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

George-Levi Gayle[*]   Andrew Shephard[†]

April 15, 2016

Abstract

This paper develops an empirical approach to optimal income taxation design within an equilibrium collective marriage market model. 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 for married couples we allow for an arbitrary form of tax jointness.

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] In

---

[*]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.

contrast, the majority of OECD countries have individual income taxation where each individual is taxed separately based on his/her 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, whilst 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). 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.[4] When the particapation 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.

By taking the married unit as given the optimal nonlinear tax system analyses in Kleven, Kreiner and Saez (2007, 2009) ignores the distortionary effect of couple's taxation

[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 a form of 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 both marriage and time allocation outcomes.

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

[4]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. See also Brett (2007), Cremer, Lozachmeur and Pestieau (2012), Frankel (2014), and Immervoll et al. (2011).

on who gets married and to whom they get married. It is also ill equipped to analyze the distortionary effect of taxation on the intra-household allocation of resources. Moreover, the primary/secondary earner asymmetry ignores the potential role of the tax system in inducing specialization in couples.[5]

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 behavioral 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 within household allocation of resources? 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 a 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, and which features both intensive and extensive labour supply margins, home production time, and both public and private good consumption. As in Galichon, Kominers and Weber (2014) we allow for utilities to be imperfectly transferable across spouses. In this environment we provide sufficient conditions for the existence and uniqueness of equilibrium, demonstrate identification, and describe computationally efficient ways to both solve and estimate the model.

Using data from the American Community Survey and the American Time Use Sur-

---

[5]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).

[6]Other papers that integrate a collective time allocation model within an empirical marriage-matching model include Chiappori, Costa Dias and Meghir (2015) who consider an equilibrium model of education and marriage with labour supply and consumption in a transferable utility model, and Choo and Seitz (2013).

vey we structurally estimate our equilibrium model, exploiting variation across markets in terms of both tax and transfer policies, and population vectors. We then use our estimated model directly to examine problems related to the optimal design of the tax system, while acknowledging that taxes may distort labour supply and time allocation decisions, as well as marriage market outcomes, and the within household decision process. Our taxation design problem is based on an *individualistic* social welfare function, with inequality both within and across households adversely affecting social welfare. 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.

The remainder of the paper proceeds as follows. In Section 2 we describe our equilibrium model of marriage, consumption, and time allocation. Section 3 introduces the analytical framework that we use to study taxation design within our equilibrium collective model. Section 4.2 describes our microeconometric specification, while Section 4 discusses the data and estimation procedure, as well as detailing our main estimation results. In Section 5 we present our main optimal taxation design results, both allowing for very general forms for the tax schedule, as well as forms which restrict the form of jointness. Finally, Section 6 concludes.

2   A model of marriage and time allocation

We present an empirical model of marriage-matching and intrahousehold 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 market unless such a distinction is important and proceed to describe the problem for a given market. In such a market there are I types of men and J types of women. The population vector of men is given by $\mathcal{M}$, whose element m_i denotes the measure of type i males. Similarly, the population vector of women is given by $\mathcal{F}$, 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 occurs 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.[7] 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 position. 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.

2.1 Time allocation problem

We first describe the decision 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 one), and a non-marketable public good produced with home time.

2.1.1 Time allocation problem: single individuals

Consider a single male of type i. His total time endowment is L_0 and he chooses the time allocation vector $\mathbf{a}^i = (\ell^i, h^i_w, h^i_Q)$ comprising hours of leisure ℓ^i, market work time h^i_w, and home production time h^i_Q, 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 $L_0 = \ell^i + h^i_w + h^i_Q$.[8] Associated

---

[7]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.

[8]A description of the choice set used in our empirical implementation, together with the parameterisation of the utility function and the complete stochastic structure, are provided in Section 4.2.

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,

$$q^i = y^i + w^i h_w^i - T(w^i h_w^i, y^i; \mathbf{X}^i) - FC(h_w^i; \mathbf{X}^i), \tag{2a}$$
$$Q^i = A_{i0}(\mathbf{X}^i) \cdot h_Q^i. \tag{2b}$$

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 $A_{i0}(\mathbf{X}^i)$ may depend upon both own type, and demographic characteristics.

The solution to this constrained utility maximization problem is described by the incentive compatible time allocation vector $\mathbf{a}_{i0}^{i\star}(w^i, y^i, \mathbf{X}^i, \epsilon^i; T)$, 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 $v_{i0}^i(w^i, y^i, \mathbf{X}^i, \epsilon^i; T)$. The decision problem for single women of type j is described similarly and yields the indirect utility function $v_{0j}^j(w^j, y^j, \mathbf{X}^j, \epsilon^j; T)$.

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). An important economic benefit of marriage is given by the publicness of some consumption. We assume that the home produced good (that is produced by combining male and female home time) is public within the household, which both members may consume equally.[9] Consider an $\langle i, j \rangle$ couple and let λ_{ij} denote the Pareto weight on female utility in such

---

[9]Since individuals do not internalise the externality associated with the home good when making their marriage decision, the equilibrium of the marriage market without taxes is not efficient.

a union.[10] The household chooses a time allocation vector for each adult, as well as determining 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 their own time allocation, and not on the time allocation of their 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:

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

subject to:

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

$$q^j = s_{ij} \cdot q, \quad (4b)$$

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

In turn, this set of equality constraints describe i) that *total* family consumption of the private good equals family net income with the tax schedule here allowed to depend very generally on the labour market earnings of both spouses,[11] less any fixed work-related costs; ii) the wife receives the endogenous consumption share $0 \leq s_{ij} \leq 1$ of the private good; 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 family demographics.

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

---

[10] 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. The parameterization of the utility function in our empirical implementation will imply a very close connection between the Pareto weight and the endogenous consumption share (see Section 4.2).

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

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 position, 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:

$$U_{i0}^i(T) = \mathbb{E}[v_{i0}^i(w^i, y^i, \mathbf{X}^i, \epsilon^i; T)],$$
$$U_{0j}^j(T) = \mathbb{E}[v_{0j}^j(w^j, y^j, \mathbf{X}^j, \epsilon^j; T)],$$

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 system T and a candidate Pareto weight λ_{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 λ_{ij}, while the expected value of his wife is strictly increasing in λ_{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 male g of type i 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 which 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 σ_θ. 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:

$$p_{ij}^i(T,\boldsymbol{\lambda}^i) = \Pr[U_{ij}^i(T,\lambda_{ij}) + \theta_{ij}^i > \max\{U_{ij}^i(T,\lambda_{ih}) + \theta_{ih}^i, U_{i0}^i(T) + \theta_{i0}^i\} \quad \forall h \neq j]$$
$$= \frac{\mu_{ij}^d(T,\boldsymbol{\lambda}^i)}{m_i} = \frac{\exp[U_{ij}^i(T,\lambda_{ij})/\sigma_\theta]}{\exp[U_{i0}^i(T)/\sigma_\theta] + \sum_{h=1}^{J} \exp[U_{ih}^i(T,\lambda_{ih})/\sigma_\theta]}, \tag{7}$$

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 $\mu_{ij}^d(T,\boldsymbol{\lambda}^i)$ is the measure of type i males who "demand" type j females (the conditional choice probabilities $p_{ij}^i(T,\boldsymbol{\lambda}^i)$ multiplied by the measure of men of type i). Women also receive idiosyncratic payoffs associated with the different marital states 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:

$$p_{ij}^{j}(T, \boldsymbol{\lambda}^{j}) = \frac{\mu_{ij}^{s}(T, \boldsymbol{\lambda}^{j})}{f_{j}} = \frac{\exp[U_{ij}^{j}(T, \lambda_{ij})/\sigma_{\theta}]}{\exp[U_{0j}^{j}(T)/\sigma_{\theta}] + \sum_{g=1}^{I}\exp[U_{gj}^{j}(T, \lambda_{gj})/\sigma_{\theta}]}, \quad (8)$$

where $\boldsymbol{\lambda}^{j} = [\lambda_{1j}, \dots, \lambda_{Ij}]^{\mathsf{T}}$ is the $I \times 1$ vector of Pareto weights for a type j female, and $\mu_{ij}^{s}(T, \boldsymbol{\lambda}^{j})$ 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 $I \times J$ matrix of Pareto weights $\boldsymbol{\lambda} = [\boldsymbol{\lambda}^{1}, \boldsymbol{\lambda}^{2}, \dots, \boldsymbol{\lambda}^{J}]$ 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,

$$\mu_{ij}(T, \boldsymbol{\lambda}) = \mu_{ij}^{d}(T, \boldsymbol{\lambda}^{i}) = \mu_{ij}^{s}(T, \boldsymbol{\lambda}^{j}) \quad \forall i = 1, \dots, I, j = 1, \dots, J. \quad (9)$$

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 their 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.[12] 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:*

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

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

Proof. See Appendix A. □

In Appendix C we describe the numerical algorithm that we apply when solving for an equilibrium of the marriage given any tax and transfer system T. We also note

---

[12]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).

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.

In what follows, we use $G_{i0}^i(w^i, \mathbf{X}^i, \epsilon^i)$ and $G_{0j}^j(w^j, \mathbf{X}^j, \epsilon^j)$ 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 $G_{ij}(\mathbf{w}, \mathbf{X}, \epsilon)$. It is also necessary to describe the endogenous distribution of idiosyncratic payoffs for individuals within a given marital position. These differ from the unconditional $EV(0, \sigma_\theta)$ distribution for the population as a whole, because individuals non-randomly select into different marital positions on the basis of these. They are therefore also a function of tax policy. We let $H_{i0}^i(\theta^i; T)$ denote the cumulative distribution function of these payoffs amongst single type i males and similarly define $H_{0j}^j(\theta^j; T)$ for single type j females. Amongst married men and women in an $\langle i, j \rangle$ match these are given by $H_{ij}^i(\theta^i; T)$ and $H_{ij}^j(\theta^j; T)$ 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

$Y(\cdot)$. The social welfare function is defined as the sum of these transformed utilities:

$$\mathcal{W}(T) = \underbrace{\sum_i \mu_{i0}(T) \int Y\left[v^i_{i0}(w^i, y^i, \boldsymbol{X}^i, \boldsymbol{\epsilon}^i; T) + \theta^i\right] dG^i_{i0}(w^i, \boldsymbol{X}^i, \boldsymbol{\epsilon}^i) \, dH^i_{i0}(\theta^i; T)}_{\text{single men}}$$

$$+ \underbrace{\sum_j \mu_{0j}(T) \int Y\left[v^j_{0j}(w^j, y^j, \boldsymbol{X}^j, \boldsymbol{\epsilon}^j; T) + \theta^j\right] dG^j_{0j}(w^j, \boldsymbol{X}^j, \boldsymbol{\epsilon}^j) \, dH^j_{0j}(\theta^j; T)}_{\text{single women}}$$

$$+ \underbrace{\sum_{i,j} \mu_{ij}(T) \int Y\left[v^i_{ij}(\mathbf{w}, \mathbf{y}, \mathbf{X}, \boldsymbol{\epsilon}; T, \lambda_{ij}(T)) + \theta^i\right] dG_{ij}(\mathbf{w}, \mathbf{X}, \boldsymbol{\epsilon}) \, dH^i_{ij}(\theta^i; T)}_{\text{married men}}$$

$$+ \underbrace{\sum_{i,j} \mu_{ij}(T) \int Y\left[v^j_{ij}(\mathbf{w}, \mathbf{y}, \mathbf{X}, \boldsymbol{\epsilon}; T, \lambda_{ij}(T)) + \theta^j\right] dG_{ij}(\mathbf{w}, \mathbf{X}, \boldsymbol{\epsilon}) \, dH^j_{ij}(\theta^j; T)}_{\text{married women}}. \quad (11)$$

The maximization of $\mathcal{W}(T)$ is subject to a number of constraints. Firstly 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 is our formulation of the problem through the inclusion of the indirect utility functions. Second, individual's optimally select into different marital positions 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 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, there is the requirement that an exogenously determined revenue amount $\overline{T}$ is

raised, as given by the revenue constraint:

$$\mathcal{R}(T) = \underbrace{\sum_i \mu_{i0}(T) \int R_{i0}^i(w^i, y^i, \mathbf{X}^i, \epsilon^i; T) \, dG_{i0}^i(w^i, \mathbf{X}^i, \epsilon^i)}_{\text{revenue from single men}}$$

$$+ \underbrace{\sum_j \mu_{0j}(T) \int R_{0j}^j(w^j, y^j, \epsilon^j, \mathbf{X}^j; T) \, dG_{0j}^j(w^j, \mathbf{X}^j, \epsilon^j)}_{\text{revenue from single women}}$$

$$+ \underbrace{\sum_{i,j} \mu_{ij}(T) \int R_{ij}(\mathbf{w}, \mathbf{y}, \mathbf{X}, \epsilon; T, \lambda_{ij}(T)) \, dG_{ij}(\mathbf{w}, \mathbf{X}, \epsilon)}_{\text{revenue from married couples}} \geq \overline{T}, \qquad (12)$$

where $R_{i0}^i(w^i, y^i, \mathbf{X}^i, \epsilon^i)$ describes the amount of revenue raised from a single type i male given w^i, y^i, $\mathbf{X}^i$, and ϵ^i, and that his time allocation decision is optimal given T. We similarly define $R_{0j}^j(w^j, y^j, \epsilon^j, \mathbf{X}^j; T)$ 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.

Note that taxes affect the problem in the following way. First, they have a direct effect on welfare and revenue holding behavior and the marriage market fixed. Second, there is a behavioral effect such that time allocations within any given match change affecting both welfare and revenue. Third, there is a marriage market effect that changes who marries with 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. Firstly, we use data from the 2006 *American Community Survey* (ACS). This provides us with information on education, marital patterns, demographics, incomes, and labour supply. We supplement this with pooled *American Time Use Survey* (ATUS) data, which we use to construct broad measure of home time for individuals sampled in the pre-recession period (2002–2007).[13] Following

---

[13]The American Time Use Survey (ATUS) is a nationally representative cross-sectional time-use survey launched in 2003 by the U.S. Bureau of Labor Statistics (BLS). The ATUS interviews a randomly selected

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.[14] Home production hours contains core home production, activities related to home ownership, obtaining goods and services, and care of other adults. It also contains childcare hours that measure all time spent by the individual caring for, educating, or playing with their children.[15]

For both men and women we define three broad education groups for our analysis: high school and below, less than four year college ("some college"), four year college and above. These constitute the individual *types* for the purposes of marriage market matching.[16] Our sample is restricted to single individuals who are aged 25–35 (inclusive). For married couples, we include all individuals where the reference person householder (as defined by the Census Bureau) belongs to this same age band.[17]

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.[18] 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), including both federal and state tax rates (including the Earned Income Tax Credit), and supplemented with

---

individual age 15 and older from a subset of the households that have completed their eighth and final interview for the Current Population Survey (CPS), the U.S. monthly labor force survey.

[14]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.

[15]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 that is completed by survey respondents. We adjust the sample weights so we continue to have a uniform distribution of week days following our sample selection. This is a common adjustment. See, e.g. Frazis and Stewart (2007).

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

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

[18]We do not use a finer level of market disaggregation due to sample size and computational considerations. There are nine Census Bereau divisions: New England (Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont); Mid-Atlantic (New Jersey, New York, and Pennsylvania); East North Central (Illinois, Indiana, Michigan, Ohio, and Wisconsin); West North Central (Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, and South Dakota); South Atlantic (Delaware, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, Washington D.C., and West Virginia); East South Central (Alabama, Kentucky, Mississippi, and Tennessee); West South Central (Arkansas, Louisiana, Oklahoma, and Texas); West Mountain (Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, and Wyoming); Pacific (Alaska, California, Hawaii, Oregon, and Washington).

detailed program rules for major welfare programs. The inclusion of 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 Appendix E.

4.2 Empirical specification

In Section 4.5 we will see that there are important differences in labour supply and the time spent on home production activities for men and women. Moreover, there are large differences between those who are single and those who are married (and with whom). 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 public good consumption. Preferences are unchanged by the state of 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, the share 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 λ_{ij}.[19] In the case that $\sigma_q = 1$ this 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

---

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

presence of dependent children in the household.[20] For singles, the demographic transition 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 the productivity of home time.

The home 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_Q^i and h_Q^j, as well as a match specific term $A_{ij}(\mathbf{X})$ that determines the overall efficiency of production within an $\langle i, j \rangle$ match for a household with demographics characteristics $\mathbf{X}$. That is:

$$\widetilde{Q}_{ij}(h_Q^i, h_Q^j; \mathbf{X}) = A_{ij}(\mathbf{X}) \times (h_Q^i)^\alpha (h_Q^j)^{1-\alpha}, \tag{14}$$

In our application we are restrictive in the specification of the match specific component. For all married households without children we set $A_{ij}(\mathbf{X}) = 1$. For married households with children we restrict the match specific component in an $\langle i, j \rangle$ match to be of the form $\widetilde{A}_j \times B_j^{\mathbb{1}[i=j]}$. The parameter B_j captures potential complementarity in the home production technology for similar individuals.[21]

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.

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 and

---

[20]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. 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.

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

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,[22] there are a total of 30 discrete time allocation alternatives for individuals, and $30^2 = 900$ discrete alternatives for couples.

The state specific errors associated with the discrete individual time allocation decisions $\epsilon_{\mathbf{a}^i}$ and $\epsilon_{\mathbf{a}^j}$ are assumed Type-I extreme value, with the scale parameter σ_ϵ. 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.[23]

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 what is 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 the scale of the state specific errors.

---

[22]Restricting the choice set in this way is of little consequence as we only remove alternatives that are never practically chosen.

[23]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 $(\mathbf{w}, \mathbf{y}, \mathbf{X}, \epsilon)$ 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 $\tau_\epsilon > 0$ which varies with all possible joint discrete time alternatives. The probability of a given joint time allocation conditional on is given by the usual conditional Logit form. As the smoothing parameter $\tau_\epsilon \to 0$ we get the unsmoothed simulated frequency.

4.3.1 Identifying the wife's Pareto weight from marriage

[To be completed]

Chiappori and Ekeland (2009), Blundell, Chiappori and Meghir (2005) and Browning, Chiappori and Lewbel (2013).

[To be completed]

Below we will show the identification of the wife's Pareto weights using equilibrium restrictions the from the marriage market under very mild conditions.

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

Proof. See Appendix F. □

The strongest assumption for the identification of the wife's Pareto weights is that the idiosyncratic marital payoffs is distributed Type-I extreme value with an unknown (albeit common) scale parameter. However, in the ITU matching models this assumption is needed for existence and uniqueness of equilibrium.[24]

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 home production time. Then, under the maintained assumption that while the budget set and home technology may differ by marital status individual preferences do not, we use our knowledge of the Pareto weight (identified from marriage market equilibrium conditions and market variation in population vectors), together with information on the time allocation behaviour of married couples to identify the home production technology for individuals in couples.[25] These

---

[24]See Galichon and Salanié (2015) for discussion of the importance the Type-I extreme value assumption for identification of matching models and its generalisation.

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

objects imply identification of the expected values in any given marriage market position. 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 F.

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 complete description of all the moments used is provided in Appendix G.

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 position and in each market. That is, equation 9 holds.[26] In practice, this equilibrium constraints procedure 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 Appendix C. Letting β denote the $B \times 1$ parameter vector, our estimation problem may be formally described as:

$$[\widehat{\beta}, \lambda(\widehat{\beta})] = \underset{\beta, \lambda}{\arg \min} \left[\mathbf{m_{sim}}(\beta, \lambda) - \mathbf{m_{data}} \right]^\mathsf{T} \mathbf{W} \left[\mathbf{m_{sim}}(\beta, \lambda) - \mathbf{m_{data}} \right]$$

$$\text{s.t. } \mu_{ijk}^d(\beta, \lambda_k^i) = \mu_{ijk}^s(\beta, \lambda_k^j) \quad \forall i = 1, \dots, I, j = 1, \dots, J, k = 1, \dots, K,$$

where λ defines the stacked $(I \times J \times K)$ vector of Pareto weights in all markets, $\mathbf{m_{data}}$ is the $M \times 1$ vector of empirical moments, $\mathbf{m_{sim}}(\beta, \lambda)$ is the model moment vector given β and an arbitrary (i.e. potentially non-equilibrium) vector of Pareto weights λ. Finally, we have that $\mathbf{W}$ 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

---

[26]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.

of the diagonal variance-covariance matrix of the empirical moments.[27] The solution to this estimation problem is such that $\widehat{\boldsymbol{\lambda}} = \boldsymbol{\lambda}(\widehat{\boldsymbol{\beta}})$.[28]

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. A more complete characterization (including the parameter estimates) is provided in Appendix H.

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. 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 time for both men and women in different marriage market positions, 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 is observed in the case of women.

---

[27]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 $r \gg 1$.

[28]The variance matrix of our estimator is given by:

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

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
Men		–	0.140 [0.129]	0.128 [0.107]	0.122 [0.087]
	High school and below	0.159 [0.133]	0.139 [0.148]	0.066 [0.065]	0.026 [0.043]
	Some college	0.113 [0.111]	0.037 [0.033]	0.087 [0.100]	0.051 [0.045]
	College and above	0.119 [0.079]	0.013 [0.020]	0.037 [0.045]	0.154 [0.178]

Notes: Table shows 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 2006 American Community Survey using sample selection as detailed in Section 4.1.
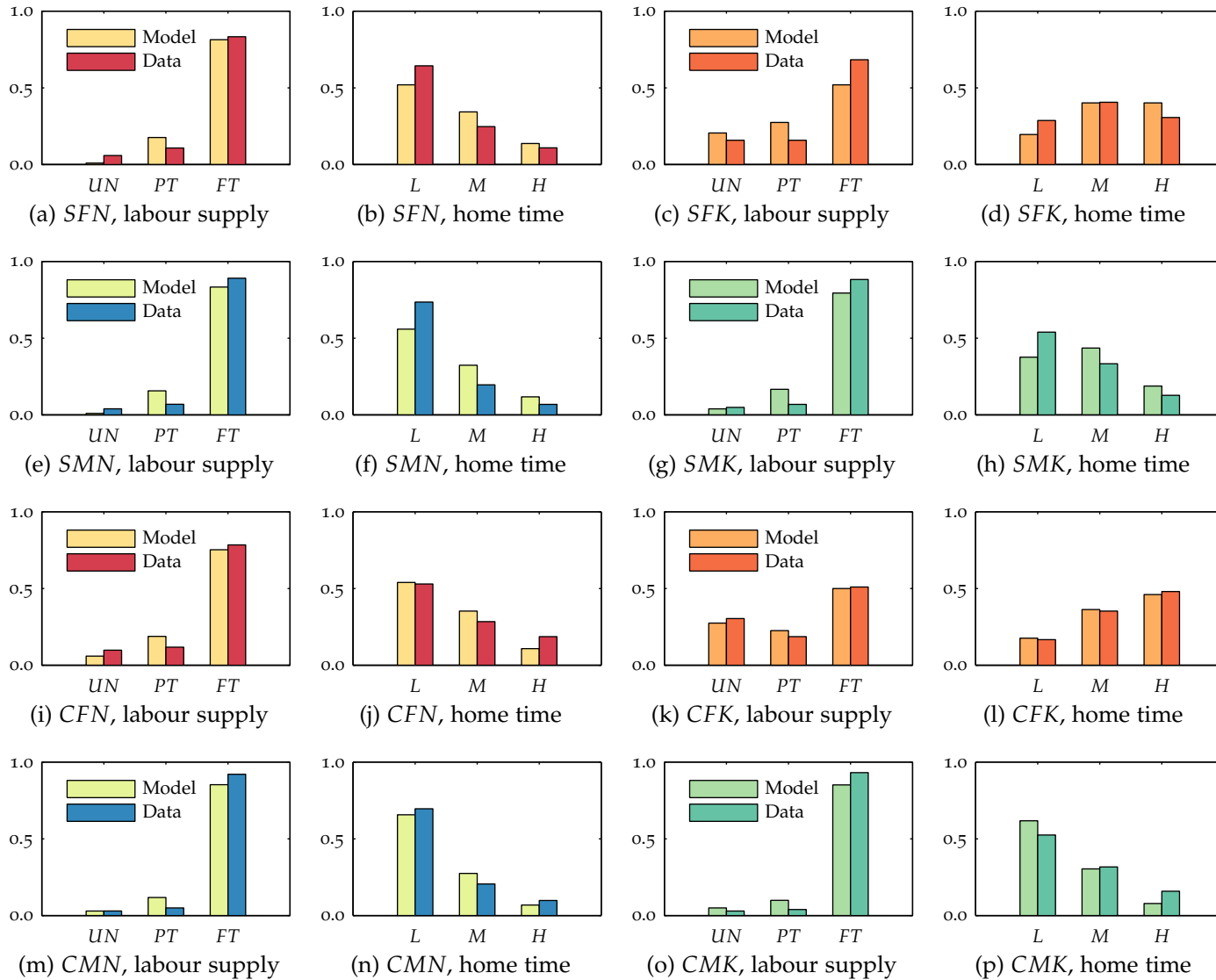
Figure 1: Figure shows empirical and predicted frequencies of work and home time, aggregated over types and conditional on marital status, gender, and children. We label using the convention that *S* (*C*) identifies singles (couples); *F* (*M*) identifies women (men); *N* (*K*) identifies childless (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).
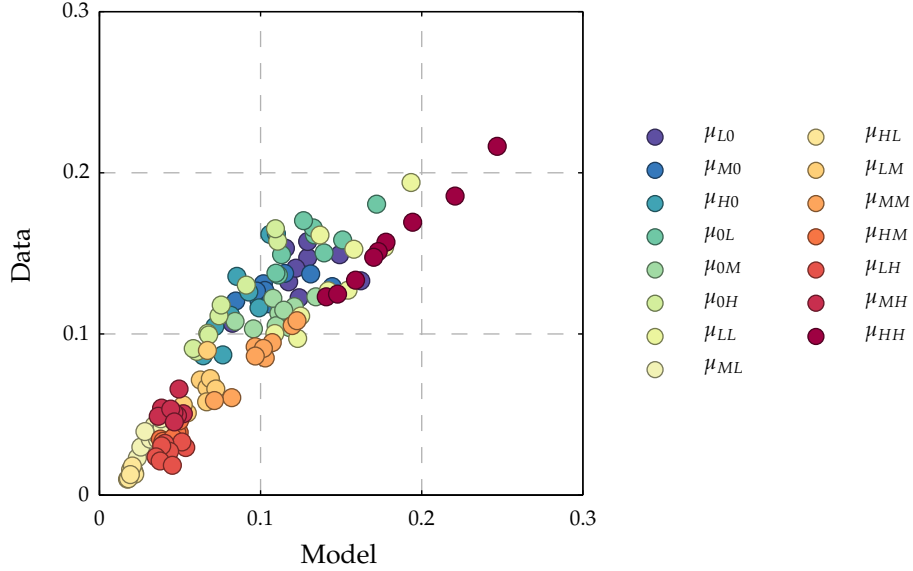
Figure 2: Market variation, marriage market matching function. Figure shows elements of the empirical and predicted marriage market matching function. Each market corresponds to a Census Bureau-designated division.

Our estimation targets a number of moments conditional on market, with our semi-parametric identification result reliant upon the presence of such 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 positions. Again, we observe a strong clustering of points around the diagonal.

An important object of interest is the Pareto weight, and how this varies at the level of the match and across markets. The Pareto weights implied by our model estimates are presented in Table 2. There are some important features from the table. Firstly, the female weight is increasing when she is more educated relative to her husband. For example, a college educated woman receives (on average) a share of 0.46 if she is married a man who has 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 increased to 0.61. Second, there is an asymmetric gender impact of differences in education: 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, which here reflects the joint impact

(a) Female: Unconditional work hours   (b) Male: Unconditional work hours

● Single: high school and below   ● Single: college and above   ● Couple: some college
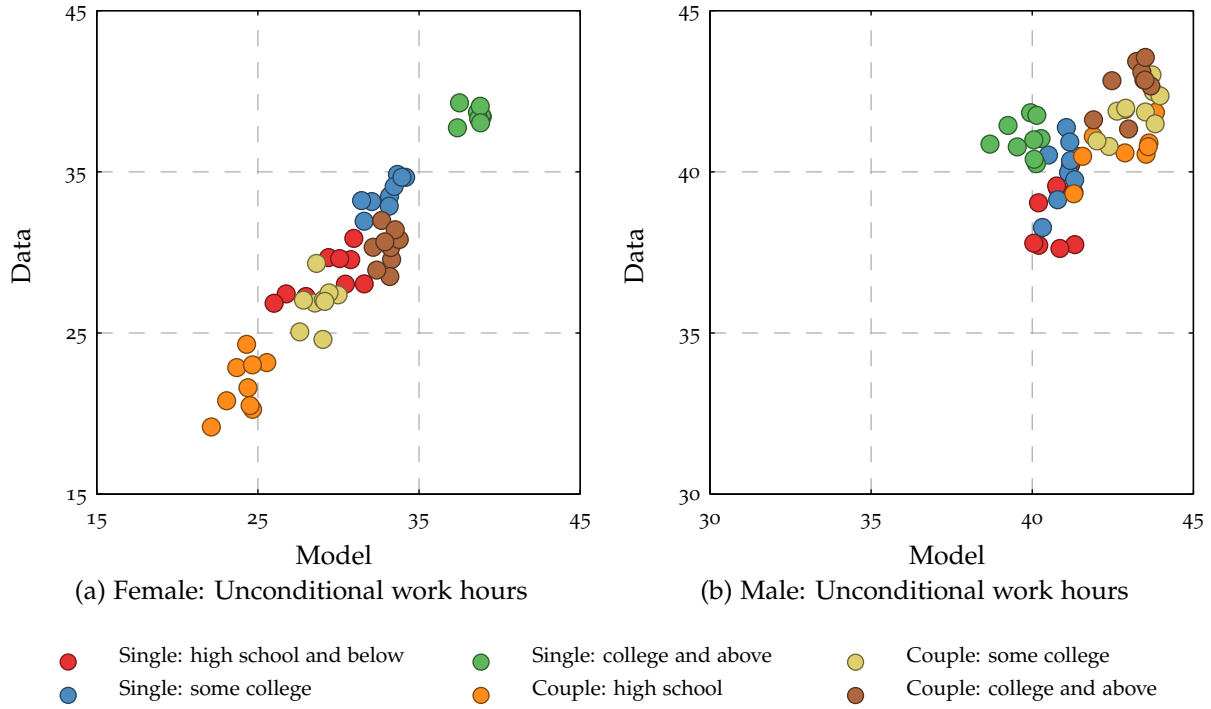● Single: some college   ● Couple: high school   ● Couple: college and above

Figure 3: Market variation, labour supply. Figure shows empirical and predicted mean unconditional work hours and employment of men and women by education level and market. Each 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	0.509 [0.490–0.538]	0.544 [0.530–0.558]	0.605 [0.584–0.616]
	Some college	0.430 [0.411–0.456]	0.491 [0.483–0.505]	0.548 [0.528–0.555]
	College and above	0.326 [0.303–0.354]	0.380 [0.368–0.402]	0.459 [0.455–0.466]

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.

of variation in taxes and the population vectors.

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 adult in a couple household means that we are perturbing the tax schedule as we move in a single dimension.[29] 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 that describe how employment, work hours, and home time respond as the wage of their spouse is varied.[30]

Our labour supply elasticities suggest that women are more responsive to changes in their own wage (both on the intensive and extensive margin) 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). The evidence on cross-wage labour supply effects is more limited, although the results here are consistent with the findings of e.g. Blau and Kahn (2007). In the same table 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 spouses wage is increased. 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 of the marriage market. This

---

[29]Starting from a fully joint system (as is true in our estimation exercise), and for any given joint time allocation decision, this 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).

[30]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. For cross wage conditional work hours elasticities, we condition on being employed both before and after the net-wage increase.

Table 3: Simulated elasticities

	Married		Single	
	Men	Women	Men	Women
Work hours				
Own-wage elasticity	0.16	0.23	0.05	0.07
Cross-wage elasticity	-0.10	-0.17	–	–
Participation				
Own-wage elasticity	0.07	0.29	0.02	0.19
Cross-wage elasticity	-0.04	-0.16	–	–
Home hours				
Own-wage elasticity	-0.21	-0.19	-0.06	-0.13
Cross-wage elasticity	0.12	0.13	–	–

Notes: All elasticities simulated under 2006 federal and state tax and transfer systems, aggregated over markets, and holding the marriage market fixed. Elasticities are calculated by increasing the individual' net wage rate by 1% (own-wage elasticity) or the net wage of their spouse by 1% (cross-wage elasticity) as described in the main text. *Participation* elasticities measure the percentage increase in the employment rate; *work hours* elasticities measure the percentage increase in hours of work amongst workers in the base system; *home hours* elasticities measure the percentage increase in total home time hours.

comparative static exercise implies a marriage market elasticity of -0.12. This falls into the range of estimates in the literature that has examined the impact of taxation on marriage decisions, which often finds 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 these cases we quantify the welfare loss relative to our more general benchmark specification.

    The results presented in this section assume a single marriage market, with the pop-

ulation 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:

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

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

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

$$\mathcal{W}_{ij}^i(T) = \int_{\theta^i} \int_{\mathbf{w},\mathbf{X},\epsilon} Y[v_{ij}^i(\mathbf{w},\mathbf{y},\mathbf{X},\epsilon;T,\lambda_{ij}) + \theta^i]\,dG_{ij}(\mathbf{w},\mathbf{y},\mathbf{X})\,dH_{ij}^i(\theta^i)$$

$$= p_{ij}^i(T)^{-\delta\sigma_\theta}\Gamma(1 - \delta\sigma_\theta)\int_{\mathbf{w},\mathbf{X},\epsilon} \frac{\exp[\delta v_{ij}^i(\mathbf{w},\mathbf{y},\mathbf{X},\epsilon;T,\lambda_{ij})]}{\delta}\,dG_{ij}(\mathbf{w},\mathbf{X},\epsilon) - \frac{1}{\delta},$$

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

$$\mathcal{W}_{ij}^i(T) = \gamma - \sigma_\theta \log p_{ij}^i(T) + U_{ij}^i(T,\lambda_{ij})$$

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 positions, and for married women, single men and single women.

A proof of this proposition is provided in Appendix B. As part of that proof, we characterize the distribution of the idiosyncratic payoffs for individuals who select into a given marital position.[31] This result allows us to decompose the welfare function contributions in to parts that reflect the distribution of idiosyncratic utility payoffs from

---

[31]This is a related, but distinct result compared to Proposition 1 in Blundell and Shephard (2012). That proposition does not apply to the welfare contribution conditional on a given marital state as (for individuals in couples) the maximisation problem of the household is not synonymous with the maximisation of the individual utility function.

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 that we use 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 of N ordered tax brackets $0 = n_1 < n_2 < \ldots < n_N$ that apply to the earnings of a given individual. We assume, but do not require, that these brackets are the same for both members in a married couple, and also for singles. 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 $\beta_T = [\mathbf{t}_N, \text{vec}(\mathbf{T}_{N \times N})]$.

The tax parameter vector $\mathbf{t}_{N \times 1}$ and tax matrix $\mathbf{T}_{N \times N}$ define tax liabilities at levels of 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 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.[32] Given this interpolating function we write the tax schedule for married couples at arbitrary earnings as $T(z_1, z_2)$, where z_1 and z_2 are henceforth used to denote the labour earnings of the two spouses. For a single individual with earnings z we have $T(z)$. Note that in our illustrations we do not include demographics as a conditioning variable.

In our application we set $N = 10$ with the earnings nodes (expressed in dollars per week in 2006 prices) as $\{0, 200, 400, 650, 950, 1300, 1700, 2200, 2800, 3500\}$. Thus, we have

---

[32]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.

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 EITC payments) and net of welfare transfers. We solve the optimal design problem numerically. Given our parameterisation 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 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 $\mu_{ij}^d(T, \boldsymbol{\lambda}^i) = \mu_{ij}^s(T, \boldsymbol{\lambda}^j)$ 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 4a we present the joint (net-income) budget constraint for both singles and married couples, calculated under the 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 while different are constant. As the earnings of either spouse change in any direction such that we enter 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 then informative about the degree of tax jointness. In particular, the edges of the three dimensional surface are clearly seen not to be parallel. To better illustrate the implied degree of tax jointness, in Figure 4b we show the associated marginal tax rate of a given individual, as the earnings of their spouse is fixed at different levels. We note a number of features. Firstly, we can see that there exists a broadly progressive structure; second, marginal tax rates are close to zero (or negative) at low earnings; third marginal tax rates tend to be lower the higher is the earnings of ones spouse. This third feature, as we discuss further below, is the *negative jointness* result described in Kleven,

Kreiner and Saez (2009).

In Figure 5 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 $\delta = 0$ (as seen in Figure 4) we have i) higher transfers when not working; ii) lower marginal tax rates (pure tax credits) at low earnings; 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 their 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 the couple.[33] These are monotonically declining in earnings as we move in both directions. Moreover, given the curvature of the utility function, there is a considerable redistributive motive even in the $\delta = 0$ case.

[To be completed]

5.3 Understanding

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 are provided in the Appendix.

[To be completed]

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 characterise the tax schedule with a given revenue requirement by solving the same constrained welfare maximisation problem as before. Second, in order to quantify the cost of these restricted forms we consider the dual problem. That is, we now maximise the revenue raised from our tax system, subject

---

[33]See the accompanying table note for details regarding the calculation of these.

(a) Net income function

(b) Marginal tax rate conditional on spousal earnings

Figure 4: 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 when we conditional on alternative values of spousal earnings.

(a) Net income function

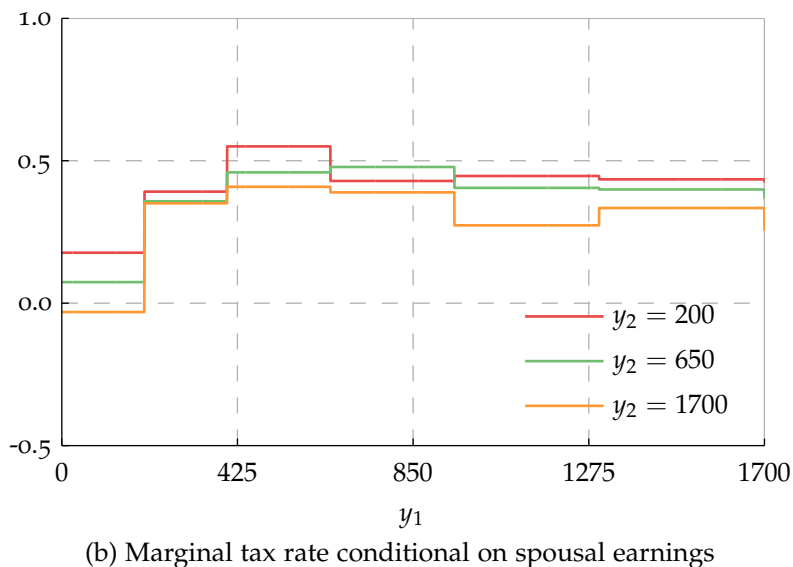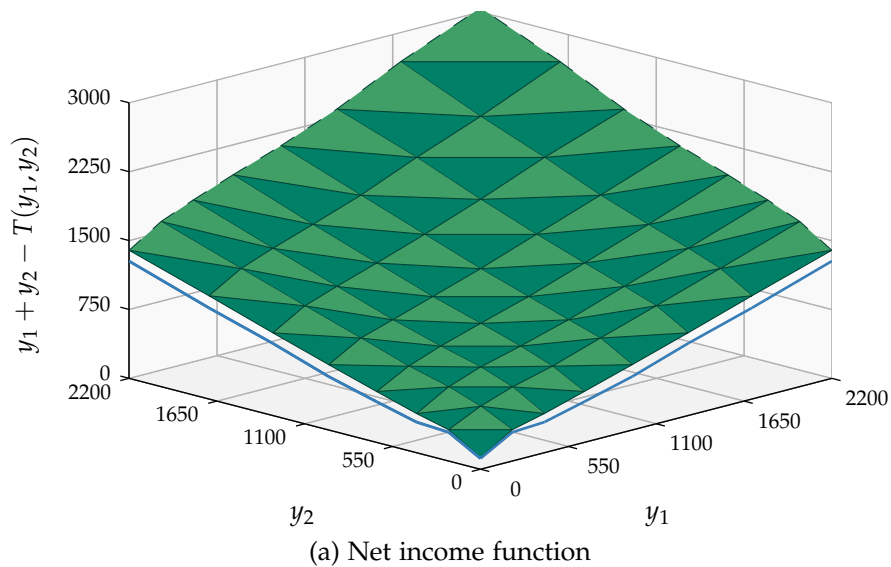(b) Marginal tax rate conditional on spousal earnings
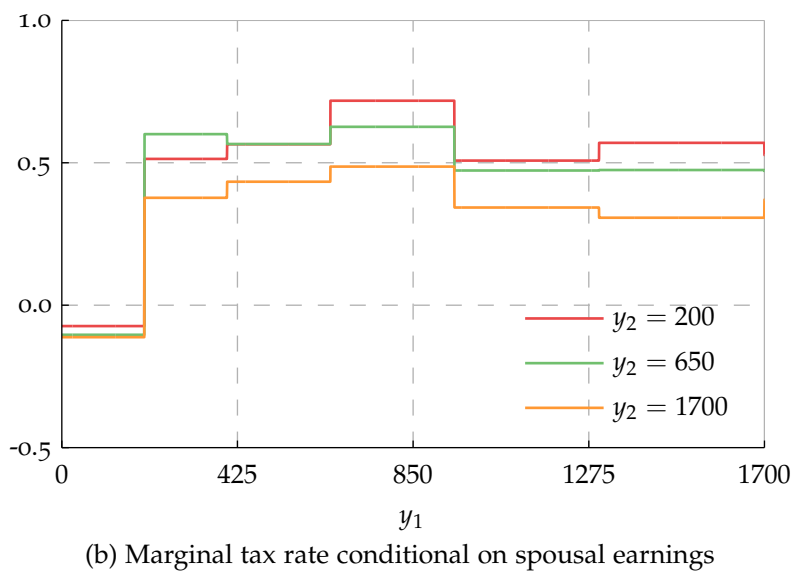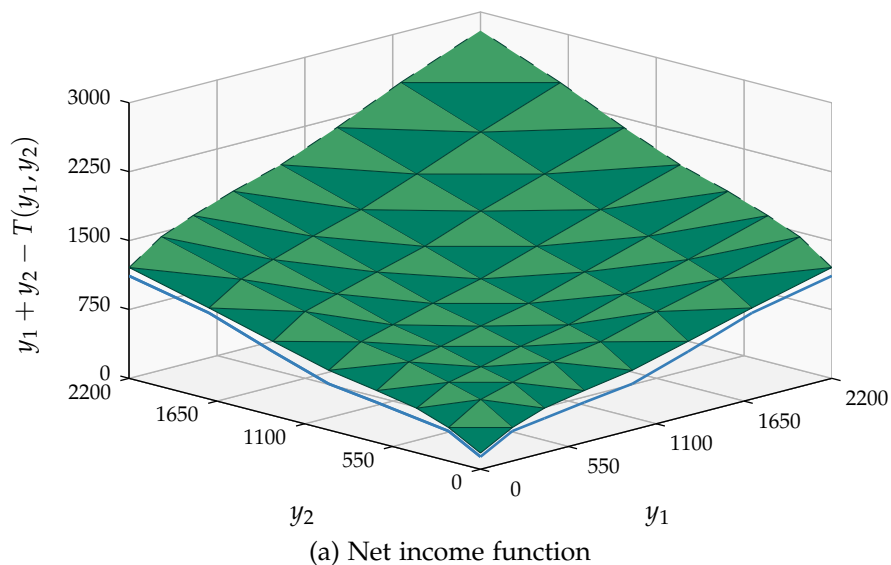
Figure 5: Optimal tax schedule with $\delta = -1$. In panel (a) we show net-income as a function of labour earnings for both single individuals (blue line) and couples (three dimensional surface). Marginal tax rates for both spouses (while potentially different) are constant within each of the shaded triangles. In panel (b) we show the implied structure of marginal tax rates when we conditional on alternative values of spousal earnings.

Table 4: Social welfare weights under optimal system

		Lowest earnings range								
		0–200	200–400	400–650	650–950	950–1300	1300–1700	1700–2200	2200–2800	2800+
$\delta = 0$:	0–200	2.584 [0.512]	–	–	–	–	–	–	–	–
	200–400	2.011 [4.755]	1.458 [4.853]	–	–	–	–	–	–	–
	400–650	1.584 [6.347]	1.218 [9.633]	1.034 [4.470]	–	–	–	–	–	–
	650–950	1.290 [6.233]	1.024 [7.449]	0.883 [6.849]	0.766 [2.722]	–	–	–	–	–
	950–1300	1.027 [5.140]	0.844 [5.067]	0.743 [4.325]	0.652 [3.492]	0.562 [1.057]	–	–	–	–
	1300–1700	0.825 [4.547]	0.695 [3.694]	0.621 [2.940]	0.553 [2.300]	0.482 [1.241]	0.416 [0.335]	–	–	–
	1700–2200	0.650 [2.626]	0.559 [1.977]	0.509 [1.651]	0.459 [1.350]	0.410 [0.703]	0.359 [0.359]	0.311 [0.103]	–	–
	2200–2800	0.491 [0.840]	0.432 [0.600]	0.400 [0.540]	0.366 [0.427]	0.337 [0.211]	0.302 [0.098]	0.268 [0.053]	0.227 [0.008]	–
	2800+	0.377 [0.153]	0.341 [0.097]	0.320 [0.093]	0.295 [0.083]	0.276 [0.042]	0.252 [0.018]	0.230 [0.008]	0.197 [0.002]	0.173 [0.000]
$\delta = -1$:	0–200	3.294 [0.616]	–	–	–	–	–	–	–	–
	200–400	2.219 [6.849]	1.547 [5.357]	–	–	–	–	–	–	–
	400–650	1.618 [8.711]	1.181 [10.588]	0.941 [4.383]	–	–	–	–	–	–
	650–950	1.298 [5.979]	0.958 [7.349]	0.769 [6.347]	0.629 [2.425]	–	–	–	–	–
	950–1300	0.964 [3.879]	0.726 [4.660]	0.592 [3.940]	0.487 [3.159]	0.381 [0.976]	–	–	–	–
	1300–1700	0.755 [3.605]	0.589 [3.441]	0.486 [2.739]	0.406 [2.187]	0.320 [1.227]	0.266 [0.342]	–	–	–
	1700–2200	0.491 [2.179]	0.394 [1.893]	0.335 [1.613]	0.287 [1.328]	0.237 [0.672]	0.203 [0.378]	0.165 [0.112]	–	–
	2200–2800	0.318 [0.702]	0.270 [0.570]	0.239 [0.510]	0.208 [0.409]	0.179 [0.208]	0.159 [0.107]	0.133 [0.060]	0.108 [0.008]	–
	2800+	0.194 [0.132]	0.177 [0.101]	0.162 [0.098]	0.145 [0.086]	0.127 [0.044]	0.115 [0.019]	0.100 [0.009]	0.084 [0.003]	0.068 [0.000]

Highest earnings range (row label for both panels)

Notes: Table presents average social welfare weights and joint probability mass under the optimal system for alternative δ 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.

to the incentive and marriage market equilibrium constraints, *and* the requirement that the level of social welfare achieved is at least that as obtained from our more general specification from Section 5.2. We consider the following restricted 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 on 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 is $T(z_1, z_2) = 2 \times \widetilde{T}(z_1/2 + z_2/2)$, with the same tax schedule 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 aggregate income: $T(z_1, z_2) = \overline{T}(z_1 + z_2)$.

We now discuss our results. In Figure 6 we present the implied marginal rate structure in the $\delta = 0$ case.[34] As in Figure 4, we have constructed these conditional on alternative spousal earnings levels. 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, it does imply higher marginal 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 different spousal earnings levels.

There is also an important impact of these alternative tax policies on the marriage market. In Table 5 we present the marriage market matching function and the equilibrium (i.e. market clearing) Pareto weights that are associated with these. Relative to our

---

[34]The broad patterns described here also apply in the $\delta = -1$ case.

most general specification, the fraction of single individuals is higher by 1.7 percentage points in our independent taxation specification, 2.2 percentage points higher in the joint taxation specification with income splitting, and 13.8 percentage points higher when we have joint taxation with income aggregation. The Pareto weights (which we recall are defined as the weight on female utility in the household problem) also vary across these specifications. With few exceptions, we obtain *lower* weights on female utility in these more restrictive specifications. The differences are most pronounced when we consider joint taxation with income splitting or aggregation.

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. That is, we now maximise 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 as obtained from our more general specification from Section 5.2. The differences in revenue raised can then be interpreted as the cost of the more restrictive structures considered here.

[To be completed]

5.5 The importance of the marriage market

In the simulations presented in Section 5.2 we saw that under the range of government preference parameters considered, that the implied marital sorting pattern was relatively close to that from our estimated model. To better understand the importance of the marriage market in determining the optimal structure of taxes and transfers we perform the following exercise. We resolve for the optimal structure holding the entire vector of Pareto weights, marriage market positions, and distributions of idiosyncratic payoffs fixed at their values from the corresponding optimum from the previous section. 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.

[To be completed]
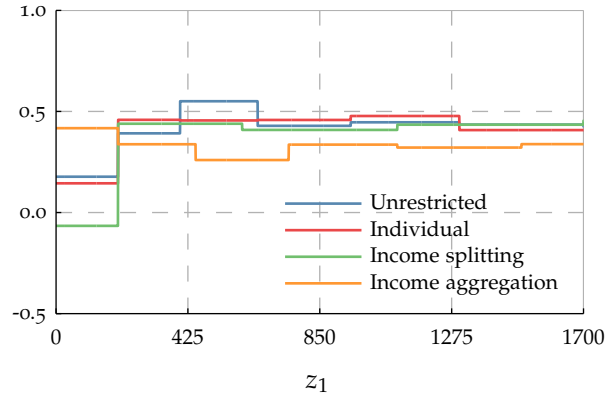
(a) $z_2 = 200$
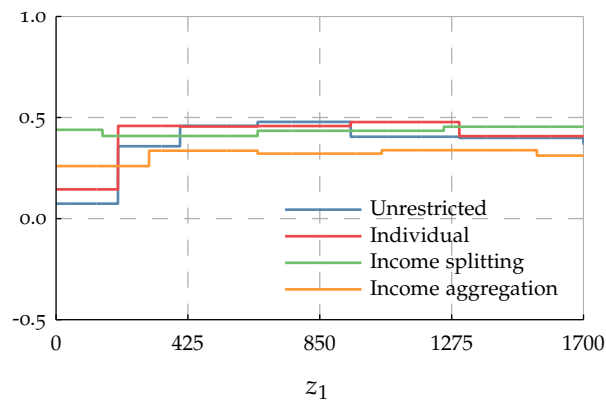
(b) $z_2 = 650$
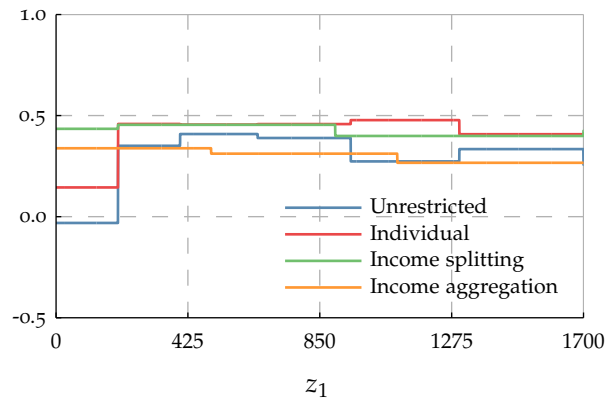
(c) $z_2 = 1700$

Figure 6: 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. *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).

Table 5: Marriage matching function

			Women		
			High school and below	Some college	College and above
(a). Unrestricted					
		–	0.117	0.091	0.093
		–	–	–	–
	High school and below	0.130	0.151	0.069	0.039
		–	(0.481)	(0.521)	(0.587)
Men	Some college	0.096	0.038	0.109	0.045
		–	(0.412)	(0.479)	(0.539)
	College and above	0.074	0.024	0.049	0.175
		–	(0.324)	(0.384)	(0.471)
(b). Independent					
		–	0.089	0.096	0.132
		–	–	–	–
	High school and below	0.113	0.175	0.068	0.034
		–	(0.460)	(0.507)	(0.581)
Men	Some college	0.105	0.040	0.106	0.037
		–	(0.390)	(0.465)	(0.532)
	College and above	0.099	0.025	0.048	0.150
		–	(0.305)	(0.372)	(0.465)
(c). Income splitting					
		–	0.081	0.099	0.142
		–	–	–	–
	High school and below	0.111	0.182	0.065	0.031
		–	(0.446)	(0.501)	(0.580)
Men	Some college	0.107	0.041	0.106	0.035
		–	(0.379)	(0.457)	(0.530)
	College and above	0.104	0.026	0.049	0.144
		–	(0.296)	(0.366)	(0.460)
(d). Income aggregation					
		–	0.151	0.134	0.154
		–	–	–	–
	High school and below	0.179	0.127	0.054	0.030
		–	(0.434)	(0.498)	(0.591)
Men	Some college	0.137	0.031	0.087	0.033
		–	(0.363)	(0.445)	(0.532)
	College and above	0.122	0.021	0.043	0.136
		–	(0.279)	(0.351)	(0.454)

Notes: Table shows marriage matching function under alternative tax schedule specifications. Black numbers correspond to elements of the marriage market matching function; blue numbers in parenthesis are the Pareto weights. *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).

6 Summary and conclusion

This paper has developed an empirical approach to optimal income taxation design within an equilibrium collective marriage market model. Our analysis centred around a parsimonious micro-econometric time allocation model, which was estimated using American Community Survey and American Time Use Survey data. We showed that the model is able to jointly explain labour supply, home time, and marriage market patterns. Moreover, it was able to successfully explain how these vary across markets.

[To be completed]

Appendices

A Proof of proposition 1

We assume that the distribution $G_{ij}(\mathbf{w}, \mathbf{y}, \mathbf{X}, \epsilon)$ 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 ℓ, 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:

$$ED_{ij}(\boldsymbol{\lambda}) = \mu_{ij}^d(\boldsymbol{\lambda}^i) - \mu_{ij}^s(\boldsymbol{\lambda}^j), \quad \forall i = 1, \dots, I, j = 1, \dots, J. \tag{15}$$

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. $ED_{ij}(\boldsymbol{\lambda}^*) = 0, \forall i = 1, \dots, I, j = 1, \dots, J$. Equilibrium uniqueness implies that there is a single vector that achieves this. Under our regularity conditions we have that: (i) $U_{ij}^i(\lambda_{ij})$ *and* $U_{ij}^i(\lambda_{ij})$ *are continuously differentiable in* λ_{ij}; (ii) $\partial U_{ij}^i(\lambda_{ij})/\partial \lambda = -\frac{\lambda_{ij}}{(1-\lambda_{ij})} \partial U_{ij}^j(\lambda_{ij})/\partial \lambda <$ 0; (iii) $\lim_{\lambda_{ij} \to 0} ED_{ij}(\lambda_{ij}, \boldsymbol{\lambda}_{-ij}) > 0$, and; (iv) $\lim_{\lambda_{ij} \to 1} ED_{ij}(\lambda_{ij}, \boldsymbol{\lambda}_{-ij}) < 0$.

A.1 Properties of the excess demand functions

We now state further properties of the excess demand functions, which we apply when we provide our proof of existence and uniqueness. Note that as we vary λ_{ij} our excess

demand functions must satisfy:

$$\frac{\partial ED_{ij}(\boldsymbol{\lambda})}{\partial \lambda_{ij}} = \underbrace{\frac{\partial \mu_{ij}^d(\boldsymbol{\lambda}^i)}{\partial \lambda_{ij}}}_{-} - \underbrace{\frac{\partial \mu_{ij}^s(\boldsymbol{\lambda}^j)}{\partial \lambda_{ij}}}_{+} < 0, \tag{16a}$$

$$\frac{\partial ED_{ik}(\boldsymbol{\lambda})}{\partial \lambda_{ij}} = \underbrace{\frac{\partial \mu_{ik}^d(\boldsymbol{\lambda}^i)}{\partial \lambda_{ij}}}_{+} - \underbrace{\frac{\partial \mu_{ik}^s(\boldsymbol{\lambda}^k)}{\partial \lambda_{ij}}}_{-} > 0; \ k \neq j, \tag{16b}$$

$$\frac{\partial ED_{kj}(\boldsymbol{\lambda})}{\partial \lambda_{ij}} = \underbrace{\frac{\partial \mu_{kj}^d(\boldsymbol{\lambda}^k)}{\partial \lambda_{ij}}}_{+} - \underbrace{\frac{\partial \mu_{kj}^s(\boldsymbol{\lambda}^j)}{\partial \lambda_{ij}}}_{-} > 0; k \neq i, \tag{16c}$$

$$\frac{\partial ED_{kl}(\boldsymbol{\lambda})}{\partial \lambda_{ij}} = \underbrace{\frac{\partial \mu_{kl}^d(\boldsymbol{\lambda}^k)}{\partial \lambda_{ij}}}_{0} - \underbrace{\frac{\partial \mu_{kl}^s(\boldsymbol{\lambda}^l)}{\partial \lambda_{ij}}}_{0} = 0; \ k \neq i, l \neq j. \tag{16d}$$

Note that equation (16d) is a consequence of the IIA property of the Type-I extreme value distribution and is therefore critical for our proof of the uniqueness of equilibrium.

A.2 Existence

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

$$\boldsymbol{\Gamma}(\boldsymbol{\lambda}) = \psi \cdot \mathbf{ED}(\boldsymbol{\lambda}) + \boldsymbol{\lambda}, \tag{17}$$

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 non-decreasing 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 $U_{ij}^k(\lambda_{ij})$ for $k = i, j$ is derived from the time allocation problem described in the main text then one has proven the existence of equilibrium. It is therefore sufficient to show that one can construction a $\boldsymbol{\Gamma}(\boldsymbol{\lambda}) = \psi \cdot \mathbf{ED}(\boldsymbol{\lambda}) + \boldsymbol{\lambda}$ such that:

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

2. $\boldsymbol{\Gamma}(\boldsymbol{\lambda})$ 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 (16a) to (16d). $\qquad\square$

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

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 $\psi_{ij}^- = \min\{-\lambda_{ij}/ED_{ij}(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \in BC_{ij}^-\}$.[35] Continuity of $ED_{ij}(\boldsymbol{\lambda})$ implies that both ψ_{ij}^+ and ψ_{ij}^- exist and are **strictly** positive. Then for all $\psi_{ij} \in (0, \min\{\psi_{ij}^+, \psi_{ij}^-\})$ we have that

$$0 \leq \psi_{ij} ED_{ij}(\boldsymbol{\lambda}) + \lambda_{ij} \leq 1 \tag{18}$$

$\qquad\square$

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

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 $\psi_{ij} \in (0, 1/D_{ij})$ one has the following:

$$\frac{\partial[\psi_{ij} ED_{ij}(\boldsymbol{\lambda}) + \lambda_{ij}]}{\partial \lambda_{ij}} = \psi_{ij} \frac{\partial ED_{ij}(\boldsymbol{\lambda})}{\partial \lambda_{ij}} + 1 \geq -\psi_{ij} D_{ij} + 1 > 0, \tag{19a}$$

$$\frac{\partial[\psi_{ij} ED_{ij}(\boldsymbol{\lambda}) + \lambda_{ij}]}{\partial \lambda_{kj}} = \psi_{ij} \frac{\partial ED_{ij}(\boldsymbol{\lambda})}{\partial \lambda_{kj}} \geq 0 \text{ for } k \neq i, \tag{19b}$$

$$\frac{\partial[\psi_{ij} ED_{ij}(\boldsymbol{\lambda}) + \lambda_{ij}]}{\partial \lambda_{il}} = \psi_{ij} \frac{\partial ED_{ij}(\boldsymbol{\lambda})}{\partial \lambda_{il}} \geq 0 \text{ for } l \neq j, \tag{19c}$$

which follows from equations (16a) to (16d). Let:

$$\overline{\psi} = \min\{\min\{\psi_{11}^+, \psi_{11}^-\}, \ldots, \min\{\psi_{IJ}^+, \psi_{IJ}^-\}, 1/2D_{11}, \ldots, 1/2D_{IJ}\}. \tag{20}$$

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

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

---

[35] 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 ∞.

A.3 Uniqueness of Equilibrium

Uniqueness follows from the differentiability of $\Gamma(\boldsymbol{\lambda})$. That is given that $[0,1]^{I \times J}$ is closed and connected if $\Gamma(\boldsymbol{\lambda})$ is differentiable almost everywhere. Then if for some $\beta \in (0,1)$ we have that $\|\Gamma(\boldsymbol{\lambda}) - \Gamma(\boldsymbol{\lambda}')\| \leq \beta \|\boldsymbol{\lambda} - \boldsymbol{\lambda}'\|$ for all $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}'$ and some norm $\|.\|$, then by the contraction mapping theorem there exists a unique fixed $\boldsymbol{\lambda}^* \in [0,1]^{I \times J}$ such that $\Gamma(\boldsymbol{\lambda}^*) = \boldsymbol{\lambda}^*$. 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:

$$0 \geq \frac{\partial \Gamma_{ij}(\boldsymbol{\lambda})}{\partial \lambda_{ij}} = \psi \cdot \frac{\partial ED_{ij}(\boldsymbol{\lambda})}{\partial \lambda_{ij}} + 1 < 1, \tag{21}$$

since $\psi > 0$ and from equation (16a) we have that $\partial ED_{ij}(\boldsymbol{\lambda})/\partial \lambda_{ij} < 0$. Moreover, by construction:

$$\frac{1}{2} \geq \frac{\partial \Gamma_{ij}(\boldsymbol{\lambda})}{\partial \lambda_{ik}} \geq 0. \tag{22}$$

And the IIA property in (16d) implies that

$$\frac{\partial \Gamma_{ij}(\boldsymbol{\lambda})}{\partial \lambda_{kl}} = 0. \tag{23}$$

Therefore, Γ is a contraction, since, by the mean value theorem,

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

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

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

$\square$

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 position. 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 θ_j that has scale parameter σ_θ. We now characterize the distribution of θ_j conditional on j being chosen. Letting $p_j = \left(\sum_k \exp[(U_k - U_j)/\sigma_\theta] \right)^{-1}$ denote the associated conditional choice probability it follows that:

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

Hence, the distribution of the idiosyncratic payoff conditional on option j being optimal, is also extreme value with the common scale parameter σ_θ 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 form of the utility transformation function (equation **??**), 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 position 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 analytical character-isation 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:

$$\int_{\theta_j} \exp(\delta\theta_j) \, \mathrm{d}H_j(\theta_j) = \frac{1}{\sigma_\theta} \int_{\theta_j} \exp(\delta\theta_j) \exp(-[\theta_j + \sigma_\theta \log p_j]/\sigma_\theta) e^{-\exp(-[\theta_j + \sigma_\theta \log p_j]/\sigma_\theta)} \, \mathrm{d}\theta_j$$

$$= \exp(-\delta\sigma_\theta \log p_j) \int_0^\infty t^{-\delta\sigma_\theta} \exp(-t) \, \mathrm{d}t$$

$$= p_j^{-\delta\sigma_\theta} \Gamma(1 - \delta\sigma_\theta).$$

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 position is simply given by:

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

$$= \gamma - \sigma_\theta \log p_j + \int_{\mathbf{Z}_j} v(\mathbf{Z}_j) \, \mathrm{d}G_j(\mathbf{Z}_j),$$

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 Marriage market numerical algorithm

In this Appendix we describe the iterative algorithm that we use to calculate the market clearing vector of Pareto weights. We first note that using the conditional choice prob-

abilities from equation 7 we are able to write the quasi-demand equation of type i men for type j spouses as:

$$\sigma_\theta \times \left[\ln \mu_{ij}^d(T, \boldsymbol{\lambda}^i) - \ln \mu_{i0}^d(T, \boldsymbol{\lambda}^i)\right] = U_{ij}^i(T, \lambda_{ij}) - U_{i0}^i(T). \qquad (24)$$

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:

$$\sigma_\theta \times \left[\ln \mu_{ij}^s(T, \boldsymbol{\lambda}^j) - \ln \mu_{0j}^s(T, \boldsymbol{\lambda}^j)\right] = U_{ij}^j(T, \lambda_{ij}) - U_{0j}^j(T). \qquad (25)$$

The algorithm proceeds as follows:[36]

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 $j = 1, \dots, J$.

2. Taking the difference of the quasi-demand (equation 24) and the quasi-supply (equation 25) functions for each $\langle i, j \rangle$ submarriage market, and imposing the market clearing condition $\mu_{ij}^d(T, \boldsymbol{\lambda}^i) = \mu_{ij}^s(T, \boldsymbol{\lambda}^j)$ we obtain:

$$\sigma_\theta \times \left[\ln \mu_{0j}^s - \ln \mu_{i0}^d\right] = U_{ij}^i(T, \lambda_{ij}) - U_{i0}^i(T) - \left[U_{ij}^j(T, \lambda_{ij}) - U_{0j}^j(T)\right], \qquad (26)$$

which given the single measures μ_{i0}^d and μ_{0j}^s (and the tax schedule T) is only a function of the Pareto weight for that submarriage-market λ_{ij}. Given our assumptions on the utility functions there exists a unique solution to equation 26. 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 $\mu_{i0}^d(T)$ and $\mu_{0j}^s(T)$ from *Step 1*. These 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 $U_{ij}^i(T, \lambda)$ and $U_{ij}^j(T, \lambda)$ for each marital match combination $\langle i, j \rangle$ on a fixed grid of Pareto weights $\lambda \in \boldsymbol{\lambda}^{\mathrm{grid}}$ with $\inf[\boldsymbol{\lambda}^{\mathrm{grid}}] \gtrsim 0$ and $\sup[\boldsymbol{\lambda}^{\mathrm{grid}}] \lesssim 1$. We may then replace $U_{ij}^i(T, \lambda)$ and $U_{ij}^j(T, \lambda)$ with an approximating parametric function, so that no expected

---

[36]This is similar to the method described in independent work by Galichon, Kominers and Weber (2014).
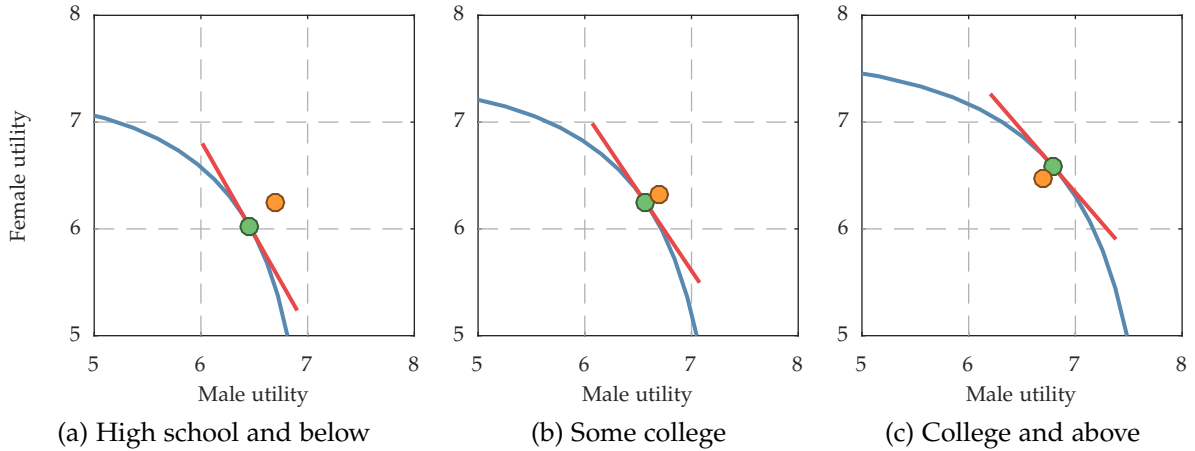
Figure 7: Expected utility possibility frontier. The figure shows the expected utility possibility frontier for college and above men when married to women with different schooling levels. The figure is obtained from the estimated model with empirical tax and transfer system and is calculated under the 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.

values are actually evaluated within the iterative algorithm.[37]

D   UPF

E   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.[38] Our measure of taxes includes both state and federal Earned Income Tax Credit (EITC) programmes, and we also account for Food Stamps and the Temporary Assistance for Needy Families (TANF)

---

[37]Calculating the expected values within a match are (by many orders of magnitude) the most computationally expensive part of our algorithm. An implication of this is that if there are multiple markets K, and each market $k \leq K$ only differs by the population vectors $\mathcal{M}_k$ and $\mathcal{F}_k$ and/or the demographic transition function, then the computational cost in obtaining the equilibrium for all K markets is *approximately independent* of the number of markets K considered. In our application we also have market variation in taxes and transfers so we are not able to exploit this property.

[38]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).

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 ι 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 Earned Income Tax Credit (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 are calculated with the NBER 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 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 this we (locally) modify the tax schedule in these events.[39]

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.

---

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

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.[40]

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 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:

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

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 $Y_{TANF}^{\widehat{\max}}[\mathbf{X}_{\iota k}]$ 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.[41]

---

[40]In practice the Food Stamp Program also has a gross-earnings and net-earnings income test. These require that earnings are below some threshold that is related to the Federal Poverty Line 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 taxes/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.

[41]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 40. We also do not consider the time limits in eligibility.

F   Identification

F.1   Proof of Proposition 2

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

$$\ln \mu_{ij}(T, \boldsymbol{\lambda}) - \ln \mu_{i0}(T, \boldsymbol{\lambda}^i) = \frac{U_{ij}^i(T, \lambda_{ij}) - U_{i0}^i(T)}{\sigma_\theta}, \tag{27a}$$

$$\ln \mu_{ij}(T, \boldsymbol{\lambda}) - \ln \mu_{0j}(T, \boldsymbol{\lambda}^j) = \frac{U_{ij}^j(T, \lambda_{ij}) - U_{0j}^j(T)}{\sigma_\theta}. \tag{27b}$$

The left hand side of equations 27a and 27b is data 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 λ_{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:

$$\frac{\partial \left[\ln \mu_{ij}(T, \boldsymbol{\lambda}) - \ln \mu_{i0}(T, \boldsymbol{\lambda}^i) \right]}{\partial m_i} = \frac{1}{\sigma_\theta} \frac{\partial U_{ij}^i(T, \lambda_{ij})}{\partial \lambda_{ij}} \frac{\partial \lambda_{ij}}{\partial m_i}, \tag{28a}$$

$$\frac{\partial \left[\ln \mu_{ij}(T, \boldsymbol{\lambda}) - \ln \mu_{0j}(T, \boldsymbol{\lambda}^j) \right]}{\partial m_i} = \frac{1}{\sigma_\theta} \frac{\partial U_{ij}^j(T, \lambda_{ij})}{\partial \lambda_{ij}} \frac{\partial \lambda_{ij}}{\partial m_i}. \tag{28b}$$

Taking the ratio of the partial derivatives in equations 28a and 28b we define:

$$z_{ij} = \frac{\partial U_{ij}^i(T, \lambda_{ij}) / \partial \lambda_{ij}}{\partial U_{ij}^j(T, \lambda_{ij}) / \partial \lambda_{ij}}.$$

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

F.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 using the household time allocation patterns we can identify the home production technology for couple households. The following assumptions are used in the proof of identification in this section. While some of them are easily relaxed, however, for clarity, easy of analysis, and because they directly relevant for the empirical and optimal design analysis in the rest of the paper 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, σ_ϵ.*

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

$$u^i(\ell^i, q^i, Q^i, \mathbf{X}^i) = u_q^i(q^i, \mathbf{X}^i) + u_\ell^i(\ell^i, \mathbf{X}^i) + u_Q^i(Q^i, \mathbf{X}^i).$$

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

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

Assumption ID-5. *There exist a element of $\mathbf{X}^i$, $\mathbf{X}_r^i$, such that $\mathbf{X}_r^i$ affects $A(\mathbf{X}^i)$ but not $u_Q^i(Q^i, \mathbf{X})$. Also there exist an $\mathbf{X}_*^i$ such that $A(\mathbf{X}_*^i) = 1$.*

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

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

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

F.2.1 Step 1: The identification using the singles problem

Consider the decision problem of a single male of type i. Let $\mathcal{A}^i = \{1, \ldots, \overline{A}^i\}$ be the 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:

$$\log \left[\frac{\mathbb{P}(a)}{\mathbb{P}(1)} \right] = \frac{\widehat{u}^i(a) - \widehat{u}^i(1)}{\sigma_\epsilon}. \tag{29}$$

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 29 with respect to w^i and using Assumption ID-2 yields:

$$\frac{\partial \log \left[\mathbb{P}(a)/\mathbb{P}(1) \right]}{\partial w} = \frac{1}{\sigma_\epsilon} \cdot \frac{\partial u_q^i(q^i(a); \mathbf{X}^i)}{\partial q} \cdot \left[1 - \frac{\partial T(w^i h_w^i(a), y^i; \mathbf{X}^i)}{\partial w} \right] \cdot h_w^i(a), \tag{30}$$

where $q^i(a)$ and $h_w^i(a)$ 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 Assumption ID-3 and ID-7, the scale coefficient for the state specific errors σ_ϵ is identified. Hence, the marginal utility of private consumption is identified. Integrating equation 30 and combining with equation 29, then implies that the sum, $u_\ell^i(\ell^i; \mathbf{X}^i) + u_Q^i(Q^i; \mathbf{X}^i)$, is identified up to a normalizing constant. Then for each level of feasible home hours, both $u_\ell^i(\ell^i; \mathbf{X}^i)$ and $u_Q^i(Q^i; \mathbf{X}^i)$ are identified by varying the level of the level of market hours with either home time or leisure fixed. Under Assumption ID-5 the home efficiency parameter $A_{i0}(\mathbf{X}^i)$ is identified by comparing $u_Q^i(Q^i(a); \mathbf{X}^i)$ across different values of $\mathbf{X}^i$.

F.2.2 Step 2: Identification of marriage home production function.

In Step 1 we show that the sub-utilities are identified up to a normalizing constant, without loss of generality, the location normalization is set to zero throughout the rest of the proof. Consider a household of type $\langle i, j \rangle$ with the time allocation set $\mathcal{A}_{ij} = \{1, \dots, \overline{A}\}$, $\overline{A} = \overline{A}^i \times \overline{A}^j$, 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 the time allocation $a \in \mathcal{A}_{ij}$. Let $\varepsilon_a^{ij} = (1 - \lambda_{ij}) \times \epsilon_a^i + \lambda_{ij} \times \epsilon_a^j$, and define $\mathcal{G}_{ij}^a(\cdot)$ to be the joint cumulative distribution function of $[\varepsilon_a^{ij} - \varepsilon_1^{ij}, \dots, \varepsilon_a^{ij} - \varepsilon_{a-1}^{ij}, \varepsilon_{a+1}^{ij} - \varepsilon_a^{ij}, \dots, \varepsilon_a^{ij} - \varepsilon_{\overline{A}}^{ij}]$. For each $a \in \{1, \dots, \overline{A} - 1\}$

define:

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

with $\widehat{\mathbf{u}}^{ij} = [\widehat{u}_1^{ij} - \widehat{u}_A^{ij}, \ldots, \widehat{u}_{A-1}^{ij} - \widehat{u}_A^{ij}]^\mathsf{T}$ defining the $(\overline{A} - 1)$ vector of utility differences, and let $\mathcal{Q}(\widehat{\mathbf{u}}^{ij}) = [\mathcal{Q}_1(\widehat{\mathbf{u}}^{ij}), \ldots, \mathcal{Q}_{A-1}(\widehat{\mathbf{u}}^{ij})]^\mathsf{T}$ 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.[42] Given that the distribution of ϵ is known and λ_{ij} is identified then 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,

$$\Delta_{ij}(a) = \widehat{\mathbf{u}}_{[a]}^{ij} - (1 - \lambda_{ij}) \times [u_\ell^i(\ell^i(a^i); \mathbf{X}^i) + u_q^i((1 - s_{ij}(a; \lambda_{ij})) \cdot q(a); \mathbf{X}^i)]$$
$$- \lambda_{ij} \times [u_\ell^j(\ell^j(a^j); \mathbf{X}^j) + u_q^i(s_{ij}(a; \lambda_{ij}) \cdot q(a); \mathbf{X}^j)].$$

The identification arguments from Step 1 imply that $u_q^i(q^i; \mathbf{X}^i)$ and $u_q^j(q^j; \mathbf{X}^j)$ are known. The marriage market equilibrium conditions from Proposition 2 λ_{ij} is identified. These, together with Assumption ID-2 and Assumption ID-4 imply that $s_{ij}(a; \lambda_{ij})$ is also known. Thus, it follows that $\Delta_{ij}(a)$ is identified. Finally, the definition of $\widehat{u}^{ij}(a)$ and Assumption ID-2 imply:

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

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

G   Moment list

In this appendix we list the complete set of estimation moments (total of X moments). The fit of the model is described in Section 4.5 from the main text. Recall that there are nine markets and three education groups (types) for both men and women.

[To be completed]

---

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

H Additional parameter and results tables

In Table 6 we present the estimates from our model, together with the accompanying standard errors. These are obtained from the estimation procedure described in Section 4 from the main text.

Table 6: Parameter estimates

	Estimate	Standard error
Log-wage offers:		
Male, high school and below: mean	2.611	0.003
Male, high school and below: s.d.	0.486	0.003
Male, some college: mean	2.848	0.003
Male, some college: s.d.	0.479	0.003
Male, college: mean	3.218	0.004
Male, college: s.d.	0.530	0.003
Female, high school and below: mean	2.112	0.004
Female, high school and below: s.d.	0.565	0.004
Female, some college: mean	2.449	0.004
Female, some college: s.d.	0.543	0.003
Female, college: mean	2.907	0.003
Female, college: s.d.	0.512	0.003
Preference parameters:		
Leisure scale	1.038	0.070
Home good scale	0.149	0.035
Leisure curvature	0.665	0.073
Home good curvature	-0.069	0.033
Fixed costs (kids)	86.934	1.964
Marital shock, s.d.	0.135	0.005
State specific error, s.d.	0.286	0.006
Home production technology:		
Male production share	0.050	0.007
Single productivity (no children),high school and below	2.000	0.472
Single productivity (no children),some college	3.189	0.705
Single productivity (no children),college	8.120	1.763
Male productivity (children)	9.714	2.279
Female productivity (children), high school and below	13.021	3.041
Female productivity (children), some college	14.553	3.427
Female productivity (children), college	15.906	3.779
HH productivity (children) female, high school and below	4.403	0.947
HH productivity (children) female, some college	7.158	1.596
HH productivity (children) female, college	4.068	0.894
HH productivity (children) educational homogamy, h/school and below	2.102	0.097
HH productivity (children) educational homogamy, some college	1.168	0.016
HH productivity (children) educational homogamy, college	2.949	0.180

Notes: All parameters estimated simultaneously using a moment based estimation procedure as detailed in Section 4 from the main text. See Footnote 28 for a description of the method used to calculate standard errors. All incomes are expressed in dollars per-week in average 2006 prices.

Figure 8: Marginal tax rate schedules by State

(a) Single, no children

(b) Single, children

(c) Couples, no children

(d) Couples, children
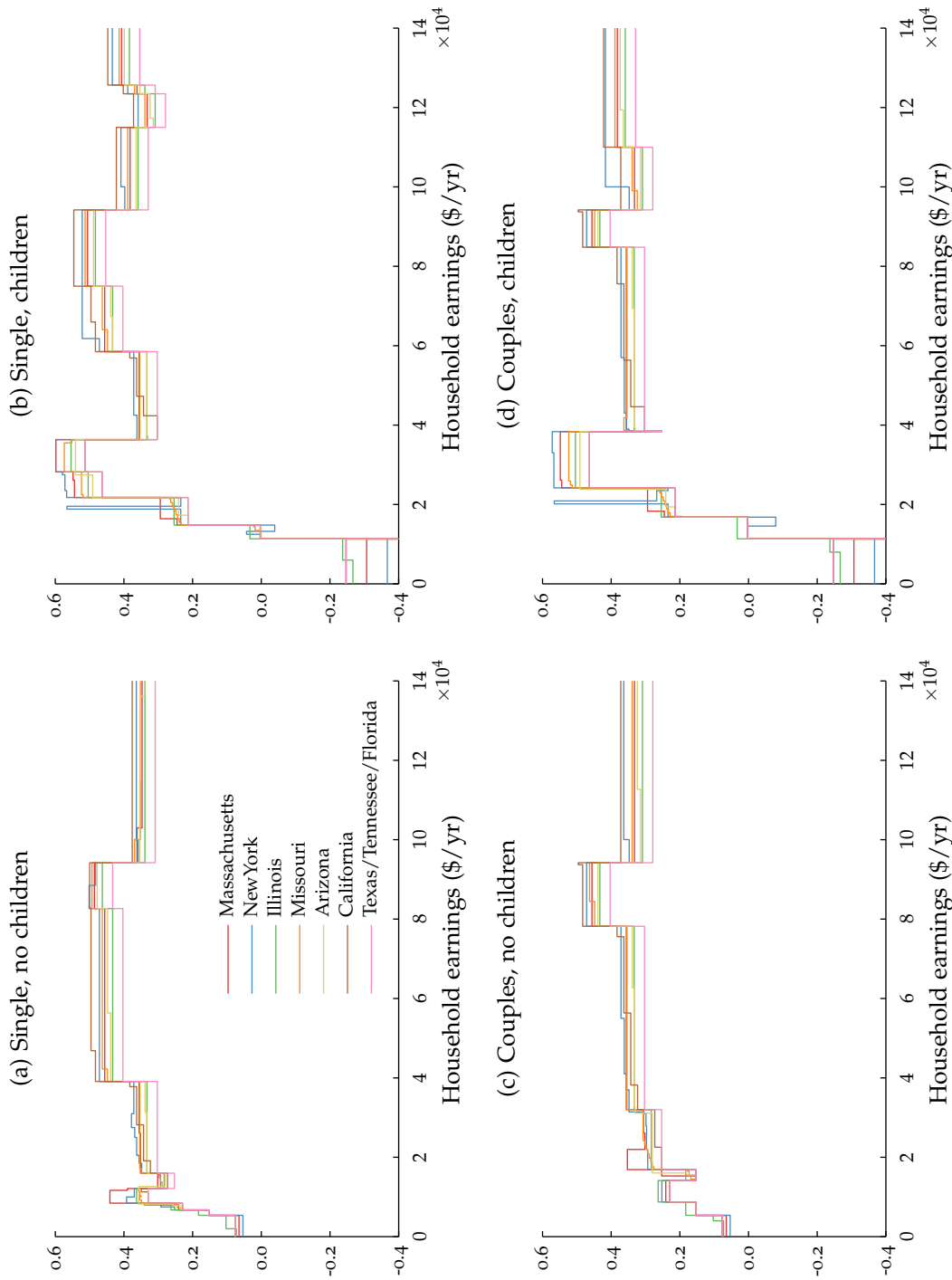
Notes: Figure shows the effective combined Federal and State marginal tax rate, including income tax, payroll tax, and Earned Income Tax Credit. Figure calculated using TAXSIM under 2006 tax systems, assumes that there are no other sources of income, and (when present) there are two children. For couples we assume married filing jointly status; for singles with children we assume head-of-household status. Presented schedules correspond to most populous States in respective Census Bureau-designated division; Texas, Tennessee, and Florida do not have any State Income Tax or State EITC and so are combined.

References

Aguiar, Mark, and Erik Hurst. 2007. "Measuring Trends in Leisure: The Allocation of Time Over Five Decades." *The 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.

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." *The 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." *The 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. Oxford University Press for Institute for Fiscal Studies.

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

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

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, and Ivar Ekeland. 2009. "The Microeconomics of Efficient Group Behavior: Identification." *Econometrica*, 77(3): 763–799.

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." *Structural Econometric Models*, , ed. Matthew Shum, Chapter 15, 291–336.

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." *The 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.

Donni, Olivier. 2003. "Collective household labor supply: nonparticipation and income taxation." *Journal of Public Economics*, 87(5–6): 1179–1198.

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." *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches*, 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*.

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

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

Hotz, V. Joseph, and Robert A. Miller. 1993. "Conditional Choice Probabilities and the Estimation of Dynamic Models." *The 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.

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." *The Review of Economic Studies*, 78(1): 328–355.

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. Oxford University Press for Institute for Fiscal Studies.

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

Piketty, Thomas, and Emmanuel Saez. 2013. "Chapter 7 - Optimal Labor Income Taxation." In *handbook of public economics, vol. 5*. Vol. 5 of *Handbook of Public Economics*, , ed. Raj Chetty Martin Feldstein Alan J. Auerbach and Emmanuel Saez, 391–474. Elsevier.

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

Vermeulen, Frederic. 2005. "A collective model for female labour supply with non-participation and taxation." *Journal of Population Economics*, 19(1): 99–118.