ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS

WORKING PAPER SERIES

# Optimal Taxation, Marriage, Home Production, and Family Labor Supply

| | |
|---|---|
| **Authors** | George-Levi Gayle, and Andrew Shepherd |
| **Working Paper Number** | 2016-010B |
| **Revision Date** | July 2016 |
| **Citable Link** | https://doi.org/10.20955/wp.2016.010 |
| **Suggested Citation** | Gayle, G.-L., Shepherd, A., 2016; Optimal Taxation, Marriage, Home Production, and Family Labor Supply, Federal Reserve Bank of St. Louis Working Paper 2016-010. URL https://doi.org/10.20955/wp.2016.010 |

Federal Reserve Bank of St. Louis, Research Division, P.O. Box 442, St. Louis, MO 63166

# Optimal Taxation, Marriage, Home Production, and Family Labour Supply

George-Levi Gayle* Andrew Shephard†

July 6, 2016

**Abstract**

An empirical approach to optimal income taxation design is developed within an equilibrium collective marriage market model with imperfectly transferable utility. Taxes distort labour supply and time allocation decisions, as well as marriage market outcomes, and the within household decision process. Using data from the American Community Survey and American Time Use Survey, we structurally estimate our model and explore empirical design problems. We consider the optimal design problem when the planner is able to condition taxes on marital status, as in the U.S. tax code, but we allow the schedule for married couples to have an arbitrary form of tax jointness. Our results suggest that the optimal tax system for married couples is characterized by negative jointness, although the welfare gains from this jointness are shown to be quite modest.

# 1 Introduction

Tax and transfer policies often depend on family structure, with the tax treatment of married and single individuals varying significantly both across countries and over time. In the United States there is a system of joint taxation where the household is taxed based on total family income. Given the progressivity of the tax system, it is not neutral with respect to marriage and both large marriage penalties and marriage bonuses coexist.[1]

*Federal Reserve Bank of St. Louis and Department of Economics, Washington University in St. Louis, Campus Box 1208, One Brookings Drive, St. Louis, MO 63130. Email: ggayle@wustl.edu.

†Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104. E-mail: asheph@econ.upenn.edu.

[1] A marriage penalty is said to exist when the tax liability for a married couple exceeds the total tax liability of unmarried individuals with the same total income. The reverse is true for a marriage bonus. While married couples in the United States have the option of "Married Filing Jointly" or "Married Filing Separately", the latter is very different from the tax schedule that unmarried individuals face.

In contrast, the majority of OECD countries tax individuals separately based on each individual's income. In such a system, married couples are treated as two separate individuals, and hence there is no subsidy or tax on marriage.[2] But what is the appropriate choice of tax unit and how *should* individuals and couples be taxed? A large and active literature concerns the optimal design of tax and transfer policies. In an environment where taxes affect the economic benefits from marriage, such a design problem has to balance redistributive objectives with efficiency considerations while recognizing that the structure of taxes may affect who gets married, and to whom they get married, as well as the intra-household allocation of resources.

Following the seminal contribution of Mirrlees (1971), a large theoretical literature has emerged that studies the optimal design of tax schedules for single individuals. This literature casts the problem as a one-dimensional screening problem, recognizing the asymmetry of information that exists between agents and the tax authorities.[3] The analysis of the optimal taxation of couples has largely been conducted in environments where the form of the tax schedule is restricted to be linearly separable, but with potentially distinct tax rates on spouses (see Boskin and Sheshinski, 1983, Apps and Rees, 1988, 1999, 2007, and Alesina, Ichino and Karabarbounis, 2011, for papers in this tradition).[4] A much smaller literature has extended the Mirrleesian approach to study the optimal taxation of couples as a two-dimensional screening problem. Most prominently, Kleven, Kreiner and Saez (2009) consider a unitary model of the household, in which the primary earner makes a continuous labour supply decision (intensive only margin) while the secondary worker makes a participation decision (extensive only margin), and characterize the optimal form of tax jointness.[5] When the participation of the secondary earner provides a signal of the couple being better off, the tax rate on secondary earnings is shown to be *decreasing* with primary earnings.

[2] This is a oversimplification of actual tax systems. Even though many countries have individual income tax filing, there are often other ways in which tax jointness may emerge. For example, transfer systems often depend on family income and certain allowances may be transferable across spouses. See Immervoll et al. (2009) for an evaluation of the tax-transfer treatment of married couples in Europe. Our estimation incorporates the combined influence of taxes and transfers on marriage and time allocation outcomes.

[3] See Brewer, Saez and Shephard (2010) and Piketty and Saez (2013) for recent surveys.

[4] A quantitative macroeconomic literature compares joint and independent taxation in a non-optimal taxation setting. See, e.g., Chade and Ventura (2002, 2005) and Guner, Kaygusuz and Ventura (2012).

[5] Kleven, Kreiner and Saez (2007) also present a doubly intensive model, where the both the primary and secondary earner make continuous (intensive only) labour supply choices. Immervoll et al. (2011) present a double-extensive model of labour supply, and show how tax rates vary under unitary and collective models with fixed decision weights. See also Brett (2007), Cremer, Lozachmeur and Pestieau (2012), and Frankel (2014). Note that all of these studies take the married unit as given and ignore the distortionary effect of taxation on who gets married and to whom they get married.

The theoretical optimal income taxation literature provides many important insights that are relevant when considering the design of a tax system. However, the *quantitative* empirical applicability of optimal tax theory is dependent upon a precise measurement of the key behavioural margins: How do taxes affect market work, the amount of time devoted to home production, and the patterns of specialization within the household? How do taxes influence the allocation of resources within the household? What is the effect of taxes on the decision to marry and to whom? In order to examine both the optimal degree of progressivity and jointness of the tax schedule, and to empirically quantify the importance of the marriage market in shaping these, we follow Blundell and Shephard (2012) by developing an empirical structural approach to non-linear income taxation design that centres the entire analysis around a rich micro-econometric model.

Our model integrates the collective model of Chiappori (1988, 1992) with the empirical marriage-matching model developed in Choo and Siow (2006).[6] Individuals make marital decisions that comprise extensive (to marry or not) and intensive (i.e., marital sorting) margins based on utilities that comprise both an economic benefit and an idiosyncratic non-economic benefit. The economic utilities are micro-founded and are derived from the household decision problem. We consider an environment that allows for very general non-linear income taxes, incorporates home production time, includes both public and private good consumption, and distinguishes between the intensive and extensive labour supply margins. We do not introduce an exogenous primary/secondary earner distinction.[7]

Within the household, both explicit and implicit transfers are important. The leading paradigm for modelling matching in a marriage market involves transferable utility. The assumption of transferable utility implies that all transfers within the household take place at a constant rate of exchange and hence the utility possibility frontier is linear. In this world, time allocation decisions would not depend upon the conditions of the marriage market, and taxation would not affect the relative decision weight of household

[6]Other papers that integrate a collective model within an empirical marriage-matching model include Chiappori, Costa Dias and Meghir (2015), Choo and Seitz (2013), and Galichon, Kominers and Weber (2016). Chiappori, Costa Dias and Meghir (2015) consider an equilibrium model of education and marriage with life-cycle labour supply and consumption in a transferable utility setting; Choo and Seitz (2013) consider a static model with family labour supply and estimate a semi-parametric version of their model; Galichon, Kominers and Weber (2016) provide an empirical application where they estimate a matching model with consumption, allowing for imperfectly transferable utility (as we consider here).

[7]The large growth in female labour force participation has made the traditional distinction between primary and secondary earners much less clear. Women now make up around half of the U.S. workforce, with an increasing fraction of households in which the female is the primary earner. See, e.g., Blau and Kahn (2007) and Gayle and Golan (2012).

members.[8] As in the general framework presented in Galichon, Kominers and Weber (2014, 2016), we therefore allow for utilities to be imperfectly transferable across spouses, thus generating a non-linear utility possibility frontier. In this environment we provide sufficient conditions for the existence and uniqueness of equilibrium in terms of the model primitives, demonstrate semi-parametric identification, and describe a computationally efficient way to estimate the model using an equilibrium constraints approach.

Using data from the American Community Survey (ACS) and the American Time Use Survey (ATUS) we structurally estimate our equilibrium model, exploiting variation across markets in terms of both tax and transfer policies and population vectors. We show that the model is able to jointly explain labour supply, home time, and marriage market patterns. Moreover, it is able to successfully explain the variation in these outcomes across markets, with the behavioural implications of the model shown to be consistent with the existing empirical evidence.

We use our estimated model directly to examine problems related to the optimal design of the tax system by developing an extended Mirrlees framework. Our taxation design problem is based on an *individualistic* social-welfare function, with inequality both within and across households adversely affecting social-welfare. Here, taxes distort labour supply and time allocation decisions, as well as marriage market outcomes, and the within-household decision process. We allow for a very general specification of the tax schedule for both singles and married couples, that nests both individual and fully joint taxation, but also allows for very general forms of tax jointness. To preview our main findings, we find empirical support for negative jointness in the tax schedule for couples, but find that the welfare gains that this offers relative to a system of individual income taxation are relatively modest.

The remainder of the paper proceeds as follows. In Section 2 we present our equilibrium model of marriage, consumption, and time allocation, while in Section 3 we introduce the analytical framework that we use to study taxation design. In Section 4 we describes our data and empirical specification, discuss the semi-parametric identification of our model, and present our estimation procedure and results. In Section 5 we then consider the normative implications of our estimated model, both when allowing for a very general form of jointness in the tax schedule and when it is restricted. Finally, Section 6 concludes.

[8] The empirical literature has shown that policies targeting different spouses within the household may have differential effects on household demand. See, e.g., Thomas (1990), Browning et al. (1994) and Lundberg, Pollak and Wales (1997).

# 2 A model of marriage and time allocation

We present an empirical model of marriage-matching and intra-household allocations by considering a static equilibrium model of marriage with imperfectly transferable utility, labour supply, home production, and potentially joint and non-linear taxation. The economy comprises $K$ separate markets. Given that there are no interactions across markets, we suppress explicit conditioning on a market unless such a distinction is important and proceed to describe the problem for that market. In such a market there are $I$ types of men and $J$ types of women. The population vector of men is given by $\boldsymbol{\mathcal{M}}$, whose element $m_i$ denotes the measure of type-$i$ males. Similarly, the population vector of women is given by $\boldsymbol{\mathcal{F}}$, whose element $f_j$ denotes the measure of type-$j$ females. Associated with each male and female type is a utility function, a distribution of wage offers, a productivity of home time, a distribution of preference shocks, a value of non-labour income, and a demographic transition function (which is defined for all possible spousal types). While we are more restrictive in our empirical application, in principle all these objects may vary across markets. Moreover, these markets may differ in their tax system $T$ and the economic/policy environment more generally.

We make the timing assumption that the *realizations* of wage offers, preference shocks, and demographic transitions only occur following the clearing of the marriage market. There are therefore two (interconnected) stages to our analysis. First, there is the characterization of a *marriage matching function*, which is an $I \times J$ matrix $\boldsymbol{\mu}(T)$ whose $\langle i, j \rangle$ element $\mu_{ij}(T)$ describes the measure of type-$i$ males married to type-$j$ females, and which we write as a function of the tax system $T$.[9] Note that we do not allow for a cohabitation state.[10] The second stage of our analysis, which follows marriage market decisions, is then concerned with the joint time allocation and resource sharing problem for households. These two stages are linked through the decision weight in the household problem: these affect the second stage problem and so the expected value of an individual from any given marriage market pairing. These household decision (or Pareto) weights will adjust to clear the marriage market such that there is neither excess demand nor supply of any given type.

[9] Individuals may also choose to remain unmarried, and we use $\mu_{i0}(T)$ and $\mu_{0j}(T)$ to denote the respective measures of single males and females. The marriage matching function must satisfy the usual feasibility constraints. Suppressing the dependence on $T$, we require that $\mu_{i0} + \sum_j \mu_{ij} = m_i$ for all $i$, $\mu_{0j} + \sum_i \mu_{ij} = f_j$ for all $j$, and $\mu_{i0}, \mu_{0j}, \mu_{ij} \geq 0$ for all $i$ and $j$.

[10] Not allowing for cohabitation is a common assumption in the empirical marriage matching literature. For an exception see Mourifié and Siow (2014). We discuss this issue in Section 6.

## 2.1 Time allocation problem

We now describe the problem of single individuals and married couples once the marriage market has cleared. At this stage, all uncertainty (wage offers, preference shocks, and demographic transitions) has been resolved and time allocation decisions are made. Individuals have preferences defined over leisure, consumption of a market private good (whose price we normalize to 1), and a non-marketable public good produced with home time.

### 2.1.1 Time allocation problem: single individuals

Consider a single type-$i$ male. His total time endowment is $L_0$, and he chooses the time allocation vector $\mathbf{a}^i = (\ell^i, h_w^i, h_Q^i)$ comprising hours of leisure $\ell^i$, market work time $h_w^i$, and home production time $h_Q^i$, to maximize his utility. Time allocation decisions are discrete, with all feasible time allocation vectors described by the set $\mathbf{A}^i$. All allocations that belong to this set necessarily satisfy the time constraint $L_0 = \ell^i + h_w^i + h_Q^i$.[11] Associated with each possible discrete allocation is the additive state specific error $\epsilon_{\mathbf{a}^i}$. Excluding any additive idiosyncratic payoff from remaining single, the individual decision problem may formally be described by the following utility maximization problem:

$$\max_{\mathbf{a}^i \in \mathbf{A}^i} u^i(\ell^i, q^i, Q^i; \mathbf{X}^i) + \epsilon_{\mathbf{a}^i} \tag{1}$$

subject to

$$q^i = y^i + w^i h_w^i - T(w^i h_w^i, y^i; \mathbf{X}^i) - FC(h_w^i; \mathbf{X}^i), \tag{2a}$$

$$Q^i = \zeta_{i0}(\mathbf{X}^i) \cdot h_Q^i. \tag{2b}$$

Equation (2a) states that consumption of the private good is simply equal to net family income (the sum of earnings and non-labour income, minus net taxes) and less any possible fixed work of market work, $FC(h_w^i; \mathbf{X}^i) \geq 0$. These fixed costs (as in Cogan, 1981) are non-negative for positive values of working time, and zero otherwise. Equation (2b) says that total production/consumption of the home good is equal to the efficiency units of home time, where the efficiency scale $\zeta_{i0}(\mathbf{X}^i)$ may depend upon both own type (type-$i$) and demographic characteristics $\mathbf{X}^i$.

[11] The choice set used in our empirical implementation and both the parameterization of the utility function and the stochastic structure are described in Section 4.2.

The solution to this constrained utility maximization problem is described by the incentive compatible time allocation vector $\mathbf{a}^{i\star}_{i0}(w^i, y^i, \mathbf{X}^i, \epsilon^i; T)$, which upon substitution into equation (1) (and including the state-specific preference term associated with this allocation) yields the indirect utility function for type-$i$ males that we denote as $v^i_{i0}(w^i, y^i, \mathbf{X}^i, \epsilon^i; T)$. The decision problem for type-$j$ single women is described similarly and yields the indirect utility function $v^j_{0j}(w^j, y^j, \mathbf{X}^j, \epsilon^j; T)$.

### 2.1.2 Time allocation problem: married individuals

Married individuals are egoistic, and we consider a collective model that assumes an efficient allocation of intra-household resources (Chiappori, 1988, 1992).[12] An important economic benefit of marriage is given by the publicness of some consumption. The home-produced good (that is produced by combining male and female home time) is public within the household, which both members consume equally. Consider an $\langle i, j\rangle$ couple and let $\lambda_{ij}$ denote the Pareto weight on female utility in such a union.[13] The household chooses a time allocation vector for each adult and determines how total private consumption is divided between the spouses. Note that the state-specific errors $\epsilon_{\mathbf{a}^i}$ and $\epsilon_{\mathbf{a}^j}$ for any individual depend only on his/her own time allocation and not on the time allocation of his/her spouse. Moreover, the distributions of these preference terms, as well as the form of the utility function, do not change with marriage. We formally describe the household problem as

$$\max_{\mathbf{a}^i\in\mathbf{A}^i,\mathbf{a}^j\in\mathbf{A}^j,s_{ij}\in[0,1]} (1-\lambda_{ij}) \times \left[u^i(\ell^i, q^i, Q; \mathbf{X}^i) + \epsilon_{\mathbf{a}^i}\right] + \lambda_{ij} \times \left[u^j(\ell^j, q^j, Q; \mathbf{X}^j) + \epsilon_{\mathbf{a}^j}\right] \tag{3}$$

subject to

$$q = q^i + q^j = y^i + y^j + w^i h^i_w + w^j h^j_w - T(w^i h^i_w, w^j h^j_w, y^i, y^j; \mathbf{X}) - FC(h^i_w, h^j_w; \mathbf{X}), \tag{4a}$$

$$q^j = s_{ij} \cdot q, \tag{4b}$$

$$Q = \widetilde{Q}_{ij}(h^i_Q, h^j_Q; \mathbf{X}). \tag{4c}$$

In turn, this set of equality constraints describe (i) that *total* family consumption of the

[12] There are two principal ways of modelling the household in a non-unitary setting. First, there are collective (cooperative) models as we consider here, where allocations are assumed to be Pareto efficient. Second, there are strategic (non-cooperative) models based on Cournot-Nash equilibrium (e.g. Del Boca and Flinn, 2012, 1995). Donni and Chiappori (2011) provide a recent survey of non-unitary models.

[13] That the Pareto weights only depend on the types $\langle i, j\rangle$ is a consequence of our timing assumptions and efficient risk sharing within the household. See Section 2.2 for a discussion.

private good equals family net income, with the tax schedule here allowed to depend very generally on the labour market earnings of both spouses[14] less any fixed work-related costs; (ii) the wife receives the endogenous consumption share $0 \leq s_{ij} \leq 1$ of the private good; and (iii) the public good is produced using home time with the production function $\widetilde{Q}_{ij}(h^i_Q, h^j_Q; \mathbf{X})$, which may also depend upon demographic characteristics.

Letting $\mathbf{w} = [w^i, w^j]$, $\mathbf{y} = [y^i, y^j]$, $\mathbf{X} = [\mathbf{X}^i, \mathbf{X}^j]$, and $\boldsymbol{\epsilon} = [\epsilon^i, \epsilon^j]$, the solution to the household problem is the incentive compatible time allocation vectors $\mathbf{a}^{i\star}_{ij}(\mathbf{w}, \mathbf{y}, \mathbf{X}, \boldsymbol{\epsilon}; T, \lambda_{ij})$ and $\mathbf{a}^{j\star}_{ij}(\mathbf{w}, \mathbf{y}, \mathbf{X}, \boldsymbol{\epsilon}; T, \lambda_{ij})$, together with the private consumption share $s^\star_{ij}(\mathbf{w}, \mathbf{y}, \mathbf{X}, \boldsymbol{\epsilon}; T, \lambda_{ij})$. Upon substitution into the individual utility functions (and including the state-specific error associated with the individual's own time allocation decision) we obtain the respective male and female indirect utility functions $v^i_{ij}(\mathbf{w}, \mathbf{y}, \mathbf{X}, \boldsymbol{\epsilon}; T, \lambda_{ij})$ and $v^j_{ij}(\mathbf{w}, \mathbf{y}, \mathbf{X}, \boldsymbol{\epsilon}; T, \lambda_{ij})$.

## 2.2 Marriage market

We embed our time allocation model in a frictionless empirical marriage market model. As noted above, an important timing assumption is that marriage market decisions are made prior to the realization of wage offers, preference shocks, and demographic transitions. Thus, decisions are made based upon the expected value of being in a given marital pairing, together with an idiosyncratic component that we describe below.

### 2.2.1 Expected values

Anticipating our later application, we write the expected values from remaining single for a type-$i$ single male and type-$j$ single female (excluding any additive idiosyncratic payoff that we describe below) as explicit functions of the tax system $T$. These respective expected values are

$$\begin{aligned} U^i_{i0}(T) &= \mathbb{E}[v^i_{i0}(w^i, y^i, \mathbf{X}^i, \epsilon^i; T)], \\ U^j_{0j}(T) &= \mathbb{E}[v^j_{0j}(w^j, y^j, \mathbf{X}^j, \epsilon^j; T)], \end{aligned}$$

where the expectation is taken over wage offers, demographics, and the preference shocks. For married individuals, their expected values (again excluding any additive idiosyncratic utility payoffs) may similarly be written as a function of the both the tax

[14]The collective approach literature largely ignores taxation. Exceptions include Donni (2003), Lise and Seitz (2011) and Vermeulen (2005).

system $T$ and a candidate Pareto weight $\lambda_{ij}$ associated with a type $\langle i, j \rangle$ match:

$$U_{ij}^{i}(T, \lambda_{ij}) = \mathbb{E}[v_{ij}^{i}(\mathbf{w}, \mathbf{y}, \mathbf{X}, \epsilon; T, \lambda_{ij})],$$
$$U_{ij}^{j}(T, \lambda_{ij}) = \mathbb{E}[v_{ij}^{j}(\mathbf{w}, \mathbf{y}, \mathbf{X}, \epsilon; T, \lambda_{ij})].$$

Note that the Pareto weight within a match does not depend upon the realization of uncertainty. This implies full commitment and efficient risk sharing within the household. The expected value of a type-$i$ man when married to a type-$j$ woman is strictly decreasing in the wife's Pareto weight $\lambda_{ij}$, while the expected value of the wife is strictly increasing in $\lambda_{ij}$. Moreover, we also obtain an envelope condition result that relates the change in male and female expected utilities as we vary the wife's Pareto weight:

$$\frac{\partial U_{ij}^{i}(T, \lambda_{ij})}{\partial \lambda} = -\frac{\lambda_{ij}}{1-\lambda_{ij}} \times \frac{\partial U_{ij}^{j}(T, \lambda_{ij})}{\partial \lambda} < 0. \tag{5}$$

We use this relationship later when demonstrating identification of the Pareto weight.

#### 2.2.2 Marriage decision

As in Choo and Siow (2006), we assume that in addition to the systematic component of utility (as given by the expected values above) a given type-$i$ male $g$ receives an idiosyncratic payoff that is specific to him and the type of spouse $j$ that he marries but not her specific identity. These idiosyncratic payoffs are denoted $\theta_{ij}^{i,g}$ and are observed prior to the marriage decision. Additionally, each male also receives an idiosyncratic payoff from remaining unmarried that depends on his specific identity and is similarly denoted as $\theta_{i0}^{i,g}$. The initial marriage decision problem of a given male $g$ is therefore to choose to marry one of the $J$ possible types of spouses or to remain single. His decision problem is therefore

$$\max_{j} \{U_{i0}^{i}(T) + \theta_{i0}^{i,g}, U_{i1}^{i}(T, \lambda_{i1}) + \theta_{i1}^{i,g}, \ldots, U_{iJ}^{i}(T, \lambda_{iJ}) + \theta_{iJ}^{i,g}\}, \tag{6}$$

where the choice $j = 0$ corresponds to the single state.

We assume that the idiosyncratic payoffs follow the Type-I extreme value distribution with a zero location parameter and the scale parameter $\sigma_{\theta}$. This assumption implies that the proportion of type-$i$ males who would like to marry a type-$j$ female (or remain unmarried) are given by the conditional choice probabilities:

$$\begin{aligned} p^i_{ij}(T,\boldsymbol{\lambda}^i) &= \Pr[U^i_{ij}(T,\lambda_{ij}) + \theta^i_{ij} > \max\{U^i_{ij}(T,\lambda_{ih}) + \theta^i_{ih}, U^i_{i0}(T) + \theta^i_{i0}\} \quad \forall h \neq j] \\ &= \frac{\mu^d_{ij}(T,\boldsymbol{\lambda}^i)}{m_i} = \frac{\exp[U^i_{ij}(T,\lambda_{ij})/\sigma_\theta]}{\exp[U^i_{i0}(T)/\sigma_\theta] + \sum_{h=1}^{J}\exp[U^i_{ih}(T,\lambda_{ih})/\sigma_\theta]}, \end{aligned} \tag{7}$$

where $\boldsymbol{\lambda}^i = [\lambda_{i1},\ldots,\lambda_{iJ}]^\intercal$ is the $J \times 1$ vector of Pareto weights associated with different spousal options for a type-$i$ male, and $\mu^d_{ij}(T,\boldsymbol{\lambda}^i)$ is the measure of type-$i$ males who "demand" type-$j$ females (the conditional choice probabilities $p^i_{ij}(T,\boldsymbol{\lambda}^i)$ multiplied by the measure of type-$i$ men). Women also receive idiosyncratic payoffs associated with the different marital states (including singlehood) and their marriage decision problem is symmetrically defined. With identical distributional assumptions, the proportion of type-$j$ females who would like to marry a type-$i$ male is given by

$$p^j_{ij}(T,\boldsymbol{\lambda}^j) = \frac{\mu^s_{ij}(T,\boldsymbol{\lambda}^j)}{f_j} = \frac{\exp[U^j_{ij}(T,\lambda_{ij})/\sigma_\theta]}{\exp[U^j_{0j}(T)/\sigma_\theta] + \sum_{g=1}^{I}\exp[U^j_{gj}(T,\lambda_{gj})/\sigma_\theta]}, \tag{8}$$

where $\boldsymbol{\lambda}^j = [\lambda_{1j},\ldots,\lambda_{Ij}]^\intercal$ is the $I \times 1$ vector of Pareto weights for a type-$j$ female, and $\mu^s_{ij}(T,\boldsymbol{\lambda}^j)$ is the measure of type-$j$ females who would choose type-$i$ males. We also refer to this measure as the "supply" of type-$j$ females to the $\langle i,j\rangle$ sub-marriage market.

### 2.2.3 Marriage market equilibrium

An equilibrium of the marriage market is characterized by an $I \times J$ matrix of Pareto weights $\boldsymbol{\lambda} = [\boldsymbol{\lambda}^1, \boldsymbol{\lambda}^2, \ldots, \boldsymbol{\lambda}^J]$ such that for all $\langle i,j\rangle$ the measure of type-$j$ females demanded by type-$i$ men is equal to the measure of type-$j$ females supplied to type-$i$ males. That is,

$$\mu_{ij}(T,\boldsymbol{\lambda}) = \mu^d_{ij}(T,\boldsymbol{\lambda}^i) = \mu^s_{ij}(T,\boldsymbol{\lambda}^j) \quad \forall i = 1,\ldots,I, j = 1,\ldots,J. \tag{9}$$

Where we note that the equilibrium weights will depend both on the distribution of economic gains from alternative marriage market positions, the distribution of idiosyncratic marital payoffs, and the relative scarcity of spouses of different types. Along with the usual regularity conditions, which are formally stated in Appendix A, a sufficient condition for the existence and uniqueness of a marriage market equilibrium is provided in Proposition 1. This states that the limit of individual utility is negative infinity as his/her private consumption approaches zero. Essentially, this condition allows us to make utility for any individual arbitrarily low through suitable choice of Pareto weight and will be imposed through appropriate parametric restrictions on the utility function.[15] We now state our formal existence and uniqueness proposition.

**Proposition 1.** *If the idiosyncratic marriage market payoffs follow the Type-I extreme value distribution, the regularity conditions stated in Appendix A hold and the utility function satisfies*

$$\lim_{q^i \to 0} u^i(\ell^i, q^i, Q; \mathbf{X}^i) = \lim_{q^j \to 0} u^j(\ell^j, q^j, Q; \mathbf{X}^j) = -\infty, \tag{10}$$

*then an equilibrium of the marriage market exists and is unique.*

*Proof.* See Appendix A. □

Our proof is based on constructing excess demand functions and then showing that a unique Walrasian equilibrium exists. This is the same approach used by Galichon, Kominers and Weber (2016) under a more general heterogeneity structure. A proof using the marriage matching function and Type-I extreme value errors is presented in Galichon, Kominers and Weber (2014). In Online Appendix E we describe the numerical algorithm and approximation methods that we apply when solving for the equilibrium of the marriage market given any tax and transfer system $T$. In that appendix we also note important properties regarding how the algorithm scales as the number of markets is increased.

# 3 Optimal taxation framework

In this section we present the analytical framework that we use to study tax reforms that are optimal under a social-welfare function. The social planners problem is to choose a tax system $T$ to maximize a social-welfare function subject to a revenue requirement, the individual/household incentive compatibility constraints, and the marriage market equilibrium conditions. The welfare function is taken to be *individualistic*, and is based on individual maximized (incentive compatible) utilities following both the clearing of the marriage market, and the realizations of wage offers, state-specific preferences, and demographic transitions. Note that inequality both within and across households will adversely affect social welfare.

[15]Similar limiting properties of the utility function are common in the literature on collective household models, to rule out corner solutions. See, e.g., Donni (2003).

In what follows, we use $G^i_{i0}(w^i, \mathbf{X}^i, \boldsymbol{\epsilon}^i)$ and $G^j_{0j}(w^j, \mathbf{X}^j, \boldsymbol{\epsilon}^j)$ to respectively denote the single type-$i$ male and single type-$j$ female joint cumulative distribution functions for wage offers, state-specific errors, and demographic transitions. The joint cumulative distribution function within an $\langle i, j\rangle$ match is similarly denoted $G_{ij}(\mathbf{w}, \mathbf{X}, \boldsymbol{\epsilon})$. It is also necessary to describe the endogenous distribution of idiosyncratic payoffs for individuals in a given marital pairing. These differ from the unconditional $EV(0, \sigma_\theta)$ distribution for the population as a whole, because individuals non-randomly select into different marital pairings. They are therefore also a function of tax policy. We let $H^i_{i0}(\theta^i; T)$ denote the cumulative distribution function of these payoffs amongst single type-$i$ males and similarly define $H^j_{0j}(\theta^j; T)$ for single type-$j$ females. Among married men and women in an $\langle i, j\rangle$ match, these are given by $H^i_{ij}(\theta^i; T)$ and $H^j_{ij}(\theta^j; T)$, respectively. We provide a theoretical characterization of these distributions in Appendix B.

Our simulations will consider the implications of alternative redistributive preferences for the planner, which we will capture through the utility transformation function $\Upsilon(\cdot)$.[16] The social-welfare function is defined as the sum of these transformed utilities:

$$
\begin{aligned}
\mathcal{W}(T) = {} & \underbrace{\sum_i \mu_{i0}(T) \int \Upsilon\left[v^i_{i0}(w^i, y^i, \mathbf{X}^i, \boldsymbol{\epsilon}^i; T) + \theta^i\right] \mathrm{d}G^i_{i0}(w^i, \mathbf{X}^i, \boldsymbol{\epsilon}^i)\, \mathrm{d}H^i_{i0}(\theta^i; T)}_{\text{single men}} \\
& + \underbrace{\sum_j \mu_{0j}(T) \int \Upsilon\left[v^j_{0j}(w^j, y^j, \mathbf{X}^j, \boldsymbol{\epsilon}^j; T) + \theta^j\right] \mathrm{d}G^j_{0j}(w^j, \mathbf{X}^j, \boldsymbol{\epsilon}^j)\, \mathrm{d}H^j_{0j}(\theta^j; T)}_{\text{single women}} \\
& + \underbrace{\sum_{i,j} \mu_{ij}(T) \int \Upsilon\left[v^i_{ij}(\mathbf{w}, \mathbf{y}, \mathbf{X}, \boldsymbol{\epsilon}; T, \lambda_{ij}(T)) + \theta^i\right] \mathrm{d}G_{ij}(\mathbf{w}, \mathbf{X}, \boldsymbol{\epsilon})\, \mathrm{d}H^i_{ij}(\theta^i; T)}_{\text{married men}} \\
& + \underbrace{\sum_{i,j} \mu_{ij}(T) \int \Upsilon\left[v^j_{ij}(\mathbf{w}, \mathbf{y}, \mathbf{X}, \boldsymbol{\epsilon}; T, \lambda_{ij}(T)) + \theta^j\right] \mathrm{d}G_{ij}(\mathbf{w}, \mathbf{X}, \boldsymbol{\epsilon})\, \mathrm{d}H^j_{ij}(\theta^j; T)}_{\text{married women}} . \qquad (11)
\end{aligned}
$$

The maximization of $\mathcal{W}(T)$ is subject to a number of constraints. First, there are the usual incentive compatibility constraints that require that time allocation and consumption decisions for individuals and households are optimal given $T$. We embed this requirement in our formulation of the problem through the inclusion of the indirect utility functions.

[16] Note that in general, this formulation implies that the planner is weighting individual utilities differently relative to the household (as determined by the market clearing vector of Pareto weights).

Second, individuals optimally select into different marital pairings based upon expected values and their realized idiosyncratic payoffs (equation (6)). Third, we obtain a marriage market equilibrium so that given $T$ there is neither an excess demand or excess supply of spouses in each sub-marriage market (equation (9)). In Proposition 1 we provide sufficient conditions for the existence and uniqueness of a marriage market equilibrium given $T$. Fourth, it is required that an exogenously determined revenue amount $\overline{T}$ is raised, as given by the revenue constraint:

$$
\begin{aligned}
\mathcal{R}(T) = & \underbrace{\sum_i \mu_{i0}(T) \int R^i_{i0}(w^i, y^i, \mathbf{X}^i, \epsilon^i; T)\, \mathrm{d}G^i_{i0}(w^i, \mathbf{X}^i, \epsilon^i)}_{\text{revenue from single men}} \\
& + \underbrace{\sum_j \mu_{0j}(T) \int R^j_{0j}(w^j, y^j, \epsilon^j, \mathbf{X}^j; T)\, \mathrm{d}G^j_{0j}(w^j, \mathbf{X}^j, \epsilon^j)}_{\text{revenue from single women}} \\
& + \underbrace{\sum_{i,j} \mu_{ij}(T) \int R_{ij}(\mathbf{w}, \mathbf{y}, \mathbf{X}, \epsilon; T, \lambda_{ij}(T))\, \mathrm{d}G_{ij}(\mathbf{w}, \mathbf{X}, \epsilon) \geq \overline{T}}_{\text{revenue from married couples}}, \qquad (12)
\end{aligned}
$$

where $R^i_{i0}(w^i, y^i, \mathbf{X}^i, \epsilon^i; T)$ describes the tax revenue raised from an optimizing type-$i$ single male given $w^i$, $y^i$, $\mathbf{X}^i$, $\epsilon^i$, and the tax system $T$. We similarly define $R^j_{0j}(w^j, y^j, \epsilon^j, \mathbf{X}^j; T)$ for single type-$j$ women, and $R_{ij}(\mathbf{w}, \mathbf{y}, \mathbf{X}, \epsilon; T, \lambda_{ij}(T))$ for married $\langle i, j \rangle$ couples.

Taxes affect the problem in the following ways. First, they have a direct effect on welfare and revenue holding behaviour and the marriage market fixed. Second, there is a behavioural effect such that time allocations within a match change and affect both welfare and revenue. Third, there is a marriage market effect that changes who marries whom, the allocation of resources within the household (through adjustments in the Pareto weights), and the distribution of the idiosyncratic payoffs within any given match.

# 4 Data, identification and estimation

## 4.1 Data

We use two data sources for our estimation. First, we use data from the 2006 ACS which provides us with information on education, marital patterns, demographics, incomes, and labour supply. We supplement this with pooled ATUS data, which we use to construct a broad measure of home time for individuals sampled in the pre-recession period (2002–2007).[17] Following Aguiar and Hurst (2007) and Aguiar, Hurst and Karabarbounis (2012), we segment the total endowment of time into three broad mutually exclusive time-use categories: work activities, home production activities, and leisure activities.[18] Home production hours contains core home production, activities related to home ownership, obtaining goods and services, care of other adults, and childcare hours that measure all time spent by an individual caring for, educating, or playing with his/her child(ren).[19]

For both men and women we define three broad education groups for our analysis: high school and below, some college (less than four years of college), and college and above (a four-year or advanced degree). These constitute the individual *types* for the purposes of marriage market matching.[20] Our sample is restricted to single individuals ages 25–35 (inclusive). For married couples, we include all individuals where the reference householder (as defined by the Census Bureau) belongs to this same age band.[21]

Our estimation allows for market variation in the population vectors and the economic environment (taxes and transfers). We define a market at the level of the Census Bureau-designated division, with each division comprising a small number of states.[22] Within these markets, we calculate accurate tax schedules (defined as piecewise linear functions of family earnings) prior to estimation using the National Bureau of Economic Research TAXSIM calculator (see Feenberg and Coutts, 1993). These tax schedules include both federal and state tax rates (including the Earned Income Tax Credit) supplemented with detailed program rules for major welfare programs. The inclusion of

[17]The ATUS is a nationally representative cross-sectional time-use survey launched in 2003 by the U.S. Bureau of Labor Statistics. The ATUS interviews randomly selected individuals age 15 and older from a subset of the households that have completed their eighth and final interview for the Current Population Survey, the U.S. monthly labour force survey.

[18]See Aguiar, Hurst and Karabarbounis (2012) for a full list of the time-use categories contained in the ATUS data and a description of how there are categorized.

[19]We use sample weights when constructing empirical moments from each data source. Measures of home time from ATUS are constructed based on a 24-hour time diary completed by survey respondents. We adjust the sample weights so we continue to have a uniform distribution of weekdays following our sample selection. This is a common adjustment. See, e.g., Frazis and Stewart (2007).

[20]This type of educational categorization is standard in the marriage market literature. Papers that have used similar categories include Choo and Siow (2006), Choo and Seitz (2013), Goussé, Jacquemet and Robin (2015), Chiappori, Iyigun and Weiss (2009), and Chiappori, Salanié and Weiss (2014).

[21]Similar age selections are common in the literature. See Chiappori, Iyigun and Weiss (2009), Chiappori, Salanié and Weiss (2014), and Galichon and Salanié (2015) for examples.

[22]There are nine U.S. Census Bureau divisions. We do not use a finer level of market disaggregation due to sample size and computational considerations. An alternative feasible approach (but at the loss of sample size) would be to estimate the model on a subset of states.

welfare benefits is important as it allows us to better capture the financial incentives for lower-income households. We describe our implementation of these welfare rules and the calculation of the combined tax and transfer schedules in Online Appendix D.

## 4.2 Empirical specification

In Section 4.5 we will see that there are important differences between men and women in labour supply and the time spent on home production activities. Moreover, there are large differences between those who are single and those who are married (and to whom married). Our aim is to construct a credible and parsimonious model of time allocation decisions that can well describe these facts.

All the estimation and simulation results presented here assume individual preferences that are separable in the private consumption good, leisure, and the public good consumption. Preferences are unchanged by the marriage, and similarly do not vary with worker type (education), gender, or other demographic characteristics. Specifically,

$$u(\ell, q, Q; \mathbf{X}) = \frac{q^{1-\sigma_q} - 1}{1 - \sigma_q} + \beta_\ell \frac{\ell^{1-\sigma_\ell} - 1}{1 - \sigma_\ell} + \beta_Q \frac{Q^{1-\sigma_Q} - 1}{1 - \sigma_Q}. \tag{13}$$

This preference specification allows us to derive an analytical expression for the private good consumption share $s_{ij}$ for any joint time allocation in the household (i.e., the solution to equation (3)). Given our parameterization, $s_{ij}$ is independent of the total household private good consumption and is tightly connected to the Pareto weight. We have

$$s_{ij}(\lambda_{ij}) = \left[ 1 + \left( \frac{\lambda_{ij}}{1 - \lambda_{ij}} \right)^{-1/\sigma_q} \right]^{-1},$$

which is clearly increasing in the female weight $\lambda_{ij}$.[23] In the case that $\sigma_q = 1$ this expression reduces to $s_{ij}(\lambda_{ij}) = \lambda_{ij}$. To ensure that the sufficient conditions required for the existence and uniqueness of a marriage market equilibrium are satisfied (as described in Proposition 1), we require that $\sigma_q \geq 1$.

In our empirical application the demographic characteristics $\mathbf{X}$ will correspond to the presence of dependent children in the household.[24] For singles, the demographic transi-

[23] In the case where the private good curvature parameter $\sigma_q$ varies across spouses, the endogenous consumption share $s_{ij}$ will also be a function of household private good consumption.

[24] The model we have presented here does not have a cohabitation state. For individuals with children who were observed to be cohabiting, we treat them as both a single man and single women with children.

tion process depends on gender and own type. For married couples they depend on both own type and spousal type. These transition processes are estimated non-parametrically by market. Demographics (children) enter the model in the follow ways. First, children directly enter the empirical tax schedule, $T$. Second, children may affect the fixed work related costs (see equations (2a) and (4a)) with fixed costs restricted to be zero for individuals without children. Third, as we now describe, the presence of children may affect home time productivity.

The home time productivity of singles without children is restricted to be the same for both men and women. It may vary with education type. We allow this productivity to vary by gender for individuals with children. For married couples, we assume a Cobb-Douglas home production technology that depends on the time inputs of both spouses, $h^i_Q$ and $h^j_Q$, as well as a match specific term $\zeta_{ij}(\mathbf{X})$ that determines the overall efficiency of production within an $\langle i, j\rangle$ match and with demographics characteristics $\mathbf{X}$. That is

$$\widetilde{Q}_{ij}(h^i_Q, h^j_Q; \mathbf{X}) = \zeta_{ij}(\mathbf{X}) \times (h^i_Q)^{\alpha}(h^j_Q)^{1-\alpha}, \tag{14}$$

In our application, we restrict the specification of the match specific component. For all married households without children, we set $\zeta_{ij}(\mathbf{X}) = 1$. For married households with children, we restrict the match specific component in an $\langle i, j\rangle$ match to be of the form $\widetilde{\zeta}_j \times \vartheta_j^{\mathbb{1}[i=j]}$. The parameter $\vartheta_j$ captures potential complementarity in the home production technology for similar individuals.[25,26]

In addition to the home technology, individual heterogeneity also enters our empirical specification through market work productivity. Log-wage offers are normally distributed, with the parameters of the distribution an unrestricted function of both gender and the level of education.[27]

We define the time allocation sets $\mathbf{A}^i$ and $\mathbf{A}^j$ symmetrically for all individuals. The total time endowment $L_0$ is set equal to 112 hours per week. To construct these sets, we assume that both leisure and home time have a non-discretionary component (4 hours

This means that individuals in such unions are treated as if they are not able to enjoy the public good quality of home time. For the purposes of calculating tax liabilities, we only allow cohabiting women to claim children as a dependent.

[25]These restrictions were informed by first estimating a more general specification.

[26]Absent a measurement system for home produced output, preferences for the home produced good are indistinct from the production technology. For example, the parameter $\sigma_Q$ may reflect curvature in the utility or returns to scale in the production process.

[27]For simplicity we assume the realizations of wages within the household are independent conditional on male and female type, but this may be relaxed.

and 12 hours, respectively), and then define the residual discrete grid comprising 9 equi-spaced values. A unit of time is therefore given by $(112 - 12 - 4)/(9 - 1) = 12$ hours. Restricting market work and (discretionary) home time to be no more than 60 hours per week, there are a total of 30 discrete time allocation alternatives for individuals and $30^2 = 900$ discrete alternatives for couples.

The state-specific errors $\boldsymbol{\epsilon}_{\mathbf{a}^i}$ and $\boldsymbol{\epsilon}_{\mathbf{a}^j}$ associated with the individual time allocation decisions are Type-I extreme value with the scale parameter $\sigma_\epsilon$. The marriage decision depends upon the expected value of a match. For couples, the maximization problem of the household is not the same as the utility maximization problem of an individual. As a result, the well-known convenient results for expected utility and conditional choice probabilities in the presence of extreme value errors (see, e.g., McFadden, 1978) do not apply for married individuals. We therefore evaluate these objects numerically.[28]

## 4.3 Identification

The estimation will be of a fully specified parametric model. It is still important to explore non-/semi-parametric identification of the model because it indicates the source of variation in the data that is filtered through the economic model that gives rise to the parameter estimates, versus which parameter estimates arise from the functional form imposed in estimation. Here we explore semi-parametric identification. Using the marriage market equilibrium conditions and variation in the population vectors across markets, we prove identification of the wife's Pareto weight. Then using observations on the time allocation decisions of single and married individuals, we prove identification of the primitives of the model, i.e., the utility function, home production technology, and parameters of the distributions of state-specific errors.

### 4.3.1 Identifying the wife's Pareto weight from marriage

The literature on the identification of collective models largely focuses on the identification and estimation of the sharing rule. While knowledge of the sharing rule is useful in

[28]We approximate the integral over these preference shocks through simulation. To preserve smoothness of our distance metric (in estimation), as well as the welfare and revenue functions (in our design simulations), we employ a Logistic smoothing kernel. Conditional on $(\mathbf{w}, \mathbf{y}, \mathbf{X}, \boldsymbol{\epsilon})$ and the match $\langle i, j \rangle$ this assigns a *probability* of any given joint allocation being chosen by the household. We implement this by adding an extreme value error with scale parameter $\tau_\epsilon > 0$ that varies with all possible joint discrete time alternatives. The probability of a given joint time allocation is given by the usual conditional Logit form. As the smoothing parameter $\tau_\epsilon \to 0$, we get the unsmoothed simulated frequency.

answering a large set of empirical questions, for the purposes of our empirical taxation design exercise, it is the set of model primitives and the household decision weights that are important. In the context of a collective model with both public and private goods, Blundell, Chiappori and Meghir (2005) and Browning, Chiappori and Lewbel (2013) show that if there exists a distribution factor, then both the model primitives and the household decision weights are identified. With such a model embed in an equilibrium marriage market setting, the existence of a distribution factor becomes synonymous with variation across marriage markets. Below we show how the marriage market equilibrium conditions, together with market variation, allow us to identify the household decision weight under very mild conditions.

**Proposition 2.** *Under the conditions stated in Proposition 1, and with sufficient market variation in population vectors, the wife's Pareto weight is identified.*

*Proof.* See Appendix C. □

The strongest assumption for the identification of the wife's Pareto weights is that the idiosyncratic marital payoffs are distributed Type-I extreme value with an unknown scale parameter.[29] This distributional assumption, is however, used at every stage of our analysis. In particular, it was used when establishing the existence and uniqueness of equilibrium and in the computation of equilibrium (see Section 2.2). It is also used later when theoretically characterizing the contribution of these marital payoffs to the social-welfare function in our optimal taxation application (see Section 5).

#### 4.3.2 Identifying the other primitives

The identification of the utility function, the home production technology, and the scale of the state-specific error distribution follows directly from standard semi-parametric identification results for discrete choice models (see Matzkin, 1992, 1993), here modified to reflect the joint-household decision problem. The observed time allocation decisions of single individuals is first used to identify the utility function, the scale of the state-specific errors, and the efficiency of single individual's home production time. Then, under the maintained assumption that while the budget set and home technology may differ by marital status but individual preferences do not, we use our knowledge of the Pareto weight (whose identification is discussed above) together with information on the

[29]See Galichon and Salanié (2015) for semi-parametric identification results in transferable utility matching models with more general heterogeneity structures.

time allocation behaviour of married couples to identify the home production technology for couples.[30] The wage offer distributions are identified as several exclusion restrictions needed for identification arise naturally in our framework (e.g. children and spouse characteristics affect labour force participation but not wages).[31] These objects imply identification of the expected values in any given marriage market pairing. The observed population vectors and marriage market matching function then imply identification of the scale of the idiosyncratic marital payoff. A formal description of our identification arguments, together with the required assumptions, is presented in Appendix C.

## 4.4 Estimation

We estimate our model with a moment based procedure, constructing a rich set of moments that are pertinent to household time allocation decisions and marital sorting patterns. A description of all the moments used is provided in Online Appendix F.

We employ an equilibrium constraints (or MPEC) approach to our estimation (Su and Judd, 2012). This requires that we augment the estimation parameter vector to include the complete vector of Pareto weights for each market. Estimation is then performed with $I \times J \times K$ non-linear equality constraints that require that there is neither excess demand nor supply for individuals in any marriage market pairing and in each market. That is, equation (9) holds.[32] In practice, this equilibrium constraints approach is much quicker than a nested fixed-point approach (which would require that we solve the equilibrium for every candidate model parameter vector in each market) and is also more accurate as it does not involve the solution approximation step that we describe in Online Appendix E. Letting $\boldsymbol{\beta}$ denote the $B \times 1$ parameter vector, our estimation problem may be formally described as

$$[\widehat{\boldsymbol{\beta}}, \boldsymbol{\lambda}(\widehat{\boldsymbol{\beta}})] = \arg\min_{\boldsymbol{\beta}, \boldsymbol{\lambda}} [\mathbf{m_{sim}}(\boldsymbol{\beta}, \boldsymbol{\lambda}) - \mathbf{m_{data}}]^{\intercal} \mathbf{W} [\mathbf{m_{sim}}(\boldsymbol{\beta}, \boldsymbol{\lambda}) - \mathbf{m_{data}}]$$
$$\text{s.t. } \mu^d_{ijk}(\boldsymbol{\beta}, \boldsymbol{\lambda}^i_k) = \mu^s_{ijk}(\boldsymbol{\beta}, \boldsymbol{\lambda}^j_k) \quad \forall i = 1, \ldots, I, j = 1, \ldots, J, k = 1, \ldots, K,$$

where $\boldsymbol{\lambda}$ defines the stacked $(I \times J \times K)$ vector of Pareto weights in all markets, $\mathbf{m_{data}}$ is

[30] The assumption that preferences are unchanged by marriage is used extensively in the literature. See Browning, Chiappori and Lewbel (2013), Couprie (2007), and Lewbel and Pendakur (2008), among others.

[31] See Das, Newey and Vella (2003).

[32] Given our definition of a market and the number of male/female types, this involves $3 \times 3 \times 9 = 81$ additional parameters and non-linear equality constraints.

the $M \times 1$ vector of empirical moments, $\mathbf{m_{sim}}(\boldsymbol{\beta}, \boldsymbol{\lambda})$ is the model moment vector given $\boldsymbol{\beta}$ and an arbitrary (i.e., potentially non-equilibrium) vector of Pareto weights $\boldsymbol{\lambda}$. Finally, $\mathbf{W}$ defines an $M \times M$ positive definite weighting matrix. Given the well-known problems associated with the use of the optimal weighting matrix (Altonji and Segal, 1996), we choose $\mathbf{W}$ to be a diagonal matrix, whose element is proportional to the inverse of the diagonal variance-covariance matrix of the empirical moments.[33] The solution to this estimation problem is such that $\widehat{\boldsymbol{\lambda}} = \boldsymbol{\lambda}(\widehat{\boldsymbol{\beta}})$.[34]

## 4.5 Estimation results

We now provide a brief overview of the results of our initial estimation exercise, focusing upon the fit of the model to some of the most salient features of the data as well as the behavioural implications of our model estimates. Parameter estimates are provided in Online Appendix G.

In Table 1 we show the fit to marital sorting patterns across all markets and can see that the while we slightly under predict the incidence of singlehood for college educated individuals, in general the model is capable of well replicating empirical marital sorting patterns. Consistent with the data, we obtain strong assortative mating on education. Recall that we do not have any parameter at the match level than can be varied to fit marital patterns independently of the time allocation behaviour. In Figure 1 we present the marginal distributions of market and home time for both men and women in different marriage market pairings, and by the presence of children (here aggregated over own and spousal types and markets). The model is able to generate the most salient features of the data: relative to single women, married women work less and have higher home time, with the differences most pronounced for women with children. There are much smaller differences in both labour supply and home time between single and married men. Men with children have higher home time than men without children, although the difference is much smaller than observed for women.

[33]Our empirical moments are calculated using two data sources that have very different sample sizes. Consequently, the empirical moments from the ACS are estimated with much greater precision than are those from the ATUS. To allow those from the ATUS to have a meaningful influence in our estimation, we scale the corresponding elements of $\mathbf{W}$ by a fixed factor $r \gg 1$.

[34]The variance matrix of our estimator is given by

$$\left[\mathbf{D_m^{\intercal} W D_m}\right]^{-1} \mathbf{D_m^{\intercal} W \Sigma W^{\intercal} D_m} \left[\mathbf{D_m^{\intercal} W D_m}\right]^{-1},$$

where $\boldsymbol{\Sigma}$ is the $M \times M$ covariance matrix of the empirical moments, and $\mathbf{D_m} = \partial \mathbf{m_{sim}}(\boldsymbol{\beta}, \boldsymbol{\lambda}(\boldsymbol{\beta}))/\partial \boldsymbol{\beta}$ is the $M \times B$ derivative matrix of the moment conditions with respect to the model parameters at $\boldsymbol{\beta} = \widehat{\boldsymbol{\beta}}$.

Table 1: Empirical and predicted marital sorting patterns

| | | Women | | | |
|---|---|---|---|---|---|
| | | | High school and below | Some college | College and above |
| | | – | 0.140 | 0.128 | 0.122 |
| | | | [0.133] | [0.110] | [0.091] |
| Men | High school and below | 0.159 | 0.139 | 0.066 | 0.026 |
| | | [0.140] | [0.146] | [0.063] | [0.040] |
| | Some college | 0.113 | 0.037 | 0.087 | 0.051 |
| | | [0.112] | [0.032] | [0.100] | [0.044] |
| | College and above | 0.119 | 0.013 | 0.037 | 0.154 |
| | | [0.082] | [0.019] | [0.045] | [0.177] |

*Notes:* Table shows the empirical and simulated marriage market matching function, aggregated over all marriage markets. The statistic in brackets corresponds to the simulated value given the model estimates. Empirical frequencies are calculated with the 2006 ACS using sample selection as detailed in Section 4.1.

Our estimation targets a number of moments conditional on market, with our semi-parametric identification result reliant upon the presence of market variation. In Figure 2 we show how well the model can explain market variation in marital sorting patterns. Each data point represents an element of the marriage market matching function in a given market, and we observe a strong concentration of the points around the diagonal, indicating a good model fit. In Figure 3 we illustrate the fit to cross-market unconditional work hours for men and women by type and in different marriage market pairings. Again, we observe a strong clustering of points around the diagonal.

Important objects of interest are the Pareto weights and how they vary at the level of the match and across markets. The Pareto weights implied by our model estimates are presented in Table 2 and we note important features. First, the female weight is increasing when a woman is more educated relative to her spouse. For example, a college educated woman receives (on average) a share of 0.46 if she is married to a man with the same level of education. For a woman of the same education type to be willing to marry a high school educated male, her share must be 0.61. Second, there is an asymmetric gender impact of education differences: with the exception of the lowest education match, we always have that $\lambda_{ij} + \lambda_{ji} < 1$. Third, there is dispersion in these weights across markets, reflecting the joint impact of variation in taxes and the population vectors.

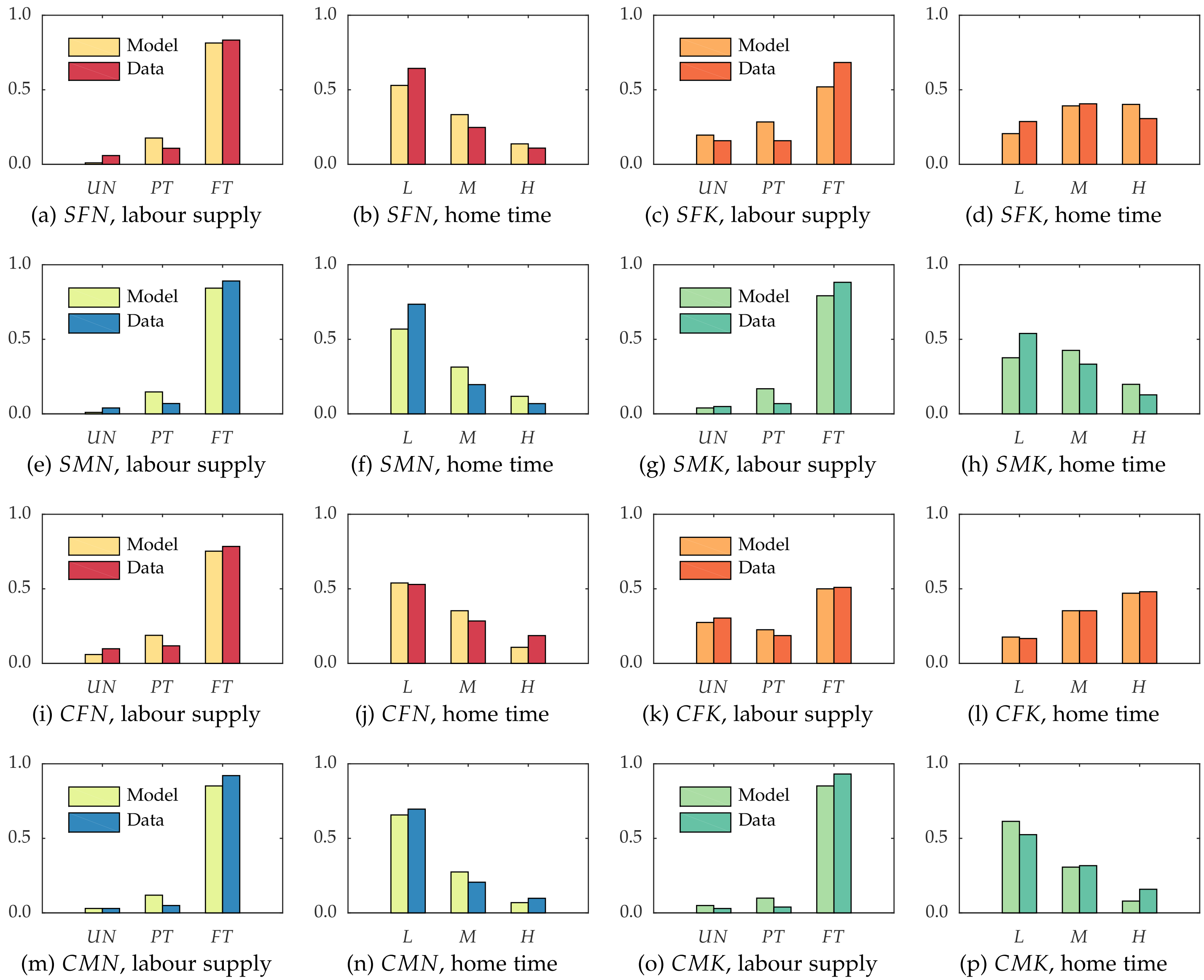



Figure 1: Figure shows empirical and predicted frequencies of work and home time, aggregated over types and conditional on marital status, gender, and children. $S$ ($C$) identifies singles (couples); $F$ ($M$) identifies women (men); $N$ ($K$) identifies childless (children). $UN$ is non-employment; $PT$ is part-time (12, 24 hours); $FT$ is full-time (36, 48, 60 hours). $L$ is low home time (4, 16 hours); $M$ is medium home time (28, 40 hours); $H$ is high home time (52, 64 hours).

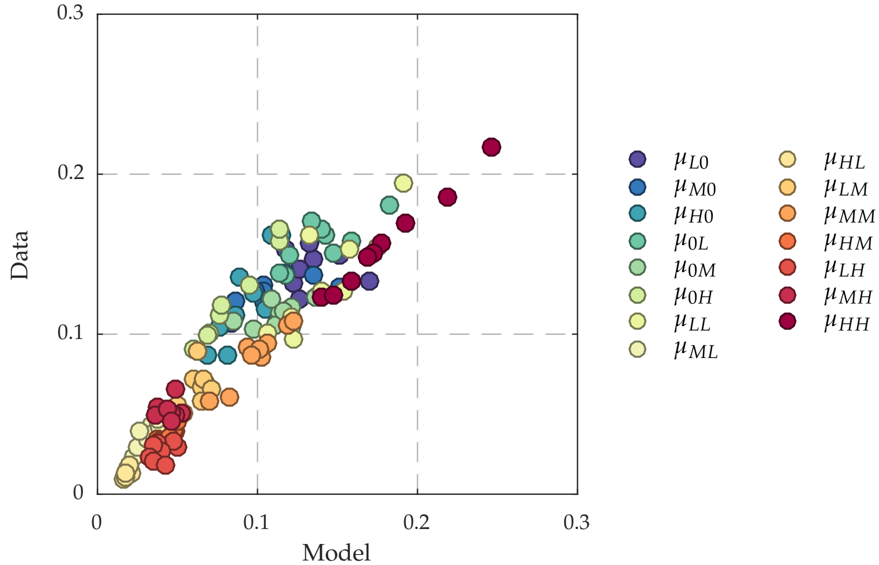



Figure 2: Figure shows elements of the empirical and predicted marriage market matching function. A market corresponds to a Census Bureau-designated division.

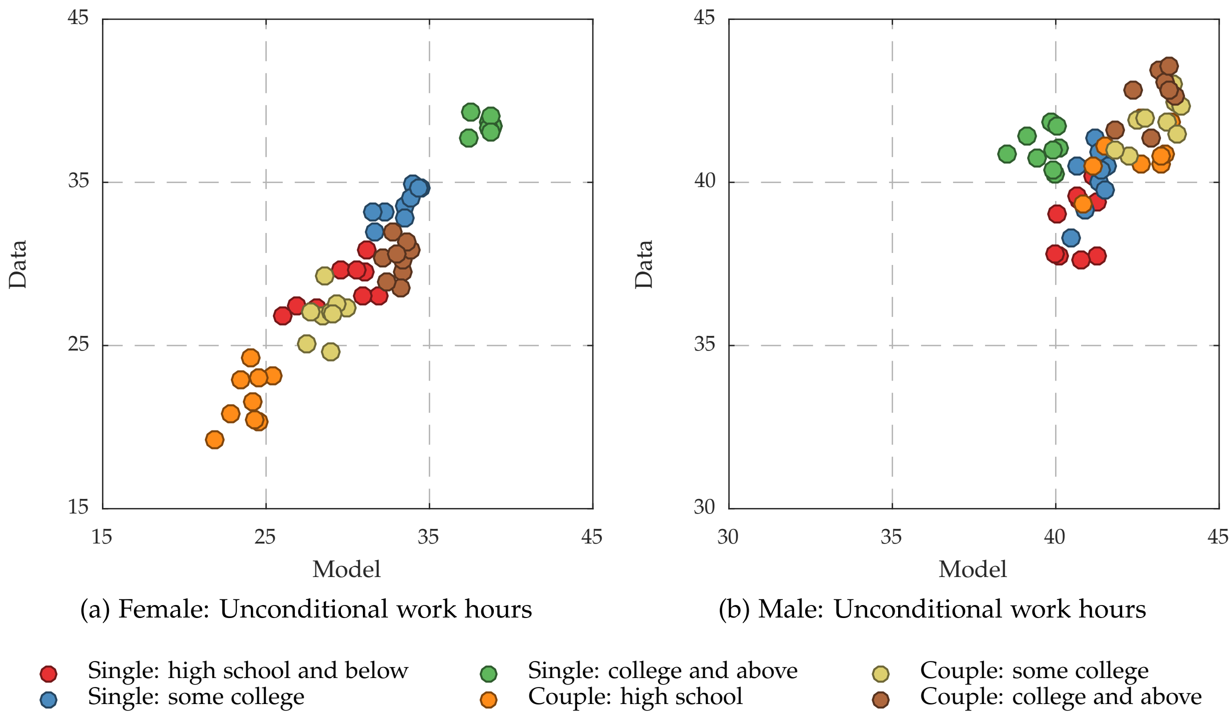



Figure 3: Figure shows empirical and predicted mean unconditional work hours of men and women by education and market. A market corresponds to a Census Bureau-designated division.

Table 2: Pareto weight distribution

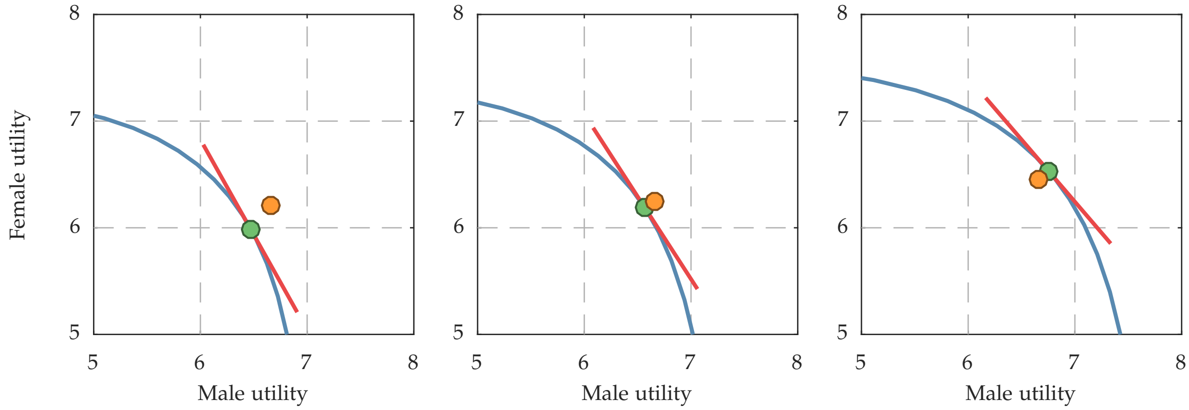
| | | Women | | |
|---|---|---|---|---|
| | | High school and below | Some college | College and above |
| Men | High school and below | 0.516 | 0.545 | 0.613 |
| | | [0.498–0.545] | [0.531–0.558] | [0.594–0.622] |
| | Some college | 0.441 | 0.493 | 0.558 |
| | | [0.422–0.466] | [0.485–0.508] | [0.542–0.565] |
| | College and above | 0.333 | 0.380 | 0.465 |
| | | [0.312–0.359] | [0.369–0.401] | [0.461–0.471] |

*Notes:* Table shows the distribution of Pareto weights under the 2006 federal and state tax and transfer systems. The numbers in black correspond to the average weight across markets (weighted by market size) within an $\langle i, j \rangle$ match. The range in brackets provides the range of values that we estimate across markets.

There are both economic and non-economic gains from marriage. In Figure 4 we present the empirical expected utility possibility frontier in marriages where the male has a college degree or higher and the education level of the female is varied. The patterns for other matches are similar. The expected utility possibility frontier, which is highly non-linear, shifts out as we increase the schooling level of the woman, and only in the joint "college and above" matches does the expected value of marriage exceed that of singlehood. Heterogamous marriages are therefore primarily explained by the non-economic gains.

While the following optimal design exercise directly uses the behavioural model developed in Section 2, to help understand the implications of our parameter estimates for time allocation decisions, we simulate elasticities under the actual 2006 tax systems for different family types. All elasticities are calculated by increasing the net wage rate while holding the marriage market fixed and correspond to uncompensated changes. In the presence of a non-separable tax schedule, increasing the net wage of a given married adult means that we are perturbing the tax schedule as we move in a single dimension.[35] The results of this exercise are shown in Table 3. For single individuals we report employment, conditional work hours, and home time elasticities in response to changes in their own wage. For married individuals we additionally report *cross-wage* elasticities

[35]Starting from a fully joint system (as is true in our estimation exercise) and for any given joint time allocation decision, this perturbation is equivalent to first taxing the spouse whose net wage is not varied on the original joint tax schedule and then reducing marginal tax rates for subsequent earnings (as then applied to the earnings of their spouse, whose net wage we are varying).




(a) High school and below (b) Some college (c) College and above

Figure 4: The figure shows the expected utility possibility frontier in marriages where the male has a college degree or higher and the education level of the female is varied. The figure is obtained from the estimated model with the empirical tax and transfer system and is calculated under the New England market. The green point in each panel indicates the expected utilities in the sub-marriage market given the market clearing Pareto weights. The orange point indicates the expected utilities in the single state.

that describe how employment, work hours, and home time respond as the wage of his/her spouse is varied.[36]

Our labour supply elasticities suggest that women are more responsive to changes in their own wage (both on the intensive and extensive margins) than are men. The same pattern is true with respect to changes in the wage of their partner. However, own-wage elasticities are always larger (in absolute terms) than are cross-wage elasticities. The own-wage hours and participation elasticities that we find are very much consistent with the range of estimates in the labour supply literature (see, e.g., Meghir and Phillips, 2010).[37] The evidence on cross-wage labour supply effects is more limited, although the results here are consistent with the estimates from the literature (see Blau and Kahn, 2007, Devereux, 2004, and Heim, 2009). Also in Table 3 we report home hours elasticities, which suggest that individuals substitute away from home time for a given uncompensated change in their wage and substitute towards home time when their spouse's wage is in-

[36]Own-wage conditional work-hours elasticities condition on being employed in the base system. As we increase the net wage of an individual (holding that of any spouse fixed) their employment is necessarily weakly increasing. Cross-wage conditional work-hours elasticities condition on being employed both before and after the net wage increase.

[37]As shown by Saez (2002) and Laroque (2005), differences in labour supply responsiveness at the intensive and extensive margins can have important implications for taxation design.

Table 3: Simulated elasticities

| | Married | | Single | |
|---|---|---|---|---|
| | Men | Women | Men | Women |
| *Work hours* | | | | |
| Own-wage elasticity | 0.17 | 0.24 | 0.04 | 0.08 |
| Cross-wage elasticity | -0.11 | -0.19 | – | – |
| *Participation* | | | | |
| Own-wage elasticity | 0.08 | 0.30 | 0.03 | 0.18 |
| Cross-wage elasticity | -0.05 | -0.17 | – | – |
| *Home hours* | | | | |
| Own-wage elasticity | -0.21 | -0.21 | -0.06 | -0.14 |
| Cross-wage elasticity | 0.12 | 0.13 | – | – |

*Notes:* All elasticities are simulated under 2006 federal and state tax/transfer systems, aggregated over markets, and hold the marriage market fixed. Elasticities are calculated by increasing an individual's net wage rate by 1% (own-wage elasticity) or the net wage of his/her spouse by 1% (cross-wage elasticity). *Participation* elasticities measure the percentage increase in the employment rate; *work hours* elasticities measure the percentage increase in hours of work among workers; and *home hours* elasticities measure the percentage increase in total home time.

creased. The same tax-induced home time pattern was reported in Gelber and Mitchell (2011).

We also simulate elasticities related to the impact of taxes on the marriage market. We consider a perturbation whereby we increase the marriage penalty/decrease the marriage bonus by 1% and then resolve for the equilibrium. This comparative static exercise implies a marriage market elasticity of -0.14. This result falls into the range of estimates in the literature that has examined the impact of taxation on marriage decisions, which often find what are considered modest (but statistically significant) effects. See, e.g., Alm and Whittington (1999) and Eissa and Hoynes (2000).

# 5 Optimal taxation of the family

In this section we consider the normative implications when we adopt a social-welfare function with a set of subjective social-welfare weights. There are two main stages to our analysis. Firstly, we consider the case where we do not restrict the form of jointness permitted in our choice of tax schedule for married couples. Under alternative assumptions on the degree of inequality aversion, we empirically characterize the form of the optimal tax system and show the importance of the marriage market in determining this. Second, we consider the choice of tax schedules when it is restricted to be either fully joint for married couples or completely independent. In both cases we quantify the welfare loss relative to our more general benchmark specification.

The results presented in this section assume a single market, with the population vectors for men and women defined as those corresponding to the aggregate. We consider the following form for the utility transformation function in our social-welfare function

$$\Upsilon(v;\theta) = \frac{e^{\delta v} - 1}{\delta}, \tag{15}$$

which is the same form as considered in the applications in, e.g., Mirrlees (1971) and Blundell and Shephard (2012). Under this specification $-\delta = -\Upsilon''(v;\theta)/\Upsilon'(v;\theta)$ is the coefficient of absolute inequality aversion, with $\delta = 0$ corresponding to the linear case (by L'Hôpital's rule).

This utility transformation function has useful properties, and in conjunction with the additivity of the idiosyncratic marital payoffs permits us to obtain the following result:

**Proposition 3.** *Consider a married type-i male in an $\langle i, j\rangle$ marriage. The contribution of such individuals to $\mathcal{W}(T)$ in equation (11) for $\delta < 0$ is given by*

$$\begin{aligned}\mathcal{W}^i_{ij}(T) &= \int_{\theta^i}\int_{\mathbf{w},\mathbf{X},\boldsymbol{\epsilon}} \Upsilon[v^i_{ij}(\mathbf{w},\mathbf{y},\mathbf{X},\boldsymbol{\epsilon};T,\lambda_{ij}) + \theta^i]\,\mathrm{d}G_{ij}(\mathbf{w},\mathbf{y},\mathbf{X})\,\mathrm{d}H^i_{ij}(\theta^i) \\ &= p^i_{ij}(T)^{-\delta\sigma_\theta}\Gamma(1-\delta\sigma_\theta)\int_{\mathbf{w},\mathbf{X},\boldsymbol{\epsilon}} \frac{\exp[\delta v^i_{ij}(\mathbf{w},\mathbf{y},\mathbf{X},\boldsymbol{\epsilon};T,\lambda_{ij})]}{\delta}\,\mathrm{d}G_{ij}(\mathbf{w},\mathbf{X},\boldsymbol{\epsilon}) - \frac{1}{\delta},\end{aligned}$$

*where $\Gamma(\cdot)$ is the gamma function and $p^i_{ij}(T)$ is the conditional choice probability (equation (7)) for type-i males. For $\delta = 0$ this integral evaluates to*

$$\mathcal{W}^i_{ij}(T) = \gamma - \sigma_\theta \log p^i_{ij}(T) + U^i_{ij}(T,\lambda_{ij}),$$

*where $\gamma = -\Gamma'(1) \approx 0.5772$ is the Euler-Mascheroni constant. The form of the welfare function contribution is symmetrically defined in alternative marriage market pairings and for married women, single men and single women.*

*Proof.* See Appendix B. □

As part of our proof, we characterize the distribution of the marital idiosyncratic payoffs for individuals who select into a given marriage market pairing.[38] This result allows

[38] This is a related, but distinct, result compared with Proposition 1 in Blundell and Shephard (2012).

us to decompose the welfare function contributions into parts that reflect the distribution of idiosyncratic utility payoffs from marriage and singlehood, and that which reflects the welfare from individual consumption and time allocation decisions. It is also obviously very convenient from a computational perspective as the integral over these idiosyncratic marital payoffs does not require simulating.

## 5.1 Specification of the tax schedule

Before presenting the results from our design simulations, we first describe the parametric specification of the tax system used in our illustrations. Consider the most general case. The tax system comprises a schedule for singles (varying with earnings) and a schedule for married couples (varying with the earnings of both spouses). We exogenously define a set $\mathcal{Z}$ of $N$ ordered tax brackets $0 = n_1 < n_2 < \ldots < n_N < \infty$ that apply to the earnings of a given individual. We assume, but do not require, that these brackets are the same for each individual, married or single. Associated with each bracket point for singles is the tax level parameter vector $\mathbf{t}_{N\times 1}$. For married couples we have the tax level parameter matrix $\mathbf{T}_{N\times N}$. Consistent with real-world tax systems, we do not consider gender-specific taxation and therefore impose symmetry of the tax matrix in all our simulations. Together, our tax system is characterized by $N + N \times (N+1)/2$ tax parameters defined by the vector $\boldsymbol{\beta}_T = [\mathbf{t}_N, \text{vec}(\mathbf{T}_{N\times N})]$.

The tax parameter vector $\mathbf{t}_{N\times 1}$ and tax matrix $\mathbf{T}_{N\times N}$ define tax liabilities at earnings that coincide with the exogenously chosen tax brackets (or nodes). The tax liability for other earnings levels is obtained by fitting an interpolating function. For singles, this is achieved through familiar linear interpolation, so that the tax schedule is of a piecewise linear form. We extend this for married couples by a procedure of polygon triangulation. This procedure divides the surface into a non-overlapping set of triangles. Within each of these triangles, marginal tax rates for both spouses, while potentially different, are constant by construction.[39] Given this interpolating function, we write the tax schedule at arbitrary earnings for married couples as $T(z_1, z_2)$, where $z_1$ and $z_2$ are henceforth used to denote the labour earnings of the two spouses respectively. For a single individual with earnings $z$, and with some abuse of notation, we have $T(z)$. Note

That proposition does not apply to the welfare contribution conditional on a given marital state as (for individuals in couples) the maximization problem of the household is not synonymous with the maximization of the individual utility function.

[39]The requirement that marginal tax rates can not exceed 100% (as earnings in any feasible dimension is varied) may be incorporated by imposing $(N-1) + N \times (N-1)$ linear restrictions on the parameters.

that in our illustrations we do not condition upon demographics.

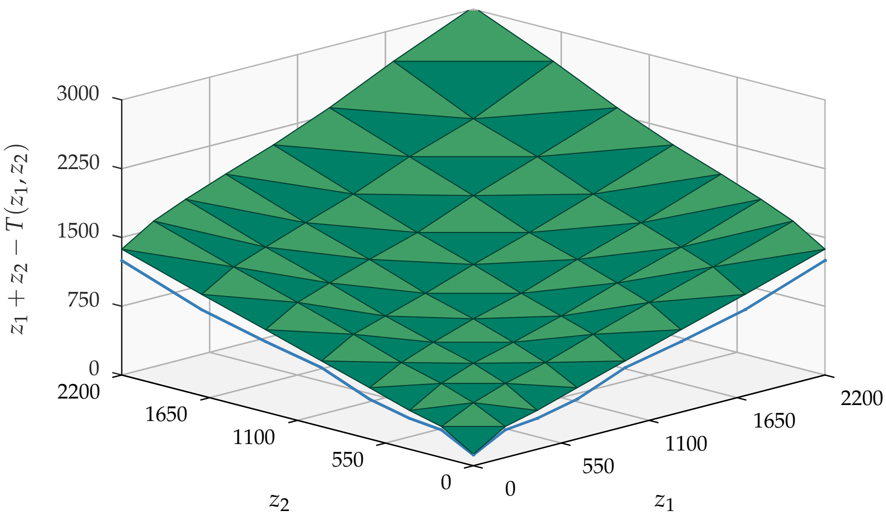
In our application, we set $N = 10$ with the earnings nodes (expressed in dollars per week in 2006 prices) as $\mathcal{Z} = \{0, 200, 400, 650, 950, 1300, 1700, 2200, 2800, 3500\}$. Thus, we have a tax system that is characterized by 65 parameters. Using our estimated model, the exogenous revenue requirement $\overline{T}$ is set equal to the expected state and federal income tax revenue (including Earned Income Tax Credits (EITC) payments) and net of welfare transfers. We solve the optimal design problem numerically. Given our parameterization of the tax schedule, we solve for the optimal tax parameter vector $\boldsymbol{\beta}_T$ using an equilibrium constraints approach that is similar to that described in Section 4.4 in the context of estimation. This approach involves augmenting the parameter vector to include the $I \times J$ vector of Pareto weights as additional parameters and imposing the $I \times J$ equilibrium constraints $\mu^d_{ij}(T, \boldsymbol{\lambda}^i) = \mu^s_{ij}(T, \boldsymbol{\lambda}^j)$ in addition to the usual incentive compatibility and revenue constraints. This approach only involves calculating the marriage market equilibrium associated with the optimal parameter vector $\boldsymbol{\beta}^*_T$ rather than any candidate $\boldsymbol{\beta}_T$, as would be true in a nested fixed-point procedure.

## 5.2 Implications for design

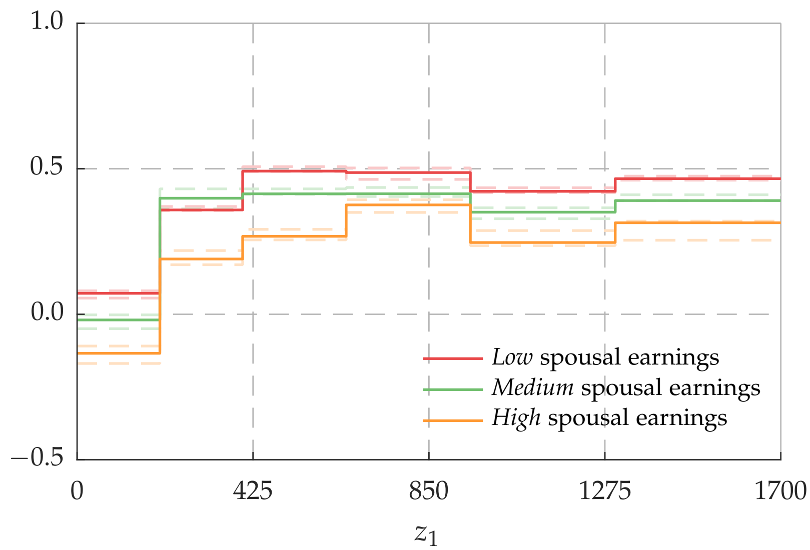
We now describe our main results. In Figure 5a we present the joint (net income) budget constraint for both singles and married couples, calculated under the government preference parameterization $\delta = 0$. For clarity of presentation, the figure has been truncated at individual earnings greater than \$2,200 a week (\$114,400 a year). The implied schedule for singles is shown by the blue line. The general flattening of this line as earnings increase indicates a broadly progressive structure for singles. In the same figure, the optimal schedule for married couples is shown by the three-dimensional surface, which is symmetric by construction (i.e., gender neutrality). Within each of the shaded triangles, the marginal tax rates of both spouses are different but constant. As the earnings of either spouse changes in any direction and enters a new triangle, marginal tax rates will potentially change. Holding constant the earnings of a given spouse, we can clearly see a progressive structure, while comparing these implied schedules at different levels of spousal earnings is informative about the degree of tax jointness.

To better illustrate the implied degree of tax jointness, in Figure 5b we show the associated marginal tax rate of a given individual as the earnings of his/her spouse is fixed at different levels.[40] Here, we also present the 95% pointwise confidence bands

[40]We present the (average) marginal tax rate for low, medium, and high spousal earnings. *Low* is

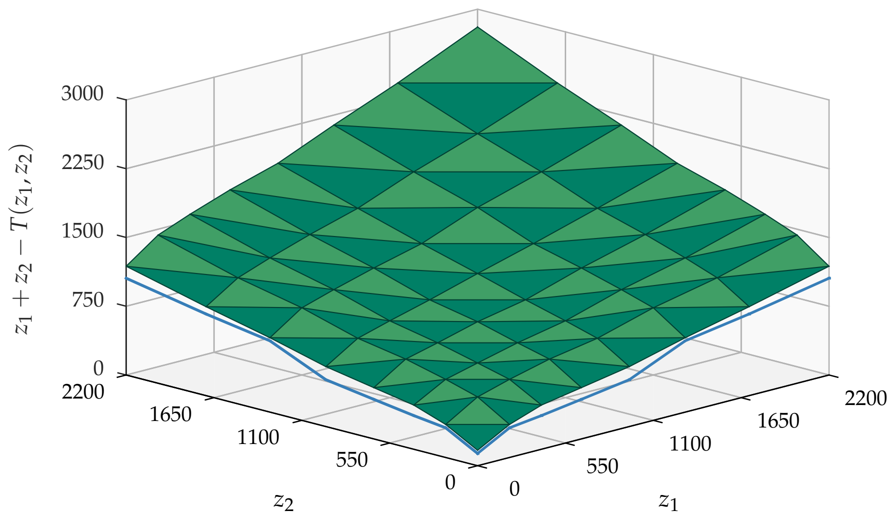



(a) Net income function

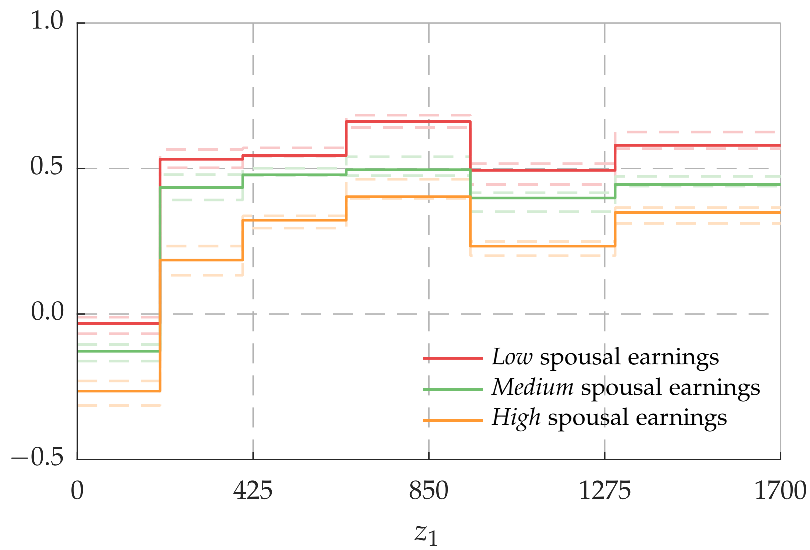



(b) Marginal tax rate conditional on spousal earnings

Figure 5: Optimal tax schedule with $\delta = 0$. In panel (a) we show net income as a function of labour earnings for both single individuals (blue line) and couples (three-dimensional surface). Marginal tax rates for both spouses (while potentially different) are constant within each of the shaded triangles. In panel (b) we show the implied structure of marginal tax rates conditional on alternative values of spousal earnings. The broken coloured lines indicate the associated 95% pointwise confidence bands. See Footnote 40 for a definition of *low*, *medium*, and *high* spousal earnings levels.




(a) Net income function




(b) Marginal tax rate conditional on spousal earnings

Figure 6: Optimal tax schedule with $\delta = -1$. In panel (a) we show net income as a function of labour earnings for both single individuals (blue line) and couples (three-dimensional surface). Marginal tax rates for both spouses (while potentially different) are constant within each of the shaded triangles. In panel (b) we show the implied structure of marginal tax rates conditional on alternative values of spousal earnings. The broken coloured lines indicate the associated 95% pointwise confidence bands. See Footnote 40 for a definition of *low*, *medium*, and *high* spousal earnings levels.

that are obtained by sampling 200 times from the distribution of parameter estimates and resolving for the optimal schedule. We note a number of features. First, we can see that there exists a broadly progressive structure; second, marginal tax rates are close to zero (or negative) at low earnings; and third marginal tax rates tend to be lower the higher the earnings of one's spouse.

In Figure 6 we repeat our analysis under an alternative parameterization for government preferences ($\delta = -1$). As we later show, this parameterization is associated with a considerably greater redistributive preference. Relative to the schedule obtained with $\delta = 0$, we have (i) higher transfers when not working; (ii) lower marginal tax rates (pure tax credits) at low earnings; and (iii) generally higher marginal tax rates with a greater degree of negative jointness (i.e., a larger difference in marginal rates as we increase the earnings of a spouse). Before commenting further on the structure of these schedules and their implications for behaviour, we first describe the underlying average social-welfare weights for these alternative government preference parameter values. These are presented in Table 4. They tell us the relative value that the government places on increasing consumption at different joint earnings levels. Given the maintained symmetry of the tax schedule, we present these welfare weights as a function of the lowest and highest earnings of a couple. These weights are monotonically declining in earnings as we move in either direction. Moreover, given the estimated curvature of the utility function, there is a considerable redistributive motive even in the $\delta = 0$ case.

The choice of tax schedule has implications for time allocation decisions, marriage market outcomes, and the distribution of resources within the household. We briefly comment on these effects when $\delta = 0$. Relative to our estimated baseline model, the most pronounced difference in labour supply behaviour is for married women: the employment rate is 87%, relative to 78% in the estimated model, while conditional work hours and home hours are both around one hour per week lower. In contrast, the employment rate for married men is approximately unchanged (96%), while work hours are two hours per week lower and home time is slightly higher. In terms of marriage market outcomes, we obtain slightly lower welfare weights/private consumption shares for women and a higher overall marriage rate (a 3-percentage-point increase).

the arithmetic average of the marginal tax rate schedule for all spousal earnings $\{z_2 | z_2 \in \mathcal{Z}, z_2 \leq 650\}$. Similarly, *medium* and *high* respectively correspond to spousal earnings levels $\{z_2 | z_2 \in \mathcal{Z}, 650 < z_2 \leq 1700\}$ and $\{z_2 | z_2 \in \mathcal{Z}, z_2 > 1700\}$.

Table 4: Social-welfare weights under optimal system

| | Highest earnings range | Lowest earnings range: 0–200 | 200–400 | 400–650 | 650–950 | 950–1300 | 1300–1700 | 1700–2200 | 2200–2800 | 2800+ |
|---|---|---|---|---|---|---|---|---|---|---|
| $\delta$ = 0: | 0–200 | 2.580 [0.434] | – | – | – | – | – | – | – | – |
| | 200–400 | 2.025 [4.332] | 1.474 [4.719] | – | – | – | – | – | – | – |
| | 400–650 | 1.576 [7.003] | 1.218 [11.001] | 1.027 [5.932] | – | – | – | – | – | – |
| | 650–950 | 1.249 [5.688] | 0.998 [7.280] | 0.861 [7.913] | 0.741 [2.704] | – | – | – | – | – |
| | 950–1300 | 0.995 [6.467] | 0.822 [5.876] | 0.726 [5.058] | 0.630 [3.853] | 0.542 [0.915] | – | – | – | – |
| | 1300–1700 | 0.770 [3.055] | 0.657 [2.773] | 0.595 [2.620] | 0.528 [2.068] | 0.464 [1.109] | 0.403 [0.318] | – | – | – |
| | 1700–2200 | 0.612 [1.493] | 0.531 [1.191] | 0.488 [1.169] | 0.437 [0.931] | 0.395 [0.490] | 0.346 [0.272] | 0.295 [0.068] | – | – |
| | 2200–2800 | 0.480 [0.764] | 0.424 [0.570] | 0.396 [0.543] | 0.361 [0.432] | 0.334 [0.205] | 0.296 [0.109] | 0.258 [0.041] | 0.225 [0.008] | – |
| | 2800+ | 0.374 [0.187] | 0.338 [0.118] | 0.317 [0.114] | 0.291 [0.098] | 0.273 [0.047] | 0.249 [0.021] | 0.222 [0.007] | 0.195 [0.003] | 0.173 [0.000] |
| $\delta$ = -1: | 0–200 | 3.343 [0.524] | – | – | – | – | – | – | – | – |
| | 200–400 | 2.231 [6.463] | 1.553 [5.447] | – | – | – | – | – | – | – |
| | 400–650 | 1.648 [9.065] | 1.191 [12.005] | 0.945 [5.601] | – | – | – | – | – | – |
| | 650–950 | 1.234 [5.272] | 0.919 [7.230] | 0.742 [7.351] | 0.588 [2.493] | – | – | – | – | – |
| | 950–1300 | 0.971 [4.785] | 0.722 [5.428] | 0.590 [4.625] | 0.473 [3.577] | 0.367 [0.873] | – | – | – | – |
| | 1300–1700 | 0.635 [2.457] | 0.505 [2.637] | 0.430 [2.515] | 0.353 [2.004] | 0.289 [1.086] | 0.234 [0.334] | – | – | – |
| | 1700–2200 | 0.443 [1.160] | 0.364 [1.160] | 0.319 [1.122] | 0.267 [0.901] | 0.228 [0.472] | 0.186 [0.290] | 0.150 [0.070] | – | – |
| | 2200–2800 | 0.297 [0.629] | 0.257 [0.525] | 0.231 [0.512] | 0.199 [0.415] | 0.177 [0.204] | 0.150 [0.115] | 0.123 [0.051] | 0.104 [0.009] | – |
| | 2800+ | 0.190 [0.162] | 0.171 [0.125] | 0.157 [0.120] | 0.138 [0.102] | 0.124 [0.049] | 0.108 [0.022] | 0.092 [0.008] | 0.081 [0.003] | 0.066 [0.000] |

*Notes:* Table presents average social-welfare weights and joint probability mass under the optimal system for alternative $\delta$ values. The probability mass is presented in brackets. Earnings are in dollars per week in 2006 prices. Welfare weights are obtained by increasing consumption in the respective joint earnings bracket (with fraction $s_{ij}(\lambda_{ij})$ of this increase in an $\langle i, j \rangle$ match accruing to the female) and calculating a derivative of the social-welfare function; weights are normalized so that the probability-mass-weighted sum under the optimal tax system is equal to unity.

## 5.3 Perturbation experiments

Designing taxes is complex. To better understand the influence of various model features on the design problem, we consider a series of perturbation experiments. Details of these, together with further simulations, are provided in Online Appendix H. Here we focus on a series of comparative static exercises that examine the role of the marriage market. In the context of their model, Kleven, Kreiner and Saez (2007) provide an informal argument that marriage distortions will tend to reduce the optimal degree of negative jointness. To examine this issue, we relax the problem by removing the $I \times J$ constraints that require zero excess demand in all marriage submarkets. We then resolve for the optimal structure holding the entire vector of Pareto weights, marriage market pairings, and distributions of idiosyncratic payoffs fixed at their values from the corresponding optimum from Section 5.2. The extent to which the optimal schedules differ (together with any imbalance in spousal type supply/demand) once the marriage market is held fixed is directly informative about the importance of the marriage market. Our results, which we present fully in the Appendix, suggest that the marriage market does have a quantitatively important influence on the design problem, even given small marriage market elasticities. We obtain a tax structure which, absent marriage market responses, implies lower net income at low earnings levels for married couples, together with lower marginal tax rates and increased negative jointness at low earnings. We also quantify the implied imbalance in the marriage market. The total excess supply/demand under this system relative to the size of the marriage market is around 7%. Again, this suggests that marriage market considerations have an important impact on the design problem.

Second, we consider how the degree of *assortative mating* influences the design problem. Frankel (2014) considered a simple binary model to analyse taxation design when couples have correlated types. In the context of uncorrelated types (as in Kleven, Kreiner and Saez, 2009) negative jointness is obtained, although this result is attenuated when the degree of exogenous assortative mating is increased. In our environment, we *endogenously* change the degree of assortative mating by introducing an additive utility component in educationally homogamous marriages. In the Online Appendix we demonstrate that the presence of assortative mating has a quantitatively important impact on the design problem. Consistent with the theoretical analysis in Frankel (2014), we find that as we induce less correlation in individual types (by making this utility component negative), the degree of negative jointness in the tax schedule for couples is increased.

## 5.4 Restrictions on the form of tax schedule jointness

Our previous analysis allowed for a very general form of jointness in the tax schedule. We now consider the design implications when the form of the jointness is restricted. There are two stages to our analysis. First, we characterize the tax schedule with a given revenue requirement by solving the same constrained welfare maximization problem as before. Second, in order to quantify the cost of these restricted forms, we consider the *dual problem*. That is, we now maximize the revenue raised from our tax system, subject to the incentive and marriage market equilibrium constraints *and* the requirement that the level of social-welfare achieved is at least that which was obtained from our unrestricted specification from Section 5.2. Here we consider the following forms for the tax schedule:

1. **Individual taxation**. In many countries there is a system of individual filing in the tax system. Under such a system, the total tax liability for a couple with earnings $z_1$ and $z_2$ is given by $T(z_1, z_2) = \widehat{T}(z_1) + \widehat{T}(z_2)$, where the function $\widehat{T}(\cdot)$ is the tax schedule that is applied to both married and single individuals.

2. **Joint taxation with income splitting**. Under a system of joint taxation with income splitting an individual is taxed upon an income measure that attributes the income of one spouse to the other. We consider equal splitting, so each household member is taxed based upon average earned income. Thus, the total tax liability for a family is $T(z_1, z_2) = 2 \times \widetilde{T}(z_1/2 + z_2/2)$, with the same tax schedule $\widetilde{T}(\cdot)$ applied to singles and couples.

3. **Joint taxation with income aggregation**. Here we maintain a common tax schedule but allow the tax liability of couples to depend upon total household earned income: $T(z_1, z_2) = \breve{T}(z_1 + z_2)$.

Full results are presented in Online Appendix I, where we show the implied marginal rate structure conditional on alternative spousal earnings levels (as in in Figures 5b and 6b). The rate schedule in the case of independent taxes does not, by definition, vary with the level of spousal earnings. While the shape of the schedule is broadly similar (relative to the unrestricted schedule) when spousal earnings are low, given our empirical finding of negative tax jointness, it does imply higher tax rates when spousal earnings are higher. Joint taxation with income splitting gives lower marginal tax rates (again, relative to the unrestricted schedule) when spousal earnings are low. At medium levels of spousal earnings, they are higher or at roughly the same level. At high levels of spousal earnings, marginal tax rates are everywhere higher. Finally, in the case of joint taxation with income aggregation, we have marginal tax rates that are higher at low earnings and lower at high earnings. This is true for the alternative spousal earnings levels. In the Online Appendix we also present the equilibrium marriage market matching functions that are associated with these alternative tax policies. Relative to the unrestricted specification, we see important changes. These changes are most pronounced when we consider joint taxation with income aggregation: the marriage rate is 14 percentage points lower, while the diagonal of the matrix becomes less dominant (i.e., less assortative mating).

The tax schedules derived are revenue equivalent to our most general specification but imply a reduction in social-welfare. We now quantify this welfare loss. To this end, we consider the dual problem of the planner as described above. The differences in revenue raised with the same social-welfare target can be interpreted as the cost of the more restrictive tax instruments. Individual taxation implies a welfare loss that is equivalent to around 1% of revenue; joint taxation with income splitting implies a 4% loss, while income aggregation implies a 9% loss. All numbers are slightly larger in the case of greater redistributive preference ($\delta = -1$) but the ranking remains the same. Thus, while we our most general specification did imply that the optimal system was characterized by negative jointness, the actual welfare gains from introducing this jointness appear quite modest.

# 6 Summary and conclusion

We have presented a micro-econometric equilibrium marriage matching model with labour supply, public home production, and private consumption. Household decisions are made cooperatively and, as in the general framework presented in Galichon, Kominers and Weber (2014, 2016), utility is imperfectly transferable across spouses. We provide sufficient conditions on the primitives of the model in order to obtain existence and uniqueness of equilibrium. Semi-parametric identification results are presented, and we show how the marriage market equilibrium conditions, together with market variation, allow us to identify the household decision weight.

Using an equilibrium constraints approach, we then estimate our model using American Community Survey and American Time Use Survey data, while incorporating detailed representations of the U.S. tax and transfer systems. We show that the model is able to jointly explain labour supply, home time, and marriage market patterns. Moreover, it is able to successfully explain the variation in these outcomes across markets, with the behavioural implications of the model shown to be consistent with the existing empirical evidence.

Our estimated model is then embedded within an extended Mirrlees framework to explore empirical taxation design problems. Our design exercise concerns the simultaneous choice of a tax schedule for singles and for married couples, recognizing that taxes may affect outcomes including who marries with whom and the allocation of resources within the household. For married couples, we allow for a very general form of the tax schedule and find empirical support for negative tax jointness (Kleven, Kreiner and Saez, 2009). Importantly, we also find that the welfare gain that such a system offers relative to the optimal system under fully independent taxation is relatively modest.

We believe that this paper represents an important step in placing both the family, and the marriage market, at the heart of the taxation design problem. Common with much of the empirical marriage matching literature, we do not consider cohabitation. But cohabitation is increasing in prevalence. If tax authorities do not recognize cohabitation, the ability for couples to cohabit introduces a form of tax avoidance. Our environment is static, with an irrevocable marriage decision. Marriage has an important life-cycle component and introduces many complex dynamic considerations related to the insurance that marriage may provide and the risk that different marriages may be exposed to. Taxes also affect other outcomes, such as education, that are relevant for the marriage decision. The exploration of such considerations is left for future work.

# Appendices

## A Proof of Proposition 1

We assume that the distribution $G_{ij}(\mathbf{w}, \mathbf{y}, \mathbf{X}, \boldsymbol{\epsilon})$ is absolutely continuous and twice continuously differentiable. The individual utility functions $u^i(\ell^i, q^i, Q; \mathbf{X}^i)$ and $u^j(\ell^j, q^j, Q; \mathbf{X}^j)$ are assumed increasing and concave in $\ell$, $q$, and $Q$, and with $\lim_{q^i \to 0} u^i(\ell^i, q^i, Q; \mathbf{X}^i) = \lim_{q^j \to 0} u^j(\ell^j, q^j, Q; \mathbf{X}^j) = -\infty$. To proceed we define the *excess demand function* as

$$ED_{ij}(\boldsymbol{\lambda}) = \mu_{ij}^d(\boldsymbol{\lambda}^i) - \mu_{ij}^s(\boldsymbol{\lambda}^j), \quad \forall i = 1, \ldots, I, j = 1, \ldots, J.$$

Here and in what follows, we suppress the dependence of the excess demand functions (and other objects) on the tax system $T$. Equilibrium existence is synonymous with the excess demand for all types being equal to zero at some vector $\boldsymbol{\lambda}^* \in [0,1]^{I\times J}$, i.e., $ED_{ij}(\boldsymbol{\lambda}^*) = 0$, $\forall i = 1,\ldots,I, j = 1,\ldots,J$. Equilibrium uniqueness implies that there is a single vector that achieves this result.[41] Under our regularity conditions, we have that: (i) $U^i_{ij}(\lambda_{ij})$ *and* $U^i_{ij}(\lambda_{ij})$ *are continuously differentiable in* $\lambda_{ij}$; (ii) $\partial U^i_{ij}(\lambda_{ij})/\partial\lambda = -\frac{\lambda_{ij}}{(1-\lambda_{ij})}\partial U^j_{ij}(\lambda_{ij})/\partial\lambda < 0$; (iii) $\lim\limits_{\lambda_{ij}\to 0} ED_{ij}(\lambda_{ij}, \boldsymbol{\lambda}_{-ij}) > 0$; and (iv) $\lim\limits_{\lambda_{ij}\to 1} ED_{ij}(\lambda_{ij}, \boldsymbol{\lambda}_{-ij}) < 0$.

## A.1 Properties of the excess demand functions

We now state further properties of the excess demand functions. We have

$$\frac{\partial ED_{ij}(\boldsymbol{\lambda})}{\partial\lambda_{ij}} = \frac{\partial\mu^d_{ij}(\boldsymbol{\lambda}^i)}{\partial\lambda_{ij}} - \frac{\partial\mu^s_{ij}(\boldsymbol{\lambda}^j)}{\partial\lambda_{ij}} < 0, \tag{A.1a}$$

$$\frac{\partial ED_{ik}(\boldsymbol{\lambda})}{\partial\lambda_{ij}} = \frac{\partial\mu^d_{ik}(\boldsymbol{\lambda}^i)}{\partial\lambda_{ij}} - \frac{\partial\mu^s_{ik}(\boldsymbol{\lambda}^k)}{\partial\lambda_{ij}} > 0;\ k \neq j, \tag{A.1b}$$

$$\frac{\partial ED_{kj}(\boldsymbol{\lambda})}{\partial\lambda_{ij}} = \frac{\partial\mu^d_{kj}(\boldsymbol{\lambda}^k)}{\partial\lambda_{ij}} - \frac{\partial\mu^s_{kj}(\boldsymbol{\lambda}^j)}{\partial\lambda_{ij}} > 0; k \neq i, \tag{A.1c}$$

$$\frac{\partial ED_{kl}(\boldsymbol{\lambda})}{\partial\lambda_{ij}} = \frac{\partial\mu^d_{kl}(\boldsymbol{\lambda}^k)}{\partial\lambda_{ij}} - \frac{\partial\mu^s_{kl}(\boldsymbol{\lambda}^l)}{\partial\lambda_{ij}} = 0;\ k \neq i, l \neq j, \tag{A.1d}$$

where equation (A.1d) follows from the IIA property of the Type-I extreme value distribution.

## A.2 Existence

To prove existence we construct a function, $\boldsymbol{\Gamma}(\boldsymbol{\lambda})$, as

$$\boldsymbol{\Gamma}(\boldsymbol{\lambda}) = \psi \cdot \mathbf{ED}(\boldsymbol{\lambda}) + \boldsymbol{\lambda},$$

for $\psi > 0$, which maps $[0,1]^{I\times J}$ onto $[0,1]^{I\times J}$. Then by Tarski's theorem, if $\boldsymbol{\Gamma}(\boldsymbol{\lambda})$ is nondecreasing in $\boldsymbol{\lambda}$, there exists a $\boldsymbol{\lambda}^* \in [0,1]^{I\times J}$ such that $\boldsymbol{\lambda}^* = \boldsymbol{\Gamma}(\boldsymbol{\lambda}^*)$. However, $\boldsymbol{\lambda}^* = \psi \cdot \mathbf{ED}(\boldsymbol{\lambda}^*) + \boldsymbol{\lambda}^*$ iff $\mathbf{ED}(\boldsymbol{\lambda}^*) = 0$. Assuming that $U^k_{ij}(\lambda_{ij})$ for $k = i, j$ is derived from the time allocation problem described in the main text, then one has proven the existence

[41] Reformulating their matching model as a demand system, Galichon, Kominers and Weber (2016) also use the properties of the excess demand function to provide a proof of existence and uniqueness with a more general heterogeneity structure.

of equilibrium. It is therefore sufficient to show that one can construct a $\mathbf{\Gamma}(\boldsymbol{\lambda}) = \psi \cdot \mathbf{ED}(\boldsymbol{\lambda}) + \boldsymbol{\lambda}$ such that

1. $\psi \cdot \mathbf{ED}(\boldsymbol{\lambda}) + \boldsymbol{\lambda} \in [0,1]^{I\times J}$
2. $\mathbf{\Gamma}(\boldsymbol{\lambda})$ is non-decreasing in $\boldsymbol{\lambda}$.

**Lemma 1.** *The excess demand functions are continuously differentiable with $\mathbf{ED}(\mathbf{0}_{I\times J}) \succcurlyeq \mathbf{0}$ and $\mathbf{ED}(\mathbf{1}_{I\times J}) \preccurlyeq \mathbf{0}$.*

*Proof of Lemma 1.* The continuously differentiability follows directly from the regularity conditions described above. $\mathbf{ED}(\mathbf{0}_{I\times J}) \succcurlyeq 0$ and $\mathbf{ED}(\mathbf{1}_{I\times J}) \preccurlyeq 0$ follow from our regularity conditions along with equations (A.1a)–(A.1d). □

**Lemma 2.** *For all $\langle i,j\rangle$ there exist a $\psi_{ij} > 0$ such that $0 \le \Gamma_{ij}(\boldsymbol{\lambda}) \le 1$.*

*Proof of Lemma 2.* For each $\langle i,j\rangle$, define the sets $BC^+_{ij} = \{\boldsymbol{\lambda} \in [0,1]^{I\times J} : ED_{ij}(\boldsymbol{\lambda}) > 0\}$ and $BC^-_{ij} = \{\boldsymbol{\lambda} \in [0,1]^{I\times J} : ED_{ij}(\boldsymbol{\lambda}) < 0\}$. Then define $\psi^+_{ij} = \min\{(1-\lambda_{ij})/ED_{ij}(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \in BC^+_{ij}\}$ and $\psi^-_{ij} = \min\{-\lambda_{ij}/ED_{ij}(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \in BC^-_{ij}\}$.[42] Continuity of $ED_{ij}(\boldsymbol{\lambda})$ implies that both $\psi^+_{ij}$ and $\psi^-_{ij}$ exist and are positive. Then for all $\psi_{ij} \in (0, \min\{\psi^+_{ij}, \psi^-_{ij}\})$, we have that

$$0 \le \psi_{ij} ED_{ij}(\boldsymbol{\lambda}) + \lambda_{ij} \le 1.$$

□

**Lemma 3.** *There exist a $\psi > 0$ such that $\mathbf{0}_{I\times J} \preccurlyeq \mathbf{\Gamma}(\boldsymbol{\lambda}) \preccurlyeq \mathbf{1}_{I\times J}$ and $\partial\mathbf{\Gamma}(\boldsymbol{\lambda})/\partial\lambda_{ij} \succcurlyeq \mathbf{0}_{I\times J}$.*

*Proof of Lemma 3.* Let $D_{ij} = \max_{k,l} \max_{\boldsymbol{\lambda}} \{|\partial ED_{ij}(\boldsymbol{\lambda})/\partial\lambda_{kl}| : \boldsymbol{\lambda} \in [0,1]^{I\times J}\}$. Then for each $\langle i,j\rangle$ and for all $\psi_{ij} \in (0, 1/D_{ij})$

$$\begin{aligned}
\frac{\partial[\psi_{ij} ED_{ij}(\boldsymbol{\lambda}) + \lambda_{ij}]}{\partial\lambda_{ij}} &= \psi_{ij}\frac{\partial ED_{ij}(\boldsymbol{\lambda})}{\partial\lambda_{ij}} + 1 \ge -\psi_{ij}D_{ij} + 1 > 0,\\
\frac{\partial[\psi_{ij} ED_{ij}(\boldsymbol{\lambda}) + \lambda_{ij}]}{\partial\lambda_{kj}} &= \psi_{ij}\frac{\partial ED_{ij}(\boldsymbol{\lambda})}{\partial\lambda_{kj}} \ge 0 \text{ for } k \ne i,\\
\frac{\partial[\psi_{ij} ED_{ij}(\boldsymbol{\lambda}) + \lambda_{ij}]}{\partial\lambda_{il}} &= \psi_{ij}\frac{\partial ED_{ij}(\boldsymbol{\lambda})}{\partial\lambda_{il}} \ge 0 \text{ for } l \ne j,
\end{aligned}$$

[42] Although $BC^+_{ij}$ and $BC^-_{ij}$ are not compact the minimum still exist over these sets because as we approach the "open part" of the set, the objective goes to $+\infty$.

which follows from equations (A.1a)–(A.1d). Let

$$\overline{\psi} = \min\{\min\{\psi_{11}^{+}, \psi_{11}^{-}\}, \ldots, \min\{\psi_{IJ}^{+}, \psi_{IJ}^{-}\}, 1/2D_{11}, \ldots, 1/2D_{IJ}\}.$$

Now choose any $\psi \in (0, \overline{\psi})$ and define $\mathbf{\Gamma}(\boldsymbol{\lambda}) = \psi \cdot \mathbf{ED}(\boldsymbol{\lambda}) + \boldsymbol{\lambda}$. We now have $\mathbf{\Gamma} : [0,1]^{I\times J} \to [0,1]^{I\times J}$ with $\partial\mathbf{\Gamma}(\boldsymbol{\lambda})/\partial\lambda_{ij} \succcurlyeq \mathbf{0}_{I\times J}$ for all pairs $\langle i, j\rangle$. $\square$

Thus, from Lemma 3 Tarski's conditions are satisfied and an equilibrium exists.

## A.3 Uniqueness of equilibrium

Uniqueness follows from the differentiability of $\mathbf{\Gamma}(\boldsymbol{\lambda})$. Lemma 4 proves that if $\mathbf{\Gamma}(\boldsymbol{\lambda})$ is differentiable almost everywhere then $\mathbf{\Gamma}(\boldsymbol{\lambda})$ is a contraction on $[0,1]^{I\times J}$ and by the contraction mapping theorem there exists a unique fixed $\boldsymbol{\lambda}^* \in [0,1]^{I\times J}$ such that $\mathbf{\Gamma}(\boldsymbol{\lambda}^*) = \boldsymbol{\lambda}^*$. However, $\boldsymbol{\lambda}^* = \psi \cdot \mathbf{ED}(\boldsymbol{\lambda}^*) + \boldsymbol{\lambda}^*$ iff $\mathbf{ED}(\boldsymbol{\lambda}^*) = 0$, therefore $\boldsymbol{\lambda}^*$ is also the unique equilibrium to our model.

**Lemma 4.** *Under the regularity conditions, there is a unique equilibrium.*

*Proof of Lemma 4.* For notational ease let $\Gamma_{ij}(\boldsymbol{\lambda})$ be defined as

$$\Gamma_{ij}(\boldsymbol{\lambda}) = \psi \cdot ED_{ij}(\boldsymbol{\lambda}) + \lambda_{ij}.$$

From the proof of Lemma 3 we know that

$$0 \geq \frac{\partial\Gamma_{ij}(\boldsymbol{\lambda})}{\partial\lambda_{ij}} = \psi \cdot \frac{\partial ED_{ij}(\boldsymbol{\lambda})}{\partial\lambda_{ij}} + 1 < 1,$$

since $\psi > 0$ and from equation (A.1a) we have that $\partial ED_{ij}(\boldsymbol{\lambda})/\partial\lambda_{ij} < 0$. Moreover, by construction

$$\frac{1}{2} \geq \frac{\partial\Gamma_{ij}(\boldsymbol{\lambda})}{\partial\lambda_{ik}} \geq 0.$$

And the IIA property in equation (A.1d) implies that

$$\frac{\partial\Gamma_{ij}(\boldsymbol{\lambda})}{\partial\lambda_{kl}} = 0.$$

Therefore, $\mathbf{\Gamma}$ is a contraction since by the mean value theorem,

$$\left|\Gamma_{ij}(\boldsymbol{\lambda}) - \Gamma_{ij}(\boldsymbol{\lambda}')\right| \leq \left\|\nabla\Gamma_{ij}(\widetilde{\boldsymbol{\lambda}})\right\| \left\|\boldsymbol{\lambda} - \boldsymbol{\lambda}'\right\| < \beta \left\|\boldsymbol{\lambda} - \boldsymbol{\lambda}'\right\|.$$

Where $\|\cdot\|$ is the sup norm, $\widetilde{\boldsymbol{\lambda}}$ is a point on the line between $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}'$, and $\beta$ is a number less than 1 such that the absolute values of the derivatives of $\boldsymbol{\Gamma}$ are less than $\beta$. This result implies that

$$\left\|\boldsymbol{\Gamma}(\boldsymbol{\lambda}) - \boldsymbol{\Gamma}(\boldsymbol{\lambda}')\right\| \leq \beta \left\|\boldsymbol{\lambda} - \boldsymbol{\lambda}'\right\|.$$

□

# B Proof of Proposition 3

In this Appendix we derive the contribution of the marital shocks within each match to the social-welfare function. We proceed in two steps. First, we characterize the *distribution* of martial preference shocks within a particular match, recognizing the non-random selection into a given pairing. Second, given this distribution, we obtain the adjustment term using our specification of the utility transformation function.

Consider the first step. For brevity of notation, here we let $U_j$ denote the expected utility of a given individual from choice/spousal type $j$. Associated with each alternative $j$ is an extreme value error $\theta_j$ that has scale parameter $\sigma_\theta$. We now characterize the distribution of $\theta_j$ conditional on $j$ being chosen. Letting $p_j = (\sum_k \exp[(U_k - U_j)/\sigma_\theta])^{-1}$ denote the associated conditional choice probability. it follows that

$$\begin{aligned}
\Pr[\theta_j < x | j = \arg\max_k U_k + \theta_k] &= \frac{1}{\sigma_\theta p_j} \int_{-\infty}^{x} \prod_{k \neq j} \exp\left(-e^{-\frac{\theta_j + U_j - U_k}{\sigma_\theta}}\right) \exp\left(-e^{-\frac{\theta_j}{\sigma_\theta}}\right) e^{-\frac{\theta_j}{\sigma_\theta}} \, \mathrm{d}\theta_j \\
&= \frac{1}{\sigma_\theta p_j} \int_{-\infty}^{x} \exp\left(-e^{-\frac{\theta_j}{\sigma_\theta}} \sum_k e^{-\frac{U_j - U_k}{\sigma_\theta}}\right) e^{-\frac{\theta_j}{\sigma_\theta}} \, \mathrm{d}\theta_j \\
&= \frac{1}{\sigma_\theta p_j} \int_{-\infty}^{x} \exp\left(-e^{-\frac{\theta_j}{\sigma_\theta}} p_j^{-1}\right) e^{-\frac{\theta_j}{\sigma_\theta}} \, \mathrm{d}\theta_j \\
&= \exp\left(-e^{-\frac{\theta_j + \sigma_\theta \log p_j}{\sigma_\theta}}\right).
\end{aligned}$$

Hence, the distribution of the idiosyncratic payoff conditional on $j$ being optimal is extreme value with the scale parameter $\sigma_\theta$ and the shifted location parameter $-\sigma_\theta \log p_j$.

**Marital payoff adjustment term:** $\delta < 0$

Now consider the second step when $\delta < 0$. Using the utility transformation function (equation (15)) and letting $\mathbf{Z}_j$ denote the entire vector of post-marriage realizations in choice $j$ (wages, preference shocks, demographics), it follows that the contribution to social-welfare of an individual in this marital pairing may be written in the form

$$\int_{\theta_j}\int_{\mathbf{Z}_j} \Upsilon[v_j(\mathbf{Z}_j)+\theta_j]\,\mathrm{d}G_j(\mathbf{Z}_j)\,\mathrm{d}H_j(\theta_j) = \int_{\theta_j}\exp(\delta\theta_j)\,\mathrm{d}H_j(\theta_j)\int_{\mathbf{Z}_j}\frac{\exp[\delta v(\mathbf{Z}_j)]}{\delta}\,\mathrm{d}G_j(\mathbf{Z}_j) - \frac{1}{\delta},$$

where we have suppressed the dependence on the tax system $T$.

We now complete our proof in the $\delta < 0$ case by providing an analytic characterization of the integral term over the idiosyncratic marital payoff. Using the result that $\theta_j|_{j=\arg\max_k U_k+\theta_k} \sim EV(-\sigma_\theta \log p_j, \sigma_\theta)$ from above, we have

$$\begin{aligned}
\int_{\theta_j}\exp(\delta\theta_j)\,\mathrm{d}H_j(\theta_j) &= \frac{1}{\sigma_\theta}\int_{\theta_j}\exp(\delta\theta_j)\exp(-[\theta_j+\sigma_\theta\log p_j]/\sigma_\theta)e^{-\exp(-[\theta_j+\sigma_\theta\log p_j]/\sigma_\theta)}\,\mathrm{d}\theta_j \\
&= \exp(-\delta\sigma_\theta\log p_j)\int_0^\infty t^{-\delta\sigma_\theta}\exp(-t)\,\mathrm{d}t \\
&= p_j^{-\delta\sigma_\theta}\Gamma(1-\delta\sigma_\theta).
\end{aligned}$$

The second equality performs the change of variable $t = \exp(-[\theta_j + \sigma_\theta \log p_j]/\sigma_\theta)$, and the third equality uses the definition of the Gamma function. Since we are considering cases where $\delta < 0$, this integral will converge.

**Marital payoff adjustment term:** $\delta = 0$

The proof when $\delta = 0$ follows similarly. Here the contribution to social-welfare of a given individual in a given marital pairing is simply given by

$$\begin{aligned}
\int_{\theta_j}\int_{\mathbf{Z}_j} \Upsilon[v_j(\mathbf{Z}_j)+\theta_j]\,\mathrm{d}G_j(\mathbf{Z}_j)\,\mathrm{d}H_j(\theta_j) &= \int_{\theta_j}\theta_j\,\mathrm{d}H_j(\theta_j) + \int_{\mathbf{Z}_j} v(\mathbf{Z}_j)\,\mathrm{d}G_j(\mathbf{Z}_j) \\
&= \gamma - \sigma_\theta\log p_j + \int_{\mathbf{Z}_j} v(\mathbf{Z}_j)\,\mathrm{d}G_j(\mathbf{Z}_j),
\end{aligned}$$

with the second equality using the above result for the distribution of marital shocks within a match and then just applying the well-known result for the expected value of the extreme value distribution with a non-zero location parameter.

# C Identification

## C.1 Proof of Proposition 2

Consider a given market $k \leq K$. From the conditional choice probabilities (equations (7) and (8)) and imposing market clearing $\mu^d_{ij}(T, \boldsymbol{\lambda}^i) = \mu^j_{ij}(T, \boldsymbol{\lambda}^j) = \mu_{ij}(T, \boldsymbol{\lambda})$ we have that

$$\ln \mu_{ij}(T, \boldsymbol{\lambda}) - \ln \mu_{i0}(T, \boldsymbol{\lambda}^i) = \frac{U^i_{ij}(T, \lambda_{ij}) - U^i_{i0}(T)}{\sigma_\theta}, \tag{C.1a}$$

$$\ln \mu_{ij}(T, \boldsymbol{\lambda}) - \ln \mu_{0j}(T, \boldsymbol{\lambda}^j) = \frac{U^j_{ij}(T, \lambda_{ij}) - U^j_{0j}(T)}{\sigma_\theta}. \tag{C.1b}$$

The left-hand side of equations (C.1a) and (C.1b) are obtained from the empirical marriage matching function and is therefore identified. Now consider variation in this object as we vary population vectors. Importantly, variation in population vectors has no impact on the value of the single state and only affects the value in marriage through its influence on the Pareto weight $\lambda_{ij}$. That is, such variation serves as a *distribution factor* (see Bourguignon, Browning and Chiappori, 2009). From a marginal perturbation in, e.g., $m_i$ we obtain[43]

$$\frac{\partial \left[\ln \mu_{ij}(T, \boldsymbol{\lambda}) - \ln \mu_{i0}(T, \boldsymbol{\lambda}^i)\right]}{\partial m_i} = \frac{1}{\sigma_\theta} \frac{\partial U^i_{ij}(T, \lambda_{ij})}{\partial \lambda_{ij}} \frac{\partial \lambda_{ij}}{\partial m_i}, \tag{C.2a}$$

$$\frac{\partial \left[\ln \mu_{ij}(T, \boldsymbol{\lambda}) - \ln \mu_{0j}(T, \boldsymbol{\lambda}^j)\right]}{\partial m_i} = \frac{1}{\sigma_\theta} \frac{\partial U^j_{ij}(T, \lambda_{ij})}{\partial \lambda_{ij}} \frac{\partial \lambda_{ij}}{\partial m_i}. \tag{C.2b}$$

Taking the ratio of the partial derivatives in equations (C.2a) and (C.2b) we define

$$\pi_{ij} = \frac{\partial U^i_{ij}(T, \lambda_{ij}) / \partial \lambda_{ij}}{\partial U^j_{ij}(T, \lambda_{ij}) / \partial \lambda_{ij}}.$$

We proceed by combining the definition of $z_{ij}$ with our envelope condition result (equation (5)) which requires that $(1 - \lambda_{ij}) \cdot \partial U^i_{ij}(T, \lambda_{ij}) / \partial \lambda_{ij} + \lambda_{ij} \cdot \partial U^j_{ij}(T, \lambda_{ij}) / \partial \lambda_{ij} = 0$. It immediately follows that $\lambda_{ij} = \pi_{ij} / (\pi_{ij} - 1)$, which establishes identification.

[43] In practice as we increase, e.g., the number of college educated males, we will necessarily be decreasing the number of males with some college or high school education. This does not change the logic of what follows.

## C.2 Identification of utility and home production functions

The identification proof will proceed in two steps. First, we demonstrate identification of the time allocation problem for single individuals. Second, we show how we use the household time allocation patterns to identify the home production technology for married couples. The following assumptions are used in the proof of identification in this section. While some of them are easily relaxed, for reasons of clarity and ease of exposition, and because they relate directly to the empirical and optimal design analysis, these assumptions are maintained here. We also only consider identification of the model without the fixed cost of labour force participation, as it adds nothing to the analysis.

**Assumption ID-1.** *The state-specific errors, $\epsilon_{\mathbf{a}^i}$ are distributed Type-I extreme value with location parameter zero and an unknown scale parameter, $\sigma_\epsilon$.*

**Assumption ID-2.** *The systematic utility function is additively separable in leisure, $\ell^i$, private consumption, $q^i$, and home goods, $Q^i$. That is*

$$u^i(\ell^i, q^i, Q^i, \mathbf{X}^i) = u^i_q(q^i, \mathbf{X}^i) + u^i_\ell(\ell^i, \mathbf{X}^i) + u^i_Q(Q^i, \mathbf{X}^i).$$

**Assumption ID-3.** *There is a known private consumption level $\widehat{q}$ such that $\partial u^i_q(\widehat{q}, \mathbf{X}^i)/\partial q = 1$.*

**Assumption ID-4.** *$u^i_Q(Q^i, \mathbf{X}^i)$ is monotonically increasing in $Q$, i.e. $\partial u^i_Q(Q^i, \mathbf{X}^i)/\partial Q > 0$.*

**Assumption ID-5.** *There exist an element of $\mathbf{X}^i$, $\mathbf{X}^i_r$, such that $\mathbf{X}^i_r$ affects $\zeta_{i0}(\mathbf{X}^i)$ but not $u^i_Q(Q^i, \mathbf{X})$. Also there exists an $\mathbf{X}^i_*$ such that $\zeta_{i0}(\mathbf{X}^i_*) = 1$.*

**Assumption ID-6.** *The support of $Q$ is the same for both single individuals and married couples.*

**Assumption ID-7.** *Conditional on work hours $h^i_w$, the tax schedule $T$ is differentiable in earnings, with $\partial T(w^i h^i_w, y^i; \mathbf{X}^i)/\partial w h_w \neq 1$.*

**Assumption ID-8.** *The utility of function of the private good, $u^i_q(q^i, \mathbf{X}^i)$, is monotonically increasing and quasi-concave in $q^i$.*

### C.2.1 Step 1: The identification using the singles problem

Consider the problem of a single type-$i$ male. Let $\mathcal{A}^i = \{1, \ldots, \overline{A}^i\}$ be an index representation set of time allocation alternatives, with $\widehat{u}^i(a)$ denoting the systematic part of utility associated with alternative $a \in \mathcal{A}^i$ (where the dependence on conditioning variables is suppressed for notational compactness). Without loss of generality, let $a = 1$ be the choice where the individual does not work and has the lowest level of home hours. Under Assumption ID-1, well-known results imply that the following holds:

$$\log\left[\frac{\mathbb{P}(a)}{\mathbb{P}(1)}\right] = \frac{\widehat{u}^i(a) - \widehat{u}^i(1)}{\sigma_\epsilon}, \tag{C.3}$$

where the conditional choice probabilities $\mathbb{P}(\cdot)$ should be understood as being conditional on $[y^i, w^i, \mathbf{X}^i, T]$. Taking the partial derivative of equation (C.3) with respect to $w^i$ and using Assumption ID-2 yields

$$\frac{\partial \log\left[\mathbb{P}(a)/\mathbb{P}(1)\right]}{\partial w} = \frac{1}{\sigma_\epsilon} \cdot \frac{\partial u^i_q(q^i(a); \mathbf{X}^i)}{\partial q} \cdot \left[1 - \frac{\partial T(w^i h^i_w(a), y^i; \mathbf{X}^i)}{\partial w h_w}\right] \cdot h^i_w(a), \tag{C.4}$$

where $q^i(a)$ and $h^i_w(a)$ are the respective private consumption and market work hours associated with the allocation $a$. The conditional choice probabilities and the marginal tax rates are known and hence, given Assumptions ID-3 and ID-7, the scale coefficient for the state-specific errors $\sigma_\epsilon$ is identified. Hence, the marginal utility of private consumption is identified. Integrating equation (C.4) and combining with equation (C.3) implies that the sum $u^i_\ell(\ell^i; \mathbf{X}^i) + u^i_Q(Q^i; \mathbf{X}^i)$ is identified up to a normalizing constant. Then for each level of feasible home hours, both $u^i_\ell(\ell^i; \mathbf{X}^i)$ and $u^i_Q(Q^i; \mathbf{X}^i)$ are identified by varying the level of market hours and fixing either home time or leisure. Under Assumption ID-5, the home efficiency parameter $\zeta_{i0}(\mathbf{X}^i)$ is identified by comparing $u^i_Q(Q^i(a); \mathbf{X}^i)$ across different values of $\mathbf{X}^i$.

### C.2.2 Step 2: Identification of marriage home production function.

In Step 1 we show that the subutilities are identified up to a normalizing constant. Without loss of generality, we set the location normalization to be zero in what follows. Consider a $\langle i, j\rangle$ household with the time allocation set $\mathcal{A}_{ij} = \{1, \ldots, \overline{A}\}$, $\overline{A} = \overline{A}^i \times \overline{A}^j$, and let $\widehat{u}^{ij}(a) = (1 - \lambda_{ij}) \times \widehat{u}^i(a) + \lambda_{ij} \times \widehat{u}^j(a)$ denote the systematic part of household utility associated with $a \in \mathcal{A}_{ij}$. Let $\varepsilon^{ij}_a = (1 - \lambda_{ij}) \times \epsilon^i_a + \lambda_{ij} \times \epsilon^j_a$, and define $\mathcal{G}^a_{ij}(\cdot)$ to be the joint cumulative distribution function of $[\varepsilon^{ij}_a - \varepsilon^{ij}_1, \ldots, \varepsilon^{ij}_a - \varepsilon^{ij}_{a-1}, \varepsilon^{ij}_{a+1} - \varepsilon^{ij}_a, \ldots, \varepsilon^{ij}_a - \varepsilon^{ij}_A]$. For each $a \in \{1, \ldots, \overline{A} - 1\}$, define

$$\mathbb{P}(a) = \mathcal{Q}_j(\widehat{\mathbf{u}}^{ij}) \equiv \mathcal{G}^a_{ij}(\widehat{u}^{ij}_a - \widehat{u}^{ij}_1, \ldots, \widehat{u}^{ij}_a - \widehat{u}^{ij}_{a-1}, \widehat{u}^{ij}_a - \widehat{u}^{ij}_{a+1}, \ldots, \widehat{u}^{ij}_a - \widehat{u}^{ij}_A),$$

with $\widehat{\mathbf{u}}^{ij} = [\widehat{u}^{ij}_1 - \widehat{u}^{ij}_{\overline{A}}, \ldots, \widehat{u}^{ij}_{\overline{A}-1} - \widehat{u}^{ij}_{\overline{A}}]^\intercal$ defining the $(\overline{A}-1)$ vector of utility differences, and let $\mathcal{Q}(\widehat{\mathbf{u}}^{ij}) = [\mathcal{Q}_1(\widehat{\mathbf{u}}^{ij}), \ldots, \mathcal{Q}_{A-1}(\widehat{\mathbf{u}}^{ij})]^\intercal$ define a $(\overline{A}-1)$ dimensional vector function. Then, by Proposition 1 of Hotz and Miller (1993), the inverse of $\mathcal{Q}(\widehat{\mathbf{u}}^{ij})$ exists.[44] Given that the distribution of $\epsilon$ is known and $\lambda_{ij}$ is identified, the inverse of $\mathcal{Q}(\widehat{\mathbf{u}}^{ij})$ is known. Hence, the vector $\widehat{\mathbf{u}}^{ij} = \mathcal{Q}^{-1}(\mathbb{P}(1), \ldots, \mathbb{P}(\overline{A}-1))$ is identified. Define

$$\begin{aligned}\Delta_{ij}(a) = \widehat{\mathbf{u}}^{ij}_{[a]} - (1-\lambda_{ij}) \times [u^i_\ell(\ell^i(a^i); \mathbf{X}^i) + u^i_q((1 - s_{ij}(a;\lambda_{ij})) \cdot q(a); \mathbf{X}^i)] \\ - \lambda_{ij} \times [u^j_\ell(\ell^j(a^j); \mathbf{X}^j) + u^i_q(s_{ij}(a;\lambda_{ij}) \cdot q(a); \mathbf{X}^j)].\end{aligned}$$

The arguments from Step 1 imply that $u^i_q(q^i; \mathbf{X}^i)$ and $u^j_q(q^j; \mathbf{X}^j)$ are known. From Proposition 2 we have that $\lambda_{ij}$ are identified. These, together with Assumption ID-2 and Assumption ID-4, imply that $s_{ij}(a;\lambda_{ij})$ is also known. Thus, identification of $\Delta_{ij}(a)$ follows. Finally, the definition of $\widehat{u}^{ij}(a)$ and Assumption ID-2 imply

$$\Delta_{ij}(a) = (1-\lambda_{ij}) \times u^i_Q(\widetilde{Q}_{ij}(h^i_Q(a), h^j_Q(a); \mathbf{X}), \mathbf{X}^i) + \lambda_{ij} \times u^j_Q(\widetilde{Q}_{ij}(h^i_Q(a), h^j_Q(a); \mathbf{X}), \mathbf{X}^j).$$

The subutility function of the public good does not depend on $\mathbf{w}$. Therefore, once we observe different values of these two variables, $u^j_Q(\widetilde{Q}_{ij}(h^i_Q(a), h^j_Q(a); \mathbf{X}), \mathbf{X}^j)$ and $u^i_Q(\widetilde{Q}_{ij}(h^i_Q(a), h^j_Q(a); \mathbf{X}), \mathbf{X}^i)$ are identified. Finally, under Assumption ID-4 the inverse of $u^i_Q$ and $u^j_Q$ exist and hence $\widetilde{Q}_{ij}(h^i_Q(a^i), h^j_Q(a^j); \mathbf{X})$ is identified.

# References


**Aguiar, Mark, and Erik Hurst.** 2007. "Measuring Trends in Leisure: The Allocation of Time Over Five Decades." *Quarterly Journal of Economics*, 122(3): 969–1006.

**Aguiar, Mark, Erik Hurst, and Loukas Karabarbounis.** 2012. "Recent Developments in the Economics of Time Use." *Annual Review of Economics*, 4(1): 373–397.

**Alesina, Alberto, Andrea Ichino, and Loukas Karabarbounis.** 2011. "Gender-Based Taxation and the Division of Family Chores." *American Economic Journal: Economic Policy*, 3(2): 1–40.


[44] Notice that $\varepsilon^{ij}_a$ is not i.i.d. However, independence is not a required condition of the Hotz and Miller (1993) proposition.

**Alm, James, and Leslie A. Whittington.** 1999. "For Love or Money? The Impact of Income Taxes on Marriage." *Economica*, 66(263): 297–316.

**Altonji, Joseph G., and Lewis M. Segal.** 1996. "Small-Sample Bias in GMM Estimation of Covariance Structures." *Journal of Business & Economic Statistics*, 14(3): 353–366.

**Apps, Patricia F., and Ray Rees.** 1988. "Taxation and The Household." *Journal of Public Economics*, 35(3): 355–369.

**Apps, Patricia F., and Ray Rees.** 1999. "On the Taxation of Trade Within and Between Households." *Journal of Public Economics*, 73(2): 241–263.

**Apps, Patricia F., and Ray Rees.** 2007. "The Taxation of Couples." Working Paper.

**Blau, Francine D., and Lawrence M. Kahn.** 2007. "Changes in the Labor Supply Behavior of Married Women: 1980–2000." *Journal of Labor Economics*, 25(3): 393–438.

**Blundell, Richard, and Andrew Shephard.** 2012. "Employment, Hours of Work and the Optimal Taxation of Low-Income Families." *Review of Economic Studies*, 79(2): 481–510.

**Blundell, Richard, Pierre-André Chiappori, and Costas Meghir.** 2005. "Collective Labor Supply with Children." *Journal of Political Economy*, 113(6): 1277–1306.

**Boskin, Michael J., and Eytan Sheshinski.** 1983. "Optimal Tax Treatment of the Family: Married Couples." *Journal of Public Economics*, 20(3): 281–297.

**Bourguignon, François, Martin Browning, and Pierre-André Chiappori.** 2009. "Efficient Intra-Household Allocations and Distribution Factors: Implications and Identification." *Review of Economic Studies*, 76(2): 503–528.

**Brett, Craig.** 2007. "Optimal Nonlinear Taxes for Families." *International Tax and Public Finance*, 14(3): 225–261.

**Brewer, Mike, Emmanuel Saez, and Andrew Shephard.** 2010. "Means Testing and Tax Rates on Earnings." In *Dimensions of Tax Design: The Mirrlees Review*, ed. James Mirrlees, Stuart Adam, Timothy Besley, Richard Blundell, Stephen Bond, Robert Chote, Malcolm Grammie, Paul Johnson, Gareth Myles and James Poterba, 90–173. Oxford University Press for Institute for Fiscal Studies.

**Browning, Martin, François Bourguignon, Valérie Lechene, and Pierre-André Chiappori.** 1994. "Income and Outcomes: A Structural Model of Intrahousehold Allocation." *Journal of Political Economy*, 102(6): 1067–1096.

**Browning, Martin, Pierre-André Chiappori, and Arthur Lewbel.** 2013. "Estimating Consumption Economies of Scale, Adult Equivalence Scales, and Household Bargaining Power." *Review of Economic Studies*.

**Chade, Hector, and Gustavo Ventura.** 2002. "Taxes and Marriage: A Two-Sided Search Analysis." *International Economic Review*, 43(3): 965–985.

**Chade, Hector, and Gustavo Ventura.** 2005. "Income Taxation and Marital Decisions." *Review of Economic Dynamics*, 8(3): 565–599.

**Chiappori, Pierre-André.** 1988. "Rational Household Labor Supply." *Econometrica*, 56(1): pp. 63–90.

**Chiappori, Pierre-André.** 1992. "Collective Labor Supply and Welfare." *Journal of Political Economy*, 100(3): pp. 437–467.

**Chiappori, Pierre-André, Bernard Salanié, and Yoram Weiss.** 2014. "Partner Choice and the Marital College Premium." Working Paper.

**Chiappori, Pierre-André, Monica Costa Dias, and Costas Meghir.** 2015. "The Marriage Market, Labor Supply and Education Choice." Working Paper.

**Chiappori, Pierre-André, Murat Iyigun, and Yoram Weiss.** 2009. "Investment in Schooling and the Marriage Market." *American Economic Review*, 99(5): 1689–1713.

**Choo, Eugene, and Aloysius Siow.** 2006. "Who Marries Whom and Why." *Journal of Political Economy*, 114(1): pp. 175–201.

**Choo, Eugene, and Shannon Seitz.** 2013. "The Collective Marriage Matching Model: Identification, Estimation, and Testing." In *Structural Econometric Models* Vol. 31 of *Advances in Econometrics*, ed. Matthew Shum and Eugene Choo, 291–336. Emerald Group Publishing Limited.

**Cogan, John F.** 1981. "Fixed Costs and Labor Supply." *Econometrica*, 49(4): pp. 945–963.

**Couprie, Hélène.** 2007. "Time Allocation within the Family: Welfare Implications of Life in a Couple." *Economic Journal*, 117(516): 287–305.

**Cremer, Helmuth, Jean-Marie Lozachmeur, and Pierre Pestieau.** 2012. "Income Taxation of Couples and the Tax Unit Choice." *Journal of Population Economics*, 25(2): 763–778.

**Das, Mitali, Whitney K. Newey, and Francis Vella.** 2003. "Nonparametric Estimation of Sample Selection Models." *Review of Economic Studies*, 70(1): 33–58.

**Del Boca, Daniela, and Christopher Flinn.** 2012. "Endogenous Household Interaction." *Journal of Econometrics*, 166(1): 49–65. Annals Issue on "Identification and Decisions", in Honor of Chuck Manski's 60th Birthday.

**Del Boca, Daniela, and Christopher J. Flinn.** 1995. "Rationalizing Child-Support Decisions." *American Economic Review*, 85(5): 1241–1262.

**Devereux, Paul J.** 2004. "Changes in Relative Wages and Family Labor Supply." *Journal of Human Resources*, 39(3): 696–722.

**Donni, Olivier.** 2003. "Collective Household Labor Supply: Nonparticipation and Income Taxation." *Journal of Public Economics*, 87(5–6): 1179–1198.

**Donni, Olivier, and Pierre-André Chiappori.** 2011. "Nonunitary Models of Household Behavior: A Survey of the Literature." In *Household Economic Behaviors*, ed. A. J. Molina, 1–40. New York: Springer New York.

**Eissa, Nada, and Hilary Williamson Hoynes.** 2000. "Tax and Transfer Policy, and Family Formation: Marriage and Cohabitation." University of California, Berkeley, mimeo.

**Feenberg, Daniel, and Elisabeth Coutts.** 1993. "An Introduction to the TAXSIM Model." *Journal of Policy Analysis and Management*, 12(1): 189–194.

**Frankel, Alexander.** 2014. "Taxation of Couples under Assortative Mating." *American Economic Journal: Economic Policy*, 6(3): 155–77.

**Frazis, Harley, and Jay Stewart.** 2007. "Where Does the Time Go? Concepts and Measurement in the American Time Use Survey." In *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches.* Vol. 67 of *Studies in Income and Wealth*, ed. Ernst R. Berndt and Charles R. Hulten, 73–97. University of Chicago Press.

**Galichon, Alfred, and Bernard Salanié.** 2015. "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models." Working Paper.

**Galichon, Alfred, Scott D. Kominers, and Simon Weber.** 2014. "An Empirical Framework for Matching with Imperfectly Transferable Utility." *Social Science Research Network Working Paper Series*.

**Galichon, Alfred, Scott D. Kominers, and Simon Weber.** 2016. "Costly Concessions: An Empirical Framework for Matching with Imperfectly Transferable Utility." *Social Science Research Network Working Paper Series*.

**Gayle, George-Levi, and Limor Golan.** 2012. "Estimating a Dynamic Adverse-Selection Model: Labour-Force Experience and the Changing Gender Earnings Gap 1968–1997." *Review of Economic Studies*, 79(1): 227–267.

**Gelber, Alexander M., and Joshua W. Mitchell.** 2011. "Taxes and Time Allocation: Evidence from Single Women and Men." *Review of Economic Studies*.

**Goussé, Marion, Nicolas Jacquemet, and Jean-Marc Robin.** 2015. "Marriage, Labor Supply, and Home Production." Working Paper.

**Guner, Nezih, Remzi Kaygusuz, and Gustavo Ventura.** 2012. "Taxation and Household Labour Supply." *The Review of Economic Studies*, 79(3): 1113–1149.

**Heim, Bradley T.** 2009. "Structural Estimation of Family Labor Supply with Taxes: Estimating a Continuous Hours Model Using a Direct Utility Specification." *Journal of Human Resources*, 44(2): 350–385.

**Hotz, V. Joseph, and Robert A. Miller.** 1993. "Conditional Choice Probabilities and the Estimation of Dynamic Models." *Review of Economic Studies*, 60(3): 497–529.

**Immervoll, Herwig, Henrik Jacobsen Kleven, Claus Thustrup Kreiner, and Nicolaj Verdelin.** 2009. "An Evaluation of the Tax-Transfer Treatment of Married Couples in European Countries." IZA Discussion Paper No. 3965.

**Immervoll, Herwig, Henrik Jacobsen Kleven, Claus Thustrup Kreiner, and Nicolaj Verdelin.** 2011. "Optimal Tax and Transfer Programs for Couples with Extensive Labor Supply Responses." *Journal of Public Economics*, 95(11–12): 1485–1500. Special Issue: International Seminar for Public Economics on Normative Tax Theory.

**Kleven, Henrik Jacobsen, Claus Thustrup Kreiner, and Emmanuel Saez.** 2007. "The Optimal Income Taxation of Couples as a Multi-dimensional Screeing Problem." CESifo Working Paper 2092.

**Kleven, Henrik Jacobsen, Claus Thustrup Kreiner, and Emmanuel Saez.** 2009. "The Optimal Income Taxation of Couples." *Econometrica*, 77(2): 537–560.

**Laroque, Guy.** 2005. "Income Maintenance and Labor Force Participation." *Econometrica*, 73(2): 341–376.

**Lewbel, Arthur, and Krishna Pendakur.** 2008. "Estimation of Collective Household Models with Engel Curves." *Journal of Econometrics*, 147(2): 350–358.

**Lise, Jeremy, and Shannon Seitz.** 2011. "Consumption Inequality and Intra-household Allocations." *Review of Economic Studies*, 78(1): 328–355.

**Lundberg, Shelly J., Robert A. Pollak, and Terence J. Wales.** 1997. "Do Husbands and Wives Pool Their Resources? Evidence from the United Kingdom Child Benefit." *Journal of Human Resources*, 32(3): 463–480.

**Matzkin, Rosa L.** 1992. "Nonparametric and Distribution-Free Estimation of the Binary Threshold Crossing and the Binary Choice Models." *Econometrica*, 60(2): 239–270.

**Matzkin, Rosa L.** 1993. "Nonparametric Identification and Estimation of Polychotomous Choice Models." *Journal of Econometrics*, 58(1): 137–168.

**McFadden, Daniel.** 1978. "Modeling the Choice of Residential Location." In *Spatial Interaction Theory and Planning Models*, ed. A. Karlqvist, L. Lundqvist, F. Snickars and J. Weibull, 75–96. North Holland.

**Meghir, Costas, and David Phillips.** 2010. "Labour Supply and Taxes." In *Dimensions of Tax Design: The Mirrlees Review*, ed. James Mirrlees, Stuart Adam, Timothy Besley, Richard Blundell, Stephen Bond, Robert Chote, Malcolm Grammie, Paul Johnson, Gareth Myles and James Poterba, 202–274. Oxford University Press for Institute for Fiscal Studies.

**Mirrlees, J. A.** 1971. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies*, 38(2): 175–208.

**Mourifié, Ismael, and Aloysius Siow.** 2014. "Cohabitation versus Marriage: Marriage Matching with Peer Effects." Working Paper.

**Piketty, Thomas, and Emmanuel Saez.** 2013. "Optimal Labor Income Taxation." In *Handbook of Public Economics* Vol. 5, ed. Alan J. Auerbach, Raj Chetty, Martin Feldstein and Emmanuel Saez, 391–474. Elsevier.

**Saez, Emmanuel.** 2002. "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses." *Quarterly Journal of Economics*, 117(3): 1039–1073.

**Su, Che-Lin, and Kenneth L. Judd.** 2012. "Constrained Optimization Approaches to Estimation of Structural Models." *Econometrica*, 80(5): 2213–2230.

**Thomas, Duncan.** 1990. "Intra-Household Resource Allocation: An Inferential Approach." *Journal of Human Resources*, 25(4): 635–664.

**Vermeulen, Frederic.** 2005. "A Collective Model for Female Labour Supply with Nonparticipation and Taxation." *Journal of Population Economics*, 19(1): 99–118.

# Supplement to "Optimal Taxation, Marriage, Home Production, and Family Labour Supply"

In these online appendices we (i) describe our empirical tax and transfer schedule implementation; (ii) describe the iterative algorithm and solution approximation methods for calculating the marriage market equilibrium; (iii) describe the set of targeted estimation moments; (iv) present additional parameter and results tables; (v) provide more details on the set of perturbation exercises; and (vi) present results on tax schedule restrictions.

# D Empirical tax and transfer schedule implementation

In this appendix we describe our implementation of the empirical tax and transfer schedules for our estimation exercise. Since some program rules will vary by U.S. state, here we are explicit in indexing the respective parameters by market.[45] Our measure of taxes includes both state and federal Earned Income Tax Credit (EITC) programmes, and we also account for the Food Stamps Program and the Temporary Assistance for Needy Families (TANF) program. It does not include other transfers and non-income taxes such as sales and excises taxes. In addition to market, the tax schedules that we calculate also vary with marital status and with children. We assume joint filing status for married couples. For singles with children we assume head of household filing status.

Consider (a married or single) household $\iota$ in market $k$, with household earnings $E_{\iota k} = \mathbf{h}_w^{\iota k} \cdot \mathbf{w}_{\iota k}$ and demographic characteristics $\mathbf{X}_{\iota k}$. As before, the demographic conditioning vector comprises marital status and children. The total net tax liability for such a household is given by $T_{\iota k} = \widetilde{T}_{\iota k} - Y_{TANF}^{\iota k} - Y_{FSP}^{\iota k}$, where $\widetilde{T}_{\iota k}$ is the (potentially negative) tax liability from income taxes and the EITC, $Y_{TANF}^{\iota k}$ and $Y_{FSP}^{\iota k}$ are the respective (non-negative) amounts of TANF and Food Stamps.

### Income taxes and EITC

Our measure of income taxes $\widetilde{T}_{\iota k}$ includes both federal and state income taxes, as well as federal and state EITC. These taxes are calculated with the National Bureau of Economic Research TAXSIM calculator, as described in Feenberg and Coutts (1993). Prior to estimation, we calculate schedules for all markets and for all family types. We assume joint

[45] Since our definition of a market is at a slightly more aggregated level than the state level, we apply the state tax rules that correspond to the most populous state within a defined market (Census Bureau-designated division).

filing status for married couples. In practice, only around 2% of married couples choose to file separate tax returns. For singles with children, we assume head-of-household filing status. Note that certain state rules may imply discontinuous changes in tax liabilities following a marginal change in earnings. To avoid the technical and computational issues that are associated with such a change we (locally) modify the tax schedule in these events.[46]

#### Food Stamp Program

Food Stamps are available to low-income households both with and without children. For the purposes of determining the entitlement amount, net household earnings are defined as

$$N_{FSP}^{\iota k} = \max\{0, E_{\iota k} + Y_{TANF}^{\iota k} - D_{FSP}[\mathbf{X}_{\iota k}]\},$$

where $Y_{TANF}^{\iota k}$ is the dollar amount of TANF benefit received by this household (see below), and $D_{FSP}[\mathbf{X}_{\iota k}]$ is the standard deduction, which may vary with household type. The dollar amount of Food Stamp entitlement is then given by

$$Y_{FSP}^{\iota k} = \max\{0, Y_{FSP}^{\max}[\mathbf{X}_{\iota k}] - \tau_{FSP} \times N_{FSP}^{\iota k}\},$$

where $Y_{FSP}^{\max}[\mathbf{X}_{\iota k}]$ is the maximum food stamp benefit amount for a household of a given size and $\tau_{FSP} = 0.3$ is the phase-out rate.[47]

#### TANF

TANF provides financial support to families with children. Given the static framework we are considering, we are not able to incorporate certain features of the TANF program, notably the time limits in benefit eligibility (see Chan, 2013). For the purposes of

[46]These discontinuities are typically small. Our modification procedure involves increasing/decreasing marginal rates in earnings tax brackets just below the discontinuity.

[47]In practice, the Food Stamp Program also has a gross-earnings and net-earnings income test. These require that earnings are below some threshold related to the federal poverty level for eligibility (see, e.g., Chan, 2013). For some families, these eligibility rules would mean that there may be a discontinuous fall in entitlement (to zero) as earnings increase. While these rules are straightforward to model, we do not incorporate them for the same reason we do not allow discontinuities in the combined income tax/EITC schedule. We also assume a zero excess shelter deduction in our calculations and do not consider asset tests. Incorporating asset tests (even in a dynamic model) is very challenging as there exist very specific definitions of *countable* assets that do not correspond to the usual assets measure in life-cycle models.

entitlement calculation, we define net household earnings as

$$N_{TANF}^{\iota k} = \max\{0, (1 - R_{TANF}^{k}) \times (E_{\iota k} - D_{TANF}^{k}[\mathbf{X}_{\iota k}])\},$$

where the *dollar* earnings disregard $D_{TANF}^{k}[\mathbf{X}_{\iota k}]$ varies by market and household characteristics. The market-level *percent* disregard is given by $R_{TANF}^{k}$. The dollar amount of TANF entitlement is then given by

$$Y_{TANF}^{\iota k} = \min\{Y_{TANF}^{\max}[\mathbf{X}_{\iota k}], \max\{0, r_{TANF}^{k} \times (\widehat{Y_{TANF}^{\max}}[\mathbf{X}_{\iota k}] - N_{TANF}^{\iota k})\}\}.$$

Here $Y_{TANF}^{\max}[\mathbf{X}_{\iota k}]$ defines the maximum possible TANF receipt in market $k$ for a household with characteristics $\mathbf{X}_{\iota k}$, while $\widehat{Y_{TANF}^{\max}}[\mathbf{X}_{\iota k}]$ defines what is typically referred to as the *payment standard*. The ratio $r_{TANF}^{k}$ is used in some markets to adjust the total TANF amount.[48]

# E Marriage market numerical algorithm

In this appendix we describe the iterative algorithm and the solution approximation method that we use to calculate the market clearing vector of Pareto weights. The algorithm is based on that presented in Galichon, Kominers and Weber (2014, 2016). We first note that using the conditional choice probabilities from equation (7) we are able to write the quasi-demand equation of type-$i$ men for type-$j$ spouses as:

$$\sigma_\theta \times \left[\ln \mu_{ij}^{d}(T, \boldsymbol{\lambda}^{i}) - \ln \mu_{i0}^{d}(T, \boldsymbol{\lambda}^{i})\right] = U_{ij}^{i}(T, \lambda_{ij}) - U_{i0}^{i}(T). \tag{E.1}$$

Similarly, the conditional choice probabilities for females from equation (8) allows us to express the quasi-supply equation of type-$j$ women to the $\langle i, j\rangle$ submarket as:

$$\sigma_\theta \times \left[\ln \mu_{ij}^{s}(T, \boldsymbol{\lambda}^{j}) - \ln \mu_{0j}^{s}(T, \boldsymbol{\lambda}^{j})\right] = U_{ij}^{j}(T, \lambda_{ij}) - U_{0j}^{j}(T). \tag{E.2}$$

The algorithm proceeds as follows:

1. Provide an initial guess of the measure of both single males $0 < \mu_{i0}^{d} < m_i$ for $i = 1, \dots I$, and single females $0 < \mu_{0j}^{s} < f_j$ for $j = 1, \dots, J$.

[48]For reasons identical to those discussed in the case of Food Stamps, we do not consider the similar gross and net income eligibility rules that exist for TANF, as well as the corresponding asset tests. See Footnote 47. We also do not consider eligibility time limits.

2. Taking the difference of the quasi-demand (equation (E.1)) and the quasi-supply (equation (E.2)) functions for each $\langle i, j \rangle$ submarriage market and imposing the market clearing condition $\mu^d_{ij}(T, \boldsymbol{\lambda}^i) = \mu^s_{ij}(T, \boldsymbol{\lambda}^j)$ we obtain

$$\sigma_\theta \times \left[\ln \mu^s_{0j} - \ln \mu^d_{i0}\right] = U^i_{ij}(T, \lambda_{ij}) - U^i_{i0}(T) - \left[U^j_{ij}(T, \lambda_{ij}) - U^j_{0j}(T)\right], \tag{E.3}$$

which given the single measures $\mu^d_{i0}$ and $\mu^s_{0j}$ (and the tax schedule $T$) are only a function of the Pareto weight for that submarriage-market $\lambda_{ij}$. Given our assumptions on the utility functions, there exists a unique solution to equation (E.3). This step therefore requires solving for the root of $I \times J$ univariate equations.

3. From *Step 2*, we have a matrix of Pareto weights $\boldsymbol{\lambda}$ given the single measures $\mu^d_{i0}$ and $\mu^s_{0j}$ from *Step 1*. These measures can be updated by calculating the conditional choice probabilities (equation (7) and equation (8)). The algorithm returns to Step 2 and repeats until the vector of single measures for both males and females has converged.

In practice, we are able to implement this algorithm by first evaluating the expected utilities $U^i_{ij}(T, \lambda)$ and $U^j_{ij}(T, \lambda)$ for each marital match combination $\langle i, j \rangle$ on a fixed grid of Pareto weights $\lambda \in \boldsymbol{\lambda}^{\text{grid}}$ with $\inf[\boldsymbol{\lambda}^{\text{grid}}] \gtrapprox 0$ and $\sup[\boldsymbol{\lambda}^{\text{grid}}] \lessapprox 1$. We may then replace $U^i_{ij}(T, \lambda)$ and $U^j_{ij}(T, \lambda)$ with an approximating parametric function so that no expected values are actually evaluated within the iterative algorithm.

Note that calculating the expected values within a match are (by many orders of magnitude) the most computationally expensive part of the algorithm. An implication of this is that if there are $K$ markets, and each market $k \leq K$ only differs by the population vectors $\boldsymbol{\mathcal{M}}_k$ and $\boldsymbol{\mathcal{F}}_k$ and/or the demographic transition functions, the computational cost in obtaining the equilibrium for all $K$ markets is *approximately independent* of the number of markets $K$ considered. This property is true given that the initial evaluation of expected values on $\boldsymbol{\lambda}^{\text{grid}}$ is independent of the market in this case.[49]

# F Estimation moments

In this appendix we describe the set of targeted estimation moments. Recall that there are nine markets ($K = 9$) and three education groups/types for both men ($I = 3$) and

[49] We do not exploit this property in our application, as we also have market variation in taxes and transfers.

women ($J = 3$) in our empirical application. The first set of moments relate to the marriage market. Within each market, we describe the number of single men and women by own education, and married households by joint education ($K \times [I + J + I \times J]$ moments). The second set of moments describe labour supply patterns. By market, gender, marital status and own education, we describe mean conditional work hours and employment rates ($K \times 4 \times [I + J]$ moments); aggregating over markets, we describe the fraction of individuals in non-employment/part-time/full-time status by gender, marital status, the presence of children, and own/joint education level (for singles/couples respectively) ($6 \times [I + J] + 12 \times I \times J$ moments); the mean and standard deviation of conditional work hours is described by gender, marital status, and own education, while mean conditional hours for married men and women are also described by joint education levels ($8 \times [I + J] + 2 \times I \times J$ moments). The third set of moments describe accepted wages and conditional earnings. The mean and standard deviation of accepted log-wages are described by gender, marital status, and own education ($4 \times [I + J]$ moments); the mean and standard deviation of conditional earnings are described by the same conditioning variables (again, $4 \times [I + J]$ moments). The fourth set of moments relate to unconditional home time. Similar to labour supply, we describe the fraction of individuals with low/medium/high unconditional home hours by gender, marital status, the presence of children, and own/joint education level (for singles/couples respectively) ($6 \times [I + J] + 12 \times I \times J$ moments); the mean and standard deviation of unconditional home hours is described by gender, marital status, and own education, while mean unconditional home hours for married men and women are also described by joint education levels ($8 \times [I + J] + 2 \times I \times J$ moments). In total, we have 765 moments.

# G Parameter results tables

In Table 5 we present the estimates from our model, together with the accompanying standard errors. These results are obtained from the estimation procedure described in Section 4.

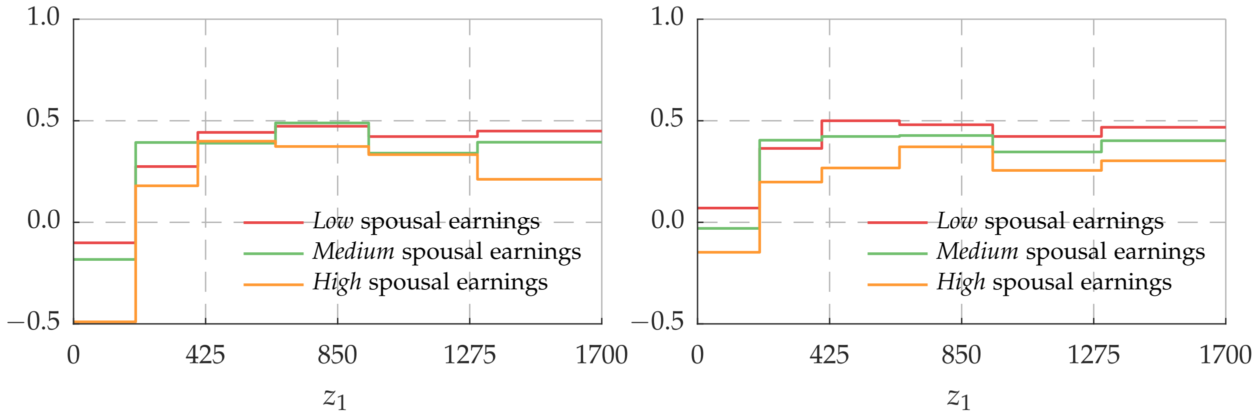
# H Perturbation experiments

This appendix provides further details and results for the perturbation comparative static exercises discussed in Section 5.3.

Table 5: Parameter estimates

| | Estimate | Standard error |
|---|---|---|
| *Log-wage offers:* | | |
| Male, high school and below: mean | 2.562 | 0.003 |
| Male, high school and below: s.d. | 0.461 | 0.002 |
| Male, some college: mean | 2.802 | 0.003 |
| Male, some college: s.d. | 0.446 | 0.003 |
| Male, college: mean | 3.253 | 0.007 |
| Male, college: s.d. | 0.575 | 0.005 |
| Female, high school and below: mean | 2.128 | 0.004 |
| Female, high school and below: s.d. | 0.551 | 0.003 |
| Female, some college: mean | 2.442 | 0.004 |
| Female, some college: s.d. | 0.515 | 0.003 |
| Female, college: mean | 2.915 | 0.003 |
| Female, college: s.d. | 0.503 | 0.002 |
| *Preference parameters:* | | |
| Leisure scale | 0.992 | 0.064 |
| Home good scale | 0.204 | 0.039 |
| Leisure curvature, $\sigma_\ell$ | 0.722 | 0.070 |
| Home good curvature, $\sigma_Q$ | -0.186 | 0.038 |
| Fixed costs (kids) | 87.351 | 1.716 |
| Marital shock, s.d. | 0.111 | 0.004 |
| State specific error, s.d. | 0.286 | 0.005 |
| *Home production technology:* | | |
| Male production share | 0.050 | 0.006 |
| Single productivity (no children),high school and below | 1.693 | 0.287 |
| Single productivity (no children),some college | 1.876 | 0.314 |
| Single productivity (no children),college | 5.983 | 0.954 |
| Male productivity (children) | 7.087 | 1.269 |
| Female productivity (children), high school and below | 8.716 | 1.529 |
| Female productivity (children), some college | 9.573 | 1.689 |
| Female productivity (children), college | 10.654 | 1.909 |
| HH productivity (children) female, high school and below | 3.699 | 0.586 |
| HH productivity (children) female, some college | 5.278 | 0.868 |
| HH productivity (children) female, college | 3.417 | 0.553 |
| HH productivity (children) educational homogamy, high school and below | 1.771 | 0.058 |
| HH productivity (children) educational homogamy, some college | 1.112 | 0.011 |
| HH productivity (children) educational homogamy, college | 2.341 | 0.110 |

*Notes:* All parameters estimated simultaneously using a moment based estimation procedure as detailed in Section 4 from the main text. See Footnote 34 for a description of the method used to calculate standard errors. All incomes are expressed in dollars per-week in average 2006 prices.




(a) Fixed marriage market

(b) Equilibrium marriage market

Figure 7: Optimal tax schedule with fixed and equilibrium marriage market under $\delta = 0$. Figure shows marginal tax rates under alternative assumptions on the marriage market, conditional on spousal earnings, $z_2$. See Footnote 40 for a definition of *low*, *medium*, and *high* spousal earnings.

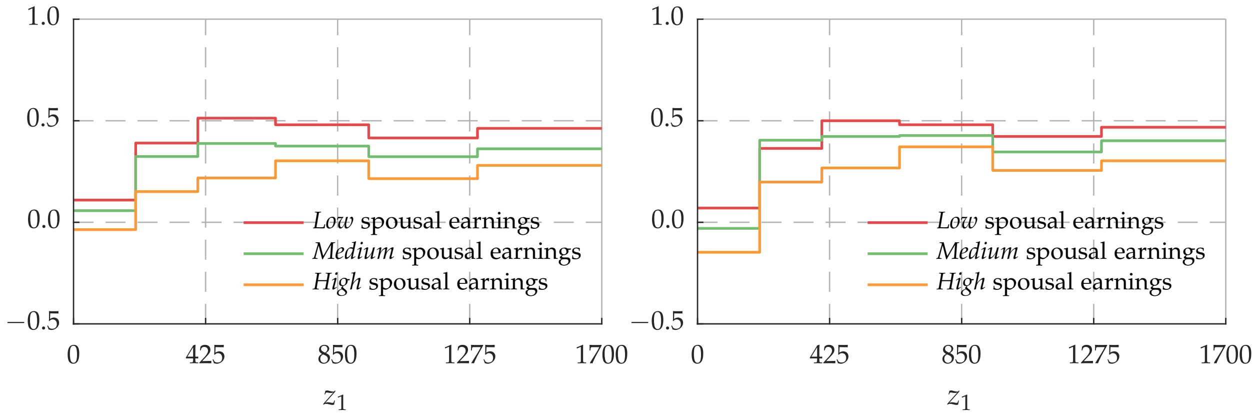
## H.1 Fixed marriage market

In examining the importance of the marriage market, we again consider the maximization of a social-welfare function $\mathcal{W}(T)$ subject to a revenue constraint $\mathcal{R}(T)$. In contrast to our main analysis, we do not allow elements of the marriage market matching function $\mu_{ij}$, the household Pareto weights $\lambda_{ij}$, or the distribution of idiosyncratic payoffs within a match $H_{ij}(\cdot)$ to depend upon the tax system $T$. Instead, when solving our constrained welfare maximization problem, we fix their values at those obtained under our general specification with a marriage market equilibrium. Figure 7 presents the implied marginal rate schedule, which is calculated under different levels of spousal earnings. In order to provide a direct visual comparison, the schedule obtained with a marriage market equilibrium is shown alongside.

To the extent that there is a non-zero marriage market elasticity, the tax schedule calculated here implies an imbalance in the marriage market. At the solution to this relaxed problem, we calculate the supply and demand in each submarriage market using equations (7) and (8). We then quantify this imbalance by constructing the normalized excess demand measure

$$\frac{\sum_{i,j} |\mu^d_{ij} - \mu^s_{ij}|}{\sum_{i,j} \mu^*_{ij}},$$

where $\mu^*_{ij}$ is the measure of type $\langle i, j \rangle$ matches obtained from the original problem with a




(a) Reduced assortative mating ($\varrho = \sigma_\theta$)

(b) Baseline assortative mating ($\varrho = 0$)

Figure 8: Optimal tax schedule with reduced assortative mating under $\delta = 0$. The figure shows marginal tax rates when assortative mating is reduced through inclusion of an additive utility cost for educationally homogamous marriages. Marginal tax rates are shown conditional on spousal earnings, $z_2$. See Footnote 40 for a definition of *low*, *medium*, and *high* spousal earnings.

marriage market equilibrium. In the $\delta = 0$ case, we obtain a value of 7%; when $\delta = -1$, we obtain a value of 8%.

## H.2 Assortative mating

The degree of assortative mating is endogenously changed by augmenting the individual utility function to include the additive payoff $\overline{\theta}_{ij}$. In what follows, we set $\overline{\theta}_{ij} = -\varrho \times \mathbb{1}[i = j]$ so that a value $\varrho > 0$ reduces the utility in educationally homogamous marriages but does not have a direct impact on the time allocation problem. In Figure 8 we show the impact that this modification has on the structure of marginal rates when $\delta = 0$. In the illustrations here, we set $\varrho = \sigma_\theta$ so the reduction in expected utility is equal in value to a one-standard-deviation idiosyncratic marital payoff. We can see that this reduction in correlation among types *increases* the degree of negative tax jointness. Table 6 presents the associated marriage market matching function.

## H.3 Labour supply responsiveness

We consider the impact of changing the responsiveness of labour supply to taxes. The most direct parameter relevant for the labour supply elasticity is the curvature parameter $\sigma_\ell$. Relative to the estimated value $\widehat{\sigma_\ell} = 0.722$, we illustrate the impact of both lower

Table 6: Marriage matching function with reduced assortative mating

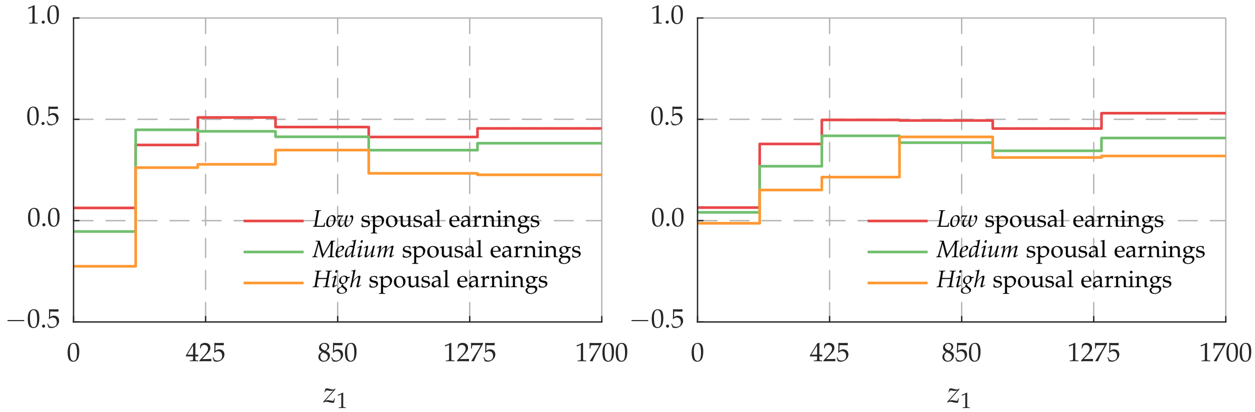
| | | | Women | | |
|---|---|---|---|---|---|
| | | | High school and below | Some college | College and above |
| *(a). Reduced assortative mating ($\varrho = \sigma_\theta$)* | | | | | |
| | | – | 0.122 | 0.077 | 0.117 |
| Men | High school and below | 0.131 | 0.101 | 0.101 | 0.056 |
| | Some college | 0.093 | 0.066 | 0.060 | 0.069 |
| | College and above | 0.091 | 0.040 | 0.080 | 0.111 |
| *(b). Baseline assortative mating ($\varrho = 0$)* | | | | | |
| | | – | 0.111 | 0.086 | 0.100 |
| Men | High school and below | 0.132 | 0.154 | 0.068 | 0.035 |
| | Some college | 0.092 | 0.040 | 0.113 | 0.043 |
| | College and above | 0.074 | 0.023 | 0.051 | 0.174 |

*Notes:* Table shows marriage matching function when assortative mating is reduced through inclusion of an additive utility cost for educationally homogamous marriages.

($\sigma_\ell = 0.6$) and higher ($\sigma_\ell = 0.9$) values and resolve for the optimal tax schedule, holding all other parameters fixed. Note that this comparative static exercise changes the responsiveness of all family members. As we increase $\sigma_\ell$, we decrease the labour supply elasticity and obtain generally higher marginal tax rates (for both singles and married couples). Results in the $\delta = 0$ case are provided in Figure 9.

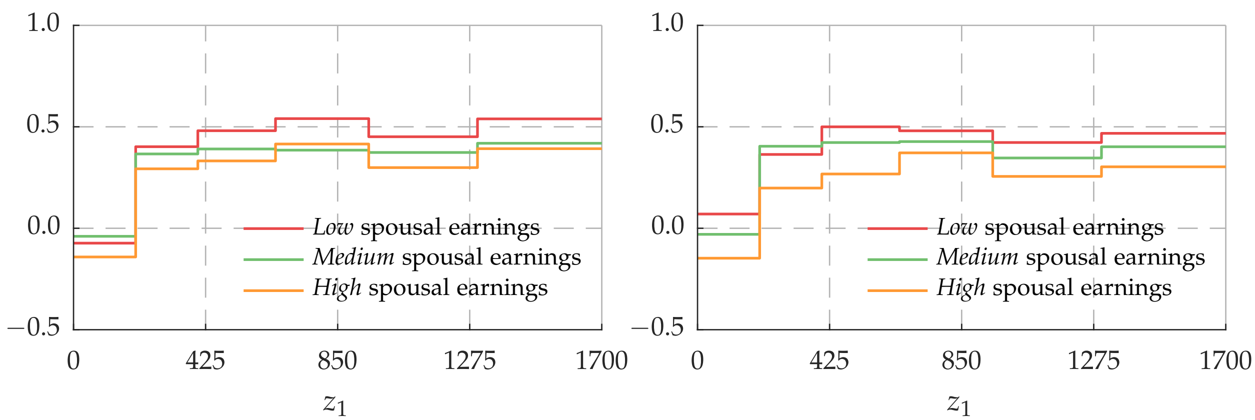
## H.4 Home time efficiency

We consider the role of home production by changing the home time efficiency parameter $\zeta$. The less productive is home time, the worse off are single individuals and individuals in couples when earnings are relatively low and so the greater is the redistributive motive. We consider the impact of scaling all home efficiency parameters by the factor $\zeta_\varrho > 0$. In Figure 10 we illustrate (when $\delta = 0$) the impact on the marginal rate structure when home time productivity is halved, i.e., $\zeta_\varrho = 0.5$. In this case, we obtain a generally more progressive tax schedule.




(a) Higher curvature ($\sigma_\ell = 0.6$) (b) Less curvature ($\sigma_\ell = 0.9$)

Figure 9: Optimal tax schedule with alternative labour responsiveness under $\delta = 0$. The figure shows marginal tax rates when the curvature of the leisure subutility function is varied. Marginal tax rates are shown conditional on spousal earnings, $z_2$. See Footnote 40 for a definition of *low*, *medium*, and *high* spousal earnings.




(a) Reduced home time efficiency ($\zeta_\varrho = 0.5$) (b) Baseline home time efficiency ($\zeta_\varrho = 1$)

Figure 10: Optimal tax schedule with reduced home time efficiency under $\delta = 0$. The figure shows marginal tax rates when the home time efficiency is uniformly reduced. Marginal tax rates are shown conditional on spousal earnings, $z_2$. See Footnote 40 for a definition of *low*, *medium*, and *high* spousal earnings.

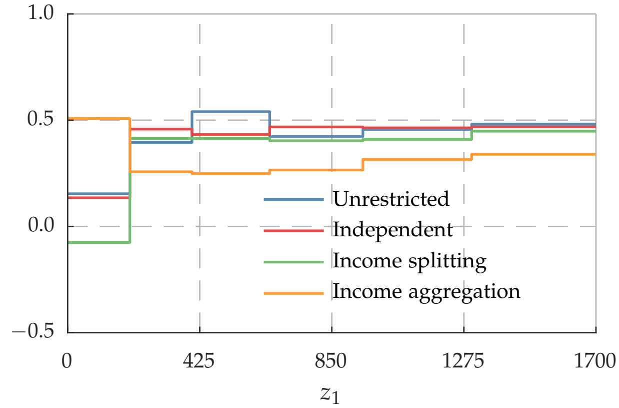
Table 7: Marriage matching function

| | | | Women | | |
|---|---|---|---|---|---|
| | | | High school and below | Some college | College and above |
| *(a). Unrestricted* | | | | | |
| | | – | 0.111 | 0.086 | 0.100 |
| Men | High school and below | 0.132 | 0.154 | 0.068 | 0.035 |
| | Some college | 0.092 | 0.040 | 0.113 | 0.043 |
| | College and above | 0.074 | 0.023 | 0.051 | 0.174 |
| *(b). Independent* | | | | | |
| | | – | 0.084 | 0.094 | 0.145 |
| Men | High school and below | 0.115 | 0.179 | 0.067 | 0.029 |
| | Some college | 0.103 | 0.042 | 0.108 | 0.035 |
| | College and above | 0.104 | 0.025 | 0.049 | 0.144 |
| *(c). Income splitting* | | | | | |
| | | – | 0.067 | 0.091 | 0.151 |
| Men | High school and below | 0.106 | 0.191 | 0.065 | 0.027 |
| | Some college | 0.100 | 0.044 | 0.111 | 0.033 |
| | College and above | 0.103 | 0.027 | 0.051 | 0.141 |
| *(d). Income aggregation* | | | | | |
| | | – | 0.148 | 0.130 | 0.165 |
| Men | High school and below | 0.183 | 0.127 | 0.053 | 0.026 |
| | Some college | 0.134 | 0.033 | 0.090 | 0.031 |
| | College and above | 0.126 | 0.021 | 0.045 | 0.130 |

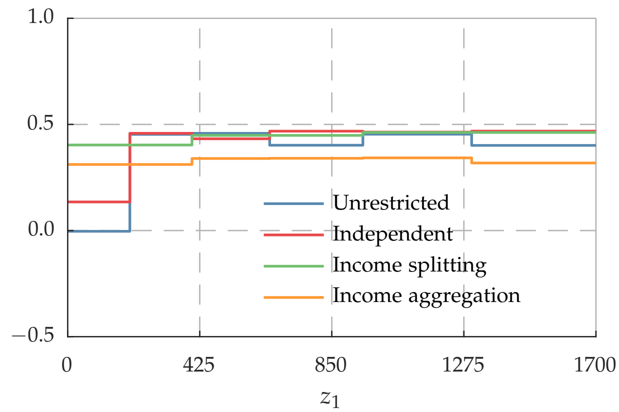
*Notes:* Table shows marriage matching function under alternative tax schedule specifications. *Unrestricted* corresponds to the schedule described in Section 5.1. *Independent*, *Income splitting*, and *Income aggregation* respectively refer to independent individual taxation, and joint taxation with income splitting and aggregation. See Section 5.4 for details.

# I Restricted tax schedules

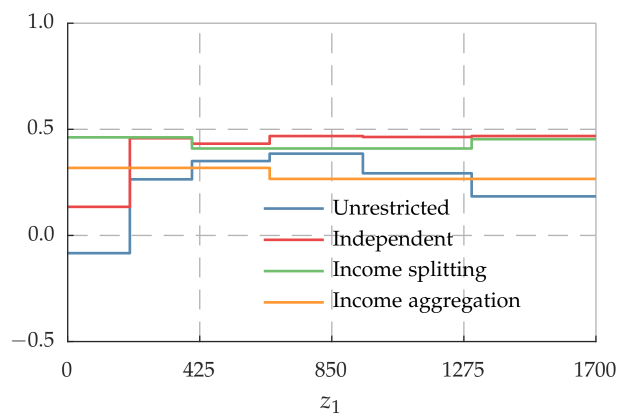
In this Appendix we describe results when the form of jointness in the tax schedule is restricted. As described in the Section 5.4 from the main text, we consider i) individual taxation; ii) joint taxation with income splitting; iii) joint taxation with income aggregation. In Figure 11 we present the implied marginal rate structure in the $\delta = 0$ case. We have constructed these conditional on alternative spousal earnings levels, but here omit confidence bands for clarity of presentation. The associated marriage matching functions are presented in Table 7.




(a) *Low* spousal earnings




(b) *Medium* spousal earnings




(c) *High* spousal earnings

Figure 11: Optimal tax schedule with restricted forms of jointness under $\delta = 0$. Figure shows marginal tax rates under alternative assumptions on the form of the tax schedule, conditional on spousal earnings, $z_2$. See Footnote 40 for a definition of *low*, *medium*, and *high* spousal earnings. *Unrestricted* corresponds to the tax schedule described in Section 5.1. *Independent*, *Income splitting*, and *Income aggregation* respectively refer to independent individual taxation, and joint taxation with income splitting and aggregation (Section 5.4).

## Supplement References

**Chan, Marc K.** 2013. "A Dynamic Model of Welfare Reform." *Econometrica*, 81(3): 941–1001.

**Feenberg, Daniel, and Elisabeth Coutts.** 1993. "An Introduction to the TAXSIM Model." *Journal of Policy Analysis and Management*, 12(1): 189–194.

**Galichon, Alfred, Scott D. Kominers, and Simon Weber.** 2014. "An Empirical Framework for Matching with Imperfectly Transferable Utility." *Social Science Research Network Working Paper Series*.

**Galichon, Alfred, Scott D. Kominers, and Simon Weber.** 2016. "Costly Concessions: An Empirical Framework for Matching with Imperfectly Transferable Utility." *Social Science Research Network Working Paper Series*.