	0.526	0.052	0.50
P(earnings<Q75)	0.750	0.786	0.047	0.76
P(earnings<Q90)	0.900	0.943	0.033	1.30
university education				
mean	2.766	2.800	0.067	0.52
var	0.185	0.113	0.034	2.09
P(earnings<Q10)	0.100	0.041	0.033	1.74
P(earnings<Q25)	0.250	0.188	0.063	0.97
P(earnings<Q50)	0.500	0.474	0.080	0.32
P(earnings<Q75)	0.750	0.785	0.065	0.54
P(earnings<Q90)	0.900	0.924	0.035	0.68

Table 20: Distr log earnings net of age effects: single women

Moment	Data	Simulated	SE data	No. SE diff
secondary education				
mean	2.756	2.706	0.052	0.95
var	0.291	0.217	0.031	2.35
P(earnings<Q10)	0.100	0.087	0.021	0.60
P(earnings<Q25)	0.250	0.256	0.038	0.17
P(earnings<Q50)	0.500	0.612	0.051	2.20
P(earnings<Q75)	0.750	0.824	0.041	1.79
P(earnings<Q90)	0.900	0.927	0.027	1.01
high school				
mean	2.966	2.788	0.064	2.75
var	0.227	0.340	0.045	2.46
P(earnings<Q10)	0.100	0.188	0.035	2.46
P(earnings<Q25)	0.250	0.522	0.058	4.65
P(earnings<Q50)	0.500	0.693	0.059	3.27
P(earnings<Q75)	0.750	0.791	0.049	0.85
P(earnings<Q90)	0.900	0.863	0.026	1.34
university education				
mean	3.101	2.968	0.052	2.53
var	0.150	0.231	0.024	3.25
P(earnings<Q10)	0.100	0.269	0.027	6.24
P(earnings<Q25)	0.250	0.420	0.053	3.18
P(earnings<Q50)	0.500	0.637	0.074	1.83
P(earnings<Q75)	0.750	0.819	0.065	1.06
P(earnings<Q90)	0.900	0.875	0.048	0.49

Table 21: Distr log earnings net of age effects: men in couples

Moment	Data	Simulated	SE data	No. SE diff
secondary education				
mean	2.916	2.944	0.016	1.73
var	0.142	0.134	0.007	1.02
P(earnings<Q10)	0.100	0.098	0.010	0.10
P(earnings<Q25)	0.250	0.231	0.018	1.02
P(earnings<Q50)	0.500	0.470	0.020	1.49
P(earnings<Q75)	0.750	0.721	0.016	1.70
P(earnings<Q90)	0.900	0.881	0.010	1.66
High School				
mean	2.973	2.941	0.016	2.04
var	0.122	0.125	0.007	0.35
P(earnings<Q10)	0.100	0.135	0.012	2.94
P(earnings<Q25)	0.250	0.294	0.018	2.44
P(earnings<Q50)	0.500	0.543	0.022	1.95
P(earnings<Q75)	0.750	0.771	0.019	1.11
P(earnings<Q90)	0.900	0.902	0.011	0.20
University				
mean	3.020	2.976	0.023	1.84
var	0.114	0.122	0.013	0.63
P(earnings<Q10)	0.100	0.164	0.020	3.24
P(earnings<Q25)	0.250	0.365	0.030	3.83
P(earnings<Q50)	0.500	0.577	0.036	2.14
P(earnings<Q75)	0.750	0.773	0.031	0.75
P(earnings<Q90)	0.900	0.901	0.017	0.07

Table 22: Distr log earnings net of age effects: women in couples

Moment	Data	Simulated	SE data	No. SE diff
secondary education				
mean	2.507	2.526	0.021	0.89
var	0.338	0.275	0.013	4.87
P(earnings<Q10)	0.100	0.051	0.009	5.35
P(earnings<Q25)	0.250	0.231	0.013	1.33
P(earnings<Q50)	0.500	0.491	0.017	0.48
P(earnings<Q75)	0.750	0.770	0.015	1.31
P(earnings<Q90)	0.900	0.911	0.011	1.06
High School				
mean	2.728	2.711	0.027	0.63
var	0.355	0.329	0.020	1.21
P(earnings<Q10)	0.100	0.042	0.012	4.72
P(earnings<Q25)	0.250	0.288	0.019	1.98
P(earnings<Q50)	0.500	0.603	0.021	4.85
P(earnings<Q75)	0.750	0.787	0.018	2.08
P(earnings<Q90)	0.900	0.877	0.012	1.89
University				
mean	3.000	2.949	0.033	1.52
var	0.243	0.265	0.021	1.04
P(earnings<Q10)	0.100	0.110	0.016	0.64
P(earnings<Q25)	0.250	0.374	0.027	4.52
P(earnings<Q50)	0.500	0.613	0.033	3.35
P(earnings<Q75)	0.750	0.776	0.028	0.92
P(earnings<Q90)	0.900	0.869	0.020	1.47

Table 23: 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	2.921	2.940	0.014	1.33
mean: spouse high school	2.978	2.948	0.023	1.29
mean: spouse univ educ	2.935	2.950	0.075	0.20
var: spouse second educ	0.149	0.137	0.007	1.64
var: spouse high school	0.121	0.128	0.008	0.77
var: spouse univ educ	0.146	0.130	0.022	0.74
High School				
mean: spouse second educ	2.941	2.905	0.016	2.28
mean: spouse high school	3.001	2.964	0.019	1.93
mean: spouse univ educ	3.083	2.999	0.035	2.36
var: spouse second educ	0.119	0.130	0.009	1.21
var: spouse high school	0.133	0.118	0.008	1.80
var: spouse univ educ	0.119	0.110	0.018	0.46
University				
mean: spouse second educ	2.988	2.975	0.049	0.27
mean: spouse high school	2.999	2.974	0.033	0.74
mean: spouse univ educ	2.978	2.981	0.029	0.13
var: spouse second educ	0.149	0.122	0.038	0.72
var: spouse high school	0.125	0.123	0.027	0.04
var: spouse univ educ	0.181	0.124	0.027	2.10

Table 24: 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	2.458	2.451	0.021	0.31
mean: spouse high school	2.534	2.526	0.026	0.33
mean: spouse univ educ	2.443	2.479	0.088	0.41
var: spouse second educ	0.308	0.300	0.011	0.69
var: spouse high school	0.295	0.290	0.017	0.31
var: spouse univ educ	0.416	0.303	0.053	2.11
High School				
mean: spouse second educ	2.670	2.675	0.039	0.12
mean: spouse high school	2.732	2.716	0.030	0.53
mean: spouse univ educ	2.784	2.696	0.049	1.78
var: spouse second educ	0.378	0.322	0.023	2.35
var: spouse high school	0.352	0.334	0.020	0.89
var: spouse univ educ	0.368	0.328	0.034	1.16
University				
mean: spouse second educ	2.958	2.877	0.101	0.79
mean: spouse high school	2.953	2.944	0.056	0.16
mean: spouse univ educ	2.986	2.952	0.037	0.90
var: spouse second educ	0.327	0.268	0.083	0.70
var: spouse high school	0.319	0.257	0.032	1.92
var: spouse univ educ	0.281	0.255	0.028	0.93

Table 25: 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				
$\ln W_m \in Q_1$, $\ln W_f \in Q_1$	0.090	0.161	0.009	7.25
$\ln W_m \in Q_1$, $\ln W_f \in Q_2$	0.096	0.147	0.010	4.71
$\ln W_m \in Q_1$, $\ln W_f \in Q_3$	0.068	0.083	0.008	1.79
$\ln W_m \in Q_2$, $\ln W_f \in Q_1$	0.142	0.152	0.010	0.99
$\ln W_m \in Q_2$, $\ln W_f \in Q_2$	0.180	0.158	0.011	1.82
$\ln W_m \in Q_2$, $\ln W_f \in Q_3$	0.116	0.095	0.011	1.89
$\ln W_m \in Q_3$, $\ln W_f \in Q_1$	0.105	0.074	0.010	2.93
$\ln W_m \in Q_3$, $\ln W_f \in Q_2$	0.121	0.081	0.011	3.69
$\ln W_m \in Q_3$, $\ln W_f \in Q_3$	0.078	0.044	0.010	3.40
Women's education: high school				
$\ln W_m \in Q_1$, $\ln W_f \in Q_1$	0.089	0.110	0.011	1.75
$\ln W_m \in Q_1$, $\ln W_f \in Q_2$	0.141	0.144	0.016	0.19
$\ln W_m \in Q_1$, $\ln W_f \in Q_3$	0.093	0.094	0.014	0.02
$\ln W_m \in Q_2$, $\ln W_f \in Q_1$	0.114	0.127	0.012	1.10
$\ln W_m \in Q_2$, $\ln W_f \in Q_2$	0.178	0.156	0.014	1.50
$\ln W_m \in Q_2$, $\ln W_f \in Q_3$	0.127	0.094	0.012	2.54
$\ln W_m \in Q_3$, $\ln W_f \in Q_1$	0.080	0.090	0.012	0.78
$\ln W_m \in Q_3$, $\ln W_f \in Q_2$	0.095	0.114	0.011	1.57
$\ln W_m \in Q_3$, $\ln W_f \in Q_3$	0.079	0.067	0.014	0.86
Women's education: university				
$\ln W_m \in Q_1$, $\ln W_f \in Q_1$	0.080	0.141	0.040	1.51
$\ln W_m \in Q_1$, $\ln W_f \in Q_2$	0.114	0.137	0.044	0.53
$\ln W_m \in Q_1$, $\ln W_f \in Q_3$	0.080	0.071	0.031	0.31
$\ln W_m \in Q_2$, $\ln W_f \in Q_1$	0.128	0.121	0.040	0.19
$\ln W_m \in Q_2$, $\ln W_f \in Q_2$	0.133	0.147	0.045	0.30
$\ln W_m \in Q_2$, $\ln W_f \in Q_3$	0.123	0.085	0.039	0.99
$\ln W_m \in Q_3$, $\ln W_f \in Q_1$	0.138	0.107	0.042	0.72
$\ln W_m \in Q_3$, $\ln W_f \in Q_2$	0.123	0.119	0.040	0.11
$\ln W_m \in Q_3$, $\ln W_f \in Q_3$	0.076	0.069	0.033	0.19

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 26: 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				
$\ln W_m \in Q_1$, $\ln W_f \in Q_1$	0.087	0.149	0.017	3.63
$\ln W_m \in Q_1$, $\ln W_f \in Q_2$	0.096	0.151	0.017	3.16
$\ln W_m \in Q_1$, $\ln W_f \in Q_3$	0.068	0.083	0.015	0.95
$\ln W_m \in Q_2$, $\ln W_f \in Q_1$	0.123	0.150	0.017	1.53
$\ln W_m \in Q_2$, $\ln W_f \in Q_2$	0.165	0.164	0.021	0.02
$\ln W_m \in Q_2$, $\ln W_f \in Q_3$	0.091	0.095	0.015	0.24
$\ln W_m \in Q_3$, $\ln W_f \in Q_1$	0.108	0.076	0.018	1.75
$\ln W_m \in Q_3$, $\ln W_f \in Q_2$	0.172	0.080	0.024	3.81
$\ln W_m \in Q_3$, $\ln W_f \in Q_3$	0.085	0.047	0.016	2.27
Women's education: high school				
$\ln W_m \in Q_1$, $\ln W_f \in Q_1$	0.073	0.099	0.011	2.19
$\ln W_m \in Q_1$, $\ln W_f \in Q_2$	0.125	0.105	0.015	1.27
$\ln W_m \in Q_1$, $\ln W_f \in Q_3$	0.065	0.063	0.011	0.18
$\ln W_m \in Q_2$, $\ln W_f \in Q_1$	0.110	0.142	0.014	2.28
$\ln W_m \in Q_2$, $\ln W_f \in Q_2$	0.167	0.167	0.019	0.04
$\ln W_m \in Q_2$, $\ln W_f \in Q_3$	0.124	0.104	0.016	1.19
$\ln W_m \in Q_3$, $\ln W_f \in Q_1$	0.125	0.110	0.016	0.88
$\ln W_m \in Q_3$, $\ln W_f \in Q_2$	0.112	0.128	0.014	1.09
$\ln W_m \in Q_3$, $\ln W_f \in Q_3$	0.095	0.078	0.015	1.10
Women's education: university				
$\ln W_m \in Q_1$, $\ln W_f \in Q_1$	0.088	0.137	0.032	1.52
$\ln W_m \in Q_1$, $\ln W_f \in Q_2$	0.089	0.145	0.021	2.54
$\ln W_m \in Q_1$, $\ln W_f \in Q_3$	0.083	0.085	0.020	0.08
$\ln W_m \in Q_2$, $\ln W_f \in Q_1$	0.081	0.118	0.024	1.48
$\ln W_m \in Q_2$, $\ln W_f \in Q_2$	0.224	0.145	0.036	2.15
$\ln W_m \in Q_2$, $\ln W_f \in Q_3$	0.147	0.080	0.025	2.63
$\ln W_m \in Q_3$, $\ln W_f \in Q_1$	0.099	0.102	0.021	0.14
$\ln W_m \in Q_3$, $\ln W_f \in Q_2$	0.104	0.118	0.021	0.66
$\ln W_m \in Q_3$, $\ln W_f \in Q_3$	0.080	0.066	0.017	0.82

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 27: 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				
$\ln W_m \in Q_1$, $\ln W_f \in Q_1$	0.069	0.199	0.041	3.13
$\ln W_m \in Q_1$, $\ln W_f \in Q_2$	0.167	0.134	0.081	0.41
$\ln W_m \in Q_1$, $\ln W_f \in Q_3$	0.027	0.061	0.019	1.68
$\ln W_m \in Q_2$, $\ln W_f \in Q_1$	0.125	0.190	0.055	1.17
$\ln W_m \in Q_2$, $\ln W_f \in Q_2$	0.216	0.148	0.052	1.31
$\ln W_m \in Q_2$, $\ln W_f \in Q_3$	0.020	0.059	0.017	2.24
$\ln W_m \in Q_3$, $\ln W_f \in Q_1$	0.146	0.098	0.061	0.78
$\ln W_m \in Q_3$, $\ln W_f \in Q_2$	0.139	0.071	0.051	1.32
$\ln W_m \in Q_3$, $\ln W_f \in Q_3$	0.083	0.036	0.037	1.27
Women's education: high school				
$\ln W_m \in Q_1$, $\ln W_f \in Q_1$	0.071	0.103	0.021	1.45
$\ln W_m \in Q_1$, $\ln W_f \in Q_2$	0.045	0.077	0.013	2.46
$\ln W_m \in Q_1$, $\ln W_f \in Q_3$	0.030	0.034	0.015	0.27
$\ln W_m \in Q_2$, $\ln W_f \in Q_1$	0.121	0.190	0.028	2.41
$\ln W_m \in Q_2$, $\ln W_f \in Q_2$	0.140	0.164	0.031	0.77
$\ln W_m \in Q_2$, $\ln W_f \in Q_3$	0.118	0.075	0.027	1.53
$\ln W_m \in Q_3$, $\ln W_f \in Q_1$	0.129	0.154	0.031	0.80
$\ln W_m \in Q_3$, $\ln W_f \in Q_2$	0.161	0.134	0.029	0.91
$\ln W_m \in Q_3$, $\ln W_f \in Q_3$	0.181	0.066	0.042	2.74
Women's education: university				
$\ln W_m \in Q_1$, $\ln W_f \in Q_1$	0.080	0.162	0.015	5.33
$\ln W_m \in Q_1$, $\ln W_f \in Q_2$	0.126	0.137	0.022	0.48
$\ln W_m \in Q_1$, $\ln W_f \in Q_3$	0.077	0.061	0.019	0.85
$\ln W_m \in Q_2$, $\ln W_f \in Q_1$	0.120	0.152	0.023	1.38
$\ln W_m \in Q_2$, $\ln W_f \in Q_2$	0.168	0.129	0.020	1.90
$\ln W_m \in Q_2$, $\ln W_f \in Q_3$	0.117	0.064	0.022	2.42
$\ln W_m \in Q_3$, $\ln W_f \in Q_1$	0.073	0.127	0.019	2.81
$\ln W_m \in Q_3$, $\ln W_f \in Q_2$	0.133	0.108	0.020	1.16
$\ln W_m \in Q_3$, $\ln W_f \in Q_3$	0.102	0.055	0.023	1.97

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 28: Distribution of log earnings among men - autocorrelation of earnings net of age effects

Moment	Data	Simulated	SE data	No. SE diff
Secondary education				
variance	0.150	0.142	0.006	1.33
1st autocorrelation	0.111	0.113	0.004	0.46
2nd autocorrelation	0.100	0.100	0.007	0.07
3rd autocorrelation	0.092	0.088	0.007	0.41
High School				
variance	0.134	0.127	0.005	1.35
1st autocorrelation	0.105	0.102	0.003	0.83
2nd autocorrelation	0.098	0.091	0.006	1.04
3rd autocorrelation	0.094	0.082	0.006	1.82
University				
variance	0.163	0.127	0.014	2.43
1st autocorrelation	0.124	0.108	0.009	1.57
2nd autocorrelation	0.106	0.099	0.016	0.46
3rd autocorrelation	0.098	0.090	0.015	0.51

Table 29: Distribution of log earnings among women - autocorrelation of earnings net of age effects

Moment	Data	Simulated	SE data	No. SE diff
Secondary education				
variance	0.308	0.294	0.008	1.58
1st autocorrelation	0.260	0.240	0.007	2.76
2nd autocorrelation	0.239	0.221	0.013	1.31
3rd autocorrelation	0.224	0.206	0.015	1.20
High School				
variance	0.351	0.331	0.012	1.61
1st autocorrelation	0.296	0.273	0.011	2.00
2nd autocorrelation	0.280	0.244	0.019	1.80
3rd autocorrelation	0.266	0.220	0.020	2.24
University				
variance	0.294	0.256	0.019	1.94
1st autocorrelation	0.237	0.201	0.014	2.61
2nd autocorrelation	0.221	0.171	0.023	2.12
3rd autocorrelation	0.210	0.145	0.024	2.66

Table 30: Regression coefficients - employment on age, men

Moment	Data	Simulated	SE data	No. SE diff
Secondary education, married				
intercept	0.801	0.824	0.046	0.48
age	0.187	0.086	0.102	0.98
age square	-0.111	-0.003	0.071	1.51
age cube	0.016	-0.016	0.015	2.16
Secondary education, singles				
intercept	0.841	0.818	0.122	0.19
age	-0.121	-0.229	0.307	0.35
age square	-0.030	0.102	0.224	0.59
age cube	0.017	-0.017	0.049	0.70
High School, married				
intercept	0.913	0.928	0.032	0.44
age	0.054	0.050	0.072	0.05
age square	-0.018	-0.018	0.051	0.01
age cube	-0.000	-0.001	0.011	0.08
High School, singles				
intercept	0.542	0.759	0.163	1.33
age	0.517	-0.018	0.353	1.51
age square	-0.301	0.009	0.251	1.24
age cube	0.056	0.004	0.055	0.93
University, married				
intercept	0.964	0.932	0.034	0.92
age	0.009	0.067	0.088	0.65
age square	-0.000	-0.042	0.066	0.63
age cube	-0.002	0.006	0.014	0.63
University, singles				
intercept	0.747	0.781	0.237	0.14
age	0.168	0.138	0.547	0.05
age square	-0.071	-0.031	0.370	0.11
age cube	0.007	-0.010	0.076	0.24

Table 31: Regression coefficients - employment on age, men

Moment	Data	Simulated	SE data	No. SE diff
Secondary education, married				
intercept	0.589	0.600	0.055	0.20
age	0.000	-0.042	0.123	0.34
age square	0.122	0.149	0.084	0.32
age cube	-0.039	-0.051	0.018	0.65
Secondary education, singles				
intercept	0.738	0.664	0.141	0.52
age	-0.377	-0.246	0.303	0.43
age square	0.265	0.155	0.209	0.53
age cube	-0.062	-0.040	0.044	0.48
High School, married				
intercept	0.819	0.804	0.056	0.26
age	-0.097	-0.012	0.128	0.66
age square	0.089	-0.014	0.092	1.13
age cube	-0.019	0.013	0.020	1.64
High School, singles				
intercept	0.701	0.652	0.189	0.26
age	0.186	0.252	0.386	0.17
age square	-0.127	-0.175	0.257	0.19
age cube	0.030	0.040	0.053	0.19
University, married				
intercept	0.917	0.834	0.071	1.16
age	-0.054	0.104	0.187	0.85
age square	-0.042	-0.139	0.143	0.67
age cube	0.023	0.043	0.032	0.63
University, singles				
intercept	0.647	0.652	0.287	0.02
age	0.427	0.284	0.648	0.22
age square	-0.199	-0.072	0.438	0.29
age cube	0.021	-0.009	0.089	0.34

Table 32: Distribution of employment among men

Moment	Data	Simulated	SE data	No. SE diff
Secondary education, married				
employment rate	0.869	0.868	0.007	0.09
% time employed < 0.2	0.069	0.018	0.008	6.09
% time employed < 0.4	0.103	0.055	0.009	4.94
% time employed < 0.6	0.133	0.146	0.011	1.07
% time employed < 0.8	0.188	0.292	0.013	7.66
Secondary education, singles				
employment rate	0.670	0.674	0.031	0.13
% time employed < 0.2	0.275	0.246	0.038	0.73
% time employed < 0.4	0.355	0.341	0.043	0.31
% time employed < 0.6	0.434	0.380	0.044	1.22
% time employed < 0.8	0.500	0.417	0.044	1.84
High School, married				
employment rate	0.944	0.949	0.005	0.96
% time employed < 0.2	0.023	0.000	0.005	4.47
% time employed < 0.4	0.032	0.001	0.006	4.91
% time employed < 0.6	0.042	0.020	0.007	3.13
% time employed < 0.8	0.073	0.107	0.008	3.93
High School, singles				
employment rate	0.811	0.774	0.026	1.40
% time employed < 0.2	0.135	0.039	0.035	2.71
% time employed < 0.4	0.184	0.138	0.039	1.17
% time employed < 0.6	0.242	0.285	0.042	1.01
% time employed < 0.8	0.359	0.423	0.046	1.39
University, married				
employment rate	0.962	0.957	0.005	0.93
% time employed < 0.2	0.006	0.000	0.004	1.37
% time employed < 0.4	0.016	0.000	0.007	2.13
% time employed < 0.6	0.026	0.007	0.009	2.02
% time employed < 0.8	0.053	0.065	0.012	0.92
University, singles				
employment rate	0.851	0.853	0.030	0.07
% time employed < 0.2	0.078	0.003	0.039	1.91
% time employed < 0.4	0.078	0.026	0.039	1.33
% time employed < 0.6	0.137	0.098	0.049	0.78
% time employed < 0.8	0.294	0.368	0.065	1.13

Table 33: Distribution of employment among women

Moment	Data	Simulated	SE data	No. SE diff
Secondary, married				
employment rate	0.712	0.667	0.009	4.56
% time employed < 0.2	0.158	0.182	0.011	2.10
% time employed < 0.4	0.239	0.286	0.012	3.88
% time employed < 0.6	0.333	0.386	0.013	4.09
% time employed < 0.8	0.441	0.502	0.013	4.48
Secondary, singles				
employment rate	0.550	0.511	0.026	1.46
% time employed < 0.2	0.338	0.342	0.031	0.12
% time employed < 0.4	0.419	0.455	0.032	1.12
% time employed < 0.6	0.494	0.565	0.033	2.11
% time employed < 0.8	0.590	0.681	0.030	2.97
High School, married				
employment rate	0.807	0.808	0.010	0.05
% time employed < 0.2	0.085	0.039	0.010	4.40
% time employed < 0.4	0.142	0.093	0.013	3.75
% time employed < 0.6	0.218	0.196	0.016	1.42
% time employed < 0.8	0.321	0.379	0.018	3.06
High School, singles				
employment rate	0.803	0.768	0.027	1.27
% time employed < 0.2	0.084	0.059	0.024	1.01
% time employed < 0.4	0.158	0.134	0.033	0.74
% time employed < 0.6	0.242	0.258	0.039	0.39
% time employed < 0.8	0.308	0.466	0.044	3.59
University, married				
employment rate	0.838	0.841	0.015	0.14
% time employed < 0.2	0.062	0.004	0.014	4.12
% time employed < 0.4	0.102	0.033	0.017	3.97
% time employed < 0.6	0.158	0.105	0.022	2.35
% time employed < 0.8	0.261	0.351	0.028	3.21
University, singles				
employment rate	0.868	0.834	0.034	1.00
% time employed < 0.2	0.071	0.011	0.041	1.46
% time employed < 0.4	0.095	0.051	0.048	0.91
% time employed < 0.6	0.119	0.164	0.049	0.92
% time employed < 0.8	0.166	0.384	0.054	3.98

Figure 5: 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).

Table 34: Matching moments - marital patterns

	data moments by (θ, s)						simulated moments by (θ, s)						No. SE diff					
	(1,1)	(1,2)	(1,3)	(2,1)	(2,2)	(2,3)	(1,1)	(1,2)	(1,3)	(2,1)	(2,2)	(2,3)	(1,1)	(1,2)	(1,3)	(2,1)	(2,2)	(2,3)
Men, marrying woman (θ, s)																		
remain single	0.121	0.171	0.097	0.020	0.042	0.007	0.128	0.181	0.118	0.021	0.047	0.011	0.65	0.50	1.45	0.22	0.48	0.60
marry (1,1)	0.118	0.086	0.012	0.187	0.069	0.009	0.100	0.077	0.000	0.179	0.065	0.000	0.67	0.43	1.99	0.40	0.46	1.91
(1,2)	0.030	0.055	0.024	0.056	0.047	0.011	0.059	0.028	0.025	0.034	0.102	0.029	2.04	1.40	0.08	1.84	3.75	3.53
(1,3)	0.004	0.008	0.012	0.006	0.004	0.006	0.000	0.009	0.007	0.000	0.015	0.030	0.61	0.03	0.72	1.58	3.10	5.26
(2,1)	0.057	0.101	0.004	0.158	0.073	0.011	0.103	0.080	0.000	0.205	0.096	0.011	2.14	1.24	1.14	2.35	2.24	0.02
(2,2)	0.021	0.041	0.009	0.058	0.094	0.038	0.034	0.015	0.016	0.020	0.065	0.020	0.90	1.50	0.86	3.25	1.90	2.59
(2,3)	0.003	0.007	0.020	0.011	0.036	0.056	0.000	0.009	0.010	0.001	0.015	0.035	0.74	0.35	1.10	2.77	4.32	3.31
Women, marrying man (θ, s)																		
remain single	0.046	0.024	0.016	0.046	0.121	0.063	0.047	0.032	0.027	0.048	0.132	0.069	0.06	0.70	0.60	0.16	1.23	0.44
marry (1,1)	0.089	0.022	0.003	0.031	0.011	0.002	0.073	0.045	0.000	0.068	0.022	0.000	0.77	1.94	0.69	2.92	1.26	0.77
(1,2)	0.065	0.041	0.006	0.055	0.022	0.004	0.053	0.019	0.008	0.050	0.010	0.006	0.77	1.46	0.34	0.48	1.20	0.85
(1,3)	0.009	0.019	0.009	0.002	0.005	0.011	0.000	0.019	0.007	0.002	0.011	0.006	1.96	0.01	0.49	0.33	1.17	0.87
(2,1)	0.308	0.092	0.009	0.185	0.068	0.013	0.258	0.056	0.000	0.257	0.023	0.001	1.60	1.91	1.74	3.37	3.30	2.62
(2,2)	0.114	0.077	0.007	0.085	0.110	0.043	0.096	0.149	0.022	0.115	0.077	0.022	1.03	3.16	2.86	2.51	1.85	3.75
(2,3)	0.014	0.018	0.010	0.013	0.045	0.066	0.000	0.046	0.044	0.012	0.027	0.043	1.98	3.43	4.68	0.23	2.26	2.93

Table 35: Distribution of education by gender

Moment	Data	Simulated	SE data	No. SE diff
Men				
Secondary	0.452	0.447	0.008	0.50
High school	0.396	0.401	0.008	0.64
University	0.151	0.150	0.005	0.19
Women				
Secondary	0.511	0.540	0.008	3.61
High school	0.346	0.331	0.007	1.84
University	0.142	0.128	0.005	2.56

Table 36: Regression coefficients: education on residual parental income by gender

Moment	Data	Simulated	SE data	No. SE diff
Men, Secondary				
parental income: residual	-0.148	-0.121	0.047	0.55
parental income: residual squared	0.022	0.007	0.008	1.76
Men, High School				
parental income: residual	0.005	0.113	0.051	2.13
parental income: residual squared	-0.001	-0.006	0.008	0.65
Women, Secondary				
parental income: residual	-0.133	-0.143	0.049	0.21
parental income: residual squared	0.045	-0.003	0.055	0.88
Women, High School				
parental income: residual	0.044	0.131	0.055	1.59
parental income: residual squared	-0.032	0.029	0.062	0.99