

Job market paper

COLLECTIVE MODELS AND THE MARRIAGE MARKET

SIMON WEBER[§]

ABSTRACT. In this paper, I develop an integrated approach to collective models and matching models of the marriage market. In the collective framework, both household formation and the intra-household allocation of bargaining power are taken as given. This is no longer the case in the present contribution, where both are endogenous to the determination of equilibrium on the marriage market. I characterize a class of “proper” collective models which can be embedded into a general matching framework with imperfectly transferable utility. In such models, the bargaining sets are parametrized by an analytical device called *distance function*, which plays a key role both for writing down the usual stability conditions and for estimation. In general, however, distance functions are not known in closed-form. I provide an efficient method for computing distance functions, that works even with the most complex collective models. Finally, I illustrate my results with a proof-of-concept application that uses PSID data. I identify the sharing rule and its distribution and consider two counterfactual experiments of women empowerment, both of which are shown to affect the sharing rule and the investment of couples in public goods.

Date: November 13, 2018.

[§]KU Leuven, simon.weber@kuleuven.be. I thank Alfred Galichon for his guidance and support. This paper has greatly benefited from comments from Laurens Cherchye, Edoardo Ciscato, Bram De Rock, Arnaud Dupuy, Marco Francesconi, Alfred Galichon, Jorge Luis García, James Heckman, Sonia Jaffe, Magne Mogstad, Jack Mountjoy, Noemi Nocera, Frederic Vermeulen, Paul Vertier and Alessandra Voena and seminar participants at the Becker Friedman Institute, CEHD, CERGE-EI, KU Leuven, LISER and University of Chicago. I am particularly grateful to the Becker Friedman Institute, the CEHD, KU Leuven and the University of Chicago for hosting me while most of this paper was written.

1. INTRODUCTION

Power matters. That is the key insight that stems from the collective model literature developed since the seminal paper of Chiappori (1988). In the collective representation of the household, partners engage in bargaining over which goods to consume or produce, and reach Pareto-efficient decisions. The allocation of power (summarized by “Pareto weights”¹) is therefore critical to understand many important household decisions such as labor market participation, time allocation within the household, public good expenditures, or fertility choices. However, it must be understood that in collective models, the Pareto weights and household formation are taken as given. In this paper, I propose to endogenize both, by merging collective models and matching models of the marriage market into a unique framework. To do so, I rely on the recent matching framework with imperfectly transferable utility (ITU) introduced in Galichon, Kominers, and Weber (2018) (GKW hereafter). My contribution is threefold. First, I characterize a class of collective models, coined “proper collective models”, that can be embedded in the ITU framework. Second, I develop the computational methods so that even the most complex proper collective models are empirically tractable. Finally, I provide an estimation strategy and an illustrative application.

The idea that the allocation of bargaining power within couples and marriage markets are somehow related is not new. It is implicit in many papers from the collective literature that use sex ratios as a distribution factor² (Chiappori, Fortin, and Lacroix 2002). Intuitively, a marriage market in which women are more scarce would see a shift of the balance of power in their favor. The same can be said about wage ratios

¹The usual way of modelling efficiency is that partners maximize a weighted sum of their utilities. These weights, called Pareto weights, are an unknown function of prices, income and power “shifters” known as distribution factors. Pareto weights provide an intuitive measure of bargaining power. An alternative, equivalent measure of power are sharing rules, which describe what share of the household resources are effectively received by one partner.

²Distribution factors are shifters that affect bargaining power, but not preferences or the budget set of the household members.

(Browning and Gørtz 2012), which is another example of commonly used distribution factors³. Implicitly, it is assumed that as women’s wage increase relatively to the wage of men, their outside option on the marriage market (be it staying single or divorcing) improves, and so does their bargaining power. In fact, Becker (1973) already pushed these ideas forward, by showing that the market structure (e.g. the relative supply of men and women) governs the division of marital surplus between partners. This suggests that there could be significant benefits from bringing together collective models and matching models of the marriage market into a unified framework. Or, to put it in Chiappori’s words, “the next step, obviously, would be an “upstream” theory that would *endogenize* both household composition (“who marries whom”) and the resulting intrahousehold allocation of power” (Chiappori 2017). This paper provides a way of achieving this goal. Therefore, in my framework, distribution factors have a priori no effect on the intra-household allocation of power. Instead, the way they influence power allocation arises naturally from the model itself.

Several papers have been successful in connecting matching and collective models, although they almost exclusively rely on a matching model with transferable utility (TU). In such models, it is assumed that some technology (money, a numeraire good, etc.) allows partners to transfer utility to each other at a constant exchange rate. This means the utility functions featured in the underlying collective model are restricted to belong to a certain class (Chiappori and Gugl 2014). On the other hand, we gain the computational tractability of the TU framework, since in that case the optimal matching on the marriage market is the one that maximizes total surplus. This allows us to use a battery of well-known computational techniques from linear programming and optimal transportation (Galichon 2016). However, it turns out that transferable utility is a demanding assumption. In that case, partners behave as a single decision maker: they maximize total utility (or surplus). Thus, public good consumption

³For an extensive list, see Browning, Chiappori, and Weiss (2014), chapter 5. Divorce laws, age and education differences, or relative income are commonly used distribution factors as well.

does not depend on the allocation of bargaining power. Power, here, does not matter (at least when it comes to explaining public good expenditures made by couples, a finding at odds with the empirical literature on collective models, see e.g. Attanasio and Lechene 2002).

Consequently, it is preferable to consider collective models in which utility is (possibly) imperfectly transferable. This means however, that we need the corresponding, fully fledged, matching framework featuring imperfectly transferable utility. In this paper, I rely on the recent work by Galichon, Kominers, and Weber (2018). They provide a general matching model with imperfectly utility and heterogeneity in tastes, that includes as particular cases matching with transferable utility, non-transferable utility, matching with taxes, etc. The GKW setting is first and foremost a matching framework. It is assumed that men (or firms) and women (or workers) meet on a frictionless marriage (or labor) market. The absence of frictions (which we retain throughout this paper) is an acceptable assumption in our case, as the focus here is on who marries whom and who gets what within couples. The framework is also completely static: divorce and remarriage are absent from the model. Upon meeting, partners bargain and choose a utility allocation⁴ (u, v) that belongs to some bargaining set \mathcal{F} . The frontier of the bargaining sets need not be a straight line of slope minus one, as would be the case with transferable utility. Nor these sets need to be convex. These sets must, however, meet a certain number of requirements, in which case they are called “proper” bargaining sets. To define equilibrium in the ITU framework, stability is used as the solution concept. In equilibrium, it must not be the case that a married individual would rather be single, or that any two individuals would rather be together than in their respective match (or singlehood state). Writing down stability conditions is not as trivial in the ITU case than it is in the transferable utility case. GKW make use of a specific parametrization of the bargaining set to write

⁴Therefore, partners make binding agreements on the marriage market. For an important alternative view with bargaining in marriage, see Pollak (2018).

down stability conditions, called *distance functions* (see section 2 for a reminder on GKW). The distance function is key for most of the results presented in that paper, and also plays an important role in estimating ITU models.

The setting introduced in GKW offers a natural way of embedding collective models into a matching framework with imperfectly transferable utility. Before doing so, however, there are two important issues to solve. First, bargaining sets are primitives in GKW, and they are assumed to be proper. However, it is unclear which class of utility functions (and collective models) will generate proper bargaining sets. Therefore, it is necessary to provide sufficient conditions on the underlying collective model for the resulting bargaining sets to be proper. Second, assuming that bargaining sets are proper, it is crucial to be able to compute distance functions for the model to be estimated. In some cases, the distance function can be derived in closed-form. However, as soon as a collective model features labor supply, public goods, time use, household production, or fertility choices, it will produce a bargaining set whose associated distance function cannot be known in closed-form (except in some trivial cases). This is a serious threat for the empirical tractability of the framework.

In this paper, I aim at addressing these concerns. I characterize collective models that will generate proper bargaining sets. I call these collective models “proper” collective models as well, because it is then possible to embed such models into the ITU machinery from GKW. I also provide a simple way of computing distance functions. The method works for any proper bargaining set and therefore is not restricted to be used with collective models, but could be applied to the other examples presented in GKW. In the case where a proper collective models has been used to generate the bargaining sets, the computation method also shows interesting connections with the collective model literature. Finally, I provide an estimation strategy based on a mathematical program with equilibrium constraints (MPEC) and illustrate the usefulness of my results with an application with non-trivial preferences. The collective model that I bring to the data is one in which partners consume a composite private good,

enjoy leisure, spend time on the labor market and invest time in housework to produce a public good. The model is estimated with PSID data. I make use of the application to show the strength of the approach advocated here. First and foremost, since Pareto weights are endogenized, we can recover them and know their distribution. Second, we can construct counterfactual experiments: I consider two cases of women empowerment, through an increase in the sex ratio (men become more abundant on the marriage market) or through an increase in women's potential wages.

An overview of the literature. The starting point of this paper are collective models. An excellent overview of the literature is provided in Browning, Chiappori, and Weiss (2014). The relevance of collective models with respect to the classical unitary framework has been illustrated in several papers, e.g. in Lundberg, Pollak, and Wales (1997) and Attanasio and Lechene (2014) for example. Collective models are particularly suited to a context in which the main source of public consumption are children. In this case, if partners have heterogenous tastes over public goods, a shift in bargaining power may influence the well-being of children. Such models have been studied in e.g. Blundell, Chiappori, and Meghir (2005) and applied in Cherchye, De Rock, and Vermeulen (2012). Aside from testing the relevance of the collective approach, most of the literature has focused on the identification of the Pareto weights/sharing rules (sometimes only up to a constant). Progress has been made in several directions. Browning and Bonke (2006) relied on direct survey evidence. Chiappori, Fortin, and Lacroix (2002) or Couprie (2007) use assignable goods (private goods whose consumption is observed for both partners) to identify the sharing rule up to a constant. Lise and Seitz (2011) identify the sharing rule parametrically using a discrete choice model of labor supply. Browning, Chiappori, and Lewbel (2013) and Lewbel and Pendakur (2008) recover the sharing rule while estimating nonlinear demand systems. Finally, several papers, e.g. Cherchye et al. (2015), have used a revealed preference approach to obtain bounds on the sharing rules.

In this paper, I take a different route by using the marriage market to endogenize Pareto weights. I draw insights from Becker’s work (Becker 1973), in which surplus sharing (utility is perfectly transferable in his paper) within couples depends endogenously on population supplies⁵. Despite these early contributions, attempts to reconcile collective models and models of the marriage market with transferable utility are only very recent. The mechanics are illustrated in Chiappori (2012), and Browning, Chiappori, and Weiss (2014) contains interesting examples. Full applications can be found in Choo and Seitz (2013) and Chiappori, Costa Dias, and Meghir (2017).

There have been very few attempts to integrate collective models into the general ITU framework. The main reference on matching with ITU for the present contribution is Galichon, Kominers, and Weber (2018), and papers cited therein. A notable contribution reconciling collective models and the marriage market is Cherchye et al. (2017), which uses a revealed preference approach to household consumption and stability on the marriage market to tighten the bounds on the sharing rule. An application of the Galichon, Kominers, and Weber (2018) framework can be found in Gayle and Shephard (2016), although that paper focuses on optimal taxation questions and contains a limited number of types.

Organization of the paper. Section 2 sets the stage by reminding the reader of the ITU matching framework. Section 3 characterizes proper collective models. In section 4, I introduce a computation method to solve general ITU models. Section 5 provides an estimation method based on a mathematical program with equilibrium constraints. In section 6, I provide an illustration with a collective model featuring private and public consumption as well as labor supplies. Finally, section 7 concludes.

⁵Becker (1973): “Theory does not take the division of output between mates as given, but rather derives it from the nature of the marriage market equilibrium”.

2. SETTING THE STAGE

2.1. The marriage market. Throughout this paper, I consider a frictionless⁶, bipartite, bilateral one-to-one marriage market. It is populated by men and women, indexed by $i \in \mathcal{I}$ and $j \in \mathcal{J}$, respectively, who may decide to form heterosexual pairs⁷. A matching is a dummy variable μ_{ij} that is equal to one if man i and woman j marry, and zero otherwise.

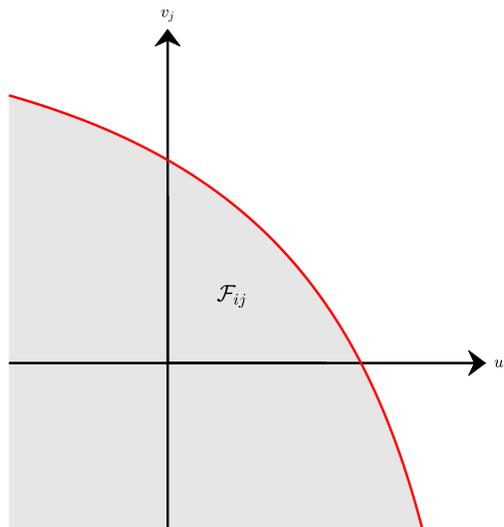
When staying single, agents receive payoffs \mathcal{U}_{i0} and \mathcal{V}_{0j} for men and women, respectively. Whenever a man i and a woman j meet on the marriage market, they bargain over utility outcomes (u, v) that lie in a bargaining set \mathcal{F}_{ij} .

2.2. Proper bargaining sets. The starting point in GKW are the bargaining sets, which serve as primitives of the model. It is assumed that these sets meet a certain number of requirements, in which case they are called *proper bargaining sets* (see definition 1 in GKW). For convenience, the definition of a proper bargaining set is reproduced below, while figure 1 shows a typical proper bargaining set.

Definition 1 (Galichon, Kominers and Weber, 2018). A bargaining set \mathcal{F}_{ij} is a *proper bargaining set* if the three following conditions are met: (i) \mathcal{F}_{ij} is closed and nonempty (ii) \mathcal{F}_{ij} is *lower comprehensive*: if $(u, v) \in \mathcal{F}_{ij}$, then $(u', v') \in \mathcal{F}_{ij}$ provided $u' \leq u$ and $v' \leq v$. (iii) \mathcal{F}_{ij} is *bounded above*: Assume $u_n \rightarrow +\infty$ and v_n bounded below then for N large enough (u_n, v_n) does not belong in \mathcal{F} for $n \geq N$; similarly for u_n bounded below and $v_n \rightarrow +\infty$.

⁶For state of the art applications of the search and matching models to the marriage market, see Goussé, Jacquemet, and Robin (2017) and Coles and Francesconi (2017).

⁷I choose to ignore same-sex couples, as they account for only 1% of couple households in the US in 2010. See Ciscato, Galichon, and Goussé (2015) for an empirical analysis of the marriage patterns of same-sex couples.

FIGURE 1. A typical proper bargaining set \mathcal{F}_{ij} 

2.3. Parametrizing the bargaining sets via distance functions. As in most matching frameworks, we will write down stability conditions in order to define equilibrium on the marriage market. However, this is not as trivial in the ITU case as it is with perfectly transferable utility. To reach our goal, we will make use of a parametrization of the bargaining sets that we call *distance functions*.

The *distance function* D is an analytical tool that describes whether a utility allocation (u, v) is a boundary point of the bargaining set or not. The distance function is not only useful to characterize the equilibrium in a simpler way, but also turns out to be crucial for estimation purposes. For a given pair (i, j) and a generic bargaining set \mathcal{F}_{ij} , the distance function is defined as

$$D_{ij}(u, v) = \min \{z \in \mathbb{R} : (u - z, v - z) \in \mathcal{F}_{ij}\}.$$

Note that, indeed, $D_{ij}(u, v) = 0$ whenever the pair of utilities (u, v) belongs to the boundary of \mathcal{F}_{ij} , $D_{ij}(u, v) \leq 0$ if the pair (u, v) is feasible, and $D_{ij}(u, v) > 0$ otherwise. This is illustrated in figure 2 for a feasible point (u_i, v_j) .

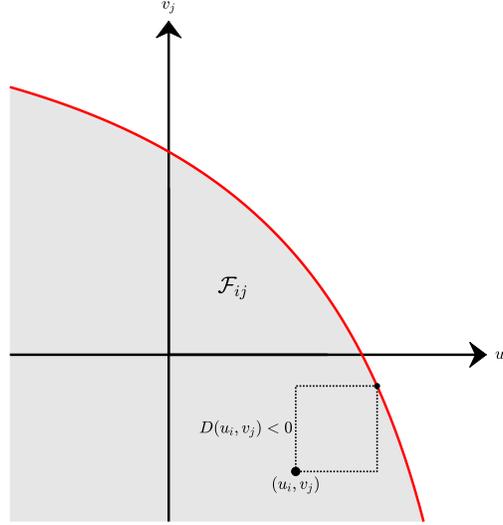


FIGURE 2. The distance function for a feasible point (u_i, v_j)

Stability conditions require that (i) for all men i and all women j , $u_i \geq \mathcal{U}_{i0}$ and $v_j \geq \mathcal{V}_{0j}$ (these are the usual participation constraints) and (ii) that for all i, j pairs, the equilibrium payoffs u_i and v_j satisfy

$$(2.1) \quad D_{ij}(u_i, v_j) \geq 0$$

Note that condition (2.1) holds with equality whenever man i and woman j are married in equilibrium. The intuition behind the stability conditions⁸ is as follows. Assume that (2.1) does not hold for some (i, j) pair, that is, $D_{ij}(u_i, v_j) < 0$ (such a point is represented in figure 2). But then, whenever man i and woman j meet, they can choose a feasible utility allocation so that both of them are strictly better off (any point in the square area in figure 2 satisfies this). Hence, man i and woman j form a blocking pair.

⁸Note that in a model with transferable utility where partners share a surplus Φ_{ij} , the distance function is $D_{ij}(u, v) = (u_i + v_j - \Phi_{ij})/2$, so that we recover the usual stability condition $u_i + v_j \geq \Phi_{ij}$.

2.4. Introducing heterogeneity in tastes. Starting from section 5 in the present paper, I will make use of the same empirical framework as in Galichon, Kominers, and Weber (2018). Namely, I will introduce *types* and heterogeneity in tastes.

Types are groups of men and women who share similar observable characteristics. The sets of types are denoted \mathcal{X} for men and \mathcal{Y} for women, while $\mathcal{X}_0 = \mathcal{X} \cup \{0\}$ and $\mathcal{Y}_0 = \mathcal{Y} \cup \{0\}$ introduce singlehood as an option (denoted $\{0\}$). I let n_x be the mass of men of type x , and m_y be the mass of women of type y . A matching is now a vector $(\mu_{xy})_{x \in \mathcal{X}, y \in \mathcal{Y}}$ that gives the mass of matches between men of type x and women of type y . The vectors $(\mu_{x0})_{x \in \mathcal{X}}$ and $(\mu_{0y})_{y \in \mathcal{Y}}$ denote the mass of single men of type x and single women of type y , respectively. Of course, a feasible matching must satisfy the following scarcity constraints

$$(2.2) \quad \begin{aligned} \mu_{x0} + \sum_y \mu_{xy} &= n_x \\ \mu_{0y} + \sum_x \mu_{xy} &= m_y \end{aligned}$$

Heterogeneity in tastes is introduced as follows. As in GKW, I assume that the payoffs received by man i and woman j are the sum of two components: (i) a systematic part, denoted U_i and V_j , respectively, that lie within the bargaining set \mathcal{F}_{ij} and over which the potential partners bargain, and (ii) an idiosyncratic part, denoted ϵ_{iy_j} and $\epsilon_{x_i j}$, respectively. In addition, it is assumed that the bargaining set only depends on the observable types, that is, $\mathcal{F}_{ij} = \mathcal{F}_{x_i y_j}$ ⁹.

One contribution of GKW is to show that we can actually focus on equilibrium where the systematic utilities, that I will denote $U_{x_i y_j}$ and $V_{x_i y_j}$, depend on the observable types of the partners. They show that doing so is in fact merely a restriction,

⁹Together, these assumptions are very similar to the “separability” assumption in the classical transferable utility framework with unobserved heterogeneity. See Galichon, Kominers, and Weber (2018) for a discussion.

as there always exists an individual equilibrium of this form. The advantage of focusing on such equilibria is that we can reduce the matching problem to a series of discrete choice problems, where each individual chooses the observable type of his potential partner.

To simplify further the analysis, I will assume that the taste shocks are i.i.d draws from an extreme value type I distribution. In this case, coined the “ITU-logit model”, it is well known that the systematic utilities can be obtained from the matching patterns using the log odds ratio formula (see McFadden (1974), and Choo and Siow (2006) for an application to the marriage market). Namely:

$$U_{xy} = \log \frac{\mu_{xy}}{\mu_{x0}} \text{ and } V_{xy} = \log \frac{\mu_{xy}}{\mu_{0y}}$$

In this setting, characterizing equilibrium is very simple. Indeed, the triple $(\mu_{xy}, U_{xy}, V_{xy})$ is an aggregate equilibrium (see definition 5 in GWK) if and only if the systematic utilities given by $U_{xy} = \log \frac{\mu_{xy}}{\mu_{x0}}$ and $V_{xy} = \log \frac{\mu_{xy}}{\mu_{0y}}$ satisfy

$$(2.3) \quad D_{xy}(\log \frac{\mu_{xy}}{\mu_{x0}}, \log \frac{\mu_{xy}}{\mu_{0y}}) = 0$$

and the matching satisfies the scarcity constraints (2.2). Equation (2.3) describes the usual feasibility/stability conditions, while the system of equations in (2.2) ensures the the equilibrium matching is feasible, given the population supplies (n_x, m_y) .

Combining these conditions, we get that in the ITU-logit framework, equilibrium is fully characterized by the system of equations

$$(2.4) \quad \begin{aligned} \mu_{x0} + \sum_y \exp(-D_{xy}(-\log \mu_{x0}, \log \mu_{0y})) &= n_x \\ \mu_{0y} + \sum_x \exp(-D_{xy}(-\log \mu_{x0}, \log \mu_{0y})) &= m_y \end{aligned}$$

In section 5, I will provide a new estimation method based on the above characterization of equilibrium.

Before going into section 3, I shall point out that while I stick to the vocabulary of marriage markets throughout the paper, my results apply to other matching markets as well, e.g. between workers and firms.

3. PROPER COLLECTIVE MODELS

Two steps are required to link collective models and models of matching with imperfectly transferable utility. First, assuming that we have some collective model in mind, we must show that the bargaining sets it generates are proper, as required by the Galichon, Kominers, and Weber (2018) matching framework. Second, since in equilibrium partners select a utility allocation that belongs to the frontier of the bargaining set, we must show that such a point is Pareto efficient. If these two conditions are met, then the collective model is said to be *proper*.

3.1. The primitives. The bargaining sets are, in GKW, the primitives of the model. In the present framework, however, the bargaining sets are not given. Instead, I assume that a specific collective model is the primitive. By that, I mean that men and women have preferences over the observable characteristics of their partner (non-economic marital gains) as well as over goods (private and possibly public goods as well) that they wish to consume (economic gains to marriage). Whenever man i meet woman j , they bargain and choose an allocation of goods ω from a feasible set Ω_{ij} , which depends on the characteristics of the partners (for example, wages and income) and technology available to them. A typical element of Ω_{ij} is an allocation $\omega = (q^a, q^b, Q)$, where q^a and q^b are vectors of private goods, consumed by the man and the woman respectively, and Q a vector of public goods. Man i and woman j receive utility $U_{ij}(q^a, Q)$ and $V_{ij}(q^b, Q)$, respectively. With some slight abuse of notation, I will drop the subscript ij in the remainder of the section, since we will be focusing on only one ij pair.

Remark 3.1 (Preferences). Note that preferences are egotistic. Note also that I am agnostic about where the goods are coming from: some could be bought on the market, some others could be produced within the household. The model can be extended to accommodate for the presence of other household members (children, for example). However, I assume that there are only two decision makers in the household, namely the husband and the wife (collective models can have more than two decision makers, see e.g. Dauphin et al. (2011)).

Thus, in this framework, the preferences of men and women (represented by the utility functions U and V) and the set of feasible allocations of goods, Ω , are primitives (which I summarize by saying that the collective model is the primitive). From them, we can derive the bargaining set as in GKW. As we shall see below, we can show that the resulting bargaining set is proper under mild assumptions. These assumptions are as follows:

Assumption 1 (Feasible Allocations). *The set of feasible good allocations, denoted Ω , is compact and convex.*

Assumption 2 (Utility Functions). *Preferences are represented by utility functions U and V that are upper semi continuous and bounded above on Ω and strictly increasing.*

Assumptions 1 and 2 are fairly standard in microeconomic theory. With the usual linear inequality budget and time constraints, the set Ω is indeed convex and compact. The assumption is still holding when some goods are home-produced and the production function is assumed to be concave. As we shall see, upper semi continuity is enough to prove closedness of the bargaining set. I also make the assumption that agents consume only “goods” (as opposed to “bads”) so that the utility is increasing in consumption. There are no negative externalities of consumption from any of the partners in this framework.

3.2. From the primitives to the bargaining sets under free disposal. As a first step, I make an assumption that will render trivial the construction of the bargaining sets and showing that they are proper, but will be relaxed later on. I do so to show that my approach (and the results that follow) is very general if one is willing to make this assumption. I assume that partners can engage in utility burning, i.e there is free disposal of utility. This assumption is discussed in the matching literature, see e.g. Galichon and Hsieh (2017).

Assumption 3 (Free disposal). *There is free disposal of utility. That is, if $u' \leq u$, $v' \leq v$ and if (u, v) is feasible, then (u', v') is feasible.*

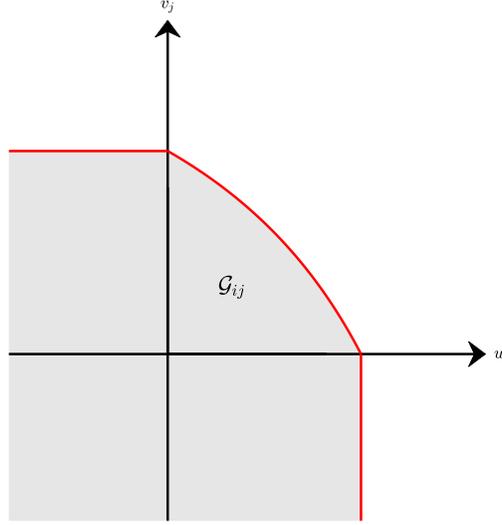
Assumption 3 means that we can construct a bargaining set, denoted \mathcal{G} , associated with (Ω, U, V) as the *utility possibility set*

$$\mathcal{G} = \{(u, v) : \exists (q^a, q^b, Q) \in \Omega : u \leq U(q^a, Q), v \leq V(q^b, Q)\}$$

By definition then, this set satisfies lower-comprehensiveness. A graphical example of the set \mathcal{G} is given in figure 3. We need only to show that this bargaining set satisfies the other requirements to be proper.

Proposition 1. *Under assumptions 1, 2 and 3, the bargaining set \mathcal{G} associated with (Ω, U, V) is a proper bargaining set*

Proof. The fact that \mathcal{G} is lower comprehensive follows directly from its definition and the assumption of free disposal. Let us show that \mathcal{G} is closed. Consider a convergent sequence $\{w_n\}_{n \in \mathbb{N}}$ in \mathcal{G} such that $w_n = (u_n, v_n) \in \mathcal{G}, \forall n \in \mathbb{N}$. For each w_n , $\exists \omega_n \in \Omega : u_n \leq U(q_n^a, Q_n), v_n \leq V(q_n^b, Q_n)$. Since Ω is compact, there exist a convergent subsequence $\{\omega_{f(n)}\}_{n \in \mathbb{N}}$ in Ω (by the Bolzano–Weierstrass theorem). From $\{\omega_{f(n)}\}_{n \in \mathbb{N}}$, construct the subsequence $\{w_{f(n)}\}_{n \in \mathbb{N}}$. For any w_k in that subsequence, $\exists \omega_k \in \Omega : u_k \leq U(q_k^a, Q_k), v_k \leq V(q_k^b, Q_k)$. Since $u_k \rightarrow u, v_k \rightarrow v$ and $\omega_k \rightarrow \omega = (q^a, q^b, Q) \in \Omega$, we see by continuity of U and V that $u \leq U(q^a, Q)$ and $v \leq V(q^b, Q)$.

FIGURE 3. The bargaining set \mathcal{G} 

Therefore $(u, v) \in \mathcal{G}$. Point (iii) in the definition of a proper bargaining set follows immediately from the fact that U and V are bounded above on Ω . ■

We have just shown that under mild conditions, we can actually construct the bargaining sets primitives in GKW as the *utility possibility sets* associated to the underlying collective model, and these sets are proper. We also know that in the ITU setting, equilibrium utilities (payoffs) belong to the boundary of the bargaining set. What can we say about these points in terms of Pareto efficiency?

Proposition 2. *Whenever the set \mathcal{G} is a proper bargaining set, a boundary point of \mathcal{G} is weakly Pareto efficient.*

Proof. The proof follows immediately from the fact that if \mathcal{G} is a proper bargaining set, it is lower comprehensive and closed. Indeed, take any (u, v) on the boundary of \mathcal{G} . If (u, v) is not weakly Pareto efficient, $\exists(u', v') \in \mathcal{G}$ such that $u' > u$ and $v' > v$, hence (u, v) is an interior point (consider the open ball of center (u, v) and radius ϵ B_ϵ where $\epsilon < \min(u' - u, v' - v)$). ■

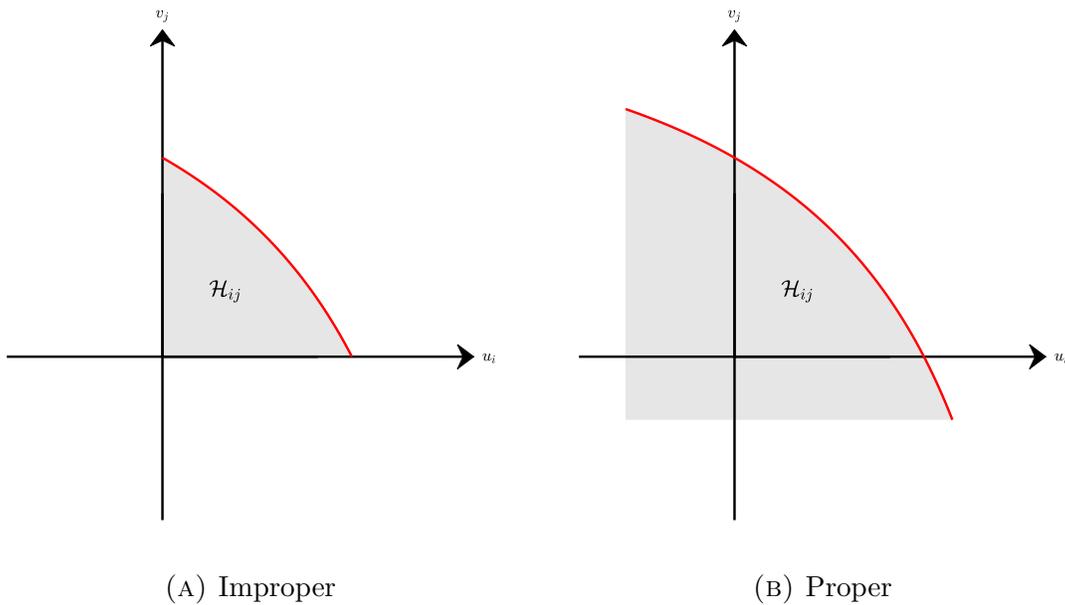
3.3. Relaxing free disposal. The assumptions made in the previous section allow us to cover a great variety of models. However, if we are to focus on a typical collective model as primitive, we must relax the free disposal assumption. I will also assume that we choose a strictly concave cardinal representation of utilities. In this section, we will have to proceed differently to construct the bargaining sets, but under a few additional assumptions it can be shown that they are proper as well.

First, suppose assumptions 1 and 2 hold, and construct the following bargaining set \mathcal{H} associated with (Ω, U, V) as the *strict utility possibility set* \mathcal{H}

$$\mathcal{H} = \{(u, v) : \exists (q^a, q^b, Q) \in \Omega : u = U(q^a, Q), v = V(q^b, Q)\}$$

Figure 4 shows two examples of a bargaining set \mathcal{H} , one being proper, the other one is not.

FIGURE 4. The bargaining set \mathcal{H}



Showing that the bargaining set \mathcal{H} is proper is, unfortunately, impossible without additional assumptions. And, once again, if we want to connect the ITU framework

to collective models, we need to provide the following two results: (i) if \mathcal{H} is a proper bargaining set, then a boundary point is Pareto efficient, and (ii) \mathcal{H} is indeed a proper bargaining set.

Let us start by showing (ii), assuming that the set \mathcal{H} is a proper bargaining set. We will make the following two assumptions:

Assumption 4 (Transferability). *If $(q^a, q^b, Q) \in \Omega$ and $q_k^a > 0$ for some private good k , then for any $q^{a'}$ such that $q_{-k}^{a'} = q_{-k}^a$, $q_k^{a'} < q_k^a$ and $(q^{a'}, q^b, Q) \in \Omega$, there is a $q^{b'}$ and some private good l such that $q_{-l}^{b'} = q_{-l}^b$, $q_l^{b'} > q_l^b$ and $(q^{a'}, q^{b'}, Q) \in \Omega$.*

Assumption 5 (Concave cardinal representation). *U and V are strictly quasi-concave*

Assumption 4 states that for any feasible allocation in which the man (or the woman) consumes a strictly positive amount of some private good, we can always find a feasible allocation that is identical except for the fact that we slightly decreased the consumption of the man for that private good and slightly increased the consumption of the woman for some private good. All in all, assumption 4 states that there is always some ways (even if imperfect) of transferring utility from one partner to the other. Assumption 5 is again very standard in the collective model literature. This will ensure that the bargaining sets are convex, which was not required in the previous section. We can now get the following result:

Proposition 3. *Whenever the strict utility possibility set \mathcal{H} is a proper bargaining set, and under assumption 1, 2, 4, and 5, a boundary point (u, v) of \mathcal{H} is strongly Pareto efficient.*

Proof. Choose any (u, v) on the boundary of \mathcal{H} attained with the allocation $\omega = (q^a, q^b, Q)$ and suppose it is not Pareto efficient. Then, there is a point $(u', v') \in \mathcal{H}$ attained for some allocation $\omega' = (q^{a'}, q^{b'}, Q')$ such that $u' \geq u$ and $v' \geq v$, with at least one strict inequality. Without loss of generality, say $u' > u$. Note that $v' > v$ yields to a contradiction since (u, v) would be an interior point (see the proof of

Proposition 2). Hence, $u' > u$ and $v' = v$. To reach u' , we need to change the amount of goods consumed by the man, so that either $q^{a'} \neq q^a$ or $Q' \neq Q$, or both.

If $Q' \neq Q$, then we can show that (u, v) is an interior point, a contradiction. Indeed, the allocation $(tq^a + (1-t)q^{a'}, tq^b + (1-t)q^{b'}, tQ + (1-t)Q') \in \Omega$ for $t \in (0, 1)$ by convexity of Ω , and gives both partners utility (u'', v'') with $u'' > u$ and $v'' > v$.

If $Q' = Q$, then $q^{a'} \neq q^a$ and there is at least one private good k for which $q_k^{a'} > q_k^a \geq 0$ since U is strictly increasing. By continuity of U and V , and Assumption 4, we can always slightly decrease $q_k^{a'}$ and slightly increase q_l^b for some private good l in a feasible way, and reach a point (u'', v'') such that $u' > u'' > u$ and $v'' > v$. Hence, (u, v) is an interior point, a contradiction. ■

Using the *strict utility possibility set* \mathcal{H} as our bargaining set is attractive because under the conditions mentioned above, the ITU framework imposes that in equilibrium partners choose a Pareto efficient outcome. Hence, the remaining piece of the puzzle is showing that we can indeed use \mathcal{H} as a bargaining set by proving that it is proper.

Assumption 6 (Vital goods). *There exists two private goods, q_1^a and q_1^b such that*

- (i) $\lim_{q_1^a \rightarrow 0^+} U(q^a, Q) = -\infty$ and $\lim_{q_1^b \rightarrow 0^+} V(q^b, Q) = -\infty$
- (ii) for any $(q^a, q^b, Q) \in \Omega$, $((q_1^{a'}, q_{-1}^a), (q_1^{b'}, q_{-1}^b), Q) \in \Omega$ whenever $q_1^{a'} \leq q_1^a$ and $q_1^{b'} \leq q_1^b$

Assumption 6 states that both men and women consume a vital private good, in the sense that if the quantity consumed tends to zero, their utility tends to minus infinity. It also states that if the household decreases the quantity of these goods, he can still at least buy the same amount of the other goods. In effect, this assumption plays the same role than the free-disposal restriction: if the couple agrees on an utility allocation (u, v) , it is always possible to reach an allocation $(u', v') : u' \leq u, v' \leq v$ by

decreasing the amount of vital goods consumed. In practice, one can think of food or water filling this role¹⁰. We now reach our final result:

Proposition 4. *Under Assumption 1, 2 and 6, the strict utility possibility set \mathcal{H} associated with (Ω, U, V) is a proper bargaining set*

Proof. Closedness and boundedness are proven exactly as in Proposition 1. We need to show that \mathcal{H} is lower comprehensive. Take $(u, v) \in \mathcal{H}$, with $u = U(q^a, Q)$ and $v = V(q^b, Q)$ for some allocation $(q^a, q^b, Q) \in \Omega$ and consider a point (u', v') such that $u' \leq u, v' \leq v$. Without loss of generality, take $u' < u$. By Assumption 2 and 6, we can find $q_1^{a'}$ such that $((q_1^{a'}, q_{-1}^{a'}, q^b, Q) \in \Omega, U((q_1^{a'}, q_{-1}^{a'}, Q) = u'$ and $V(q^b, Q) = v$.

If $v' = v$, then we stop.

If $v' < v$, there is a $q_1^{b'}$ such that $((q_1^{a'}, q_{-1}^{a'}, (q_1^{b'}, q_{-1}^{b'}), Q) \in \Omega, U((q_1^{a'}, q_{-1}^{a'}, Q) = u'$ and $V((q_1^{b'}, q_{-1}^{b'}), Q) = v'$. Therefore, $(u', v') \in \mathcal{H}$. ■

We can now combine all the results in the following corollary:

Corollary 1. *Under Assumption 1, 2, 4, 5 and 6, the bargaining set \mathcal{H} is a proper bargaining set and any point on the boundary of \mathcal{H} is strongly Pareto efficient.*

The proof of this corollary follows from putting together the previous results. For simplicity, I refer to collective models satisfying the assumptions invoked in corollary 1 as *proper collective models*. A formal definition is given below.

Definition 2. A collective model that satisfies assumptions 1, 2, 4, 5 and 6 is a *proper collective model*. Indeed, the resulting bargaining set \mathcal{H} is proper and convex and any point on the frontier is Pareto-efficient.

4. A COMPUTATION METHOD

From now on, I am assuming that we are equipped with a proper collective model. We have also chosen a convenient functional form for the utility functions of men and

¹⁰Although food preparation may be a public good, the consumption of food is private.

women. The (preference) parameters that enter the utility functions are gathered in a vector denoted θ ¹¹. In this section, I will discuss how to compute the distance function in the bargaining sets associated to such model (θ being given). In section 5, I will discuss how to estimate θ .

4.1. Computational issues behind distance functions. Unfortunately, distance function are not, in general, known in closed-form. In fact, as soon as the underlying collective model features multiple budget and time constraints, non-trivial preferences for public goods, household production or has possibly corner solutions, then it is not possible to obtain an analytical formula for the distance function. This is best illustrated by the following example.

Example 1. *Consider a case in which utility depends on the consumption of a private good q and a public good Q . All prices are set to one. The systematic utilities are as follows*

$$U_{ij}(q_{ij}^a, Q_{ij}) = \delta_{ij}^a + \log q_{ij}^a + A \log Q_{ij}$$

$$V_{ij}(q_{ij}^b, Q_{ij}) = \delta_{ij}^b + \log q_{ij}^b + B \log Q_{ij}$$

where I assume $A \neq B$. The terms δ_{ij}^a and δ_{ij}^b are non-economic gains to marriage. Normalizing prices to one and denoting Φ_{ij} the total income of pair (i, j) , the budget constraint is $q_{ij}^a + q_{ij}^b + Q_{ij} \leq \Phi_{ij}$. According to section 3, the bargaining set \mathcal{H}_{ij} associated with this model is a proper bargaining set. To compute the distance function, we make use of the first order conditions (FOCs) in the corresponding collective model. We get the standard Bowen-Lindhal-Samuelson (BLS) condition

$$Aq_{ij}^a + Bq_{ij}^b = Q_{ij}$$

Say we want to compute the distance at point (u, v) , that is, we want to find z such that $(u - z, v - z)$ belongs to the frontier. It is easily seen that the BLS condition

¹¹For example, we may assume that $U(q^a, Q) = \theta_0 + \theta_1 \log q^a + \theta_2 \log Q$, so that $\theta = (\theta_0, \theta_1, \theta_2)$.

defines Q_{ij} as an implicit function of $(u - z, v - z)$ with

$$(4.1) \quad \begin{aligned} & A \exp(u - z - \delta_{ij}^a - A \log Q_{ij}(u - z, v - z)) + \\ & B \exp(v - z - \delta_{ij}^b - B \log Q_{ij}(u - z, v - z)) = Q_{ij}(u - z, v - z) \end{aligned}$$

Unless $A = B$, there is no closed-form formula for the optimal Q_{ij} and therefore no closed-form for the distance function, which, in this case is

$$\begin{aligned} & \min_{z \in \mathbb{R}} \{ z : \exp(u - z - \delta_{ij}^a - A \log Q_{ij}(u - z, v - z)) \\ & \quad + \exp(v - z - \delta_{ij}^b - B \log Q_{ij}(u - z, v - z)) + Q_{ij}(u - z, v - z) \leq \Phi_{ij} \} \end{aligned}$$

This example brings two remarks to mind. First, there is little computational gains in writing down the collective model and making use of the FOC (such as the Bowen-Lindhal-Samuelson condition). Indeed, we still need to minimize over z and for each value of z we must solve for the optimal Q using numerical methods. Second, this example is still relatively simple because we only need to make use of one constraint (the budget constraint) to compute the distance function. We do not need worry about non negativity constraints since we work with logs, and there are no time constraints, corner solutions or public good production here. As we shall see, there are no easy way of solving more complex models.

4.2. Method. In the following, I introduce a method to compute distance functions when bargaining sets are constructed from proper collective models. In appendix B, I show that a similar method can be used whenever a bargaining set is proper (and so the method is not restricted to those bargaining sets generated by proper collective models). As we shall see below, computation relies on the very definition of the distance function but has interesting implications. While I keep things general, I will assume that the set of feasible allocations Ω_{ij} is fully characterized by R feasibility

constraints

$$(4.2) \quad g_{ij}^r(q_{ij}^a, q_{ij}^b, Q_{ij}) \leq 0, \quad r \in \{1, \dots, R\}$$

where $\{g_{ij}^r\}_{r \in \{1, \dots, R\}}$ are convex functions¹².

Problem 1. *Suppose that \mathcal{H}_{ij} is a proper bargaining set associated with some proper collective model, with preferences and set of feasible goods $(U_{ij}^\theta, V_{ij}^\theta, \Omega_{ij})$. Then we can solve*

$$(4.3) \quad \min_{z_{ij}, q_{ij}^a, q_{ij}^b, Q_{ij}} z_{ij}$$

$$s.t \quad u - z_{ij} = U_{ij}^\theta(q_{ij}^a, Q_{ij})$$

$$(4.4) \quad v - z_{ij} = V_{ij}^\theta(q_{ij}^b, Q_{ij})$$

$$g_{ij}^r(q_{ij}^a, q_{ij}^b, Q_{ij}) \leq 0, \quad r \in \{1, \dots, R\}$$

when $D_{ij}^\theta(u, v) = z_{ij}^*$, solution to the above program.

In practice, it is very fast to solve such optimization programs, and given the solution z_{ij}^* to this problem, we have $D_{ij}^\theta(u, v) = z_{ij}^*$. To ease computations, note that it is straightforward to provide the analytic gradient of the objective function to the solver, and that, in general, the analytic expression of the jacobian of the system of constraints can be obtained in analytic form as well.

Remark 4.1 (Simplifying problem 1). The above technique is fairly general, but of course, it can be combined with the approach laid out in example 1 to reduce the number of variables over which the minimization is performed. For instance, we saw in example 1 that Q_{ij} can be immediately deduced from q_{ij}^a and q_{ij}^b making use of the Bowen-Lindhal-Samuelson condition. Therefore, we can avoid minimizing over Q_{ij} in problem 1 above. In addition, solving problem 1 might not be necessary when

¹²Typically, the $\{g_{ij}^r\}_{r \in \{1, \dots, R\}}$ functions will be the usual linear budget and time constraints, but could also contain the production function of home-produced public goods.

some goods are discrete, in which case one can apply lemma 2 from GWK. Lemma 2 in that paper states that if the proper bargaining set is the union (intersection) of basic proper bargaining sets, then the distance function is the minimum (maximum) over the distance functions associated to each of these basic sets. It may well be the case that these distance functions are known in closed-form when conditioning on the discrete good, in which case problem 1 is not necessary. Therefore, lemma 2 is particularly useful for applications in which, for example, time is discrete or fertility is modelled.

Remark 4.2 (Parallel computing). Solving problem 1 is simple, but the process has to be repeated many times and this is computationally intensive. On a small marriage market with 100 men and 100 women, we must compute the distance function 10,000 times. Fortunately, the process of computing the distance function can easily be parallelized. In addition, problem 1 is solvable for multiple (i, j) pairs at once. Let us denote $\mathcal{IJ} \equiv \{1, \dots, |\mathcal{I}||\mathcal{J}|\}$ the set indexing all possible ij pairs and $\bar{\mathcal{IJ}}$ a subset of \mathcal{IJ} . To find the distance function for all ij pairs in some set $\bar{\mathcal{IJ}}$ at once, we need to minimize $\sum_{ij \in \bar{\mathcal{IJ}}} z_{ij}$ with respect to $\{z_{ij}, q_{ij}^a, q_{ij}^b, Q_{ij}\}_{ij \in \bar{\mathcal{IJ}}}$ subject to the constraints in (4.4) written for each of the $ij \in \bar{\mathcal{IJ}}$ pairs. In practice, mixing these two approaches yields the best result. That is, we compute the distance function for subsets of \mathcal{IJ} that form a partition of \mathcal{IJ} , when each subset is sent to a different worker.

4.3. Gradient of the distance function. When the number of men and women on the market is large, it is computationally costly to obtain the derivative of the distance function for each pair (i, j) using standard numerical methods. Fortunately, the following results shows how to obtain the derivative of the distance function D_{ij}^θ with respect to u, v or θ at the cost of one evaluation of the distance function.

Theorem 1. *Assume we are equipped with a proper collective model and associated proper bargaining set \mathcal{H}_{ij} . Introduce the Lagrange multipliers associated with the constraints in problem 1 as λ_1, λ_2 , and $\{\xi_r\}_{r \in \{1, \dots, R\}}$, respectively. The gradient with*

respect to u , v and θ of the distance function D_{ij} can be obtained directly when solving problem 1. Indeed, we have

$$\partial_u D_{ij}^\theta(u, v) = \lambda_1$$

$$\partial_v D_{ij}^\theta(u, v) = \lambda_2$$

$$\partial_\theta D_{ij}^\theta(u, v) = -\lambda_1 \partial_\theta U_{ij}^\theta(q_{ij}^a, Q_{ij}) - \lambda_2 \partial_\theta V_{ij}^\theta(q_{ij}^b, Q_{ij})$$

Proof. The proof follows from the Envelop Theorem, since

$$D_{ij}^\theta(u, v) = \min_{z_{ij}, q_{ij}^a, q_{ij}^b, Q_{ij}} z_{ij} \text{ s.t (4.4).}$$

■

Remark 4.3 (Gradient of the decision variables). Solving problem 1 also returns optimal values for q_{ij}^a , q_{ij}^b , and Q_{ij} . Can we compute the derivative of these values with respect to u , v or θ ? While there is no general way of proceeding, it is usually possible. We can combine the first order conditions of our minimization problem and use the implicit function theorem to recover the derivatives. I will illustrate this point with an example. Suppose that the utility functions are given by

$$U_{ij} = \delta_{ij}^a + a \log q_{ij}^a + A \log Q_{ij}$$

$$V_{ij} = \delta_{ij}^b + b \log q_{ij}^b + B \log Q_{ij}$$

and we are interested in computing the derivative of the optimal q_{ij}^a , q_{ij}^b , and Q_{ij} with respect to a . Suppose that the only constraint is $q_{ij}^a + q_{ij}^b + Q_{ij} \leq \Phi_{ij}$ where Φ_{ij} is the income of pair (i, j) . We can see that in this model, the utility possibility set \mathcal{H}_{ij} is

a proper bargaining. It is easy to see that a solution to problem 1 satisfies

$$\begin{aligned} q_{ij}^a + q_{ij}^b + Q_{ij} &= \Phi_{ij} \\ u - D_{ij}^\theta(u, v) &= a \log q_{ij}^a + A \log Q_{ij} \\ \frac{A}{a} q_{ij}^a + \frac{B}{b} q_{ij}^b &= Q_{ij} \end{aligned}$$

where the first equality comes from the budget constraint (which is binding), the second is an equality constraint in problem 1 and the third is obtained by combining the first order conditions from the underlying collective model. Therefore we have

$$\begin{pmatrix} \frac{a}{q_{ij}^a} - \frac{A}{Q_{ij}} & -\frac{A}{Q_{ij}} \\ \frac{A}{a} + 1 & \frac{B}{b} + 1 \end{pmatrix} \begin{pmatrix} \partial_a q_{ij}^a \\ \partial_a q_{ij}^b \end{pmatrix} = \begin{pmatrix} -\partial_a D_{ij}^\theta(u, v) - \log q_{ij}^a \\ \frac{A}{a^2} q_{ij}^a \end{pmatrix}$$

which we can rewrite $M \times \begin{pmatrix} \partial_a q_{ij}^a \\ \partial_a q_{ij}^b \end{pmatrix} = d$, so that

$$\begin{pmatrix} \partial_a q_{ij}^a \\ \partial_a q_{ij}^b \end{pmatrix} = M^{-1} \times d$$

and $\partial_a Q_{ij}$ follows easily.

4.4. Connection with collective models. The computation methods introduced in the previous subsection highlights an interesting connection with collective models. It can be shown that the Lagrange multipliers associated with the constraint involving the utility functions of the partners in (4.4) can be interpreted as Pareto weights. This relationship is proven in the following theorem:

Theorem 2. *Suppose we are equipped with a proper collective model and associated (proper) bargaining set \mathcal{H}_{ij} . Suppose that Ω_{ij} is characterized by a set of feasibility constraints $g_{ij}^r(q_{ij}^a, q_{ij}^b, Q_{ij}) \leq 0$, $r \in \{1, \dots, R\}$, where $\{g_{ij}^r\}$ are convex functions. Then*

(i) *given u , v and θ , the allocation $q_{ij}^{a*}, q_{ij}^{b*}, Q_{ij}^*$ solution to problem 1 is Pareto efficient.*

(ii) in addition, the Pareto weights associated to the boundary point $(u - z^*, v - z^*)$ (reached via the allocation of goods $q_{ij}^{a*}, q_{ij}^{b*}, Q_{ij}^*$) for the man and the woman are, respectively, the Lagrange multipliers λ_1^* and λ_2^* . λ_1^* and λ_2^* also satisfy $\lambda_1^* + \lambda_2^* = 1$

Proof. To prove (i), we need to show that the point $(u - z_{ij}^*, v - z_{ij}^*)$ (where z_{ij}^* solves problem 1) is a boundary point of \mathcal{H}_{ij} , and therefore it is Pareto efficient by Proposition 3. If $(u - z_{ij}^*, v - z_{ij}^*)$ is not a boundary point of \mathcal{H}_{ij} , then there is an open ball of center $(u - z_{ij}^*, v - z_{ij}^*)$ and radius ϵ completely contained in \mathcal{H}_{ij} . Therefore, the point $(u - z_{ij}^* + \frac{\epsilon}{2}, v - z_{ij}^* + \frac{\epsilon}{2})$ is also in \mathcal{H}_{ij} , but then z_{ij}^* cannot be solution to problem 1.

To prove (ii), we write the Lagrangian associated to problem 1, except that we maximize $-z_{ij}$ instead of minimizing z_{ij} :

$$\begin{aligned} \mathcal{L}(z_{ij}, q_{ij}^a, q_{ij}^b, Q_{ij}, \lambda_1, \lambda_2, \xi_r) &= -z_{ij} - \lambda_1(u - z_{ij} - U_{ij}^\theta(q_{ij}^a, Q_{ij})) \\ &\quad - \lambda_2(u - z_{ij} - V_{ij}^\theta(q_{ij}^b, Q_{ij})) \\ &\quad - \sum_r^R \xi_r g_{ij}^r(q_{ij}^a, q_{ij}^b, Q_{ij}) \end{aligned}$$

For a solution point $(z_{ij}^*, q_{ij}^{a*}, q_{ij}^{b*}, Q_{ij}^*)$, I introduce the associated nonnegative multipliers $\lambda_1^*, \lambda_2^*, \xi_r^*$. Consider now the following Lagrangian:

$$\tilde{\mathcal{L}}(q_{ij}^a, q_{ij}^b, Q_{ij}, \xi_r) = \lambda_1^* U_{ij}^\theta(q_{ij}^a, Q_{ij}) + \lambda_2^* V_{ij}^\theta(q_{ij}^b, Q_{ij}) - \sum_r^R \xi_r g_{ij}^r(q_{ij}^a, q_{ij}^b, Q_{ij})$$

and note that (i) $\tilde{\mathcal{L}}$ is a concave function of $(q_{ij}^a, q_{ij}^b, Q_{ij})$, (ii) taking $\xi_r = \xi_r^*$ and

$(q_{ij}^a, q_{ij}^b, Q_{ij}) = (q_{ij}^{a*}, q_{ij}^{b*}, Q_{ij}^*)$, we have $\xi_r^* \geq 0$ and

$$\frac{\partial \tilde{\mathcal{L}}(q_{ij}^{a*}, q_{ij}^{b*}, Q_{ij}^*, \xi_r^*)}{\partial q_{ij}^a} = 0, \quad \frac{\partial \tilde{\mathcal{L}}(q_{ij}^{a*}, q_{ij}^{b*}, Q_{ij}^*, \xi_r^*)}{\partial q_{ij}^b} = 0, \quad \text{and} \quad \frac{\partial \tilde{\mathcal{L}}(q_{ij}^{a*}, q_{ij}^{b*}, Q_{ij}^*, \xi_r^*)}{\partial Q_{ij}} = 0$$

Therefore, $(q_{ij}^{a*}, q_{ij}^{b*}, Q_{ij}^*)$ is solution to

$$\begin{aligned} & \max_{q_{ij}^a, q_{ij}^b, Q_{ij}} \lambda_1^* U_{ij}^\theta(q_{ij}^a, Q_{ij}) + \lambda_2^* V_{ij}^\theta(q_{ij}^b, Q_{ij}) \\ & \text{s.t. } g_{ij}^r(q_{ij}^a, q_{ij}^b, Q_{ij}) \leq 0, \quad r \in \{1, \dots, R\} \end{aligned}$$

Hence, $(q_{ij}^{a*}, q_{ij}^{b*}, Q_{ij}^*)$ is a Pareto efficient allocation that maximizes a social welfare function with Pareto weights λ_1^* and λ_2^* . The fact that $\lambda_1^* + \lambda_2^* = 1$ follows directly from the first order conditions of problem 1, since $\partial \mathcal{L} / \partial z_{ij} = 0 \iff -1 + \lambda_1 + \lambda_2 = 0$. ■

The proof of theorem 2 makes use of the expression of the Lagrangian of the problem of computing the distance function. The main idea is to fix the irrelevant variables and recover the expression of the Lagrangian of a collective model problem. This result is important in theory as well as in practice, since it states that we can compute the Pareto weights from the Lagrange multipliers (which will be provided by any solver).

Remark 4.4 (From bargaining sets to collective models). Theorem 2 illustrates that abstract concepts from GKW, e.g. distance functions, can be related to familiar objects of the collective framework. This is not surprising because whenever a collective model is proper, the Pareto weight associated to some boundary point of the Pareto set is in fact equal to (minus) the slope of the frontier at that particular point. And distance functions are nothing else than an alternative way of parametrizing the frontier. To see that, let us consider a proper collective model. The usual way of modelling efficiency is to solve

$$\max_{(q^a, q^b, Q) \in \Omega} \mu / (1 - \mu) U(q^a, Q) + V(q^b, Q)$$

where $\mu / (1 - \mu)$ is the Pareto weight. Or equivalently, we can solve

$$\max_{(q^a, q^b, Q) \in \Omega} V(q^b, Q)$$

subject to the constraint $U(q^a, Q) \geq u$ for some utility level u . In the latter formulation, it is clear that the Lagrange multiplier associated to the constraint $U(q^a, Q) \geq u$ is the Pareto weight $\mu/(1 - \mu)$. I denote $\Psi(u)$ the value function in this second formulation of the optimization program, for some utility level u . Note that the equation of the Pareto frontier is therefore given by $v = \Psi(u)$. Hence, applying the envelop theorem, we see that minus the slope of the efficient frontier is in fact the Pareto weight, that is $-\partial_u \Psi(u) = \mu/(1 - \mu)$. Since in this class of collective models, the Pareto frontier is strictly concave, decreasing and smooth, there is a one-to-one relationship between the Pareto weight and the position along the efficient frontier.

What theorem 2 describes is the relationship between the Pareto weight and the implicit representation of the efficient frontier given by the distance function. Indeed we have, by definition,

$$D(u, v) = 0$$

whenever (u, v) belongs to the frontier, which again implicitly defines the mapping $\Psi : u \rightarrow v$ such that:

$$D(u, \Psi(u)) = 0$$

Differentiating, we get $\partial_u D + \partial_v D \partial_u \Psi = 0$, that is

$$\mu/(1 - \mu) \equiv -\partial_u \Psi(u) = \frac{\partial_u D}{\partial_v D}$$

where we recover the result from theorem 2.

Aside from distance functions, the utility “wedge” introduced in GKW is another abstract object that proves to be connected to the notion of bargaining power. Let me first recall the definition of the wedge. For any proper bargaining set \mathcal{F} , GKW show that there exists two unique functions \mathcal{U} and \mathcal{V} ¹³ such that the set of frontier points $\{(u, v) : D(u, v) = 0\}$ is given by $\{(\mathcal{U}(w), \mathcal{V}(w)), w \in W\}$ for some open interval W in \mathbb{R} . The parameter w , called “wedge”, parametrizes the efficient frontier. Suppose

¹³These functions are $\mathcal{U}(w) = -D(0, -w)$ and $\mathcal{V}(w) = -D(w, 0)$.

now that a bargaining set \mathcal{F} is proper, and its frontier is smooth, strictly decreasing and concave. Making use of the explicit representation, we have

$$D(\mathcal{U}(w), \mathcal{V}(w)) = 0$$

for all $w \in W$. Thus $\partial_u D \partial_w \mathcal{U} + \partial_v D \partial_w \mathcal{V} = 0$, hence

$$\begin{aligned} \frac{\partial_u D}{\partial_v D} &= -\frac{\partial_w \mathcal{V}}{\partial_w \mathcal{U}} \\ \mu/(1-\mu) &= -\frac{\partial_w \mathcal{U} - 1}{\partial_w \mathcal{U}} = \frac{1}{\partial_w \mathcal{U}} - 1 \\ &= -\frac{1}{\partial_w D(0, -w)} - 1 \end{aligned}$$

Therefore, in the context of proper collective models, the wedge and the Pareto weights are two equivalent ways of parametrizing the efficient frontier. It is true, however, that the approach using wedges is much more general because it can accommodate bargaining sets whose frontier features kinks and flat segments. In these cases, the approach via Pareto weights would fail because a point on the frontier can be associated with an infinity of Pareto weights (kinks) or a Pareto weight can be associated with an infinity of points on the frontier (flat segments). In contrast, the correspondence between the wedge and the location on the frontier is always one-to-one.

5. ESTIMATION

In this section, I present a new approach to estimate a proper collective model within the ITU-logit framework (see section 6 of GKW or section 2 of the present paper for a review), based on a mathematical program with equilibrium constraints. Note that section 3 and appendix B imply that more general models could be estimated using the computation and estimation methods presented here, as long as they generate proper bargaining sets. For simplicity however, I focus on the case of proper collective models.

5.1. Parameterization. A key difference with the GKW setting is that I will use parameterizations of the systematic utilities U_{xy} and V_{xy} that are very common in the collective model literature. Specifically, I will choose a parametric proper collective model, following definition 2.

When a man of type x meets with a woman of type y , they choose vectors of private consumption, denoted q_{xy}^a and q_{xy}^b , as well as a vector of public consumption Q_{xy} . Naturally, the set of feasible allocations $(q_{xy}^a, q_{xy}^b, Q_{xy})$ must belong to some feasible set Ω_{xy} that satisfies assumption 1. For a given allocation $(q_{xy}^a, q_{xy}^b, Q_{xy})$, the partners receive the following amounts of utility

$$\mathcal{U}_{xy}^\theta(q_{xy}^a, Q_{xy}) \text{ and } \mathcal{V}_{xy}^\theta(q_{xy}^b, Q_{xy})$$

where θ is a vector of preference parameters.

I shall also specify the outside options (singlehood) for a man of type x and a woman of type y . I assume that they have access to a subset of the goods they can consume while married. They choose vectors of private and public goods, denoted (q_x^s, Q_x^s) and (q_y^s, Q_y^s) that must belong to some feasible sets Ω_{x0} and Ω_{0y} , respectively. They receive utility

$$\mathcal{U}_{x0}^\theta(q_x^s, Q_x^s) \text{ and } \mathcal{V}_{0y}^\theta(q_y^s, Q_y^s)$$

Note that in general, it will be very straightforward to solve for the optimal consumption choices (q_x^{s*}, Q_x^{s*}) and (q_y^{s*}, Q_y^{s*}) of man x and woman y when they are single. Given θ and their chosen allocation $(q_{xy}^a, q_{xy}^b, Q_{xy})$ when married, we can simply compute the systematic utilities as

$$(5.1) \quad U_{xy} = \mathcal{U}_{xy}^\theta(q_{xy}^a, Q_{xy}) - \mathcal{U}_{x0}^\theta(q_x^{s*}, Q_x^{s*}) \text{ and}$$

$$(5.2) \quad V_{xy} = \mathcal{V}_{xy}^\theta(q_{xy}^b, Q_{xy}) - \mathcal{V}_{0y}^\theta(q_y^{s*}, Q_y^{s*})$$

5.2. Estimation method. To estimate a ITU-logit model, I rely on the simple characterization of the equilibrium in that particular case. Let me introduce $u_x =$

$-\log \mu_{x0}$ and $v_y = -\log \mu_{0y}$. As a reminder, equilibrium is fully characterized by the system of equations

$$(5.3) \quad \begin{aligned} \exp(-u_x) + \sum_y \exp(-D_{xy}^\theta(u_x, v_y)) &= n_x \\ \exp(-v_y) + \sum_x \exp(-D_{xy}^\theta(u_x, v_y)) &= m_y \end{aligned}$$

Typically, we will look for the value of θ that maximizes some objective function under the constraints given by equations (5.3). The objective function, denoted F , can be a log-likelihood function as in GKW or moment-based as in Gayle and Shephard (2016).

At this point, there are two ways to proceed to estimate θ . A first solution, which is employed in GKW, is to maximize F with respect to θ , but in that case, one must solve for the (u_x^*, v_y^*) solution to (5.3) for each value of the parameters θ . Given θ , u_x^* , and v_y^* , one can then compute the value of the objective function. These steps are repeated until F is maximized. Although there are efficient algorithms to solve for the equilibrium (u_x, v_y) in system (5.3), the computational cost is extremely high when there are no analytical formula for the distance function¹⁴.

A second approach, that I suggest to use here, is to solve the following program:

$$\max_{\theta, u_x, v_y} F(\theta, u_x, v_y)$$

subject to

$$\begin{aligned} \exp(-u_x) + \sum_y \exp(-D_{xy}^\theta(u_x, v_y)) &= n_x \\ \exp(-v_y) + \sum_x \exp(-D_{xy}^\theta(u_x, v_y)) &= m_y \end{aligned}$$

¹⁴The IPFP (Iterative Projective Fitting Procedure) algorithm detailed in GKW and Galichon, Kominers, and Weber (2015) can be used to solve system (5.3). The issue, however, is that whenever the distance function is not known in closed-form, we must evaluate (numerically) the distance function many times while solving system (5.3). This is computationally very intensive.

In other words, we maximize the objective function F with respect to θ and (u_x, v_y) , under equilibrium constraints (5.3). This is very similar to the so-called MPEC approach introduced in, e.g., Su and Judd (2012). In the next section, I will rely on this second approach.

As for constructing the objective function, note that when solving problem 1 for a particular (x, y) pair, we will obtain the predicted mass of marriage for this pair of types (that would be $\exp(-z^*)$), as well as predicted consumption of private and public goods. Assuming that the marriage patterns and household decisions are observed in the data, then it is easy to construct the objective function.

6. PROOF-OF-CONCEPT APPLICATION

In this section, I provide a proof-of-concept application of the methods introduced in this paper. In particular, I consider a collective model in which the public good is produced within the household: household members combine time to produce it. I use the PSID to construct a representative marriage market of the United States. I build on an earlier note of Weber (2016) in which I used my model in a pure simulation exercise. Here, I choose a very parsimonious model that I bring to real data. Therefore, preference parameters estimates should be taken very cautiously. The purpose of the application is twofold. First, it shows how the sharing rule and its distribution are identified in these models, when in general in the collective model literature it is only possible to identify the sharing rule up to a constant using distribution factors. Second, the model is structural and allows for counterfactual experiments useful for policy recommendations, which in general is not possible with a reduced-form approach. I consider two experiments of women empowerment; one in which the mass of men is increased on the market, a second in which the wage of women increases.

6.1. Parametric specification. I shall now be more specific regarding the parameterization of the systematic utilities that appear in equation (5.1). First, I assume

that singles and couples derive utility from the consumption of a private composite good, c , and leisure, ℓ (so that $q = (c, \ell)$ in previous notations), and the public consumption of a home-produced good, Q . I normalize the price of the private composite good to 1. All agents can spend time on the labor market, in which case they earn a hourly wage denoted w . The total time endowment is T .

For singles, the public good is produced from time spent on housework, h , according to the production function

$$Q = \tilde{\zeta}h$$

Consequently, a single man of type $x \in \mathcal{X}$ faces the following maximization program

$$\begin{aligned} \max_{c_x^s, \ell_x^s, h_x^s} & a_{e(x)} \log c_x^s + \alpha_{e(x)} \log \ell_x^s + A_{e(x)} \log \tilde{\zeta} h_x^s \\ \text{s.t.} & c_x^s + (\ell_x^s + h_x^s)w_x \leq Tw_x \\ & \ell_x^s + h_x^s \leq T \end{aligned}$$

The indices $e(x)$ indicate that I allow preference parameters to vary with types. More precisely, they may depend on a subset of the observable characteristics used to form types. In this particular model, preference parameters vary with three levels of education (see below for more details). In similar fashion, a single woman of type $y \in \mathcal{Y}$ solves

$$\begin{aligned} \max_{c_y^s, \ell_y^s, h_y^s} & b_{e(y)} \log c_y^s + \beta_{e(y)} \log \ell_y^s + A_{e(y)} \log \tilde{\zeta} h_y^s \\ \text{s.t.} & c_y^s + (\ell_y^s + h_y^s)w_y \leq Tw_y \\ & \ell_y^s + h_y^s \leq T \end{aligned}$$

It is straightforward to find the optimal $(c_x^{s*}, \ell_x^{s*}, h_x^{s*})$ and $(c_y^{s*}, \ell_y^{s*}, h_y^{s*})$ from which we can compute the singlehood reservation utilities $\mathcal{U}_{x0}^\theta(c_x^{s*}, \ell_x^{s*}, h_x^{s*})$ and $\mathcal{V}_{0y}^\theta(c_y^{s*}, \ell_y^{s*}, h_y^{s*})$.

If two individuals choose to marry instead, I assume that they can both invest time in the production of the public good. Importantly, I assume that preferences

do not change with marriage, but partners have access to a new technology. This is very similar to the approach taken by Browning, Chiappori, and Lewbel (2013). Therefore, married partners receive

$$(6.1) \quad \mathcal{U}_{xy}(c_{xy}^a, \ell_{xy}^a, Q_{xy}) = \delta_{xy}^a + a_{e(x)} \log c_{xy}^a + \alpha_{e(x)} \log \ell_{xy}^a + A_{e(x)} \log Q_{xy}$$

$$(6.2) \quad \mathcal{V}_{xy}(c_{xy}^b, \ell_{xy}^b, Q_{xy}) = \delta_{xy}^b + b_{e(y)} \log c_{xy}^b + \beta_{e(y)} \log \ell_{xy}^b + B_{e(y)} \log Q_{xy}$$

where

$$Q_{xy} = \zeta (h_{xy}^a)^\eta (h_{xy}^b)^{(1-\eta)}$$

and η and $1 - \eta$ are factor shares. Time spent on housework by the man and the woman are denoted h_{xy}^a and h_{xy}^b , respectively. I assume that η is the same for all types of households¹⁵. Finally, the terms δ_{xy}^1 and δ_{xy}^2 capture non economic gains to marriage. They depend on the observable characteristics of both partners. The budget and time constraints are

$$(6.3) \quad c_{xy}^a + c_{xy}^b + (\ell_{xy}^a + h_{xy}^a)w_x + (\ell_{xy}^b + h_{xy}^b)w_y \leq T(w_x + w_y)$$

$$(6.4) \quad \ell_{xy}^a + h_{xy}^a \leq T$$

$$(6.5) \quad \ell_{xy}^b + h_{xy}^b \leq T$$

The parameters to estimate are $(a_{e(x)}, \alpha_{e(x)}, A_{e(x)}, b_{e(y)}, \beta_{e(y)}, B_{e(y)}, \eta, \delta_{xy}^a, \delta_{xy}^b)$ ¹⁶, denoted θ . Given a value of θ and a choice of $(c_{xy}^a, \ell_{xy}^a, h_{xy}^a, c_{xy}^b, \ell_{xy}^b, h_{xy}^b)$ for when man x and woman y are married, we can construct the systematic utilities simply by taking the difference $\mathcal{U}_{xy}(c_{xy}^a, \ell_{xy}^a, Q_{xy}) - \mathcal{U}_{x0}^\theta(c_x^{s*}, \ell_x^{s*}, h_x^{s*})$ and $\mathcal{V}_{xy}(c_{xy}^b, \ell_{xy}^b, Q_{xy}) - \mathcal{V}_{0y}^\theta(c_y^{s*}, \ell_y^{s*}, h_y^{s*})$.

6.2. Data. The data used for estimating the model is the PSID (Panel Study of Income Dynamics) for year 2007. The basic panel consists of 5,000 families (over

¹⁵Of course, it is possible to allow η to depend on the type of the man and the woman.

¹⁶Note that as soon as the δ_{xy}^a and δ_{xy}^b contain a constant, the parameters ζ and $\tilde{\zeta}$ cannot be separately identified from the constant.

18,000 individuals), and it is representative of the US population. The study was repeated yearly since 1968 and every two years after 1997. In this application, I use only the information available in the core dataset, since we only need basic information on the education level of individuals, their labor supply, time spent on housework, income and marital status.

To construct the sample used in this application, I applied the following selection rules. First, I keep all single men (heads of their household) aged between 42 and 55 years old and all single women (heads of their household) aged 40 – 53¹⁷. Second, I select heterosexual couples (legally married or cohabitating and heads of their household) in which the partners satisfy these age requirements. By using this older sample, the focus is on a more stable marriage market. In addition, there is less variance in the labor supply of men and women at these ages. I use information on annual labor supply and labor income to construct hourly wages (abnormally high or low wages are trimmed). When wages are not observed, I predict them using a Heckman selection model (Heckman 1979). I observe the number of years of education for each individual which allows me to categorize men and women into three broad education types: “Below high school” (BHS or E1), “High school graduate” (HS or E2) and “College” (C or E3). This standard categorization can be found in Chiappori, Costa Dias, and Meghir (2017) as well. In addition, I observe how much time men and women spend on housework. Naturally, this variable is a very imperfect measure of time spent producing household goods, since it is self reported and it is hard to know what respondents take into account when providing an answer. We know however, that this includes time spend cooking, cleaning and doing work for the house. After deleting households with missing information, I am left with a sample of about 2,000 households. I report summary statistics on the important variables in table 1. Finally, I fix the time endowment to 112 hours per week and assume that leisure is

¹⁷These age intervals were chosen because the average age difference between married men and women is two years. This way of proceeding is standard in the literature, see Chiappori, Salanié, and Weiss (2017).

TABLE 1. Descriptive statistics

	Couples			Singles			Total		
	Mean	SD	N	Mean	SD	N	Mean	SD	N
A. Men									
Age	48.66	3.7	1056	48.52	3.9	298	48.63	3.8	1354
Below high school	0.12	0.3	1056	0.14	0.4	298	0.13	0.3	1354
High school	0.39	0.5	1056	0.44	0.5	298	0.40	0.5	1354
College	0.48	0.5	1056	0.41	0.5	298	0.47	0.5	1354
Wage	21.97	14.7	1056	15.92	10.8	298	20.64	14.2	1354
Hours of work	41.43	15.7	1056	31.64	20.7	298	39.28	17.4	1354
Hours of housework	7.50	7.9	1056	9.24	7.6	298	7.89	7.9	1354
B. Women									
Age	46.74	3.7	1056	46.29	3.9	690	46.56	3.8	1746
Below high school	0.10	0.3	1056	0.16	0.4	690	0.12	0.3	1746
High school	0.36	0.5	1056	0.39	0.5	690	0.38	0.5	1746
College	0.54	0.5	1056	0.45	0.5	690	0.50	0.5	1746
Wage	15.62	10.2	1056	12.73	8.4	690	14.48	9.7	1746
Hours of work	27.72	17.4	1056	29.55	19.9	690	28.44	18.5	1746
Hours of housework	17.55	12.5	1056	12.28	10.0	690	15.47	11.8	1746

Note: this table shows summary statistics for the sample obtained after applying the selection rules mentioned in the main text. In particular, I keep men aged 42-55 and women aged 40-53.

equal to the total time endowment minus the time spent on the labor market and on housework.

As is clear from table 1, married men are two years older than their wife (this is true even without applying the age selection criteria). In general, men tend to work more, have higher wages but spend much less time on housework than women. These gaps narrow when comparing single men and single women instead of married men and married women. Women also appear to be higher educated than men (this fact is discussed in Chiappori, Salanié, and Weiss (2017)). The marriage market is also unbalanced, as men are more scarce than women. This could come from the fact that women tend to live longer than men, but more importantly that women are

overrepresented among the divorced and separated as they find it more difficult to enter a new marriage (see chapter 1 in Browning, Chiappori, and Weiss (2014)).

6.3. Identification and estimation. To estimate the model, I follow the steps described in section 5. I assume that there is a mass 1 of each type of man and woman in the sample¹⁸, that is $n_x = m_y = 1$ for all $x, y \in \mathcal{X} \times \mathcal{Y}$. Preference parameters are allowed to vary with education, as explained above.

The parameters of interest, $(a_{e(x)}, \alpha_{e(x)}, A_{e(x)}, b_{e(y)}, \beta_{e(y)}, B_{e(y)}, \eta)$, are identified from the sole observation of the marriage patterns, labor supplies, time spent on housework and wages. In a collective model with a home-produced nonmarketable good, inputs choice is driven by cost minimization. Therefore, the ratio $\eta/(1 - \eta)$ is equal to the ratio of wages (w_x/w_y) times the ratio of housework (h_{xy}^a/h_{xy}^b), which shows identification of η . The identification of the other parameters proceeds as follows. First, note that observing private consumption is not necessary¹⁹ for identification since efficiency (abstracting from corner solutions) imposes $c_{xy}^a = (a_{e(x)}/\alpha_{e(x)})w_x\ell_{xy}^a$. Second, we can combine these efficiency restrictions, the systematic utilities identified from the marriage patterns and the program for singles to identify the other parameters.²⁰

The objective function used to estimate θ is the log-likelihood constructed from the predicted frequency that each pair will form on the marriage market, as well as the predicted labor supply and housework time. Appendix A contains details about the log-likelihood. Finally, I estimate the model on a subsample of 400 households

¹⁸Here, I could adapt the model to the continuous logit framework a-la-Dupuy and Galichon (2014), but that does not change anything to the estimation, as pointed out in Galichon, Kominers, and Weber (2018).

¹⁹In most available survey data, private consumption is poorly measured. At best, only the aggregate private consumption of the couple is known. Whenever private consumption is well measured, there is no need for collective models, see e.g. Browning and Bonke (2006).

²⁰As discussed above, the systematic utilities of men and women also contain a non-economic component, δ_{xy}^a and δ_{xy}^b . I parametrize this component as a linear function of partner's characteristics. The included characteristics are: dummies for each education level, dummies for each possible level of the absolute value of the difference between the education level of the partners, age of the partners, absolute value of the age difference, and the log of wage of the partners. The parameters associated to these characteristics are obviously parametrically identified.

randomly drawn from the full sample described in table 1²¹. The subsample contains approximatively 200 couples, which is roughly equivalent to what can be found in other studies such as Cherchye, De Rock, and Vermeulen (2012) and Del Boca and Flinn (2014).

6.4. Results.

6.4.1. *Preference parameters and model fit.* To estimate the model, I use the observed marriage patterns and labor supplies and housework time of married couples and singles. To provide a better fit of the marriage patterns, the model's δ_{xy}^a and δ_{xy}^b terms includes dummies for each education type as well as a dummy when the absolute difference in education level between the man and the woman is equal to one, and one dummy for when the absolute difference in education level is two. They also depend linearly on the age of the partners, the absolute value of the difference in age and on the log-wages of the partners. Economically speaking, these components capture costs (or gains) of singlehood, as well as preferences for education and age assortativeness²².

Estimates for the main parameters are displayed in table 2²³. Women appear to have a stronger preference for leisure compared to men, which is driven by the fact that women work less than men in the data. In addition, women seem to care more about public consumption, with only slight differences across types. The estimate for η is about 1/3. The magnitude of this parameter is not surprising: indeed the

²¹Using more observations is feasible, but for the purpose of a proof-of-concept application, I kept the sample small. The model can be estimated in a reasonable amount of time with up to 1000 households.

²²These two traits are indeed the most important ones for explaining marriage patterns. See, e.g., Ciscato and Weber (2017) for the US marriage market, or Doorley, Dupuy, and Weber (2018) for Germany.

²³I only present the estimates for the parameters $(a_{e(x)}, \alpha_{e(x)}, A_{e(x)}, b_{e(y)}, \beta_{e(y)}, B_{e(y)}, \eta)$ since these are the most important ones. Not surprisingly, the other non-economic gains parameters suggest a taste for assortative mating in education and age. These estimates are available upon request.

ratio $\eta/(1 - \eta)$ is equal to the ratio of wages (w_x/w_y) times the ratio of housework (h_{xy}^a/h_{xy}^b). From the raw data, we can deduce that η should be around 0.35.

TABLE 2. Parameter Estimates

	Men			Women			
	Private C.	Leisure	Public C.	Private C.	Leisure	Public C.	Technology
	a	α	A	b	β	B	η
BHS	0.309	0.644	0.047	0.166	0.693	0.141	
HS	0.311	0.605	0.084	0.265	0.631	0.105	0.310
C	0.350	0.593	0.057	0.310	0.577	0.113	

Note: types are displayed in rows (BHS = below high school, HS = high school degree, and C = college). In the estimation, I do not impose the preference parameters for consumption, leisure and public good to sum to 1. But for better lisibility, the estimates in the table above have been rescaled.

The model fit is discussed in table 3. The model fits the data reasonably well, considering that the model is very parsimonious. The table shows the mean of predicted and observed hours of work and housework of men and women. I also provide the R^2 from linear regressions of the observed variables (hours of work or housework) on their observed counterpart.

6.4.2. *Sharing rules.* The model predicts Pareto weights for each household. However, I choose to report a more informative, alternative, monetary measure of power: the

TABLE 3. Model Fit

	$T - \frac{h_{xy}^a}{\ell_{xy}^a}$	h_{xy}^a	$T - \frac{h_{xy}^b}{\ell_{xy}^b}$	h_{xy}^b
Observed (mean)	39.103	7.092	28.177	16.064
Predicted (mean)	39.135	6.338	27.808	15.544
R-squared	0.045	0.085	0.067	0.025

This table shows the fit of the model. The first and second rows display the mean of the observed and predicted dependent variables. The third row shows the R^2 from regressing the observed dependent variables on the predicted ones.

sharing rule. In the collective model literature, sharing rules are in general preferred since their interpretation is easier: they simply express how resources (a certain amount of money) are divided among partners. For a particular pair (x, y) , denote \hat{c}_{xy}^a , \hat{c}_{xy}^b , $\hat{\ell}_{xy}^a$ and $\hat{\ell}_{xy}^b$ the predicted private consumption and leisure of the man and the woman. Similarly, introduce \hat{Q}_{xy} the predicted consumption of public good. The conditional sharing rule (from the point of view of women) is computed as follow

$$S_{xy}^{cond} = \frac{\hat{c}_{xy}^b + w_y \hat{\ell}_{xy}^b}{\hat{c}_{xy}^a + w_y \hat{\ell}_{xy}^a + \hat{c}_{xy}^b + w_y \hat{\ell}_{xy}^b}$$

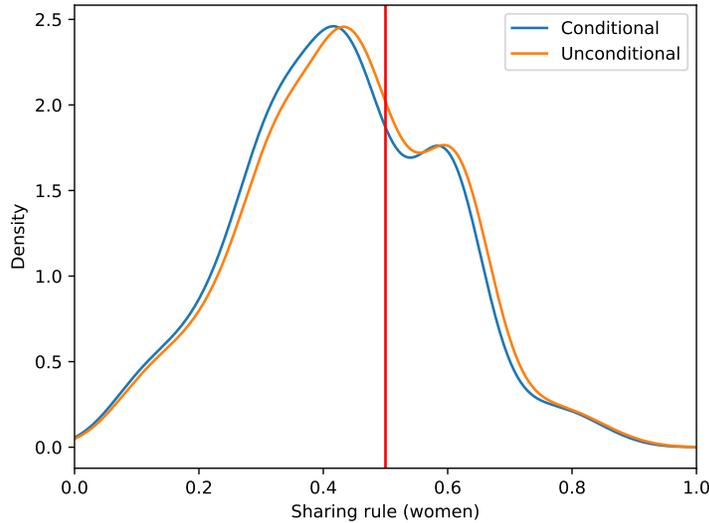
while the sharing rule is

$$S_{xy} = \frac{\hat{c}_{xy}^b + w_y \hat{\ell}_{xy}^b + P_{xy}^b \hat{Q}_{xy}}{\hat{c}_{xy}^a + w_y \hat{\ell}_{xy}^a + \hat{c}_{xy}^b + w_y \hat{\ell}_{xy}^b + P_{xy} \hat{Q}_{xy}}$$

Because the model is parametric, it is possible to recover the Lindhal prices P_{xy}^a and P_{xy}^b (where $P_{xy} \equiv P_{xy}^a + P_{xy}^b$) associated with the public good Q . Therefore, we can compute the (unconditional) sharing rule. The distribution of the estimated sharing rules for those couples who are observed in the data is showed in figure 6. The mean of the conditional sharing rule is estimated at 0.42, while it is about 0.44 for the sharing rule. Since all households produce and consume at least some public good, and that women seem to value it more, the sharing rule is higher than the conditional one. Overall, it seems that women are less favoured than men, a prediction that can be explained by the more attractive outside options of men (since they have a higher wage), and the population imbalance (the sex ratio favours men).

Sharing rules are estimated for all possible couples in the sample. In figure 7, I plot the density of the sharing rules conditional on the education level of the partners. Men's education is displayed in rows and women's education is represented in columns. The estimates are well behaved in the sense that the distributions are shifting to the right whenever the education of women increases and shifting to the left whenever the education of men increases. This is consistent with the fact that outside options

FIGURE 6. Sharing Rule



This figure shows the distribution of the conditional (the share of women’s private consumption in total private consumption) and unconditional sharing rules over all observed couples.

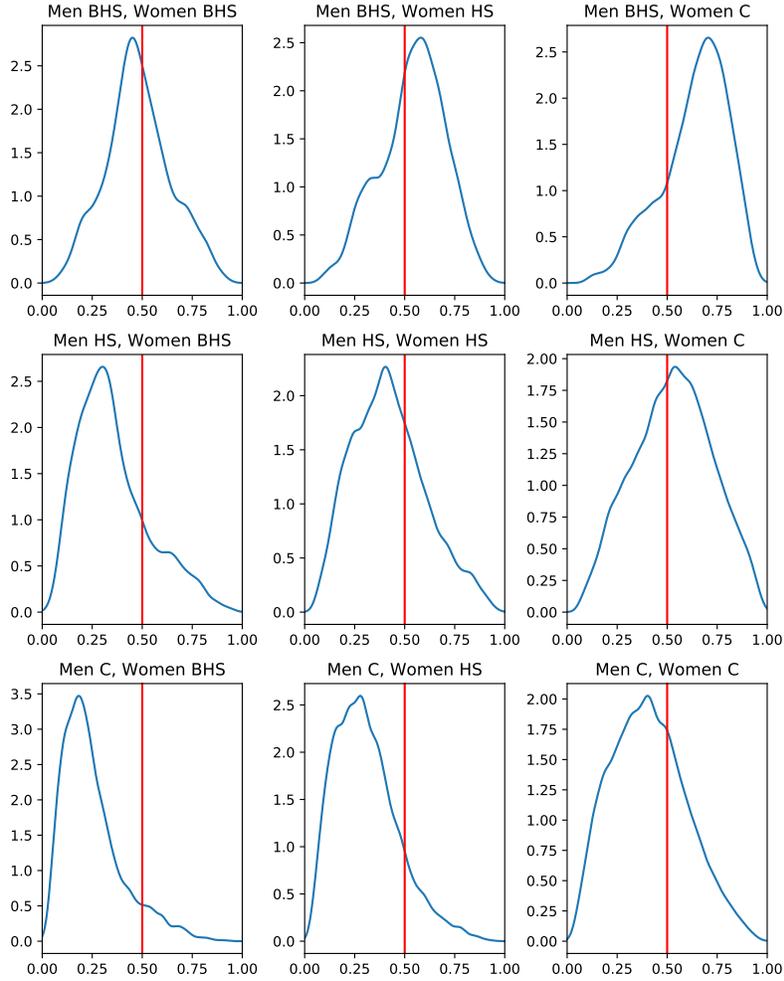
improve with education (since higher educated individuals are likely to have higher wages).

6.4.3. *Women empowerment: counterfactual experiments.* In this paragraph, I proceed with exploring two counterfactual experiments of women empowerment. This is of particular interest in an ITU matching model, even more so that in a TU model. In the latter, the sharing rule would, by definition, increase as women’s bargaining power increases. However, public good consumption would not be affected.

I consider two counterfactual experiments. In the first one, I progressively increase the number of men on the marriage market (by increasing their mass) so that the sex ratio becomes more and more balanced²⁴. As men become less scarce, we expect an

²⁴Of course, there are cases in which there could be more men than women on the marriage market, for example if the sex ratio at birth is biased in favor of boys. But in other instances, major events such as world wars would tend to increase the number of women relative to men on the marriage market. For consistency with the second experiment in which I increase women’s wage, I consider the policy of increasing the sex ratio rather than decreasing it.

FIGURE 7. Sharing Rule, by education of partners



This figure shows the distribution of the sharing rule over all (i, j) pairs, conditioning on the education levels of the husband and the wife. The husband education level is displayed in rows, the one of the wife in columns.

increase in women's bargaining power. In the second experiment, I propose to increase the wage of women. Since this improves their outside option of staying single, we can expect an improvement in the bargaining power of women.

One advantage of the present framework is that such distribution factors arise naturally from the underlying collective model and matching model. There is no need to assume that Pareto weights are given by some unknown function of prices,

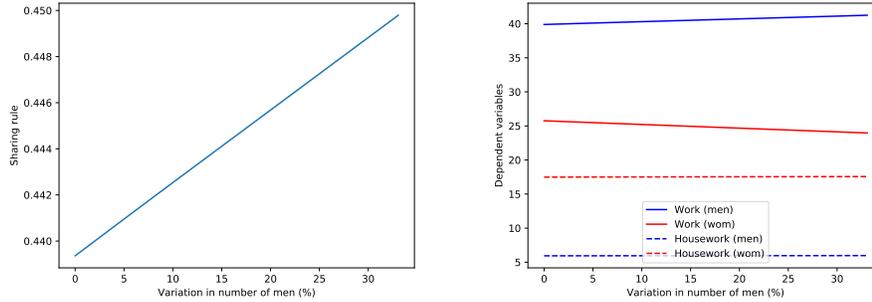
income and distribution factors. For example, it is clear that the wage ratio should affect the intra-household allocation of bargaining power in my model, because the wage ratio is explicitly related to the outside options of men and women. Changing the identity of the recipient of some nonlabor income (i.e a transfer that is no longer given to men, but to the women, as in the natural experiment studied in Lundberg, Pollak, and Wales (1997)) does change the allocation of bargaining power as well, providing this also affects the outside options of men and women. The same applies for the sex ratio, since changing population supplies changes the market structure and yields to a new equilibrium on the marriage market²⁵.

A first experiment. In the first counterfactual experiment, I gradually increase the mass of men in the population, from 0 to 10%. I also consider a final case in which the mass of men is increased up until the sex ratio is balanced on the market. For each marriage market (corresponding to each increase in the mass of men), I compute the new marriage market equilibrium by solving system (5.3) using the IPFP algorithm (see GWK). Preferences are assumed to be unchanged across each counterfactual marriage market. For a given counterfactual, the model predicts the sharing rule in each possible couple, as well as labor supply and time spent on housework. The left panel in figure 8 shows the average sharing rule²⁶ as the mass of men is increased. As expected, as men are less and less scarce on the marriage market, the sharing rule increases in favour of women, although very slightly. The right panel displays the average changes in dependent variables following the increase in the mass of men. As their bargaining power decreases, men are pushed to spend more time on the labor market and on housework, while women enjoy more leisure. Overall, the production of

²⁵This paragraph underlines the structural nature of my approach, because there is no reason for a distribution factor to affect the intra-household allocation of power unless it appears explicitly in the model.

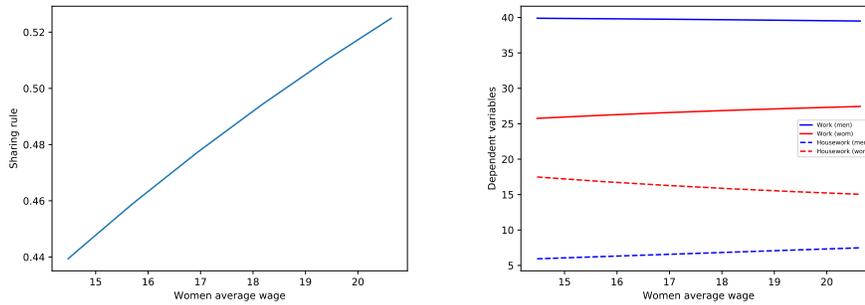
²⁶The figure shows the sharing rule and dependent variables for those couples who are observed married in the data. Of course, we could do the same exercise taking the mean or median for all possible couples, or even the full distribution.

FIGURE 8. Counterfactual sharing rule and dependent variables



This figure illustrates changes in the average sharing rule and time use dependent variables in counterfactual marriage markets. The counterfactuals correspond to a gradual increase in the mass of men.

FIGURE 10. Counterfactual sharing rule and dependent variables



This figure illustrates changes in the average sharing rule and time use dependent variables in counterfactual marriage markets. The counterfactuals correspond to a gradual increase in the wage of women.

public good is also increasing: this is in stark contrast with a model with transferable utility, in which case public good production and consumption would not change.

A second experiment. In the second experiment, I gradually increase the wage of all women up until the gender gap has, on average, disappeared. I proceed exactly as before, solving for the new marriage market equilibrium for each counterfactual. The results are presented in figure 10. The left panel shows changes in the average sharing rule as the wage of women increases. As expected, the sharing rule increases in favour

of women as their outside option of staying single becomes more attractive. As shown in the right panel of figure 10, the labor supply of women increases as their wages increase, although labor supply elasticity appears to be very small. This is because as the wage of women increases, their bargaining power increases as well, which pushes their labor supply in the other direction. The same can be said of men, whose labor supply decreases but very slightly, because their reduced bargaining power push them to spend more time on the labor market. This is a standard mechanism in the collective model literature (see Browning, Chiappori, and Weiss (2014), section 4.4.3 for a review). Finally, women (men) spend less (more) time on housework. Note that in this second experiment, it is more difficult to comment on changes in the amount of public good produced, since both bargaining power allocation and the shadow prices of time spent by women producing the public good are changing.

7. CONCLUSION

The present contribution brings together two important parts of the family economics literature: collective models, that describe the bargaining process that takes place within the household, and matching models, that have been used to model the marriage market. This is an important step forward because collective models, although very powerful, are incomplete: both household formation and the allocation of power within the household are taken as given. Modelling the marriage market jointly with household bargaining is one way of dealing with these shortcomings.

There have been a few attempts to achieve this goal using matching models with perfectly transferable utility. While this approach has produced impressive results, it is not completely satisfying. In particular, “power” does not matter in determining the public good consumption of married individuals. Therefore, I have proposed in this paper to use the matching framework with imperfectly transferable utility developed in a prior work (Galichon, Kominers, and Weber 2018) to close the gap between collective models and the marriage market. I tackle two issues that otherwise would prevent us from reaching this goal. The first is to characterize a class of collective models that produce proper bargaining sets, and therefore that can be embedded in the ITU framework. The second is to provide a way to estimate these models, based on an MPEC approach and an efficient computational method to obtain distance functions. In addition, the method underlines the connections that exist between the ITU matching framework and collective models. I provide an illustrative application using PSID data and a non-trivial model that includes private consumption, leisure and a public good that is produced from time inputs. I investigate whether a parsimonious model fit the data well, and show the strength of the approach with counterfactual experiments.

The approach developed in this paper is very general, and could be used in conjunction, or compared to, other approaches such as Cherchye et al. (2017). The model

presented in the application is voluntarily parsimonious, but could easily be extended to endogenize fertility choices and time (and money) invested in children. Finally, beyond the collective models studied in this paper, the methods developed here are applicable in other settings as well, e.g. on the labor market. These extensions are left for future research.

- Attanasio, O. P., & Lechene, V. (2014). Efficient responses to targeted cash transfers. *Journal of Political Economy*, 122(1), 178–222.
- Attanasio, O., & Lechene, V. (2002). Tests of income pooling in household decisions. *Review of Economic Dynamics*, 5(4), 720–748.
- Becker, G. S. (1973). A theory of marriage: Part I. *Journal of Political Economy*, 81(4), 813–846.
- Blundell, R., Chiappori, P.-A., & Meghir, C. (2005). Collective labor supply with children. *Journal of Political Economy*, 113(6), 1277–1306.
- Browning, M., & Bonke, J. (2006). Allocation within the household: Direct survey evidence. *Economics Series Working Papers*.
- Browning, M., Chiappori, P.-A., & Lewbel, A. (2013). Estimating consumption economies of scale, adult equivalence scales, and household bargaining power. *The Review of Economic Studies*, 80(4), 1267–1303.
- Browning, M., Chiappori, P.-A., & Weiss, Y. (2014). *Economics of the family* (C. S. of Economic Literature, Ed.). Cambridge Surveys of Economic Literature.
- Browning, M., & Gørtz, M. (2012). Spending time and money within the household*. *The Scandinavian Journal of Economics*, 114(3).
- Cherchye, L., De Rock, B., Lewbel, A., & Vermeulen, F. (2015). Sharing rule identification for general collective consumption models. *Econometrica*, 83(5), 2001–2041.
- Cherchye, L., De Rock, B., & Vermeulen, F. (2012). Married with children: A collective labor supply model with detailed time use and intrahousehold expenditure information. *American Economic Review*, 102(7), 3377–3405.
- Cherchye, L., Demuyck, T., De Rock, B., & Vermeulen, F. (2017). Household consumption when the marriage is stable. *American Economic Review*, 107(6), 1507–34.
- Chiappori, P.-A. (1988). Rational household labor supply. *Econometrica: Journal of the Econometric Society*, 63–90.

- Chiappori, P.-A. (2012). Modèles d'appariement en économie. *Revue Economique*, 63(3), 437–437.
- Chiappori, P.-A. (2017). *Matching with transfers: The economics of love and marriage*. Princeton University Press.
- Chiappori, P.-A., Costa Dias, M., & Meghir, C. (2017). The marriage market, labor supply and education choice. *Working Paper*.
- Chiappori, P.-A., Fortin, B., & Lacroix, G. (2002). Marriage market, divorce legislation, and household labor supply. *Journal of Political Economy*, 110(1), 37–72.
- Chiappori, P.-A., & Gugl, E. (2014). *Necessary and sufficient conditions for transferable utility*. Mimeo, Columbia University.
- Chiappori, P.-A., Salanié, B., & Weiss, Y. (2017). Partner choice, investment in children, and the marital college premium. *American Economic Review*, 107(8), 2109–67.
- Choo, E., & Seitz, S. (2013). The collective marriage matching model: Identification, estimation, and testing. In *Advances in econometrics* (Vol. 31, pp. 291–336). Emerald Group Publishing Limited.
- Choo, E., & Siow, A. (2006). Who marries whom and why. *Journal of Political Economy*, 114(1), 175–201.
- Ciscato, E., Galichon, A., & Goussé, M. (2015). Like attract like? a structural comparison of homogamy across same-sex and different-sex households. *Working Paper*.
- Ciscato, E., & Weber, S. (2017). The role of evolving marital preferences in growing income inequality. *Working Paper*.
- Coles, M. G., & Francesconi, M. (2017). Equilibrium search and the impact of equal opportunities for women. *IZA Discussion Paper*.
- Coupré, H. (2007). Time allocation within the family: Welfare implications of life in a couple. *The Economic Journal*, 117(516), 287–305.

- Dauphin, A., Lahga, A. E., Fortin, B., & Lacroix, G. (2011). Are children decision makers within the household? *Economic Journal*, 121(553), 871–903.
- Del Boca, D., & Flinn, C. J. (2014). Household behavior and the marriage market. *Journal of Economic Theory*, 150, 515–550.
- Doorley, K., Dupuy, A., & Weber, S. (2018). *The empirical content of marital surplus in matching models*. Institute for the Study of Labor (IZA).
- Dupuy, A., & Galichon, A. (2014). Personality traits and the marriage market. *Journal of Political Economy*, 122(6), 1271–1319.
- Galichon, A. (2016). *Optimal transport methods in economics*. Princeton University Press.
- Galichon, A., & Hsieh, Y.-W. (2017). A theory of decentralized matching markets without transfers, with an application to surge pricing. *Working Paper*.
- Galichon, A., Kominers, S. D., & Weber, S. (2015). The nonlinear Bernstein-Schrödinger equation in economics. In S. L. N. in Computer Sciences (Ed.), *Proceedings of the second conference “geometric science of information”* (Vol. 9389, pp. 51–59).
- Galichon, A., Kominers, S. D., & Weber, S. (2018). Costly concessions: An empirical framework for matching with imperfectly transferable utility. *Journal of Political Economy*, forthcoming.
- Gayle, G.-L., & Shephard, A. (2016). Optimal taxation, marriage, home production, and family labor supply. *Working Paper*.
- Goussé, M., Jacquemet, N., & Robin, J.-M. (2017). Marriage, labor supply, and home production. *Econometrica*, 85(6), 1873–1919.
- Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1), 153–161.
- Lewbel, A., & Pendakur, K. (2008). Estimation of collective household models with engel curves.pdf. *Journal of Econometrics*, 147, 350–358.

- Lise, J., & Seitz, S. (2011). Consumption inequality and intra-household allocations. *Review of Economic Studies*, 78(1), 328–355.
- Lundberg, S. J., Pollak, R. A., & Wales, T. J. (1997). Do husbands and wives pool their resources? evidence from the united kingdom child benefit. *Journal of Human Resources*, 32(3), 463–480.
- McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics*. New York: Academic Press.
- Pollak, R. A. (2018). How bargaining in marriage drives marriage market equilibrium. *Journal of Labor Economics*.
- Su, C.-L., & Judd, K. L. (2012). Constrained optimization approaches to estimation of structural models. *Econometrica*, 80(5), 2213–2230.
- Weber, S. (2016). Notes on integrating collective models and the marriage market. *Working Paper*.

APPENDIX A. LOG-LIKELIHOOD ESTIMATION

The construction of the log-likelihood follows the same step as in GKW. Our model predicts leisure and housework time for any (i, j) pair of individuals and for singles. I use this information and the assumption that housework and leisure are observed in the data with some Gaussian measurement error (with variance s_1^2 , s_2^2 , s_3^2 , and s_4^2 for the labor supply of men and women and the housework of men and women, respectively) to construct the likelihood function. Since I assumed a uniform distribution of types, we may as well reintroduce the notation (i, j) instead of types. We can form the log-likelihood as follows:

$$\begin{aligned}
\log \mathcal{L}(\theta, u_i^\theta, v_j^\theta) &= - \sum_{(i,j) \in \mathcal{C}} \left[D^\theta(u_i^\theta, v_j^\theta) + \frac{1}{2} \left(\frac{\ell_{ij}^a - \hat{\ell}_{ij}^a}{s_1} \right)^2 + \frac{1}{2} \left(\frac{\ell_{ij}^b - \hat{\ell}_{ij}^b}{s_2} \right)^2 \right] \\
&\quad - \sum_{(i,j) \in \mathcal{C}} \left[\frac{1}{2} \left(\frac{h_{ij}^a - \hat{h}_{ij}^a}{s_3} \right)^2 + \frac{1}{2} \left(\frac{h_{ij}^b - \hat{h}_{ij}^b}{s_4} \right)^2 \right] \\
&\quad - \sum_{i \in \mathcal{S}_M} \left[u_i^\theta + \frac{1}{2} \left(\frac{\ell_i^s - \hat{\ell}_i^s}{s_1} \right)^2 + \frac{1}{2} \left(\frac{h_i^s - \hat{h}_i^s}{s_3} \right)^2 \right] \\
&\quad - \sum_{j \in \mathcal{S}_F} \left[v_j^\theta + \frac{1}{2} \left(\frac{\ell_j^s - \hat{\ell}_j^s}{s_2} \right)^2 + \frac{1}{2} \left(\frac{h_j^s - \hat{h}_j^s}{s_4} \right)^2 \right] \\
&\quad - |\mathcal{I}| \log s_1 - |\mathcal{J}| \log s_2 - |\mathcal{C}| \log s_3 - |\mathcal{C}| \log s_4 \\
&\quad - \hat{N} \log(N)
\end{aligned}$$

In the above expression, ℓ_{ij}^a , ℓ_{ij}^b , h_{ij}^a and h_{ij}^b (ℓ_i^s , ℓ_j^s , h_i^s and h_j^s) denote the predicted leisure and housework time of (single) married men and women, respectively. The observed counterparts are denoted $\hat{\ell}_{ij}^a$, $\hat{\ell}_{ij}^b$, \hat{h}_{ij}^a and \hat{h}_{ij}^b ($\hat{\ell}_i^s$, $\hat{\ell}_j^s$, \hat{h}_i^s and \hat{h}_j^s for singles). \hat{N} and N are the observed and predicted number of households. The set of matched pairs (i, j) observed in the data is denoted \mathcal{C} , and \mathcal{S}_M and \mathcal{S}_F denote the set of

single men and the set of single women observed in the data, respectively. The log-likelihood is maximized over $(\theta, \{u_i\}_{i \in \mathcal{I}}, \{v_j\}_{j \in \mathcal{J}})$ subject to the constraint that the marriage market is in equilibrium (MPEC approach). Note that it is not necessary to optimize over (s_1, s_2, s_3, s_4) as it is possible to solve directly for their optimal value. For example, $s_1 = \frac{1}{|\mathcal{I}|} \left(\sum_{(i,j) \in \mathcal{C}} (\ell_{ij}^a - \hat{\ell}_{ij}^a)^2 + \sum_{i \in \mathcal{S}_M} (\ell_i^s - \hat{\ell}_i^s)^2 \right)$.

APPENDIX B. COMPUTATION METHOD FOR GENERAL PROPER BARGAINING SETS

Problem 2. *Suppose that \mathcal{G}_{ij} is a proper bargaining set associated with some collective model with free disposal, with preferences and feasible set of goods $(U_{ij}^\theta, V_{ij}^\theta, \Omega_{ij})$. Then we can solve*

$$(B.1) \quad \begin{aligned} & \min_{z_{ij}, q_{ij}^a, q_{ij}^b, Q_{ij}} z_{ij} \\ & s.t \ u - z_{ij} \leq \mathcal{U}_{ij}^\theta(q_{ij}^a, Q_{ij}) \end{aligned}$$

$$(B.2) \quad \begin{aligned} & v - z_{ij} \leq \mathcal{V}_{ij}^\theta(q_{ij}^b, Q_{ij}) \\ & g_{ij}^r(q_{ij}^a, q_{ij}^b, Q_{ij}) \leq 0, \ r \in \{1, \dots, R\} \end{aligned}$$

when $D_{ij}^\theta(u, v) = z_{ij}^*$, solution to the above program.

The solution method is the same as in the main text, except for the inequality constraints $u - z_{ij} \leq \mathcal{U}_{ij}^\theta(q_{ij}^a, Q_{ij})$ and $v - z_{ij} \leq \mathcal{V}_{ij}^\theta(q_{ij}^b, Q_{ij})$.