Marriage with Labor Supply

Nicolas Jacquemet* Jean-Marc Robin†

June 2010, *In progress*

Abstract

One of the key issues in understanding how tax policies affect labour supply is intrahousehold allocation of time and consumption. This is in particular the case of welfare benefits aimed at providing a safety net against poverty and work incentives at the same time, such as the Working Family Tax Credit programme in the UK (and the Earned Income Tax Credit in the US). The models used to address these issues typically take the household as a unit with aggregate preferences. The collective models of the family go one step further by describing intrahousehold resource allocation as a Pareto equilibrium for the exchange economy comprising each family member endowed with his/her own preferences. The limitation of the Collective framework for policy evaluation lies in the multiplicity of equilibria and the lack of a selection device that could tell not only how a welfare policy will affect resource allocation for a given sharing rule but also how it will affect the sharing rule itself. In this paper we design a search-matching models similar to that can both explain the formation of couples and select the equilibrium on the curve of optimal contracts.

Keywords: Marriage search model, Collective labor supply, Structural estimation.

JEL classification: C78, D83, J12, J22.

1 Introduction

One of the key issues in understanding how tax policies affect labour supply is intrahousehold allocation of time and consumption. This is in particular the case of welfare benefits aimed at providing a safety net against poverty and work incentives at the same time, such as the Working Family Tax Credit programme in the UK and the Earned Income Tax Credit in the US. The models used to address these issues typically take the household as a unit with aggregate preferences. The collective models of the family Chiappori (1988, 1992), Blundell, Chiappori, Magnac, and Meghir (2007) go one step further by

*Paris School of Economics and University Paris 1 Panthéon–Sorbonne ; Nicolas.Jacquemet@univ-paris1.fr ; Centre d'Economie de la Sorbonne, MSE, 106 Bd. de l'hôpital, 75013 Paris, France.

†Sciences Po, Paris and University College of London; jeanmarc.robin@sciences-po.fr ; Department of Economics, 28 rue des St Peres, 75007 Paris, France.

describing intrahousehold resource allocation as a Pareto equilibrium for the exchange economy comprising each family member endowed with his/her own preferences. The limitation of the Collective framework for policy evaluation lies in the multiplicity of equilibria and the lack of a selection device that could tell not only how a welfare policy affects resource allocation for a given sharing rule but also how it affects the sharing rule itself. In this paper we design a search-matching models in the spirit of Sattinger (1995), Lu and McAfee (1996), Shimer and Smith (2000) that can both explain the formation of couples and the selected equilibrium on the curve of optimal contracts.

The main characteristics of the model are as follows. Men and women are characterised by their wage. Each individual has an indirect utility that is a function of her own wage and of her spouse's wage via a preference externality that is created within the marriage – otherwise individuals would remain single. Nash bargaining determines the optimal transfer of resources. Single men and women randomly meet on the marriage market and the marriage is consumated if the surplus is positive.

The first modelling difficulty that we face is that we have to find a way of modelling the positive externalities generated by the marriage. A typical argument is that individuals form families to rear children. More generally, public goods and economies of scale can explain the formation of couples. We shall model these externalities in a simple way by introducing a home production process that raises the household's income.

The second difficulty is identification. We shall use data on wages and labour hours. Showing that this information is sufficient to identify the model's structure is not straightforward, and we are able to show partial results. As the model rules out idiosyncratic shocks to individuals' characteristics (wages, preferences), cross-sectional data are essentially enough for estimation. Identification is partial but constructive. We estimate the identified parameters mostly nonparametrically using data on wages, hours of work drawn from the US SIPP.

We finally use the estimated to analyse the EITC. Eissa and Hoynes (2004) find strong positive effects on the labour market participation of lone mothers but in the case of married mothers, the EITC has led to a small reduction in labour market participation – about 1 percentage point. This occurs because the credit is based on family earnings and income. To really understand this phenomenon, a collective model is required.

2 The Model

2.1 Setup

We consider a marriage market with L_m males and L_f females. The number of married couples is denoted as N and the respective numbers of single males and single females are denoted as U_m and U_f

with

$$U_m = L_m - N$$
$$U_f = L_f - N.$$

Note that this implies in particular that U_m and U_f are linked by the relation:

$$L_m - U_m = L_f - U_f = N.$$

Individuals differ in labor productivity, denoted as $x \in [x_{min}, x_{max}]$ for males and $y \in [y_{min}, y_{max}]$ for females. Let $\ell_m(x)$ and $\ell_f(y)$ denote the respective measures of males of type x and females of type y, with $L_m = \int \ell_m(x)\,\mathrm{d}x$ and $L_f = \int \ell_f(y)\,\mathrm{d}y$. The corresponding measures of wages in the sub-populations of singles are denoted as $u_m(x)$ and $u_f(y)$, with $U_m = \int u_m(x)\,\mathrm{d}x$ and $U_f = \int u_f(y)\,\mathrm{d}y$. The number of couples of type (x, y) is denoted as $n(x, y)$, with $N = \iint n(x, y)\,\mathrm{d}x\,\mathrm{d}y$ and

$$\ell_m(x) = \int n(x, y)\,\mathrm{d}y + u_m(x) \tag{1}$$
$$\ell_f(y) = \int n(x, y)\,\mathrm{d}x + u_f(y).$$

We assume that only singles search for a partner, ruling out "on-the-mariage" search. The number of meetings per period is measured by a meeting function $M(U_m, U_f)$, and $\lambda_i = \frac{M(U_m,U_f)}{U_i}$ is the instantaneous probability that a searching individual of gender i meet with a possible partner. We also denote $\lambda = \dfrac{M(U_m, U_f)}{U_m U_f}$.

All meetings do not result in a match. We assume that there exists a function $\alpha(x, y) \in [0, 1]$ indicating the probability that a match (x, y) be consumated. The matching probability is an equilibrium outcome that will be later determined. The matching set $\mathcal{M}$ is the support of distribution α.

Matches are exogenously dissolved with instantaneous probability δ .

2.2 Flow equations

In steady state flows in and out of the stocks of married couples must exactly balance each other out. This means that, for all (x, y),

$$\delta n(x, y) = u_m(x)\,\lambda_m \frac{u_f(y)}{U_f}\alpha(x, y) = \lambda u_m(x)\,u_f(y)\,\alpha(x, y). \tag{2}$$

The left-hand side is the flow of divorces. The right-hand side measures the flow of new marriages as the product of the measure of male singles of type x with the probability of a contact and the probability of drawing a female single of type y.

Integrating over x:

$$\delta \int n(x,y)\,\mathrm{d}y = \lambda u_m(x) \int u_f(y)\alpha(x,y)\,\mathrm{d}y$$

and using equation (1) to substitute $\int n(x,y)dy$ out of this equation, yields the following equilibrium conditions for u_m:

$$\delta\left[\ell_m(x) - u_m(x)\right] = \lambda u_m(x) \int u_f(y)\alpha(x,y)\,\mathrm{d}y$$

or

$$u_m(x) = \frac{\delta \ell_m(x)}{\delta + \lambda \int u_f(y)\alpha(x,y)\,\mathrm{d}y}. \tag{3}$$

Symmetrically, the equation defining the equilibrium distribution of wages in the population of single females is

$$u_f(y) = \frac{\delta \ell_f(y)}{\delta + \lambda \int u_m(x)\alpha(x,y)\,\mathrm{d}x}. \tag{4}$$

2.3 Utility flows

Individuals draw utility from consumption and leisure. Let

$$v_m(x, xT+t) = \frac{xT + t - A_m(x)}{B_m(x)} \tag{5}$$

$$v_f(y, yT+t) = \frac{yT + t - A_f(y)}{B_f(y)} \tag{6}$$

denote the indirect utility of male wage x (resp. female wage y) and nonlabour income t. T denotes total time endowment, $B_m(x)$ and $B_f(y)$ are aggregate price indices and $A_m(x)$ and $A_f(y)$ are minimum expenditures to attain a positive utility. The indirect utility function maximises the utility of consumption and leisure subject to the budget constraint $c = wh + t$, for a generic wage of w, consumption c and hours worked h (or leisure $T - h$), and to the participation constraints $c > 0$ and $T - h > 0$.

To keep the model easily tractable we rule out labour market nonparticipation. So, hours worked follow by Roy's identity as

$$h_m(x, xT+t) = T - A'_m(x) - b'_m(x)[xT + t - A_m(x)] \tag{7}$$

$$h_f(y, yT+t) = T - A'_f(y) - b'_f(y)[yT + t - A_f(y)] \tag{8}$$

where $b_m(x) = \log B_m(x)$ and $b_f(y) = \log B_f(y)$, and b'_m and b'_f denote derivatives.

2.4 Optimal rent sharing between spouses

Let $W_m(v,x)$ denote the present value of a married male of type x receiving a flow utility v from marriage, and let $W_m(x)$ denote the value for single men (derived in the next section). Equating annuities to expected income flows links values as

$$rW_m(v,x) = v + \delta\left[W_m(x) - W_m(v,x)\right]$$
$$rW_f(v,y) = v + \delta\left[W_f(y) - W_f(v,y)\right].$$

Then, define individual surpluses as

$$S_m(v,x) = W_m(v,x) - W_m(x) = \frac{v - rW_m(x)}{r+\delta}$$
$$S_f(v,y) = W_f(v,y) - W_f(y) = \frac{v - rW_f(y)}{r+\delta}.$$

They express the return to marriage as the difference between the value of a flow utility v during marriage and the value of singlehood.

We assume that spouses share ressources cooperatively using Nash bargaining, whereby transfers t_m and t_f solve

$$\max_{t_m,t_f} S_m\left(v_m(x, xT+t_m), x\right)^{\beta} S_f\left(v_f(y, yT+t_f), y\right)^{1-\beta}$$

subject to the condition

$$t_m + t_f \leq C(x,y) + z$$

where $C(x,y)+z$ is the public good that is produced in the marriage. It is supposed to be a function of wages x and y, and of a match-specific component z that is drawn from a zero-mean distribution G independently of x and y.

With quasi-linear utility functions, the solution is trivially found to be such that

$$t_m(x,y,z) - s_m(x) = \beta[C(x,y) + z - s_m(x) - s_f(y)] \tag{9}$$

and

$$t_f(x,y,z) - s_f(y) = (1-\beta)[C(x,y) + z - s_m(x) - s_f(y)] \tag{10}$$

where we denote

$$s_m(x) = B_m(x)rW_m(x) - xT + A_m(x)$$
$$s_f(y) = B_f(y)rW_f(y) - yT + A_f(y).$$

These functions are different from zero if singles can expect a return from marriage. They express the expected returns, in value, from searching for a spouse when single. In other words, they are the Average Treatment Effect of the search strategy.

Singles x and y decide to match if the overall surplus is positive, i.e.

$$s(x,y) + z > 0 \tag{11}$$

with

$$s(x,y) = C(x,y) - s_m(x) - s_f(y). \tag{12}$$

The matching probability can then be calculated as

$$\begin{aligned}\alpha(x,y) &= \Pr\{s(x,y) + z > 0 | x, y\} \\ &= 1 - G\left(-s(x,y)\right). \end{aligned} \tag{13}$$

2.5 The values for singles

The value of being single, for males, solves the option value equation:

$$\begin{aligned} rW_m(x) &= v_m(x, xT) + \lambda \iint \max\{S_m\left(v_m(x, xT + t_m(x,y,z)), x\right), 0\} \, \mathrm{d}G(z) \, u_f(y) \, \mathrm{d}y \\ &= v_m(x, xT) + \frac{\lambda}{r+\delta} \iint \max\{v_m(x, xT + t_m(x,y,z)) - rW_m(x), 0\} \, \mathrm{d}G(z) \, u_f(y) \, \mathrm{d}y. \end{aligned}$$

Equivalently,

$$s_m(x) = \frac{\lambda\beta}{r+\delta} \iint \max\{z + C(x,y) - s_m(x) - s_f(y), 0\} \, \mathrm{d}G(z) \, u_f(y) \, \mathrm{d}y. \tag{14}$$

A similar expression can be derived for females:

$$s_f(y) = \frac{\lambda(1-\beta)}{r+\delta} \iint \max\{z + C(x,y) - s_m(x) - s_f(y), 0\} \, \mathrm{d}G(z) \, u_m(x) \, \mathrm{d}x. \tag{15}$$

2.6 Equilibrium

The equilibrium is characterized as the fixed point (u_m, u_f, w_m, w_f) of the following system of equations. The first two equations determine equilibrium distributions of wages amongst singles; the last two equations determine equilibrium values.

$$u_m(x) = \frac{\ell_m(x)}{1 + \frac{\lambda}{\delta} \int u_f(y)\alpha(x,y) \, \mathrm{d}y} \tag{16}$$

$$u_f(y) = \frac{\ell_f(y)}{1 + \frac{\lambda}{\delta} \int u_m(x)\alpha(x,y) \, \mathrm{d}x} \tag{17}$$

$$s_m(x) = \frac{k\frac{\lambda}{\delta}\beta \iint \max\{z + C(x,y) - s_f(y), s_m(x)\} \, \mathrm{d}G(z) \, u_f(y) \, \mathrm{d}y}{1 + k\frac{\lambda}{\delta}\beta U_f} \tag{18}$$

$$s_f(y) = \frac{k\frac{\lambda}{\delta}(1-\beta) \iint \max\{z + C(x,y) - s_m(x), s_f(y)\} \, \mathrm{d}G(z) \, u_m(x) \, \mathrm{d}x}{1 + k\frac{\lambda}{\delta}(1-\beta) U_m} \tag{19}$$

where$k = \frac{1}{r/\delta+1}$ and

$$\begin{aligned} U_m &= \int u_m(x)\,\mathrm{d}x, \quad U_f = \int u_f(y)\,\mathrm{d}y \\ \lambda &= \frac{M(U_m, U_f)}{U_m U_f} \\ \alpha(x,y) &= 1 - G\left(s_m(x) + s_f(y) - C(x,y)\right). \end{aligned}$$

Note that equations (18), (19) rewrite equations (14), (15) so that s_m and s_f are now fixed points of contracting operators (given u_m and u_f).

Shimer and Smith (2000) prove the existence of an equilibrium for a much simpler version of the search-matching equilibrium. We shall not attempt here at extending their result. Practically, we solve the fixed point problem by value function iteration, which seems to work in most cases despite our inability to prove that the equilibrium operator is contracting.

3 Identification/Estimation

3.1 Data

We use the US Survey of Income and Program Participation (SIPP) from 1996-1999. For every year that an individual is in each panel we collect information on the labour market state at the time of the survey (quarterly), wages at the time of the survey if employed, the number of hours worked, gender, and the corresponding information for the respondent's spouse if married. Our sample is restricted to individuals who are not self-employed or in the military, between the ages of 21 and 65. We assume the environment stationary and calculate individuals' mean wages over employment spells and mean hours worked over all quarters, including non-employment spells. Thus, we somewhat reduce the transitory noise in wages and hours, and we reduce the number of labour-market non-participation spells (zeros). Then we drop all observations with zero hours worked (individuals and individuals' spouses never being employment over the 4-year period). This is definitely not a satisfactory procedure but the model cannot deal (at the moment) for both the extensive and the intensive margins of labour market participation. We also trim the 1% top and bottom wages.

In the US, the 2001 Census shows (Current Population Reports) that 30.1% of men (24.6% women) 15 years and are not married and 21 % (23.1% women) are divorcees. The median age at first marriage is 24 for men and 21.8 for women. The median age when they first divorce is 31.5 for men and 29.4 for women. The median duration of first marriages is 8.2 years (men) and 7.9 (women). The median duration between first divorce and remarriage is 3.3 years (men) and 3.5 (women), and this second mariage lasts 9.2 and 8.1 years.

About 75-80% of first marriages, depending on cohorts, reach 10 years, 60-65% 20 years, 50-60% 30 years. This is consistent a separation rate of around 2.5% per year. For second marriages 70-80%

reach 10 years, 55% 15 years and 50% 20. The separation rate is thus higher, around 3% annual. The average marriage length should thus be of 33-40 years.

The Poisson assumption that we use in the model thus does not seem to fit the data well. It seems that a large proportion of marriages never end, and those who do end relatively fast. In the estimation, we shall use a separation rate of 5% annual (average marriage duration equals 20 years), which yields $\delta = 0.0042$ per month. We shall use a discount rate of $r \simeq 2.5\%$ annual. Hence, $\frac{\lambda}{r+\delta}$ is set equal to $\frac{2}{3}\frac{\lambda}{\delta}$.

In our sample, we have $2N/(2N + U_m + U_f) = 66\%$ of the population that is married, and there is a deficit of single males vis-a-vis single females: $U_m/U_f = 66\%$ ($N = 8374, U_m = 3366, U_f = 5104$).

Let (x, y, h_m, h_f) denote an observation for a married couple and let (x, h_m^0) and (y, h_f^0) denote observations for single males and females. By definition, $h_m \equiv h_m(x, xT+t_m(x, y, z))$, $h_m^0 \equiv h_m(x, xT)$, with symmetric expressions for h_f and h_f^0. We set the maximal number of hours T equal to the upper bound of h_m, h_f in the sample, i.e. $T = 595$ hours per month.

3.2 Wage distributions

Figure 1, panel (a), shows the (kernel) densities of wages separately for male and female individuals. A clear stochastic ordering appears: married males have higher wages on average and more dispersed wages than single males, which are in the same relationship with female wages. Married females' wages are strikingly similar to single females' wages. Panel (b) displays the corresponding CDFs. The wage scale is in logs so as to emphasize the non-normality of the distributions, both tails being fatter than normal.

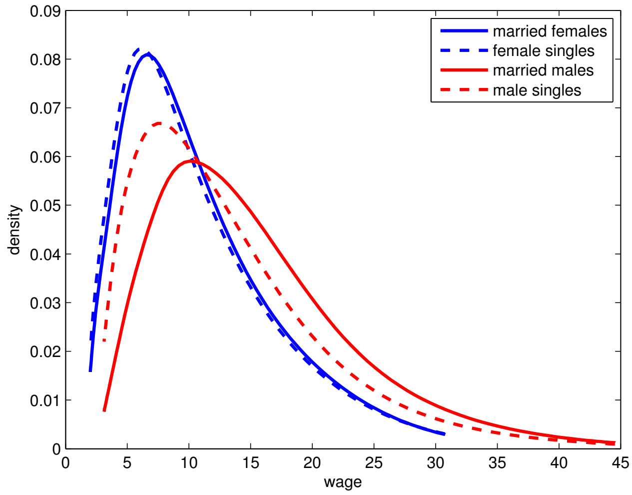
Lastly, we show the joint distribution of wages among married couples, also estimated using a Gaussian kernel (Figure 1). The most salient aspect of this distribution is that it has a very large support. Virtually no wage configuration, like a low male wage and a high female wage or vice versa, sems impossible. Moreover, wages display a relatively low correlation (31%), however positive. The wage density is indeed clearly oriented along the dominant diagonal (see the flat projection in the right panel).

This is, we believe, a very important observation. If there is sorting in economics, and more generally in social sciences, it is likely to be very imperfect. So many distinct claracteristics matter, so imperfect is the information about alternative offers, that the matching set is bound to be very wide. This is why we believe that it is important that matching models allow for both imperfection information (sequential search) and unobserved match characteristics (here modelled via a simple univariate shock z).

3.3 Hours

Figure 3 displays nonparametric kernel estimates of expected hours given own wage for singles and married individuals. Married males work more than single males, who work more than single females;

(a) Density



(b) CDF

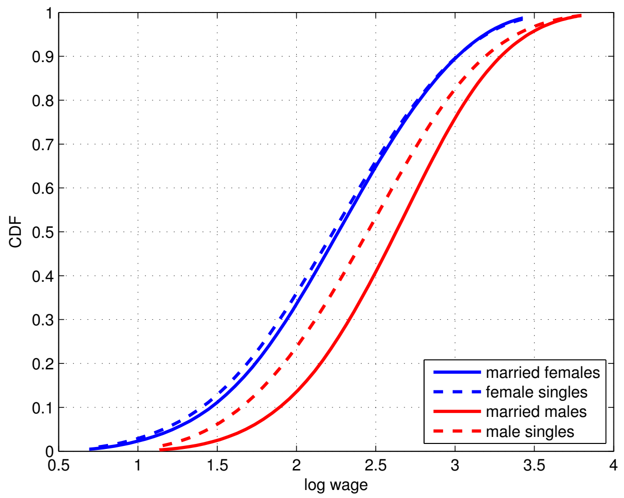


Figure 1: Wage distribution for married vs single individuals

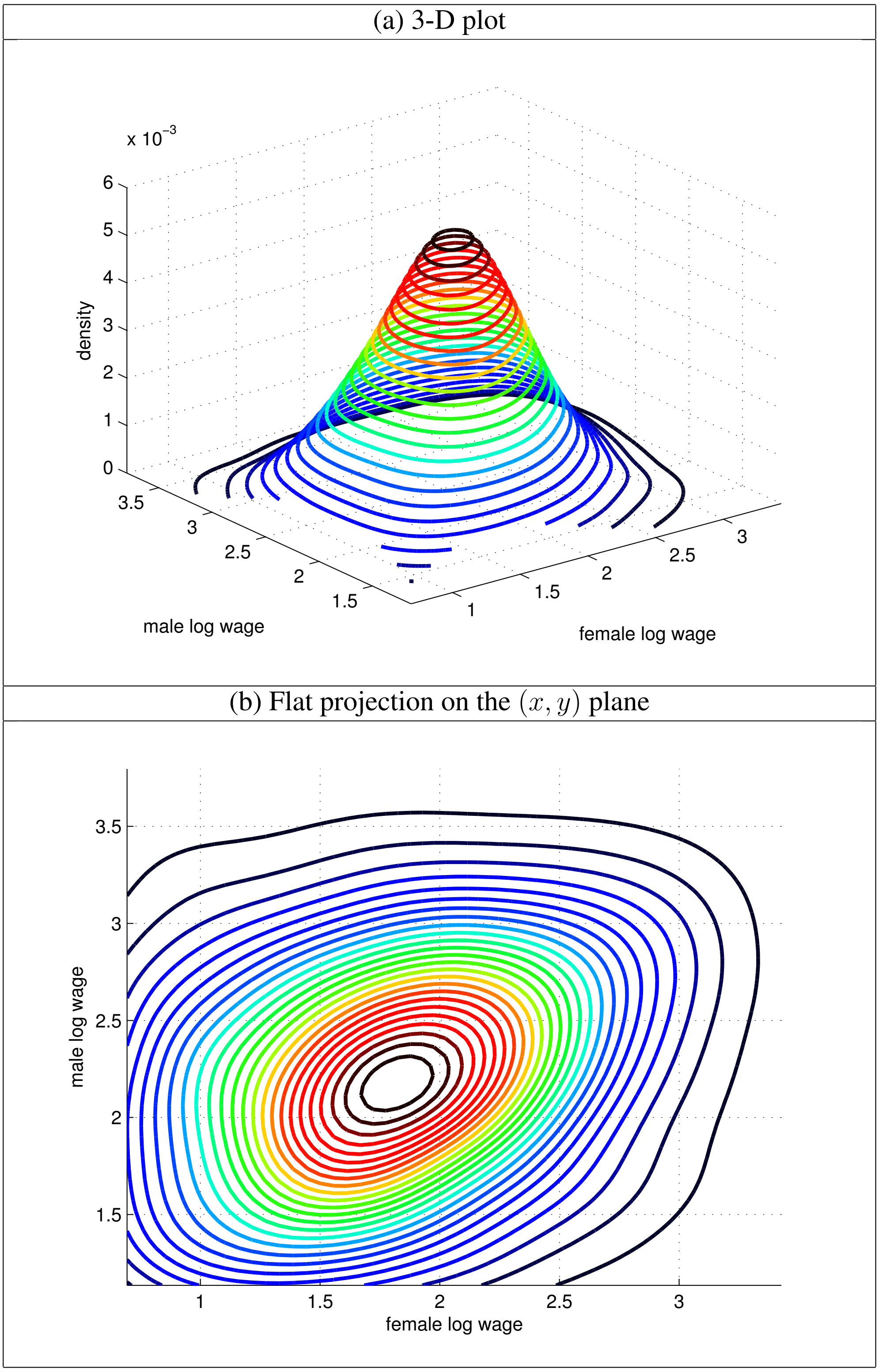


Figure 2: Joint log-wage density for married couples

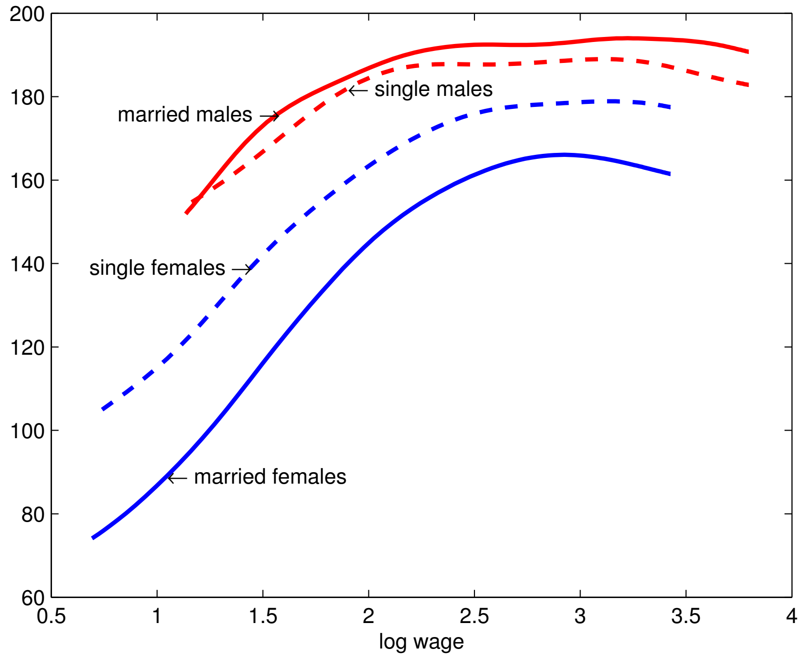


Figure 3: Mean hours supplied given own wage

single females working much more than married females. Apparently, coupling allows men to specialise in wage-work and women to specialise in household production. Figure 4 shows mean hours given own and spouse's wage for married males and females. There is little apparent effect of spouse's wage on labour supply.

4 Identification

Let the distribution of z have cdf $G(z) = \Phi(z/\sigma)$, where Φ is a distribution with mean 0 and variance 1. Let us also suppose that the support of Φ, $[v_{min}, v_{max}]$, is large enough, maybe equal to the whole real line, for the matching probability to be strictly between 0 and 1 for all (x, y):

$$0 < \alpha(x, y) = 1 - \Phi(-s(x, y)/\sigma) < 1.$$

This implies that $z_{min} = \sigma v_{min} < -s(x, y) < z_{max} = \sigma v_{max}$.

4.1 Known Φ

Then, given arbitrary choices of Φ and $\lambda/\delta > \frac{n(x,y)}{u_m(x)u_f(y)}$, the equilibrium variables can be recovered through the following steps.

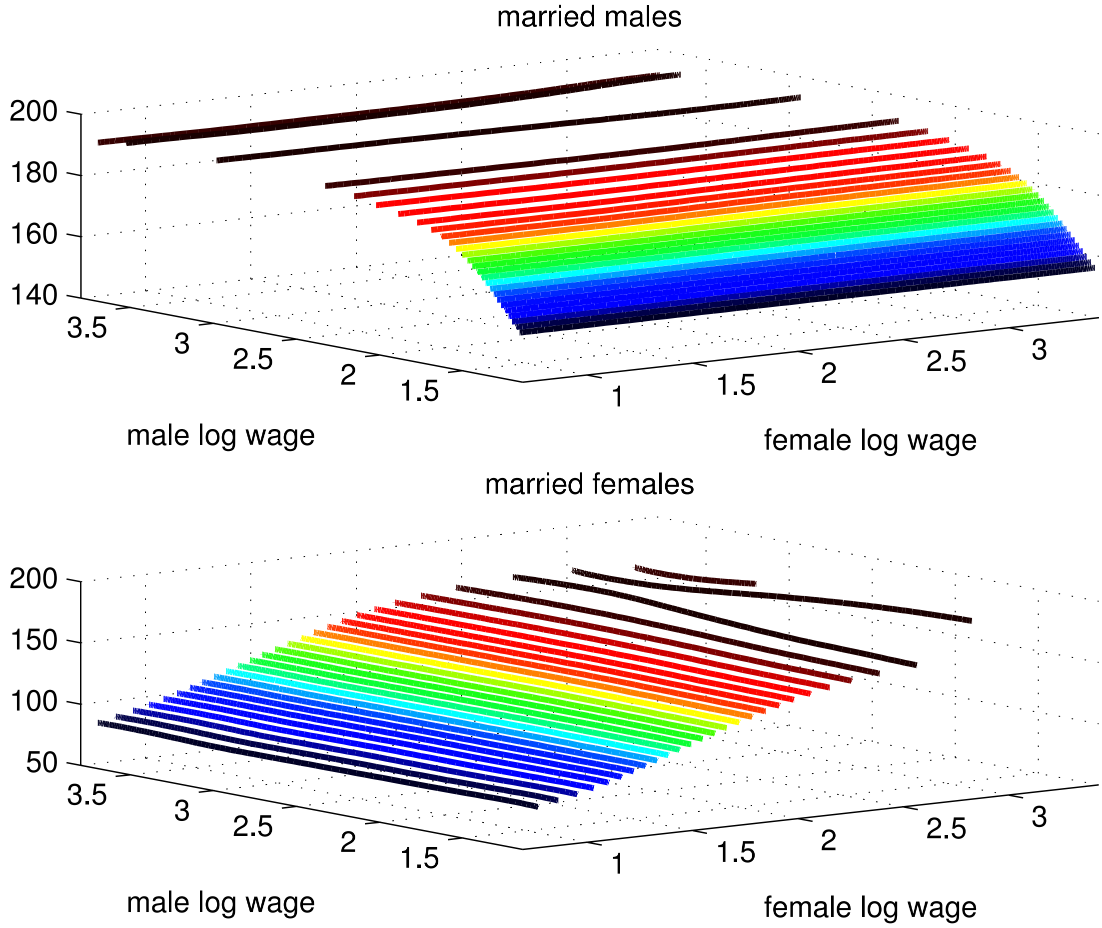


Figure 4: Hours supplied for married couples

1. Equation (2) identifies

$$\alpha(x,y) = \frac{\delta}{\lambda}\frac{n(x,y)}{u_m(x)u_f(y)}. \tag{20}$$

2. Then $s(x,y)$ can be identified as

$$s(x,y) = -\sigma\Phi^{-1}(1-\alpha(x,y)). \tag{21}$$

3. It follows from equations that(14) that

$$s_m(x) = \beta k\frac{\lambda}{\delta}\int \left(\int \max\{z+s(x,y'),0\}\,\mathrm{d}G(z)\right) u_f(y')\,\mathrm{d}y'\Big]$$

where

$$\begin{aligned}\int \max\{s(x,y)+z,0\}\,\mathrm{d}G(z) &= s(x,y)\alpha(x,y) + \int_{-s(x,y)}^{\sigma v_{max}} z\,\mathrm{d}G(z)\\ &= \sigma\left[-\alpha\Phi^{-1}(1-\alpha) + \int_{\Phi^{-1}(1-\alpha)}^{v_{max}} v\,\mathrm{d}\Phi(v)\right]\\ &= \sigma\mu_\Phi(\alpha(x,y)) \quad \text{(say).}\end{aligned}$$

Hence

$$s_m(x) = \beta\sigma k\frac{\lambda}{\delta}\int \mu_\Phi(\alpha(x,y'))u_f(y')\,\mathrm{d}y' \tag{22}$$

and by symmetry,

$$s_f(y) = (1-\beta)\sigma k\frac{\lambda}{\delta}\int \mu_\Phi(\alpha(x',y))u_m(x')\,\mathrm{d}x'. \tag{23}$$

4. Household production then follows as

$$\begin{aligned}C(x,y) &= s(x,y) + s_m(x) + s_f(y) \\ &= \sigma\Big[-\Phi^{-1}(1-\alpha(x,y)) + \beta k\frac{\lambda}{\delta}\int \mu_\Phi(\alpha(x,y'))u_f(y')\,\mathrm{d}y' \\ &\quad + (1-\beta)k\frac{\lambda}{\delta}\int \mu_\Phi(\alpha(x',y))u_m(x')\,\mathrm{d}x'\Big].\end{aligned} \tag{24}$$

5. Matching hours worked by married (fe)males with hours worked by single (fe)males on same wages, then, equations (7) and (8) imply that

$$\begin{aligned}h_m - \mathbb{E}(h_m^0|x) &= -b'_m(x)t_m(x,y,z) = -b'_m(x)[s_m(x) + \beta(s(x,y)+z)]\\ h_f - \mathbb{E}(h_f^0|y) &= -b'_f(y)t_f(x,y,z) = -b'_f(y)[s_f(y) + (1-\beta)(s(x,y)+z)]\end{aligned}$$

and, integrating over z,

$$\Delta_m(x,y) \equiv \mathbb{E}(h_m|x,y) - \mathbb{E}(h_m^0|x) = -b_m'(x)[s_m(x) + \beta\overline{s}(x,y)]$$
$$\Delta_f(x,y) \equiv \mathbb{E}(h_f|x,y) - \mathbb{E}(h_f^0|y) = -b_f'(y)[s_f(y) + (1-\beta)\overline{s}(x,y)]$$

where

$$\begin{aligned}\overline{s}(x,y) &= \mathbb{E}(s(x,y) + z|x,y,s(x,y)+z>0) \\ &= \sigma\frac{\mu_\Phi(\alpha(x,y))}{\alpha(x,y)}.\end{aligned} \tag{25}$$

(Note that hours can be measured with error if measurement errors have zero mean conditional on x,y,z.) Regressing $\Delta_m(x,y)$ on

$$\tau_m(x,y) \equiv \frac{\mu_\Phi(\alpha(x,y))}{\alpha(x,y)} + k\frac{\lambda}{\delta}\int \mu_\Phi(\alpha(x,y'))u_f(y')\,\mathrm{d}y'$$

and $\Delta_f(x,y)$ on

$$\tau_f(x,y) \equiv \frac{\mu_\Phi(\alpha(x,y))}{\alpha(x,y)} + k\frac{\lambda}{\delta}\int \mu_\Phi(\alpha(x',y))u_m(x')\,\mathrm{d}x'$$

yields consistent estimators for $-b_m'(x)\beta\sigma$ and $-b_f'(y)(1-\beta)\sigma$.

This procedure identifies $b_m'(x)\beta\sigma$ and $b_f'(y)(1-\beta)\sigma$ given Φ and λ/δ. In particular, neither β nor σ are separately identified unless we normalise the level of income effects $b_m'(x_0)$ and $b_f'(y_0)$ for two wage levels x_0 and y_0. Given such a normalisation of b_m' and b_f', or given an arbitrary choice of β and σ, $A_m(x)$ and $A_f(y)$ can be estimated by solving the linear ordinary differential equations

$$A_m'(x) + b_m'(x)A_m(x) = T - \mathbb{E}(h_m^0|x) - b_m'(x)xT \equiv k_m(x)$$
$$A_m'(y) + b_f'(y)A(y) = T - \mathbb{E}(h_f^0|y) - b_f'(y)yT \equiv k_f(y).$$

That is,

$$A_m(x) = B_m(x)\int_{x_{min}}^{x} \frac{k_m(x')}{B_m(x')}\,\mathrm{d}x'$$
$$A_f(y) = B_f(y)\int_{y_{min}}^{y} \frac{k_f(y')}{B_f(y')}\,\mathrm{d}y'$$

where

$$B_m(x) = \exp \int_{x_{min}}^{x} b'_m(x') \, \mathrm{d}x'$$
$$B_f(y) = \exp \int_{y_{min}}^{y} b'_f(y') \, \mathrm{d}y'.$$

4.2 Identification of Φ

Define hours' residuals:

$$\begin{aligned} u_m &\equiv h_m - \mathbb{E}(h_m^0|x) + b'_m(x)\Big[s_m(x) + \beta s(x,y)\Big] \\ &= h_m - \mathbb{E}(h_m^0|x) + b'_m(x)\beta\sigma\Big[-\Phi^{-1}(1-\alpha(x,y)) + k\frac{\lambda}{\delta}\int \mu_\Phi(\alpha(x,y'))u_f(y') \, \mathrm{d}y'\Big] \end{aligned}$$

and

$$u_f \equiv h_f - \mathbb{E}(h_f^0|y) + b'_f(y)(1-\beta)\sigma\Big[-\Phi^{-1}(1-\alpha(x,y)) + +k\frac{\lambda}{\delta}\int \mu_\Phi(\alpha(x',y))u_m(x') \, \mathrm{d}x'\Big]$$

By equations (14) and (15),

$$u_m = b'_m(x)\beta z$$
$$u_f = b'_f(y)(1-\beta)z.$$

Measurement errors can be added without changing the following discussion.

The second-order moments of these residuals are not informative as, for example,

$$\mathrm{Cov}\left(\frac{u_m}{b'_m(y)\beta\sigma}, \frac{u_f}{b'_f(y)(1-\beta)\sigma}\right) = \mathrm{Var}\left(\frac{z}{\sigma}|x,y,s(x,y)+z>0\right)$$

which does not depend on σ because z/σ has distribution Φ and $\frac{s(x,y)}{\sigma} = -\Phi^{-1}(1-\alpha(x,y))$, independently of σ. However, higher-order moments can also be computed, yielding more detailed information on the shape of Φ. Of course, using this information for estimation is complicated by the fact that an initial guess for Φ is needed to compute the residuals. We shall not try to follow this route as it is unlikely that making Φ a mixture of two normals, for example, is going to change significantly the results.

5 Estimation

We finally provide information on the shape of the preference parameters, the marriage externality and transfers.

5.1 Calibration of non-identified parameters

For this we need to set values to the non-identified parameters. The standard deviation of xT is 8110 and that of yT is 6165. We arbitrarily fix $\sigma = 1000$, the order of magnitude. Bargaining power is assumed to be evenly distributed between men and women, i.e. $\beta = 0.5$. Lastly, we set the probability of a (serious) dating for single males equal to one per year ($\lambda_m = 1/12$), which implies $\lambda = \lambda_m/U_f = 1.63 \times 10^{-5}$.

5.2 The matching probability

The matching probability $\alpha(x, y)$ is estimated as

$$\alpha(x, y) = \frac{\delta}{\lambda} \frac{n(x, y)}{u_m(x)u_f(y)}.$$

The densities $n(x, y)/N$, $u_m(x)/U_m$ and $u_f(y)/U_f$ are estimated by kernel density estimation with a normal kernel. We use twice the usual bandwidth to smooth the density functions in the tails. This is important as we divide n by $u_m u_f$ to get α. As in deconvolution, additional smoothing is thus required.

Figure 5 displays the shape of the matching probability function thus estimated. It is unambiguously increasing in both wages. Interestingly, isoclines seem less and less conical as wages increase (see panel (b)), which can be understood as an indication that positive assortative matching (PAM) is more pronounced for lower wages. Lastly, matching probabilities remain low and approximately constant for more than half of the distribution of wages. It is only for relatively high wages that the probability starts to raise.

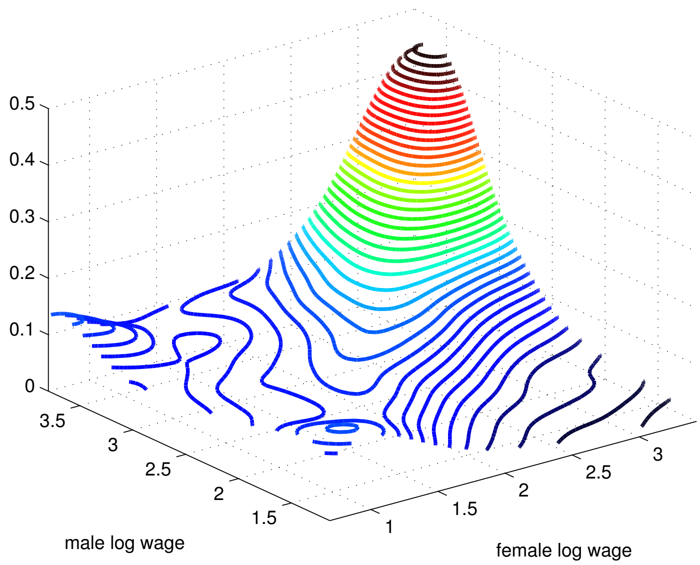
The average matching probability for a single man of type x of randomly meeting a single woman is equal to

$$\lambda_m \int \frac{u_f(y)}{U_f} \alpha(x, y)\, \mathrm{d}y$$

with $\lambda_m = \lambda U_f$ and with a similar formula for single women.[1] Figure 6 plots the implies average durations before (re)marriages. Low wage-individual have to wait for a very long time, and women more than men. The waiting time decreases with the wage. The average waiting time is 8.9 years for men and 12.1 years for women. The estimates are significantly greater than the median duration between first divorce and remarriage (for those who separate and remarry) that was 3.3 years for men and 3.5 years for women in the 2001 Census. However, these statistics do not account for censoring.

[1] Details on the numerical techniques that we use in estimation and simulation are provided in Section .

(a) 3-D plot



(b) Flat projection

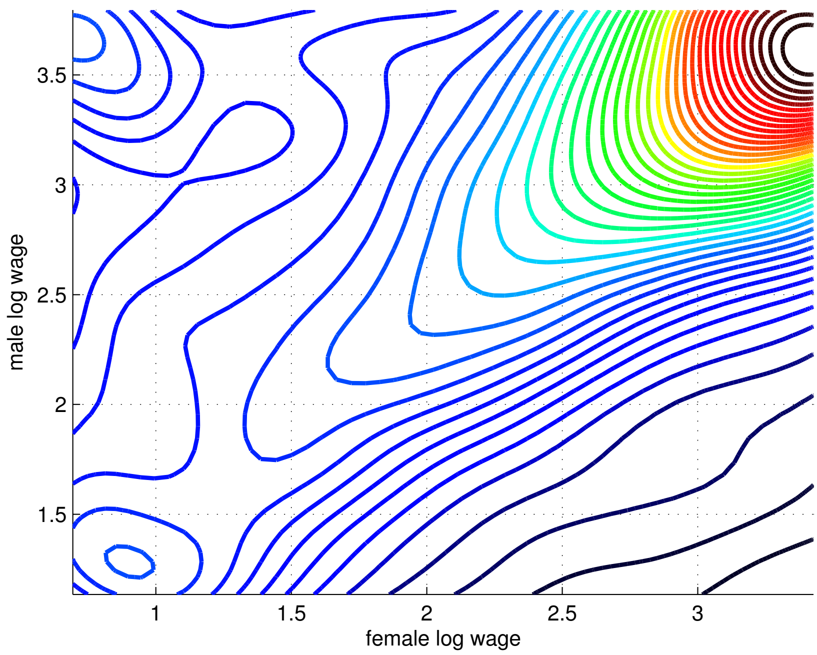


Figure 5: Non-normalised matching probabilities

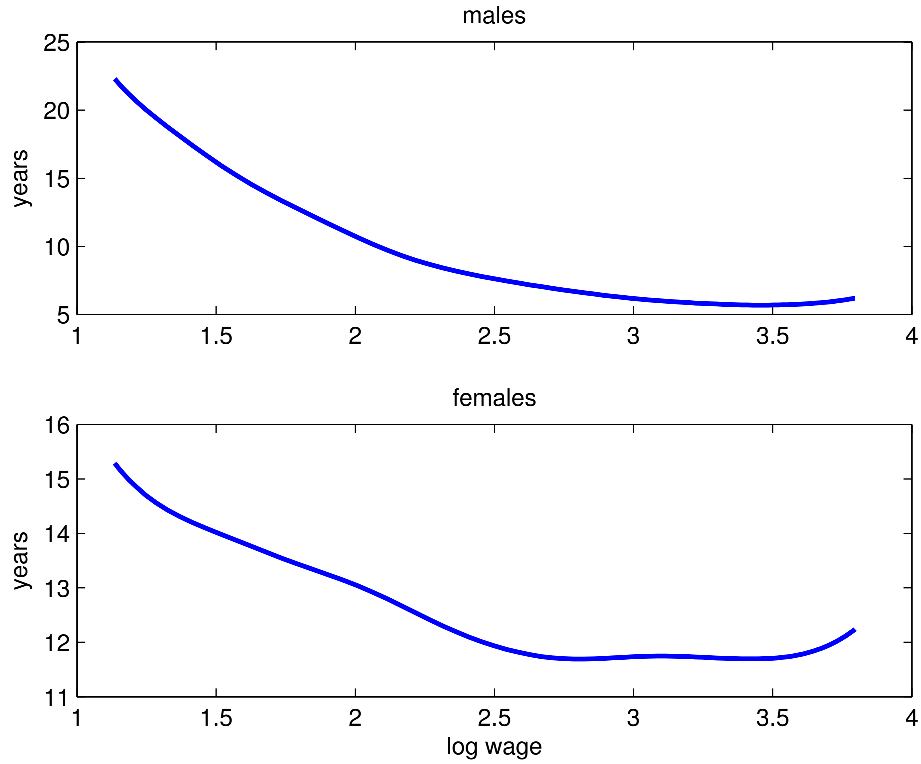


Figure 6: Average duration before marriage, by gender and type

5.3 Externality and transfers

We assume the distribution of z normal and estimate $s(x, y)$, $s_m(x)$, $s_f(y)$ and $C(x, y)$ using identifying equations (21), (22), (23) and (24) fron the identification discussion. Figure 7 shows the household production function that is estimated. Interestingly, increasing the husband's wage almost always yields a higher home production, while this is mostly the opposite for at least the lower half part of the distribution of female wages. The lowest level of the marriage externality is attained for the most improbable match with the poorest male and the richest female. At the other extreme, a poor woman complements a riche man enough to make males essentially indifferent to female wage. It is only when the woman's wage is greater than the median that the marriage externality, and the marriage probability, start to increase significantly.

Transfers can then be deduces as

$$t_m(x, y, z) = s_m(x) + \beta[s(x, y) + z]$$
$$t_f(x, y, z) = s_f(y) + (1 - \beta)[s(x, y) + z]$$

and averaging over z conditional on $s(x, y) + z > 0$ yields mean transfers

$$\bar{t}_m(x, y) = s_m(x) + \beta\bar{s}(x, y)$$
$$\bar{t}_f(x, y) = s_f(y) + (1 - \beta)\bar{s}(x, y)$$

where $\bar{s}(x, y)$ is given by equation (25).

Figure 8 displays the estimated mean transfers. Men get more (between 300 and 1000) than women (between 450 and 700). Also the male transfer is essentially independent of the female wage, and it is monotonically increasing in male wage. This is much less true for women, whose transfer does not increase with male wage when their wage is less the the median.

5.4 Preference parameters

Figure 9 shows the estimated values of $b'_m(x)B_m(x) = B'_m(x)$ and $b'_f(y)B_f(y) = B'_f(y)$, and $B_m(x)$ and $B_f(y)$. Aggregate price indices $B_m(x)$ and $B_f(y)$ are well approximated by affine functions (the variations of $b'_m(x)$ and $b'_f(y)$ is essentially noise). We estimate $\frac{\mathrm{d}B_m(x)}{\mathrm{d}x} \simeq -0.0065$ and $\frac{\mathrm{d}B_f(y)}{\mathrm{d}y} \simeq 0.024$. Figure 10 shows estimates of $A_m(x)$ and $A_f(y)$ (normalised by T) and their derivatives. A cubic approximation is shown for comparison.

For males, leisure (household production) is an inferior good whereas for women, it is a normal good. This explains why married men work more than singles and married women work less.

6 Fit

Now, we take the estimated externality function $C(x, y)$ and kernel density estimates of the unconditional wage distribution $\ell_m(x)$ and $\ell_f(y)$ and we simulate the equilibrium. Because λ should be a equilibrium parameter, we postulate a Cobb-Douglas meeting function $M(U_m, U_f) = M_0 U_m^{1/2} U_f^{1/2}$ and we estimate M_0 as $M_0 = \lambda U_m^{1/2} U_f^{1/2}$ for $\lambda = 1.63 \times 10^{-5}$, the value that we used in estimation.

The fixed point equations (16)-(19) are solved by composing the fixed point operator with itself until numerical convergence. This algorithm works well if one reduces the stepsize as in:

$$x_n = x_{n-1} + \theta[Tx_{n-1} - x_{n-1}]$$

with $\theta < 1$ (Landweber regularisation).

Using the nonparametric estimates of the joint density of wages of married couples, $n(x, y)$, and those of the densities of wages for male and female singles, $u_m(x)$ and $u_f(y)$, equation (2) implies that

$$\frac{\lambda}{\delta}\alpha(x, y) = \frac{n(x, y)}{u_m(x)u_f(y)}. \tag{26}$$

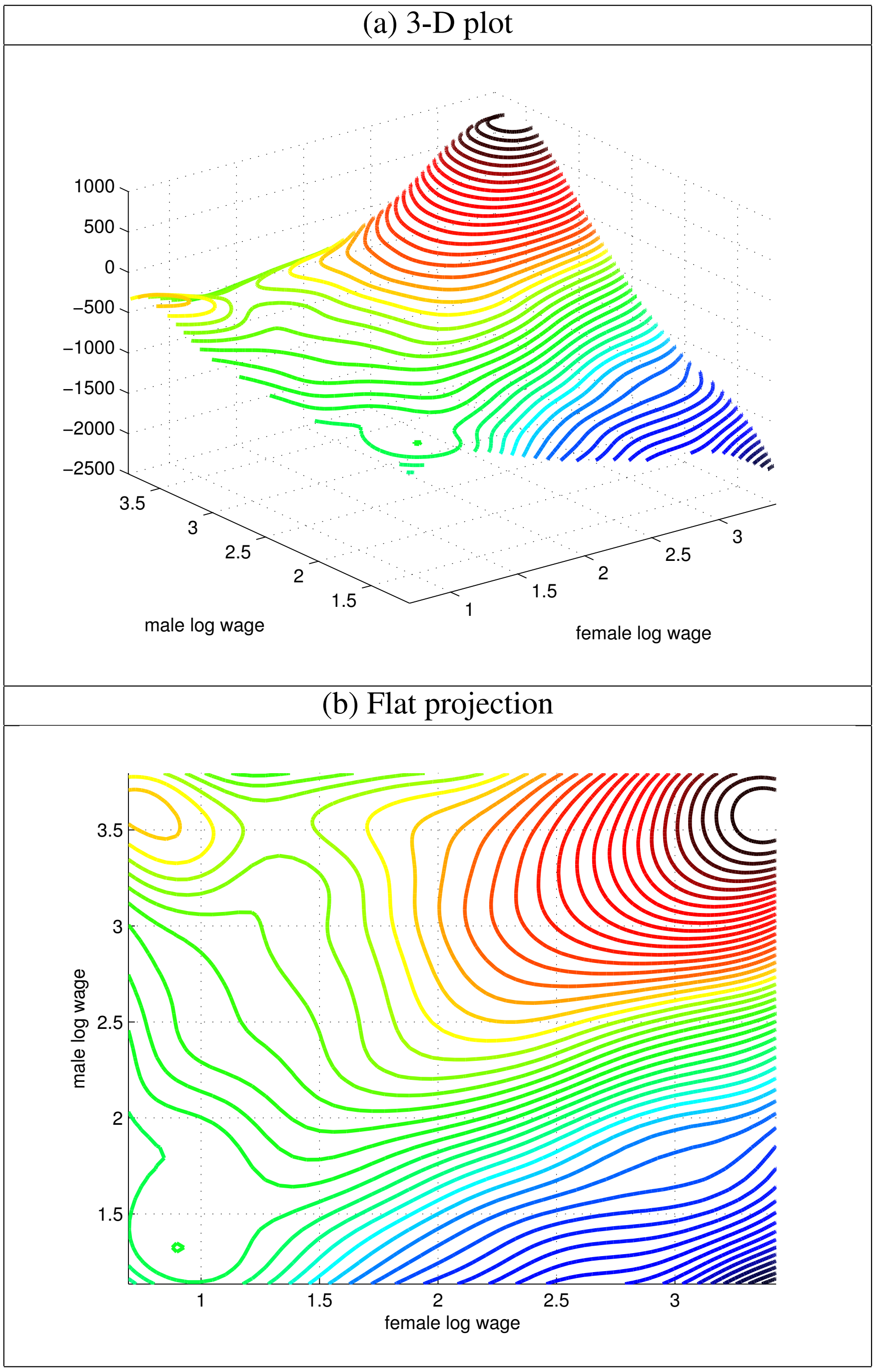


Figure 7: Externalities - Function $C(x, y)$

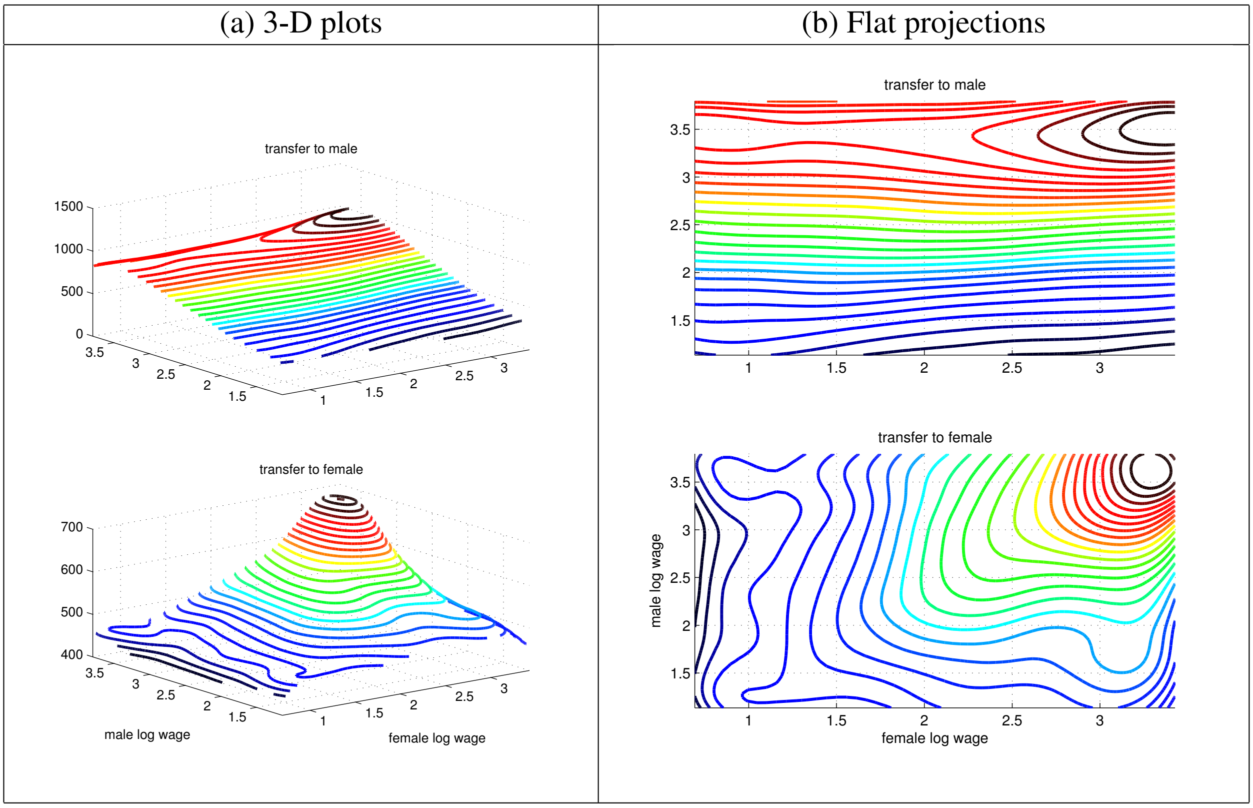


Figure 8: Mean transfers

The model then predicts that u_m and u_f are fixed points of the system

$$u_m(x) = \frac{\ell_m(x)}{1 + \frac{\lambda}{\delta}\int u_f(y)\alpha(x,y)\,\mathrm{d}y}$$
$$u_f(y) = \frac{\ell_f(y)}{1 + \frac{\lambda}{\delta}\int u_m(x)\alpha(x,y)\,\mathrm{d}x}$$

We estimate $U_m = 3340$ (instead of $U_m = 3366$) and $U_f = 5000$ (instead of $U_f = 5104$). The estimation error is esentially due to the fact that kernel density estimates on a discrete grid do not exactly sum to one (see Section (7). Figure 11 shows that steady-state restrictions yield a good fit of the wage distributions. Finally, Figure 12 shows the fit of conditional mean hours given own wage. The fit is also very good, thanks to the nonparametric estimation of b'_m and b'_f. Figures 13 and 14 show the corresponding fit for a cubic approximation of b'_m, b'_f, A_m and A_f. The fit is still satisfactory for densities (as for the matching function and the externality function) and slightly worse for mean hours.

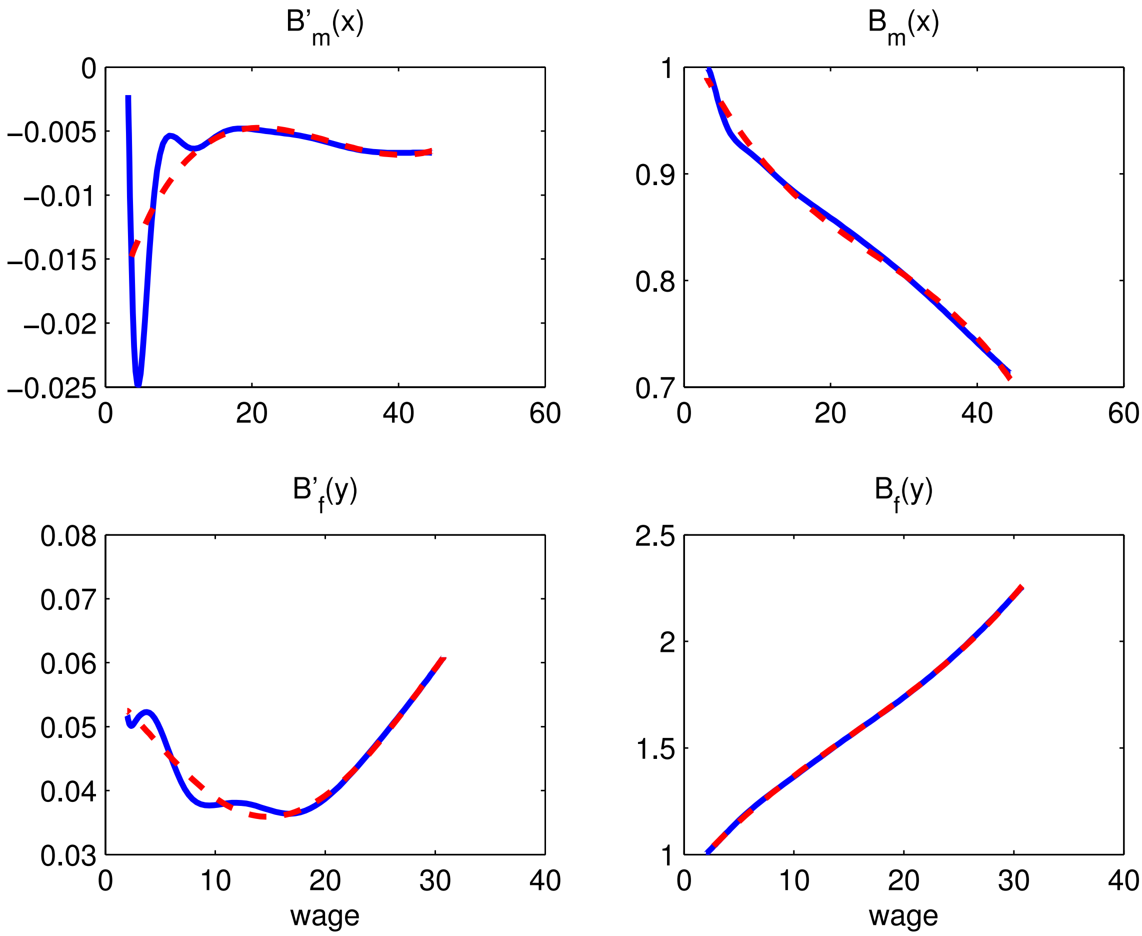


Figure 9: Preference parameters - Income effect

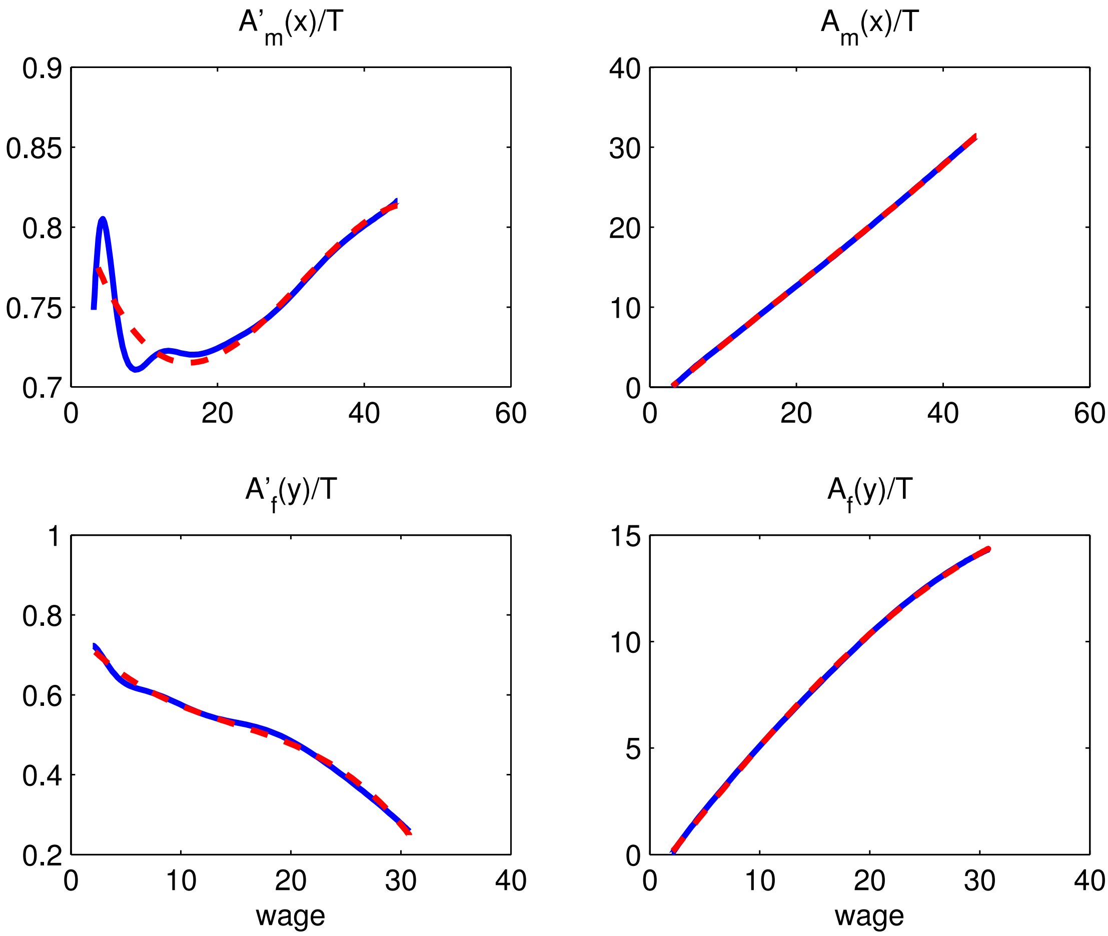


Figure 10: Preference parameters - Minimum expenditure

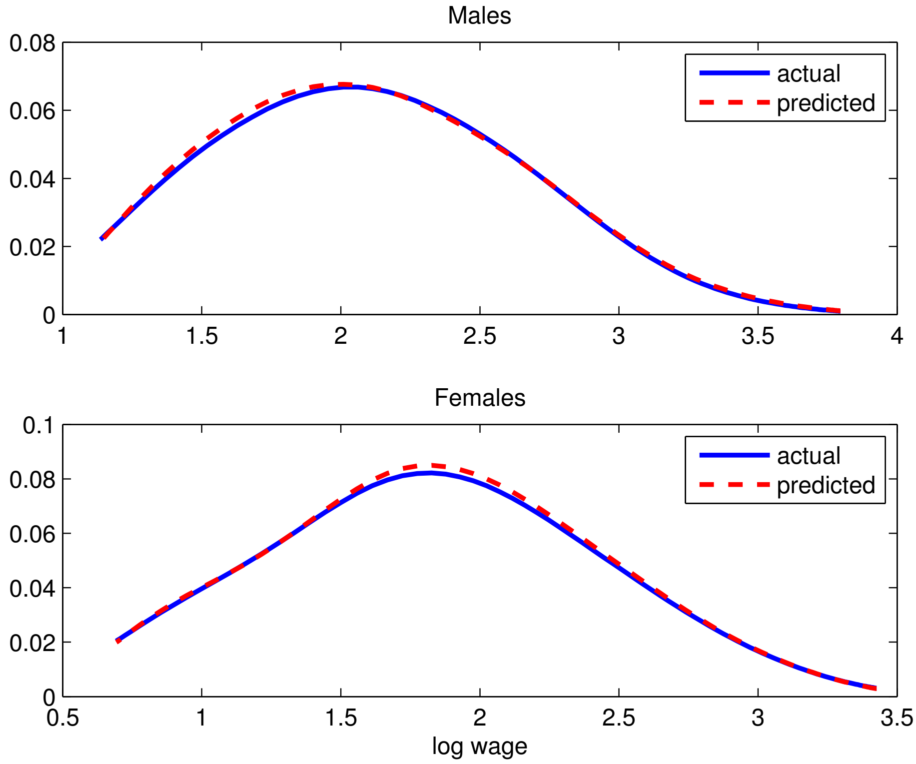


Figure 11: Fit of wage densities for singles - Nonparametric estimation

7 Numerical details

We discretise the set of wages using Chebyshev points defined as

$$x_j = \frac{x_{min} + x_{max}}{2} + \frac{x_{max} - x_{min}}{2} nodes_j$$

where, for a grid of $n+1$ (non equally distant) nodes,

$$nodes_j = \cos \frac{j\pi}{n}, \quad 0 \leq j \leq n.$$

Choosing Chebyshev polynomials to approximate smooth functions on a compact is convenient, as we can then use the Clenshaw-Curtis quadrature to approximate the integrals in the fixed-point formulas. For example,

$$\int f(x)\,\mathrm{d}x \simeq \frac{x_{max} - x_{min}}{2} \sum_{j=0}^{n} w_j f(x_j)$$

where weights w_j are easily and efficiently calculated using fast Fourier transform. The following MATLAB code can be used (Waldvogel, 2006):

```
function [nodes,wcc] = cc(n)
```

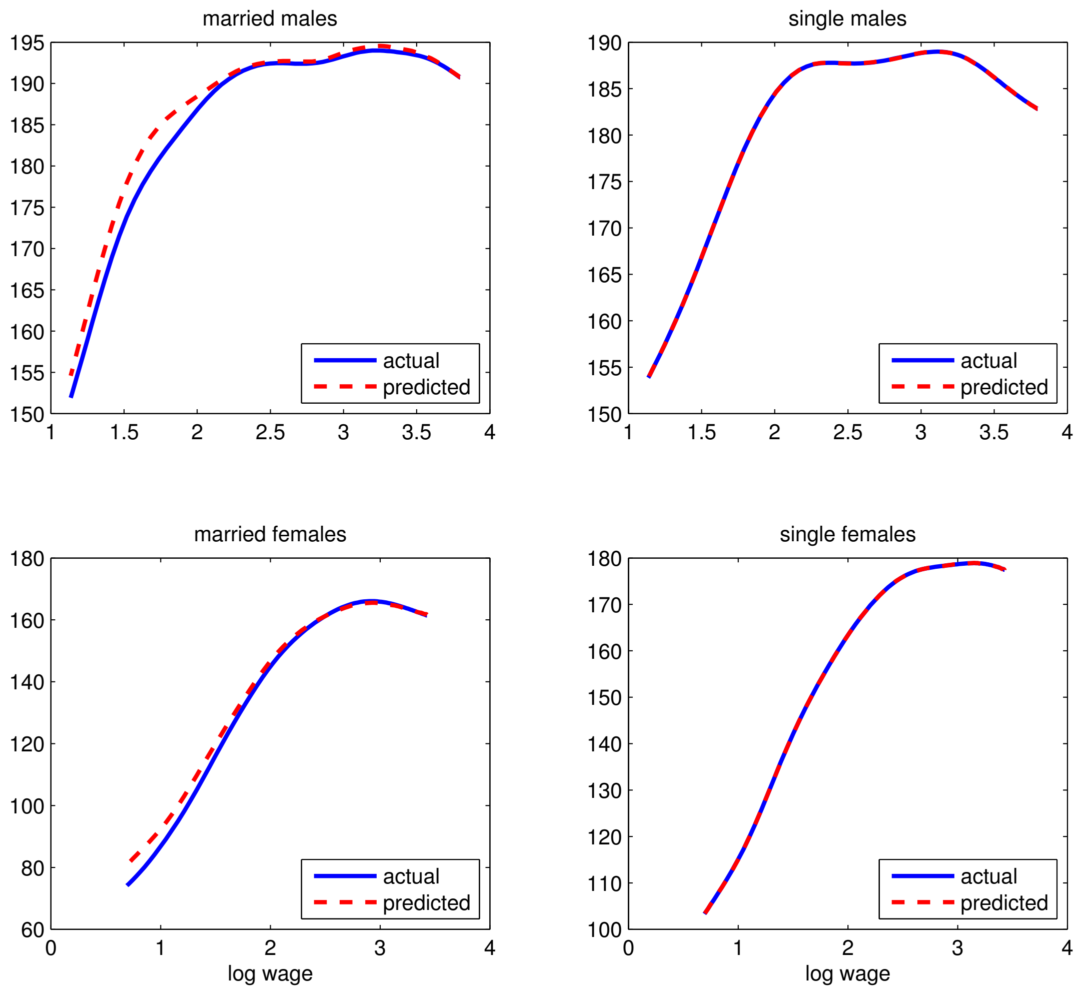


Figure 12: Fit of conditional mean hours - Nonparametric estimation

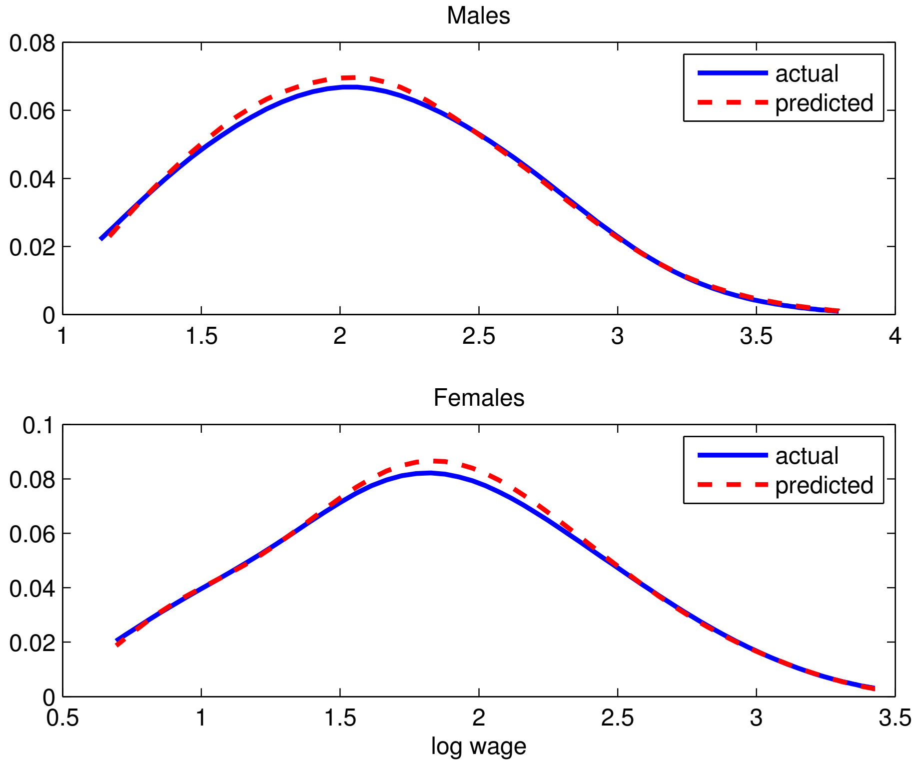


Figure 13: Fit of wage densities for singles - Cubic approximation of A, B, C

```
nodes = cos(pi*(0:n)/n);
N=[1:2:n-1]'; l=length(N); m=n-l;
v0=[2./N./(N-2); 1/N(end); zeros(m,1)];
v2=-v0(1:end-1)-v0(end:-1:2);
g0=-ones(n,1); g0(1+l)=g0(1+l)+n; g0(1+m)=g0(1+m)+n;
g=g0/(n^2-1+mod(n,2)); wcc=real(ifft(v2+g));
wcc=[wcc;wcc(1)];
```

In practice we use $n = 100$.

References

BLUNDELL, R., P.-A. CHIAPPORI, T. MAGNAC, AND C. MEGHIR (2007): "Collective Labour Supply: Heterogeneity and Non-Participation," *Review of Economic Studies*, 74(2), 417–445.

CHIAPPORI, P.-A. (1988): "Rational Household Labor Supply," *Econometrica*, 56(1), 63–90.

——— (1992): "Collective Labor Supply and Welfare," *Journal of Political Economy*, 100(3), 437–467.

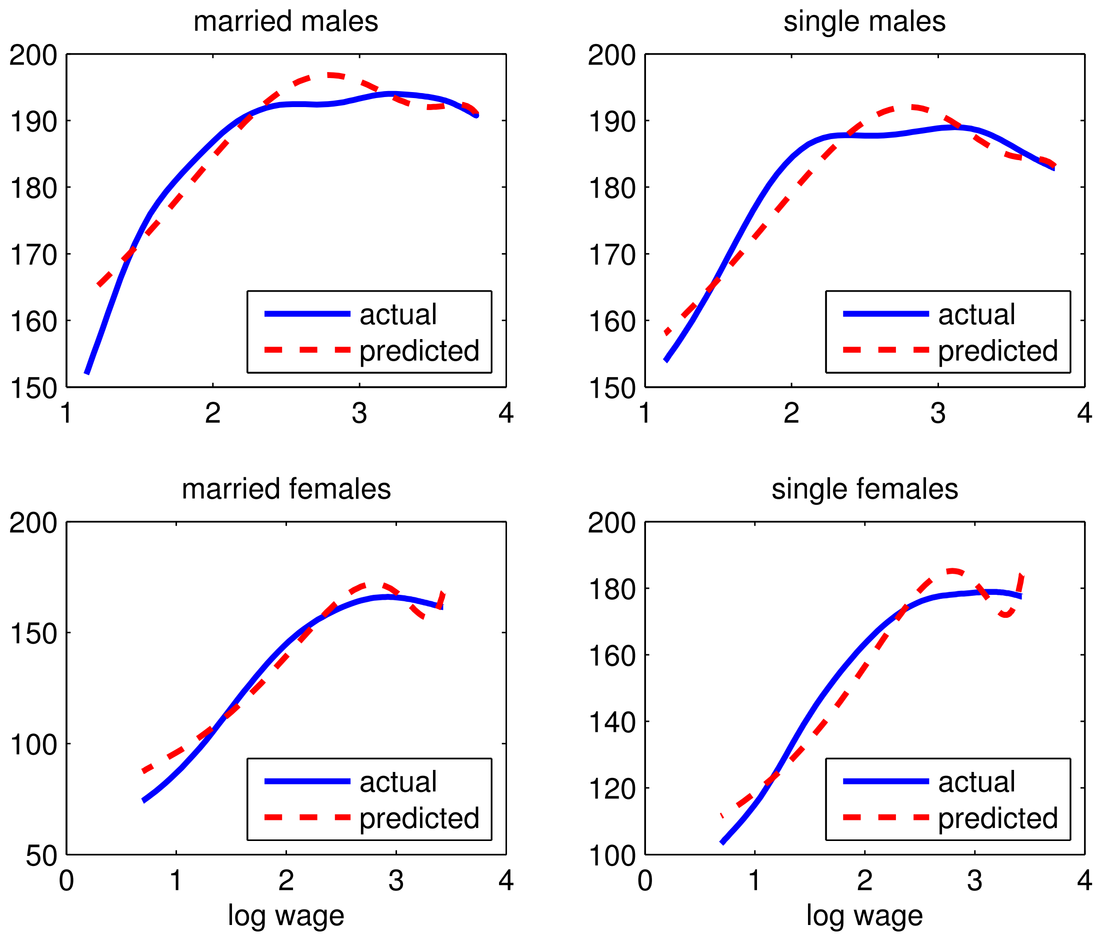


Figure 14: Fit of conditional mean hours - Cubic approximation of A, B, C

LU, M., AND R. P. MCAFEE (1996): "Matching and Expectations in a Market with Heterogeneous Agents," in *Advances in Applied Microeconomics*, ed. by M. Baye, vol. 6, pp. 121–156. JAI Press, Greenwich (CT).

SATTINGER, M. (1995): "Search and the Efficient Assignment of Workers to Jobs," *International Economic Review*, 36(2), 283–302.

SHIMER, R., AND L. SMITH (2000): "Assortative Matching and Search," *Econometrica*, 68(2), 343–369.

WALDVOGEL, J. (2006): "Fast Construction of the Fejér and Clenshaw-Curtis Quadrature Rules," *BIT Numerical Mathematic*, 46, 195–202.