Taxation of Couples: a Mirrleesian Approach for Non-Unitary Households*

Carlos E. da Costa *EPGE-FGV*

Lucas A. de Lima *Harvard University*

March 13, 2017

Abstract

We assess the scope of taxation for implementing desirable allocations in an economy in which married agents' decisions follow a bargaining procedure satisfying Nash's (1950) axioms. When one departs from a unitary view of household behavior tax schedules must play a double role: *i)* to define households' objective functions through their impact on threat points, and; *ii)* to induce the desired allocations as optimal choices for households given these objectives. We find that the *taxation principle*, which asserts that there is no loss in relying on tax schedules, fails. The set of allocations implemented by a tax schedule may be expanded by auxiliary policy instruments which: *i)* align the households' and planner's objectives, and; *ii)* manipulate taxable income elasticities. We provide conditions based on how large the transaction space is for an allocation to be implemented by a tax system alone. We also propose a test to identify (constrained) inefficient allocations based on the planner's ability to manipulate elasticities. ***Keywords:*** *Optimal Taxation; Collective Households; Nash-bargain.* ***JEL Codes:*** *D13; H21; H31.*

OPTIMAL tax theory has experienced a welcome revival in the last 15 years with the use of variational methods. Whereas the idea of perturbing tax schedules to characterize optimality conditions dates back at least to Sheshinski (1972), this approach was limited by parametric restrictions on the schedules.[1] Piketty (1997) and Saez (2002) have extended the method's scope to deal with fully non-linear tax schedules. The importance of this generalization cannot be exaggerated. The Taxation Principle – Hammond (1979,

*We thank Eduardo Azevedo, Pierre-André Chiappori, Leandro Gorno, Kory Kroft, Bernard Salanié, André Trindade, seminar participants at IPEA, Insper, the 2015 SBE, the 2016 SED, SAET and IIPF Meetings for their helpful comments. da Costa thanks CNPq project 305766/2014-7 and INCT for financial support. All errors are our responsibility.

[1] In a broader sense we can trace back this tradition to Ramsey's (1927) contribution.

1987) – guarantees that tax implementation is without loss. Deriving an optimal (unrestricted) tax schedule therefore implies finding the best possible redistributive scheme that can be conceived.[2] In this sense variational methods promise to extend Mirrlees' (1971) agenda to areas where progress has proven elusive; that including the taxation of couples.

An issue that is often overlooked is that optimal taxation theory is almost entirely written under the implicit assumption that all agents are single or that households may be treated as an individual. The purpose of this paper to take one step back and ask whether optimal redistributive policy and optimal taxation are still equivalent in a world where the multi-person nature of most households is taken seriously. If not, how does this finding affect optimal taxation? What should auxiliary policies aim?

Abstracting from marital status would not be an issue if marriage changed neither behavior nor well being. The problem is marriage does change behavior and well-being. Alternatively, the attempts to explicitly address the taxation of couples treats households as individuals, the so-called unitary approach, for which there is mounting empirical evidence rejecting it – Thomas (1990); Lundberg et al. (1997); Browning and Chiappori (1998).[3] Moreover, endowing households with preferences goes against the methodological individualism which requires aggregate behavior to be derived from individuals' choices, and under which 'household welfare' is devoid of any meaning — e.g., Browing et al. (2013); Chiappori and Meghir (2014).[4]

We adapt Mirrlees' (1971) framework to a world in which agents may differ with regards to gender and marital status, the latter endogenously defined in a marriage problem which is affected by policy. The informational structure in Mirrlees (1971) is modified by the assumption that spouses know each others' productivities. Finally, married agents decide through a bargain procedure that satisfy Nash's (1950) axioms. If an agreement is reached, spouses maximize a Nash product which threat points are determined by the equilibrium of non-cooperative disagreement games.

For such an environment, any institution – e.g. a tax system, a direct mechanism, etc. — defines not only the choice set faced by spouses in agreement but also the strategy spaces and outcome functions for the disagreement games. Thus, to compare institutions

[2]The constrained optimal allocation derived using Mirrlees's (1971) approach, can always be implemented by suitable budget sets. Hence a Taxation Principle corollary is that optimal taxation and optimal redistributive policies are synonyms.

[3]Such change in paradigm has found its way into actual policy making. *Oportunidades* in Mexico and *Bolsa Família* in Brazil are prominent examples of social programs which design explicitly acknowledges the failure of one of the well known properties of unitary models: income pooling.

[4]Of course, a unitary model may be derived from the household members' decision process as, for example, in Becker's (1981) Rotten Kid theorem. The conditions for the theorem to hold are however too restrictive – see Bergstrom (1989).

we must be able to assess both how they affect household choices given the threat points and which threat points arise as equilibria for the disagreement games.

The revelation principle — Dasgupta et al. (1979); Myerson (1979); Harris and Townsend (1981) — is used to characterize the set of implementable allocations. For singles, nothing new: agents reports their productivity and an outcome function maps announcements into transactions. As for couples, at the moment a married agent, say a wife, 'communicates with the center' she knows her own productivity, her spouses' productivity and whether the household has reached an agreement or not, same for her husband. This informational structure induces the definition of a married agent's type as: *i)* the agent's own productivity; *ii)* her (his) spouses' productivity, and; *iii)* whether the couple is in agreement or disagreement. An outcome function maps the two spouses' reports into transactions to be conducted by both. The definition of a married agent's type implies that he or she has the same type as his or her spouse. A mechanism which assigns very undesirable transactions following conflicting announcements grants the planner quite some latitude for manipulating threat points and is shown to be optimal.

If households' objectives were invariant to policy, then the classic proofs for the Taxation Principle, e.g., Hammond (1979, 1987); Guesnerie (1998), would go through. What makes this collective setting special is that households' objectives depend on threat points which are themselves dependent on policy. The equivalence between two institutions therefore depends not only on their ability to replicate choices *conditional on threat points*, but also on their ability to generate the same threat points. An obvious limitation of tax schedules which helps explain why the taxation principle fails is that it dose not distinguish between couples in agreement and couples in disagreement: the same schedule must be used by couples in either state. This is not the whole story. Not all implementable allocations can be implemented by tax systems which offer different schedules for households in agreement and households in disagreement. We provide conditions for an allocation to be implemented by such a tax system based on how large the transaction space is.

We also provide sufficient conditions for efficiency of an optimal tax schedule to be ruled out. In short, if there are reforms that increase spouses disagreement utilities without changing their equilibrium choices, then it is possible to reduce elasticities of taxable income – ETI's – and reduce inefficiencies.

Our findings have important consequences for optimal tax theory. First, since there are missing policy instruments, some departures from traditional optimal tax formulae may be explained by the attempt to mimic these instruments.

Second, behavioral responses depend on ETI's, which are not structural parameters. As we have argued before, they depend on threat points which, in turn, are affected by policy. Auxiliary policy instruments capable of reducing ETI's lead to lower efficiency losses from taxation.

Finally, the very notion that a well defined ordering which justify the properties we impose on ETI's depends on threat points being invariant to policy. This requires tax schedules not to affect threat points. Whereas this is not to be expected for tax schedules, it may be for a tax system which allows for filing options.

Thus far, we have not mentioned the impact that policy has on decision to get married. Of course, the decisions to marry and whom to marry depend on what one expects to attain in marriage. Hence, we show how the choices in the 'marriage market' enter the planner's program. Following Pollak (2016), we assume that lack of commitment precludes the use of compensatory transfers which combined with a costly reentry in marriage markets justify our using a Bargain in Marriage (BiM) approach. Under this approach, the flow utilities that are generated under any institutional environment, are the basis for a matching problem which is a variant of Gale and Shapley's (1962) classic analysis.[5]

The rest of the paper is organized as follows. After a brief related literature overview, Section 1 presents the environment. We provide an heuristic account of our findings in Section 2. In Section 3 we formalize the decision process of married couples. The main theoretical results are found in Section 4. The consequences for optimal taxation are explained in Section 5. Section 6 we consider the robustness of our finding to sensible variations of our main assumptions. Section 7 concludes.

Literature Review

Boskin and Sheshinski's (1983) pioneering work on the optimal taxation of couples adopted the unitary approach under which couples are modeled essentially as a single agents with multiple, perfectly assignable, choices of effort. They, and all early works in the taxation of couples, e.g., Apps and Rees (1988), imposed parametric restrictions on tax schedules. Recent work by Piketty (1997); Dahlby (1998); Saez (2001) have freed variational methods from these constraints and allowed the complete characterization of optimal tax schedules.

That these methods solve the Mirrlees (1971) program follows from the taxation principle — Hammond (1979, 1987); Guesnerie (1998). Unfortunately, the taxation principle was only proved under the assumption that agents are single, or if a unitary

[5]Of course we recognize that 'bargaining within marriage takes place in the shadow of marriage markets', as emphasized by Becker (1991). Still, by separating the problem in two stages, the BiM assumption helps us isolate some of the forces that will be present if commitment is not full, there remaining room for bargain after marriage.

view of the household is used.[6] The unitary approach has, however, lost a great deal of its appeal due to its poor adherence to the data — e.g., Thomas (1990); Browning and Chiappori (1998); Lundberg et al. (1997) — and its lacking a sound theoretical basis. We depart from the unitary approach and follow Manser and Brown (1980); McElroy and Horney (1981); Lundberg and Pollak (1993) in assuming that couples bargain over their preferred choices and that the outcome of such bargain satisfy Nash's (1950) axioms.

Central to our analysis is, therefore, the impact of policies on threat points.[7] We must, in this case, take a stand on how these threat points are determined. As in Lundberg and Pollak (1993); Chen and Wooley (2001) we focus on 'internal' threat points given by the utilities that spouses attain when they fail to reach an agreement and decide non-cooperatively.

Ours is a classic mechanism design application motivated by a modern view of household decisions. In this context we provide a proper definition of type that justifies the application of the Revelation Principle – Myerson (1979); Dasgupta et al. (1979); Harris and Townsend (1981) – when spouses bargain over their preferred outcome.[8] A non-standard aspect of our problem is that 'household objectives' are partially determined by the mechanism through its impact on threat points.

The endogeneity of threat points raises new issues related to behavioral responses to tax reforms. First, if a perturbation in the tax schedule affects threat points the household does not respond as if it were an individual. Its choices do not satisfy the restrictions implied by rationality, as discussed in Browning and Chiappori (1998). Second, elasticities depend on threat points and these can be manipulated to guarantee that dead weight losses are reduced or that more redistribution is possible with less 'leakages'. The idea that ETI's should be optimally chosen is also present in Kopczuk and Slemrod (2002) in a very different context. Finally, relative social welfare weights need not coincide with the household's relative marginal weights: this is called dissonance by Apps and Rees (1988). With dissonance behavior responses to tax perturbation have first order welfare effects.

[6] Golosov et al. (2014) have provided a nice example of how one may apply variational methods to the optimal taxation of couples.

[7] Since the axiomatic solution to a Nash bargain imposes efficiency, these models are collective in the sense that households can achieve efficient outcomes without specifying the process leading to such outcomes — see the discussion in Chiappori and Mazzocco (2015). Without a specification about how threat points are determined, Nash-bargain has no additional testable implications when compared to general collective models — Chiappori et al. (2012).

[8] Because spouses know each other types, on might think that the planner could use this correlation in information sets to implement better allocations as in Crémer and McLean (1988). Yet, spouses coordinate their announcements and do it efficiently in equilibrium. First best allocations are not implementable, in general.

In Immervoll et al. (2011) and Cremer et al. (2016), dissonance takes center stage.[9] In both works households weights are exogenous whereas here household welfare weights depend on threat points which can be manipulated to close the gap between household and social objectives. In this sense, the idea of dissonance parallels the behavior public finance literature distinction between the 'decision utility' and the 'experienced utility' – see Chetty (2015). In contrast with this literature the misalignment arises not due to some failure to optimize or some paternalistic concerns by the planner, as in Kanbur et al. (2006), but due to the fact that any welfare evaluation must be based on the impact it has on individuals, e.g., Browing et al. (2013).

1 Environment

The economy extends Mirrlees' (1971) by allowing a subset of agents to be married.

A continuum of individuals, of two different genders, $i = f, m$, have preferences defined over consumption, $\mathfrak{c}$, and leisure, $\mathfrak{l}$, which may depend on their gender, but are otherwise identical across individuals. For both genders, $(\mathfrak{c}, \mathfrak{l}) \in X$, where X is a closed, convex subset of R^2_+ containing 0. Preferences for a gender i agent may be represented by the utility function, $u_i(\mathfrak{c}, \mathfrak{l})$, increasing in both variables and strictly quasi-concave.

As in Mirrlees (1971), agents differ with respect to their labor market productivity, $\theta_i \in \Theta$, $i = f, m$, where Θ is a finite set.

Agents may be single or married. Each couple is identified by the pair of productivities: θ_f for the wife and θ_m for the husband. We use $\boldsymbol{\theta} = (\theta_f, \theta_m)$ for brevity. The distribution of singles of gender i is given by a measure μ^i_{s} in Θ. The distribution of couples is given by a measure $\boldsymbol{\mu}$ in $\Delta\Theta^2$. So $\boldsymbol{\mu}(\theta_f, \theta_m)$ represents how many couples with type $\boldsymbol{\theta} = (\theta_f, \theta_m)$ we have. These measures are endogenously defined as a stable matching for the marriage problem — Section 3.5.

Finally, technology is represented by a transformation function, $G : Z^2 \mapsto R$, assigning a non-positive value for feasible allocations and a positive value for unfeasible ones.

Informational Structure An agent's type is his or her private information. For singles, this means that a gender i agent is the only one to know its type, θ_i. For married agents we assume that each spouse i — for $i = f, m$ — in a couple $\boldsymbol{\theta} = (\theta_f, \theta_m)$ knows not only his or her productivity θ_i, but also his or her spouses' productivity, θ_j — for $i, j = f, m$. I.e., both the husband and the wife knows $\boldsymbol{\theta}$, and no-one else.

[9] Cremer et al. (2016), for example, have uncovered a Pigouvian term in optimal taxes that arises due the misalignment of planner and household objectives.

This assumption about the informational structure of an economy with households is consistent with the type of efficient decisions implied by the collective model in general and the Nash bargaining model in particular.[10]

Transactions An agent's consumption, $(\mathfrak{c}, \mathfrak{l})$ is observed by no one but the agent him or herself, if the agent is single. It is observed by both spouses when the agent is married. What everyone observes are the transactions made by single agents and households. A tax schedule may not, therefore, be a function of final consumption but only transactions.[11]

Singles A transaction is a vector $z \in Z \subset \mathbb{R}^2$, where Z is a subset of $\mathbb{R}^2$ containing 0 which describes everything an agents buys as positive entries and everything he or she sells as negative entries. All agents are assumed to have the same endowment, $(0, 1)$, of consumption good and time. A transaction vector $z = (c, -n)$, describes the amount of hours, n, the agent supplies and the amount of consumption goods he or she acquires in the market, c. Given the endowment $(0, 1)$ the bundles attainable by a type θ_i agent are $(\mathfrak{c}, \mathfrak{l}) \leq (c, 1 - n/\theta_i)$.

Because utility is strictly increasing in both $\mathfrak{c}$ and $\mathfrak{l}$, given z, a single agent always consumes $(\mathfrak{c}, \mathfrak{l}) = (c, 1 - n/\theta_i)$. There is one to one mapping between transactions, z, and consumption bundles, $(\mathfrak{c}, \mathfrak{l})$, which allows us to define the induced preferences over the set $\mathscr{Z}$, $\succeq_{\theta_i}$.

Couples For couples, define $\boldsymbol{z} = (z_f, z_m) \in Z^2$, where z_f are the transactions made by the wife and z_m those made by the husband.

The mapping from transactions to consumption bundles for couples is more involved than those for singles. We define a function $F^a : Z^2 \times \Theta^2 \to 2^{X^2}$ which maps the transaction realized by a couple into sets of feasible pairs of bundles that can be consumed by this couple.[12] Embedded in this function are not only the material gains from marriage but also all relevant transferability restrictions. For example, if a $\boldsymbol{\theta} = (\theta_f, \theta_m)$-couple chooses transaction $\boldsymbol{z} = (c_f, -n_f, c_m, -n_m)$, then the wife must be enjoying leisure $\mathfrak{l}_f \leq 1 - n_f/\theta_f$ and the husband, $\mathfrak{l}_m \leq 1 - n_m/\theta_m$. As for consumption goods, we may allow for gains of scale by assuming that all consumption pairs $(\mathfrak{c}_f, \mathfrak{c}_m)$ such that

[10] The idea that households can efficiently bargain for instance may not be reasonable if spouses have private information with respect to each other.

[11] The inclusion of obedience constraints – e.g., Myerson (1986) – allows one to write the program in terms of final consumption. Yet, by embedding these constraints in households' conditional preferences for transaction we make comparing institutions simpler.

[12] F^a represents the potential material gains from cohabitation. If households choose efficiently these gains are realized.

$\mathfrak{c}_f + \mathfrak{c}_m \leq \alpha(c_f + c_m)$ for $\alpha > 1$ are attainable. We write

$$F^a(\boldsymbol{z}, \boldsymbol{\theta}) := \left\{ \left(\left(\mathfrak{c}_f, \mathfrak{l}_f \right), \left(\mathfrak{c}_m, \mathfrak{l}_m \right) \right); \mathfrak{c}_f + \mathfrak{c}_m \leq \alpha \left[c_f + c_m \right]; \mathfrak{l}_f \leq 1 - \frac{n_f}{\theta_f}; \mathfrak{l}_m \leq 1 - \frac{n_m}{\theta_m} \right\}.$$

Note that consumption is transferable across agents, hence non-assignable, while leisure is not. Transactions are not, therefore, directly mapped into utilities. Instead, for any given $\boldsymbol{z} = (z_f, z_m)$ and any $\boldsymbol{\theta}$, a (conditional) utility possibility set $\mathcal{U}_{\boldsymbol{z}}(\boldsymbol{\theta})$ is defined.

1.1 Households' Decision Process

Each agent decides whether to get married and whom to marry, anticipating the flow utility he or she obtains in each possible situation. Utilities are fully anticipated at the time of marriage, and no transfers at the time of marriage or commitment to future transfers can be credibly made. We are therefore in a setting referred to as Bargain in Marriage (BiM) by Lundberg and Pollak (2009); Pollak (2016). We now describe couples' decision process leading to these utilities. In Section 3.5 we describe the first stage of this setting: the marriage market.

Couples' decide through bargains which solution satisfy a variation of Nash's (1950) axioms advanced by Zambrano (2016) to account for possible non-convexities of utility possibility sets.[13] The Nash bargaining solution is formulated in terms of a set o feasible utility pairs that households can attain by cooperating and a disagreement pair $\bar{\boldsymbol{u}} = (\bar{u}_f, \bar{u}_m)$ representing the utilities attained by spouses if an agreement is not reached. The utilities that spouses attain are a selection from the solutions to the maximization of [14]

$$\left(u_f(\mathfrak{c}_f, \mathfrak{l}_f) - \bar{u}_f \right) \left(u_m(\mathfrak{c}_m, \mathfrak{l}_m) - \bar{u}_m \right). \tag{1}$$

Whenever the utility possibility set is convex, the solution for the couple's bargain satisfying Zambrano's (2016) axioms coincides with Nash's (1950) solution.

For the program to be well defined one must explain how the disagreement utilities are determined. In family economics different possibilities have been considered. The early works of Manser and Brown (1980); McElroy and Horney (1981) took divorce to be the relevant threat points whereas Lundberg and Pollak (1993) pioneered the used of 'internal' threat points: utility attained if spouses fail reach a disagreement and behave non-cooperatively.

[13] All of Nash's (1950) original axioms but symmetry which is replaced by Zambrano's (2016) 'preference for symmetry' are retained. This axiom postulates that if a choice $\boldsymbol{u}_\alpha$ in the utility possibility set is the combination of another choice $\boldsymbol{u}_0$ and its permutation, $\boldsymbol{u}_1$, both in the utility possibility set, then neither $\boldsymbol{u}_0$ nor $\boldsymbol{u}_1$ are a solution to the Nash bargain.

[14] It is possible that we have more than one solution to this problem.

Under external threat points, redistribution across spouses that do not alter agents utilities upon divorce have no impact on household choices. Income pooling, which this approach preserves, seems to be at odds with the evidence found in Lundberg et al. (1997). There is also evidence that goes against purely 'internal' threat points, and no consensus exists on which approach best matches the data. From a purely theoretical perspective, Bergstrom (1996) using Binmore's (1985) findings regarding the nature of threat points, makes a forceful case in favor of internal threat points. Binmore (1985) adds to an otherwise standard Rubinstein bargaining game the possibility for each agent to abandon the bargain process and grab some outside option. The flow utilities attained by spouses if bargaining continues without an agreement are referred to in his work as 'impasse point' and outside options as 'breakdown points'. Threat points are shown to coincide with impasse points and to be unaltered by the presence of breakdown points. The latter only define lower bounds for the utilities that spouses may attain.[15] Under Bergstrom's (1996) interpretation, utilities attained at divorce play in household economics the role of break down points. Impasse is delayed agreement and 'burnt toast' followed by a counter proposal.[16]

We follow Bergstrom (1996) and consider two states in which a marriage can find itself: agreement or disagreement. A couple is said to be in disagreement if no common grounds for a decision has been reached. They play a non-cooperative game which generate the threat points in the form of flow utilities corresponding to Binmore's (1985) impasse point.[17] We assume that the utility loss from ending a marriage is infinite, so the breakdown points are never binding.[18]

The disagreement state does not happen in equilibrium, but spouses do understand and clearly anticipate what would happen were they not able to reach an agreement. Couples are in agreement if spouses have been able to reach common grounds on a particular ordering of transactions and on how to allocate the outcome of such transactions between themselves. In this case, spouses maximize (1).

Whether a couple is in agreement or disagreement is private information.

[15] Binmore et al. (1989) run laboratory experiments to assess whether such theoretical findings predict the actual outcome of bargains. Their findings support Binmore's (1985) theory.

[16] The idea here is that as long as the gains from marriage are divided in such a way that both parties are better off being married than being divorced, a divorce threat is not credible.

[17] Under this view the utilities that agents may attain out of marriage can only affect household's choice if one of the agents is indifferent between remaining married and breaking up with his or her spouse.

[18] That is, we assume that divorce is prohibitively costly. For robustness, we allow 'external' threat points to play a role in Section 6.2. We assume in this case that spouses may at any time (including in disagreement) unilaterally end marriage, and allow for the possibility that divorce utilities may exceed in some situations those attained in the disagreement games.

2 Heuristics for Main Results

2.1 Comparing Tax Schedules with General Mechanisms

We want to compare the set of constrained efficient allocations with the set of allocations that arise under tax schedules. To do so, we consider: *i)* a simple tax system comprised of tax schedules for singles of each gender and a tax schedule for spouses; and, *ii)* a general mechanism,

$$\mathcal{M} \equiv \left(\{\Sigma_i^s, \Sigma_i^c\}_{i=f,m}, \{g_i^s(\cdot)\}_{i=f,m}, \boldsymbol{g}^c(\cdot,\cdot) \right), \tag{2}$$

comprised of message spaces, $(\Sigma_i^s)_{i=f,m}$, for single agents, $(\Sigma_i^c)_{i=f,m}$ for married agents, and outcome functions, $(g_i^s(\cdot))_{i=f,m}$, for singles and $\boldsymbol{g}^c(\cdot,\cdot)$, for couples.

Tax schedules define the set of feasible transactions through budget sets, $\mathcal{Z}_\Psi^i := \{z_i \in Z | \psi_i(z_i) \leq 0\}$, for single agents of gender i, $i = f, m$, and $\mathcal{Z}_\Psi^c := \{\boldsymbol{z} \in Z^2 | \boldsymbol{\psi}(\boldsymbol{z}) \leq 0\}$, for couples. Under the mechanism, $\mathcal{M}$, transactions are only chosen indirectly through the messages sent to the center. The set of feasible transactions under $\mathcal{M}$, $\mathcal{Z}_\mathcal{M}^c$, are all transactions which are the outcome of some pair of messages,

$$\mathcal{Z}_\mathcal{M}^c := \left\{ \boldsymbol{z} \in Z^2 | \exists (\sigma_f, \sigma_m) \in \Sigma_f \times \Sigma_m \text{ such that } \boldsymbol{g}^c(\sigma_f, \sigma_m) = \boldsymbol{z} \right\}.$$

Analogous definitions, $\mathcal{Z}_\mathcal{M}^i$, $i = f, m$, apply to singles.

Consider a tax system, Ψ, and a mechanism, $\mathcal{M}$ such that the feasible transactions are the same, that is, $\mathcal{Z}_\mathcal{M}^c = \mathcal{Z}_\Psi^c$. We may then ask, "do the same allocations arise under the two institutions?" Since the sets of attainable transactions are the same, then, if the objective functions, (1), are the same under Ψ and under $\mathcal{M}$, we get a positive answer. By inspection, this will be the case if the threat points coincide. The question is then, does $\mathcal{Z}_\mathcal{M}^c = \mathcal{Z}_\Psi^c$ guarantee that the same threat points arise as equilibria of all households' disagreement games? If not, how do they differ?

To address these questions, we must specify the disagreement games. In disagreement spouses choose non-cooperatively: i) their transactions, under a tax system, or; ii) their announcements, under a general mechanism. The main assumption we make is that the flow utility attained by each spouse in the disagreement game depends only on the transactions realized by the couple at the Nash equilibrium for the disagreement game, *not* on the specific process leading these transactions to arise as equilibria. It is this assumption which will allow us to compare different institutions.

Formally, define for each couple, $\boldsymbol{\theta}$, a mapping Φ from transactions, $\boldsymbol{z}$, to utility pairs, $\bar{\boldsymbol{u}} = (\bar{u}_f, \bar{u}_m)$, in the interior of the set $\mathcal{U}_{\boldsymbol{z}}(\boldsymbol{\theta})$, i.e., $\Phi(\boldsymbol{z}, \boldsymbol{\theta}) = \bar{\boldsymbol{u}} \in \mathring{\mathcal{U}}_{\boldsymbol{z}(\boldsymbol{\theta})}$. This function is a reduced form aiming at capturing the set of implicit rules that are used by spouses in the non-cooperative state that arises as they struggle to find an agreement. It transforms non-cooperatively chosen transactions into spouses' consumption goods. We endow Φ with properties that are likely to prevail under any reasonable 'disagreement games' — see example on page 17.

2.2 Taxation Principle Failure

As argued, a mechanism and a tax system for which $\mathcal{Z}^c_{\mathcal{M}} = \mathcal{Z}^c_{\Psi}$ can only differ by their impact on threat points. Naturally, one can always define a mechanism which leads to the equilibrium induced by a tax schedule: just ask agents to announce their desired transactions. More interestingly, the converse need not be true, there are allocations induced by the direct mechanism that cannot be the outcome of a tax schedule.

An institution matters for the disagreement game not only through the set of available transactions but also through the strategy space faced by each spouse. In that regard, even if the available transactions are the same, $\mathcal{Z}^c_{\mathcal{M}} = \mathcal{Z}^c_{\Psi}$, the strategy spaces under which the disagreement games are played need not be the same. Under a tax schedule, a husband chooses his transactions taking his wife's transactions as given, whereas under a mechanism he chooses his announcement taking his wife's announcements as given. The space of announcements can be larger, giving the planner more degrees of freedom to assign different transactions to different couples in disagreement.

Consider a mechanism where a wife announces her desired transaction z_f. It is allowed by the planner if her husband's announced transaction z_m is that $\boldsymbol{\psi}(z_f, z_m) \leq 0$. Similarly for the husband. A first advantage that a mechanism possesses is that it may ask whether spouses are in agreement or disagreement and may adopt a different criterion of admissibility, $\bar{\boldsymbol{\psi}}(\boldsymbol{z}) \leq 0$, for couples in disagreement.[19] The first criterion, $\boldsymbol{\psi}$ is used to induce choices conditional on threat points and the second, $\bar{\boldsymbol{\psi}}$ to induce threat points. This is not all. Upon observing a pair of announcements (z_f, z_m) the tax system can only say that it is admissible, $\boldsymbol{\psi}(z_f, z_m) \leq 0$, $\boldsymbol{\psi}(z_f, z_m) > 0$ or not. In contrast, if we consider a mechanism in which each spouse is required to announce the desired transactions for both spouses, $\tilde{\boldsymbol{z}} = (\tilde{z}_f, \tilde{z}_m)$, then the planner can make (z_f, z_m) admissible only if both announcement agree. This allow the planner to better target the transactions it intends each couple to undertake in disagreement, thus enlarging the set of threat points which can be induced.

[19] This corresponds to using different schedules for couples in agreement and couples in disagreement. da Costa and Diniz (2016) argue that filing options play this role in real world tax systems. We discuss such systems in Section 5.2.

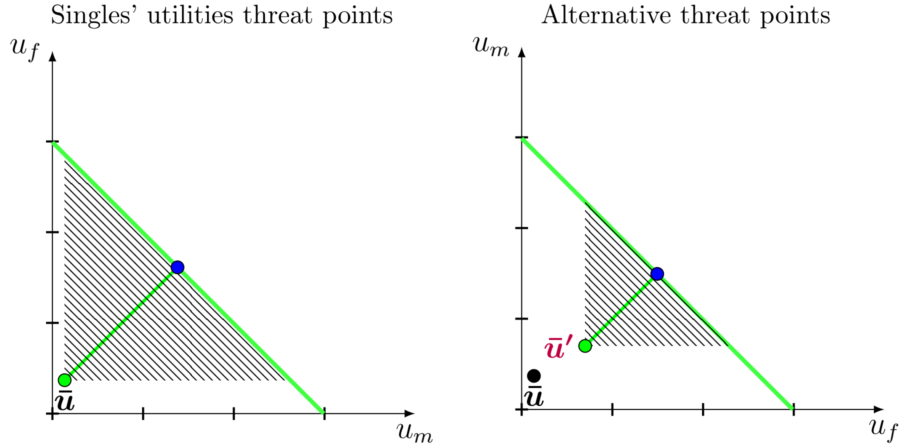

Figure 1: **Aligning Objectives** The panel in the left displays the set of individually rational utility pairs, when the threat point — green dot — is induced by the tax schedule in place. The blue dot denotes optimal choice for the couple. If the planner is able to induce a different threat point — green dot in the right panel — a new choice distribution of utilities is induced for the same transactions.

2.3 How to Improve on a Simple Tax Schedule

Transferable Utility

Having argued that the set of threat points that the planner can induce under a mechanism is larger than what is possible with a tax schedule, we now show how this can be helpful to improve upon the allocations implemented by a simple tax system, Ψ, and ask how changes in threat points may lead to better allocations.

Assume that utility is transferable across spouses. The convenient aspect of this example is that households choose transactions that maximize their surplus regardless of how it will later be shared — see Lemma 2. Given $\boldsymbol{z} = (c_f, -n_f, c_m, -n_m)$, the household solves

$$\max_{\mathfrak{c}_f} \left[\mathfrak{c}_f + h\left(1 - \frac{n_f}{\theta_f}\right) - \bar{u}_f\right]\left[c_f + c_m - \mathfrak{c}_f + h\left(1 - \frac{n_m}{\theta_m}\right) - \bar{u}_m\right],$$

where $(\bar{u}_f, \bar{u}_m)$ are the relevant threat points and $h(\cdot)$ is an increasing concave function.

If $\boldsymbol{z}$ maximizes the household's surplus under Ψ, we can represent the relevant utility possibility set for this couple as in Figure 1. The pair of utilities that arise from the Nash bargain is given by a 45 degrees line from $(\bar{u}_f, \bar{u}_m)$ to the frontier of the utility possibility set.

Assume that the planner's objective is to maximize the log of agents' utilities. In this case, the household and the planner's objectives are aligned if and only if $\bar{u}_f = \bar{u}_m$. If there is an mechanism $\mathcal{M}$, with the same set of attainable transactions, that is $\mathcal{Z}^c_{\mathcal{M}} = \mathcal{Z}^c_{\Psi}$,

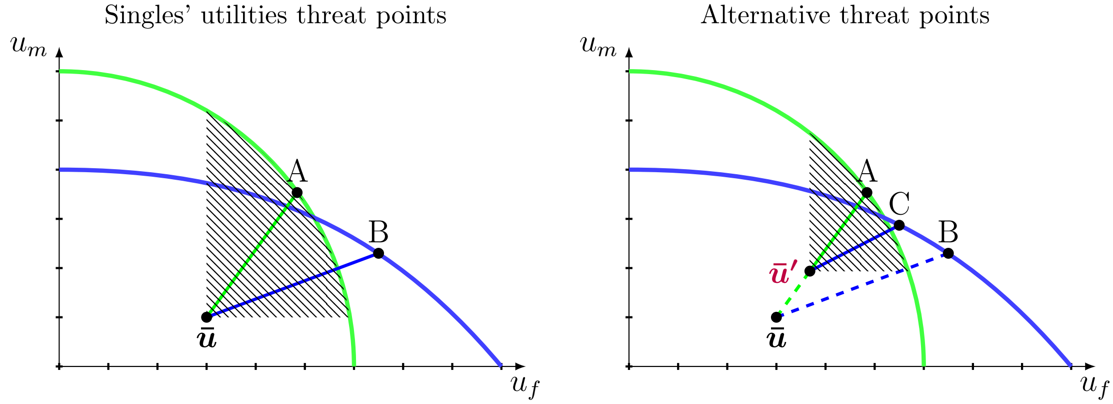

Figure 2: **Breaking Indifference:** The blue curve denotes the utility frontier for the same couple for the case in which a false report $\boldsymbol{\theta}' \neq \boldsymbol{\theta}$ is made. Point B denotes the utility pair the couple would attain in case of a lie. If the threat points changes from $\bar{\boldsymbol{u}} = (\bar{u}_f, \bar{u}_m)$ to $\bar{\boldsymbol{u}}' = (\bar{u}'_f, \bar{u}'_m)$ equilibrium choices are not changed, yet the deviation utility pair changes from B to C.

such that threat point are closer to the 45 degrees line, this would be sufficient to preserve couples' transactions, while increasing the value of the planner's objective. The right panel in Figure 1 illustrates such point.

da Costa and Diniz (2016) show that the option to file separately, present in many real world tax systems, often leads to smaller differences in the utilities attained by spouses in the disagreement game, even if it is never optimal for couples in agreement to file separately. Their work highlights one source of tax schedules' sub-optimality: its inability to separate households in agreement and disagreement. Sophisticated tax systems which allow for this separation can improve upon simple tax schedules but cannot implement all constrained efficient allocations, in general.

Partially Transferable Utility

As a second example consider the general case for which utility is only partially transferable. For a given $\boldsymbol{z} = (c_f, -n_f, c_m, -n_m)$, the conditional utility possibility set for the couple who has chosen $\boldsymbol{z}$ is the convex set bounded by the green curve in Figure 2. The set bounded by the blue curve is the conditional utility possibility set had the couple chosen $\boldsymbol{z}'$, instead. Point A in the left panel denotes the optimal choice of utilities for the household given $\boldsymbol{z}$, whereas point B denotes the preferred utility pair were the couple to choose transactions $\boldsymbol{z}'$. Finally, the green line connecting $(\bar{u}_f, \bar{u}_m)$ to point A defines the set of all potential threat points (closer to A than $\bar{\boldsymbol{u}}$) that have A as the optimal choice for this couple, given $\boldsymbol{z}$. If the couple were instead choosing $\boldsymbol{z}'$, then the blue line connecting $\bar{\boldsymbol{u}}$ to B would represent the analogous set of threat points that make B the preferred utility pair.

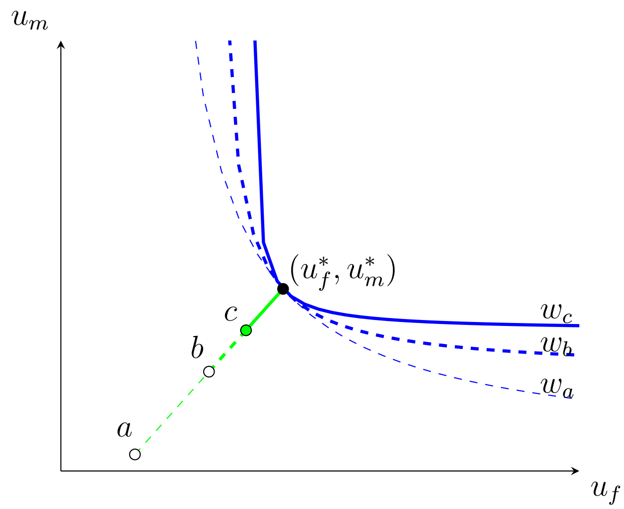

Figure 3: **Changing Elasticities.** The figure shows how the set of utility pairs which are preferred to (u_f^*, u_m^*) shrinks as we move along the curve connecting point a to point (u_f^*, u_m^*).

Assume that the couple is indifferent between A and B. It is therefore just indifferent between transactions $\boldsymbol{z}$ and $\boldsymbol{z}'$, both assumed to be feasible given the tax schedule in place.[20] Let $\boldsymbol{z}$ be the couple's choice, which we assume to be the one which raises more tax revenues and generates at least as much welfare according to the planners' metric. The planner cannot, in this case, increase this couple's tax liabilities since this would lead the couple to strictly prefer $\boldsymbol{z}'$.

Now, assume that there is $\mathcal{M}$ with $\mathcal{Z}^c_{\mathcal{M}} = \mathcal{Z}^c_{\Psi}$, which leads to a threat point along the green curve – e.g., $\bar{\boldsymbol{u}}'$ in Figure 2. Under $\mathcal{M}$, point A remains the optimal choice for the couple if it chooses $\boldsymbol{z}$ while point B is no longer optimal for the couple if it chooses $\boldsymbol{z}'$. A is now strictly better than C or any other utility pair attainable with transactions $\mathbf{z}'$; hence, $\boldsymbol{z}$ strictly preferred to $\boldsymbol{z}'$. Some room is created for a larger reform under which the set $\mathcal{Z}^c_{\mathcal{M}}$ is changed in a welfare improving direction.

By moving along the curve we are reducing the flexibility that a couple has to substitute the utility of one spouse for the other – Figure 3. We are in practice changing the relevant elasticities. This suggests a simple way to identify improvements over a tax schedule. If disagreement transactions can be strictly increased without changing the agreement utilities then, it relaxes the implementation constraints – see Corollary 1, thus creating an opportunity to improve upon the initial allocation.

Welfare improving reforms will typically combine elements of the two examples, aligning objectives and manipulating elasticities. How different threat points are induced, and what restrictions the planner faces to do it is at the essence of all that follows.

[20] Recall that axiom 'Preference for symmetry', as defined by Zambrano (2016), substitutes for 'symmetry' to deal with the non-convexity in the utility possibility sets.

3 Household Choices

To fully understand household choices we must further explain the decision process under both disagreement and agreement and describe the environment that conditions this process. For the latter we appeal to the notion of *institutional settings*, for which we use the generic denomination, $\mathcal{E}$. This stress the fact that, even when we are not using a mechanism perspective, we still need to think about the effect of our institutional settings on disagreement utilities.

3.1 Institutional Settings

An institutional setting, $\mathcal{E}$, is comprised of choice sets, $S^i_{\mathcal{E}}$, for gender i, $i = f, m$, single agents with typical element $s^i_{\mathcal{E}}$ and choice sets $S^{c,i}_{\mathcal{E}}$, with typical element $s^{c,i}_{\mathcal{E}}$, for gender i, $i = f, m$, married agents. Completing the description of $\mathcal{E}$, are functions $\zeta^i_{\mathcal{E}} : S^i_{\mathcal{E}} \mapsto Z$, and $\boldsymbol{\zeta}_{\mathcal{E}} : S^{c,f}_{\mathcal{E}} \times S^{c,m}_{\mathcal{E}} \mapsto Z \times Z$ mapping choices into transactions.

Now, $\mathcal{Z}^i_{\mathcal{E}} := \zeta^i_{\mathcal{E}}(S^i_{\mathcal{E}}) \subset Z$, $i = f, m$, and $\mathcal{Z}^c_{\mathcal{E}} := \boldsymbol{\zeta}_{\mathcal{E}}(S^{c,f}_{\mathcal{E}} \times S^{c,m}_{\mathcal{E}}) \subset Z \times Z$ define the transactions which are attainable for single agents and couples, respectively. For a gender i single agent to realize a transaction z under $\mathcal{E}$ it must choose $s \in S^i_{\mathcal{E}}$ such that $z = \zeta^i_{\mathcal{E}}(s)$, provided that such s exists. If no such s exists then, $z \notin \mathcal{Z}^i_{\mathcal{E}}$. Similarly $\boldsymbol{z}$ may be chosen by a couple if $s_f \in S^{c,f}_{\mathcal{E}}$ and $s_m \in S^{c,m}_{\mathcal{E}}$ exist such that $\boldsymbol{z} = \boldsymbol{\zeta}_{\mathcal{E}}(s_f, s_m)$. $\mathcal{Z}^c_{\mathcal{E}}$ is, therefore, the set of transactions, $\boldsymbol{z} \in Z \times Z$, which are feasible under the institutional setting $\mathcal{E}$. In all that follows we assume that for all $\mathcal{E}$, there exist $s_i \in S^i_{\mathcal{E}}$ such that $0 = \zeta^i_{\mathcal{E}}(s_i)$; for all $s_m \in S^{c,m}_{\mathcal{E}}$, $s_f \in S^{c,f}_{\mathcal{E}}$ such that $\boldsymbol{\zeta}_{\mathcal{E}}(s_f, s_m) = (0, z_m)$, and; for all $s_f \in S^{c,f}_{\mathcal{E}}$ $s_m \in S^{c,m}_{\mathcal{E}}$ such that $\boldsymbol{\zeta}_{\mathcal{E}}(s_f, s_m) = (z_f, 0)$. In words, it is always feasible for any agent to choose $z = 0$. Finally, for any couple $\boldsymbol{\theta}$, let $\mathcal{U}_{\mathcal{E}}(\boldsymbol{\theta}) := \bigcup_{\boldsymbol{z}' \in \mathcal{Z}^c_{\mathcal{E}}} \mathcal{U}_{\boldsymbol{z}'}(\boldsymbol{\theta}) \subset \mathbb{R}^2$ be the set of all attainable pair of utilities under $\mathcal{E}$.

An institutional setting is, of course, nothing but a general mechanism; just view any choice as a message regarding a desired outcome. Yet, for communication purposes it will be convenient to single out institutions that induce choices directly through budget sets, which we call tax systems.

A Simple Tax System, Ψ: A simple tax system is comprised of (possibly) gender based tax schedules for singles and a tax schedule for couples.

A tax schedule induces an allocation by defining the budgets which are available to households and single agents. For a single agent of gender i let $T_i : \mathbb{R}_+ \mapsto \mathbb{R}$ be the associated tax schedule. Then, we define the budget set, $\mathbb{B}_i$ as

$$\mathbb{B}_i := \{z; \psi_i(c, n) = \psi_i(z) = c - n - T_i(n) \leq 0\}. \tag{3}$$

In this case, $S^i_\Psi = Z$, and ζ^i_Ψ is the identity mapping, for z such that $\psi_i(z) \leq 0$ and $\zeta^i_\Psi(z) = 0$, otherwise. We also define the frontier of $\mathbb{B}_i$, $\partial\mathbb{B}_i := \{z \in \mathbb{B}_i; z' > z \Rightarrow z' \notin \mathbb{B}_i\}$.

For couples, the choice set of each spouse is $S^{c,i}_\Psi := Z$, $\boldsymbol{\zeta}_\Psi(\boldsymbol{z}) = \boldsymbol{z}$, for $\boldsymbol{z}$ such that $\boldsymbol{\psi}(\boldsymbol{z}) \leq 0$, and $\boldsymbol{\zeta}_\Psi(\boldsymbol{z}) = 0$, otherwise. The budget set, $\mathbb{B}^c$, is, therefore,

$$\mathbb{B}^c := \left\{\boldsymbol{z} \in Z^2; \boldsymbol{\psi}(\boldsymbol{z}) = c_f + c_m - n_f - n_m - T^f(n_f, n_m) - T^m(n_f, n_m) = 0\right\}. \tag{4}$$

Let also $\partial\mathbb{B}^c := \{\boldsymbol{z} \in \mathbb{B}^c; \boldsymbol{z}' > \boldsymbol{z} \Rightarrow \boldsymbol{z}' \notin \mathbb{B}^c\}$.[21]

A General Mechanism, $\mathcal{M}$: A *mechanism* or *game form*, (2), is mapped into our general notion of an institutional setting by noting that the choice sets are the message spaces, $S^i_{\mathcal{M}} = \Sigma^s_i$, $S^{c,i}_{\mathcal{M}} = \Sigma^c_i$, $i = f, m$, whereas the outcome functions are $\zeta^i_{\mathcal{M}}(\cdot) := g^s_i(\cdot)$, for singles, and $\boldsymbol{\zeta}_{\mathcal{M}}(\cdot,\cdot) := \boldsymbol{g}^c(\cdot,\cdot)$, for couples.

In principle, the outcome function could depend on messages from all agents in the economy. However we are assuming that there is a continuum of agents and a known distribution of types. Hence, assuming the outcome functions to depend only on what each agent (or couple) announces is without loss.

A Direct Mechanism, Γ: In a direct mechanism the planner asks each agent his or type. A single agent's type is his or her productivity, θ_i. For couples, a θ_f woman married to a θ_m man is said to have type (θ_f, θ_m, a) if the couple is in agreement, and a type (θ_f, θ_m, d) if the couple is in disagreement.[22]

For gender i singles, upon a report $\hat{\theta}_i$, an outcome function assigns transactions $z = \gamma^s_i(\hat{\theta}_i)$. Married agents must also report their type, which, as we have seen, is the combination of their productivity, θ_i, their spouses' productivity, θ_{-i}, for $i, -i = f, m$, and whether they are in agreement or disagreement $\iota \in \mathcal{I} := \{a, d\}$. The outcome function $\boldsymbol{\gamma} : (\Theta^2 \times \mathcal{I}) \times (\Theta^2 \times \mathcal{I})$ maps announcements of types into transactions for both spouses. Hence, $S^i_\Gamma = \Theta$, $S^{c,i}_\Gamma = \Theta^2 \times \mathcal{I}$, $\zeta^i_\Gamma(\cdot) := \gamma^s_i(\cdot)$, and, $\boldsymbol{\zeta}_\Gamma(\cdot,\cdot) := \boldsymbol{\gamma}^c(\cdot,\cdot)$.

[21] Note that, for $i = f, m.$, the tax system defines c^i through $c^i = y^i - T^i(y^f, y^m)$. For couples in agreement, only $c = c^f + c^m$ and not how it is split between c^f and c^m matters; income pooling holds conditional on holding $\bar{\boldsymbol{u}}$ fixed. A single function $T^c = T^f + T^m$ would suffice if all we were concerned about was choices in agreement. However, for couples in disagreement the exact split is relevant, so we use the general form.

[22] A married woman's type, $(\boldsymbol{\theta}, \iota) \in \Theta^2 \times \mathcal{I}$, $\mathcal{I} = \{a, d\}$, is therefore the same as her spouses'. A $(\boldsymbol{\theta}, \iota)$ woman can nonetheless be distinguished from her husband, a $(\boldsymbol{\theta}, \iota)$ man, for all purposes since gender is public information.

3.2 The Disagreement Game

Threat points are the flow utilities that arise in a non-cooperative game played by spouses while an agreement eludes them. To describe these games we define for any $\boldsymbol{\theta} \in \Theta^2$, a function, $\Phi : Z^2 \mapsto R^2$ which maps transactions, $\boldsymbol{z}$ into pair of utilities, (u_f, u_m),

$$\Phi\left(z_f, z_m; \boldsymbol{\theta}\right) \equiv \begin{pmatrix} \phi_f\left(z_f, z_m; \boldsymbol{\theta}\right) \\ \phi_m\left(z_f, z_m; \boldsymbol{\theta}\right) \end{pmatrix} = \begin{pmatrix} u_f \\ u_m \end{pmatrix}. \tag{5}$$

$\Phi\left(z_f, z_m; \boldsymbol{\theta}\right)$ summarizes the decision protocols, choice sets and outcome functions which determine how resources are used to generate consumption goods for spouses in disagreement, how transferable goods are assigned to spouses given who brought what home, what is available, how much leisure each one is consuming, and so on.[23] We endow $\Phi(\cdot, \cdot; \boldsymbol{\theta})$, with properties which are general enough to be satisfied by most sensible examples of fully specified games available while still generating testable restrictions on equilibrium behavior.

We assume $\Phi(\cdot, \cdot; \boldsymbol{\theta})$ is such that: *i)* for all $\boldsymbol{z}$, and all $\boldsymbol{\theta}$, $\Phi(\boldsymbol{z}, \boldsymbol{\theta}) \in \mathring{\mathcal{U}}_{\boldsymbol{z}}(\boldsymbol{\theta})$, where $\mathring{\mathcal{U}}_{\boldsymbol{z}}(\boldsymbol{\theta})$ is the interior of $\mathcal{U}_{\boldsymbol{z}}(\boldsymbol{\theta})$; *ii)* $\phi_f\left(\cdot, z_m; \boldsymbol{\theta}\right)$ is increasing in z_f for all z_m; *iii)* $\phi_m\left(z_f, \cdot; \boldsymbol{\theta}\right)$ is increasing in z_m for all z_f, and *iv)* if $\boldsymbol{z} > \boldsymbol{z'}$ then $\Phi(\boldsymbol{z}, \boldsymbol{\theta}) \neq \Phi(\boldsymbol{z'}, \boldsymbol{\theta})$ and $\Phi(\boldsymbol{z'}, \boldsymbol{\theta}) \not> \Phi(\boldsymbol{z}, \boldsymbol{\theta})$.

To make sense of these properties, starting with *(ii)* and *(iii)*, note that z_i is increased either by making more consumption goods available without an increase in spouse i's effort, n_i or by reducing i's effort without a reduction in resources available for consumption of both spouses. We assume that in either case this increases spouse i's utility. It may or may not increase his or her spouse's utility. Because, leisure is not transferable across spouses, a lower n_i can only reduce i's utility if his or her consumption is substantially reduced. Similarly, provided that a higher c_i does not lead to a lower $\mathfrak{c}_i$ the assumption is valid for this case as well.

As for *(i)*, agents act non-cooperatively when they are in disagreement. They will not, in general, be able to reach all points in $\mathcal{U}_{\boldsymbol{z}}(\boldsymbol{\theta})$.

Finally, *(iv)* says that if the household has strictly larger transactions, then at least one of the spouses is strictly better off.

Example 3.1. *Assume that in disagreement the household adopts the following sharing rule for the consumption goods, $\mathfrak{c}_i = s_i(\mathfrak{c}_f + \mathfrak{c}_m)$, with $s_i = c_i/(c_f + c_m)$. Consumption*

[23] By construction, disagreement is an off-equilibrium phenomenon, which observable consequences are only indirect, though the impact that threat points have on choices in agreement. Ideally we would like to use a model for the disagreement games, were it to exist, which testable implications had been scrutinized as to make it a consensual model for household behavior.

of each spouse is proportional to what each one contributes to the set of available consumption goods. We assume that spouses in disagreement are not be able to attain all material gains from cohabitation by letting the set of available allocations for spouses in disagreement be given by a function $F^d(\boldsymbol{z}, \boldsymbol{\theta}) \subset F^a(\boldsymbol{z}, \boldsymbol{\theta})$, e.g.,

$$F^d(\boldsymbol{z}, \boldsymbol{\theta}) := \{((\mathfrak{c}_f, \mathfrak{l}_f), (\mathfrak{c}_m, \mathfrak{l}_m));\ \mathfrak{c}_f + \mathfrak{c}_m \leq \beta [c_f + c_m];\\ \mathfrak{l}_f \leq 1 - n_f/\theta_f; \mathfrak{l}_m \leq 1 - n_m/\theta_m\},$$

where $1 < \beta < \alpha$ incorporates the fact that, despite some inefficiencies that arise due to disagreement, spouses might still enjoy some benefits from cohabitation.

For $\boldsymbol{z} = (c^f, -n^f, c^m, -n^m)$ we, therefore, have

$$\Phi(\boldsymbol{z}, \boldsymbol{\theta})' = \left(u_f\left(\beta c^f, 1 - \frac{n^f}{\theta^f}\right), u_m\left(\beta c^m, 1 - \frac{n^m}{\theta^m}\right)\right).$$

It is not hard to see that Φ has properties (i) to (iv).

Note how Φ summarizes through the sharing rule the division of transferable resources and through F^d the results of spouses' interaction under disagreement.

Disagreement Game Equilibria Threat points are determined by Φ evaluated at the transaction $\boldsymbol{z}$ which arises at the Nash equilibrium of this non-cooperative game. A key underlying assumption of our approach is that spouses are *not* able to commit to a strategy to be used in disagreement.[24]

Spouse i chooses an element from his/her choice set $S^{c,i}_{\mathcal{E}}$. This, along with i's spouse's choice, $s^{c,j}_{\mathcal{E}}$, is mapped into transactions by $\boldsymbol{\zeta}_{\mathcal{E}}$ and finally into i's utility by ϕ^i. The strategy space available to each spouse depends on the specific institutional arrangements, $\mathcal{E}$, that we are considering. Let, $\chi^f_{\mathcal{E}} : S^{c,m}_{\mathcal{E}} \times \Theta^2 \mapsto S^{c,f}_{\mathcal{E}}$, be defined by

$$\chi^f_{\mathcal{E}}(s^{c,m}_{\mathcal{E}}, \boldsymbol{\theta}) := \operatorname*{argmax}_{s \in S^{c,f}_{\mathcal{E}}} \phi_f\left(\boldsymbol{\zeta}_{\mathcal{E}}(s, s^{c,m}_{\mathcal{E}}); \boldsymbol{\theta}\right),$$

be the wife's reaction function for the game played under $\mathcal{E}$, that is just a function that gives the wife's best choice when her husband is choosing $s^{c,m}_{\mathcal{E}}$. $\chi^m_{\mathcal{E}}(s^{c,f}_{\mathcal{E}}, \boldsymbol{\theta})$, is defined in an analogously way for the husband. An equilibrium, $s^*_f = \chi^f_{\mathcal{E}}(s^*_m, \boldsymbol{\theta})$, $s^*_m = \chi^m_{\mathcal{E}}(s^*_f, \boldsymbol{\theta})$, for the disagreement game defines $(\bar{z}^{c,f}_{\mathcal{E}}(\boldsymbol{\theta}), \bar{z}^{c,m}_{\mathcal{E}}(\boldsymbol{\theta})) = \boldsymbol{\zeta}_{\mathcal{E}}(s^*_f, s^*_m)$, the disagreement transactions under the instutitional setting $\mathcal{E}$ for the equilibrium strategies (s^*_f, s^*_m).

Whether an equilibrium exists and, when it exists, whether it is unique depends not

[24] This rules out many different approaches for determining the threat points — see Myerson (1997) for a discussion.

only on the properties of Φ but also on those of $\mathcal{E}$. Thus, for each institution, we make assumptions directly on the game effectively played.[25]

Provided that an equilibrium does exist, $\bar{\boldsymbol{z}}_{\mathcal{E}}(\boldsymbol{\theta}) = (\bar{z}_{\mathcal{E}}^{c,f}(\boldsymbol{\theta}), \bar{z}_{\mathcal{E}}^{c,m}(\boldsymbol{\theta}))$ are the transactions that would be conducted by a couple $\boldsymbol{\theta}$ if they were not able to reach an agreement under the institutional setting $\mathcal{E}$. We finally use $\bar{\boldsymbol{u}}_{\mathcal{E}}(\boldsymbol{\theta}) = (\bar{u}_{\mathcal{E}}^{f}(\boldsymbol{\theta}), \bar{u}_{\mathcal{E}}^{m}(\boldsymbol{\theta}))$ to denote the threat points that arise under $\mathcal{E}$.

3.3 Choices in Agreement

For any $\mathcal{E}$ and any $\bar{\boldsymbol{u}} = (\bar{u}_f, \bar{u}_m)$ in the interior of $\mathcal{U}_{\mathcal{E}}(\boldsymbol{\theta})$, let

$$W(\boldsymbol{z};\boldsymbol{\theta},\bar{\boldsymbol{u}}) := \begin{cases} \max & \left(u_f\left(\mathfrak{c}_f,\mathfrak{l}_f\right) - \bar{u}_f\right)\left(u_m\left(\mathfrak{c}_m,\mathfrak{l}_m\right) - \bar{u}_m\right) \\ \text{s.t.} & \left(\left(\mathfrak{c}_f,\mathfrak{l}_f\right),\left(\mathfrak{c}_m,\mathfrak{l}_m\right)\right) \in F^a(\boldsymbol{z},\boldsymbol{\theta}) \end{cases}. \tag{6}$$

Then, for any subset A of Z^2 for which there is at least one $\boldsymbol{z} \in A$ such that $\bar{\boldsymbol{u}}$ is in the interior of $\mathcal{U}_{\boldsymbol{z}}(\boldsymbol{\theta})$ equation (6) defines for a couple $\boldsymbol{\theta}$, a utility function, $W: A \to \mathbb{R}$ parametrized by $\bar{\boldsymbol{u}}$, which represents a complete pre-order, $\succsim_{\boldsymbol{\theta},\bar{\boldsymbol{u}}}$, in $A \subset Z^2$.

Define for each environment $\mathcal{E}$, the *conditional (on $\bar{\boldsymbol{u}}$) choices* by a couple $\boldsymbol{\theta}$,

$$\hat{\boldsymbol{s}}_{\mathcal{E}}(\boldsymbol{\theta}|\bar{\boldsymbol{u}}) := \underset{s_f,s_m \in (S_{\mathcal{E}}^{c,f} \times S_{\mathcal{E}}^{c,m})}{\operatorname{argmax}} W(\boldsymbol{\zeta}_{\mathcal{E}}(s_f,s_m);\boldsymbol{\theta},\bar{\boldsymbol{u}}). \tag{7}$$

We let $\hat{\boldsymbol{z}}_{\mathcal{E}}(\boldsymbol{\theta}|\bar{\boldsymbol{u}}) := \boldsymbol{\zeta}_{\mathcal{E}}(\hat{\boldsymbol{s}}_{\mathcal{E}}(\boldsymbol{\theta},\bar{\boldsymbol{u}}))$ denote agreement transactions conditional on $\bar{\boldsymbol{u}}$.

3.4 Equilibrium Allocations Under $\mathcal{E}$

For $\bar{\boldsymbol{u}} = \bar{\boldsymbol{u}}_{\mathcal{E}}(\boldsymbol{\theta})$, i.e., when threat points are those that arise in the equilibrium of the disagreement game, we arrive at

$$\boldsymbol{s}_{\mathcal{E}}(\boldsymbol{\theta}) := \underset{s_f,s_m \in (S_{\mathcal{E}}^{c,f} \times S_{\mathcal{E}}^{c,m})}{\operatorname{argmax}} W(\boldsymbol{\zeta}_{\mathcal{E}}(s_f,s_m);\boldsymbol{\theta},\bar{\boldsymbol{u}}_{\mathcal{E}}(\boldsymbol{\theta})),$$

the equilibrium choices for spouses in a couple $\boldsymbol{\theta}$ under institutions $\mathcal{E}$.[26] Equilibrium transactions, $\boldsymbol{z}_{\mathcal{E}}(\boldsymbol{\theta})$, are defined by $\boldsymbol{z}_{\mathcal{E}}(\boldsymbol{\theta}) := \boldsymbol{\zeta}_E(s_{\mathcal{E}}^{c,f}(\boldsymbol{\theta}), s_{\mathcal{E}}^{c,m}(\boldsymbol{\theta}))$.

[25] For convenience only we have represented a pure strategy Nash equilibrium. The results are exactly the same if only mixed strategy equilibria exist and $\bar{\boldsymbol{u}}_{\mathcal{E}}(\boldsymbol{\theta})$ is the expected utilities attained in the disagreement game.

[26] Noting that $\boldsymbol{s}_{\mathcal{E}}(\boldsymbol{\theta}) = (s_{\mathcal{E}}^{c,f}(\boldsymbol{\theta}), s_{\mathcal{E}}^{c,m}(\boldsymbol{\theta}))'$ and $\boldsymbol{z}_{\mathcal{E}}(\boldsymbol{\theta}) = (z_{\mathcal{E}}^{c,f}(\boldsymbol{\theta}), z_{\mathcal{E}}^{c,m}(\boldsymbol{\theta}))'$, we have, for all $\mathcal{E}$ and all $\boldsymbol{\theta}$, $(s_{\mathcal{E}}^{c,f}(\boldsymbol{\theta}), s_{\mathcal{E}}^{c,m}(\boldsymbol{\theta})) = \hat{\boldsymbol{s}}_{\mathcal{E}}(\boldsymbol{\theta}, \bar{\boldsymbol{u}}_{\mathcal{E}}(\boldsymbol{\theta}))$, $(z_{\mathcal{E}}^{c,f}(\boldsymbol{\theta}), z_{\mathcal{E}}^{c,m}(\boldsymbol{\theta})) = \hat{\boldsymbol{z}}_{\mathcal{E}}(\boldsymbol{\theta}, \bar{\boldsymbol{u}}_{\mathcal{E}}(\boldsymbol{\theta}))$.

We may therefore define the *allocation implemented by* $\mathcal{E}$,

$$\left\{ \left(z^i_{\mathcal{E}}(\theta_i)\right)_{\theta_i, i=f,m}, \left(\boldsymbol{z}_{\mathcal{E}}(\boldsymbol{\theta})\right)_{\boldsymbol{\theta}} \right\}, \tag{8}$$

We let $\{(u^i_{\mathcal{E}}(\theta_i))_{\theta_i}, \left(u^{c,i}_{\mathcal{E}}(\boldsymbol{\theta})\right)_{\boldsymbol{\theta}}\}_{i=f,m}$, represent the associated utility profile. Recalling that $G(\cdot)$ is the transformation function that represents the economy's technology, an allocation is *feasible* if

$$G\left(\sum_{i=f,m}\left\{\sum_{\theta_i} z^i_{\mathcal{E}}(\theta_i)\mu^i_{\mathrm{s}}(\theta_i) + \sum_{\boldsymbol{\theta}} z^{c,i}_{\mathcal{E}}(\boldsymbol{\theta})\boldsymbol{\mu}(\boldsymbol{\theta})\right\}\right) \leq 0,$$

where μ^i_{s}, $i = f, m.$, and $\boldsymbol{\mu}$ are determined as a stable matching of the marriage market.

3.5 Who Marries Whom?

To decide whether to get married and whom to get married to, a type θ_f woman living in an institutional environment $\mathcal{E}$ anticipates the flow material utilities, $u^{c,f}_{\mathcal{E}}(\theta_f, \tilde{\theta}_m)$, she attains, for each possible type of man $\tilde{\theta}_m$ she may marry. She also anticipates the material utility from remaining single, $u^f_{\mathcal{E}}(\theta_f)$. This is not however all that matters for marriage. The total utility she attains from marriage is the sum of this 'material' or 'consumption' utility and a 'love' term, ξ. We assume that this 'love' term does not depend on whether the household is in agreement or disagreement, which means that the Nash product (1) is independent of love.

Before entering the marriage market, a given woman, say Miss a, draws a love vector, $\boldsymbol{\xi}_a = (\xi_a(\emptyset), (\xi_a(\theta_m))_{\theta_m}) \in R^{|\Theta|+1}$, from a distribution

$$\boldsymbol{H}_f(\xi) = \left(H^{\emptyset}_f(\xi(\emptyset)), \left\{H^{\theta_m}_f(\xi(\theta_m))\right\}_{\theta_m}\right),$$

with associated density $\boldsymbol{h}_f(\xi)$, which defines an emotional payoff $\xi_a(\theta_m)$ that she attains by marrying a type θ_m, for each θ_m, as well as an emotional payoff, $\xi_a(\emptyset)$, from remaining single. The utility that she attains if she marries a θ_m type man is $\mathfrak{v}_a(\theta_m) = u^{c,f}_{\mathcal{E}}(\eta) + \xi_a(\theta_m)$. Were she to marry, instead, a type $\hat{\theta}_m$ man she would attain $\mathfrak{v}_a(\hat{\theta}_m) = u^{c,f}_{\mathcal{E}}(\theta_f, \hat{\theta}_m) + \xi_a(\hat{\theta}_m)$. Finally, were she to remain single, she would attain $u^f_{\mathcal{E}}(\theta_f) + \xi_a(\emptyset)$. An analogous construct is defined for each man.

Note that the emotional payoff that a specific woman (man) attains depends only on the type of her (his) spouse, not on his (her) identity. This is only to allow for the use of a continuous distribution of love shocks. More importantly, have we assumed that agents cannot manipulate the information regarding their productivities to potential partners.

For each woman, a, $\boldsymbol{\xi}_a$ and $u_g^{c,f}(\boldsymbol{\theta})$ together define an ordering, $\succeq_{\theta_f,\boldsymbol{\xi}}$, of 'male' types that she would want to marry, including the option of remaining single. Since the number of types is finite the set of orderings is finite regardless of the distribution of shocks ξ. Given these orderings we define an extended type $\vartheta_i = (\theta_i, \succeq_{\theta_i,\boldsymbol{\xi}})$ for a gender i agent. The set of extended types too is a finite set. Note also that a gender $-i$ agent is indifferent between any two ϑ_i and $\hat{\vartheta}_i$ such that $\vartheta_i = (\theta_i, \succeq_{\theta_i,\boldsymbol{\xi}})$ and $\hat{\vartheta}_i = (\theta_i, \succeq_{\theta_i,\hat{\boldsymbol{\xi}}})$.

Lack of commitment implies that promises of transfers or distribution of power cannot be credibly made. We also rule out the type of pre-nuptial arrangements or pre-marital transfers that could substitute for this type of commitment. Hence, the utility attained for each potential marriage is pre-determined for each woman and each man, and does not depend on any actions that agents can make prior to marriage. We are, therefore in a setting which Pollak (2016) calls Bargain in Marriage, for which the Gale-Shapley matching model applies. The marriage problem is a pair $(\mathbf{M}, \mathbf{W})$ where $\mathbf{M}$ is the finite set of extended types for men and $\mathbf{W}$ the finite set of extended types for women. The outcome of a marriage problem is a match, i.e., a function $\nu(\vartheta_f, \vartheta_m)$ giving the fraction of women of type ϑ_f married to men of type ϑ_m, which satisfies the following feasibility constraints:

1. $\sum_{\vartheta_m} \nu(\vartheta_f, \vartheta_m) \leq 1$ for all ϑ_f
2. for all ϑ_m, $\sum_{\vartheta_f} \nu(\vartheta_f, \vartheta_m)\pi_f(\vartheta_f) \leq \pi_m(\vartheta_m)$

where $\pi_g(\vartheta_g)$ is the fraction of gender g agents of extended type ϑ_g.

We say that a matching is stable if there is no married agent who would rather be single, and there are no two married or unmarried persons who would prefer to form a new union. Theorem 1 in Azevedo and Hatfield (2015) guarantees that a stable matching exists for the marriage problem above.

A couple of things which are specific to our setting are worth mentioning. First, agents in each side of the market are indifferent between all extended types that share the same productivity. In particular, preferences are not strict in our setting. Two agents with different extended types but the same productivity are viewed by the agents of the other gender as equivalent. Second, the definition of matching used above, borrowed from Bodoh-Creed (2016) does not specify a map between agents on either side of the market, which is in contrast with the most traditional definition (e.g., Roth and Sotomayor (1990)). Within this traditional paradigm ν therefore represents an equivalent class of matchings. Third, although matching is only based on the ordering of types, ξ, is still payoff relevant, thus important from a normative perspective.

For a given stable matching let $\mathscr{H}_f^{\theta_m}(\xi(\theta_m)|\boldsymbol{\theta})$ denote the probability that a type θ_f-woman, matched with a θ_m-man such that $\boldsymbol{\theta} = (\theta_f, \theta_m)$ has a love parameter $\boldsymbol{\xi}$ with a θ_m entry no greater than $\xi(\theta_m)$. Let $\mathcal{H}_m^{\theta_f}(\xi(\theta_f)|\boldsymbol{\theta})$ be the analogous definition for a θ_m man such that $\boldsymbol{\theta} = (\theta_f, \theta_m)$, $\mathcal{H}_f^{\emptyset}(\xi(\emptyset)|\emptyset)$ the analogous definition for a θ_f woman that remains single and $\mathcal{H}_m^{\emptyset}(\xi(\emptyset)|\emptyset)$, for a θ_f man that remains single.

In this case, for a social welfare function $\Upsilon : R \to R$, write the government's objective,

$$\begin{aligned}\max \sum_{\boldsymbol{\theta}} \Bigg[\int \Upsilon(u_{\mathcal{E}}^{c,f}(\boldsymbol{\theta}) + \xi(\theta_m)) h_f^{\theta_m}(\xi(\theta_m)|\boldsymbol{\theta}) d\xi(\theta_m) \\ + \int \Upsilon(u_{\mathcal{E}}^{c,m}(\boldsymbol{\theta}) + \xi(\theta_f)) h_m^{\theta_f}(\xi(\theta_f)|\boldsymbol{\theta}) d\xi(\theta_f) \Bigg] \mu(\boldsymbol{\theta}) \\ + \sum_{\theta_f} \int \Upsilon(u_{\mathcal{E}}^{f}(\theta_f) + \xi(\emptyset)) h_f^{\emptyset}(\xi(\emptyset)|\emptyset) d\xi(\emptyset) \mu_f(\theta_f) \\ + \sum_{\theta_m} \int \Upsilon(u_{\mathcal{E}}^{m}(\theta_m) + \xi(\emptyset)) h_m^{\emptyset}(\xi(\emptyset)|\emptyset) d\xi(\emptyset) \mu_m(\theta_m),\end{aligned}$$

where lower case h are the densities associated with the distributions defined in the previous paragraph.

The conditional distributions $\mathcal{H}_i^{\theta_{-i}}(\xi(\theta_{-i})|\boldsymbol{\theta})$, $\mathcal{H}_i^{\emptyset}(\xi(\emptyset)|\emptyset)$, $i, -i = f, m.$, are endogenously determined and so are $\boldsymbol{\mu}$, μ_s^i, $i = f, m.$, for they all depend on choices which are affected by policy through its effect on $u_{\mathcal{E}}^{c,i}(\boldsymbol{\theta})$, $u_{\mathcal{E}}^{i}(\theta_i)$, $i = f, m$.

4 Implementable Allocations

In this section we formalize the heuristic examples of Section 2 and explain the latitude that the planner has in manipulating threat points. The first step toward this end is to prove that we can rely on the revelation principle to characterize the complete set of incentive-feasible allocations.

Proposition 1. *Let*

$$\left\{ \left(z_{\mathcal{M}}^i(\theta_i) \right)_{\theta_i, i=f,m}, \left(\boldsymbol{z}_{\mathcal{M}}(\boldsymbol{\theta}) \right)_{\boldsymbol{\theta}} \right\} \tag{9}$$

be an implementable allocation, i.e., an allocation for which there is a mechanism $\mathcal{M}$ for which (9) is the equilibrium allocation. There, there exists a Direct Mechanism - DM - which implements the same allocation.

Proof. See Appendix A.1. □

Note that for any given marriage problem and a given Gale-Shapley algorithm, there is a one to one mapping from the equilibrium allocation and the stable matching. Hence, the above proposition does take into account the endogeneity of all distributions.

Using Proposition 1 we can now precisely define what an incentive feasible allocation is. Start by noting that, under a direct mechanism, an announcement $(\hat{\boldsymbol{\theta}}, \iota) \in \Theta^2 \times \mathcal{I}$ by a husband and an announcement $(\hat{\boldsymbol{\theta}}', \iota') \in \Theta^2 \times \mathcal{I}$ by his wife induce transactions $\boldsymbol{\gamma}^c(\hat{\boldsymbol{\theta}}', \iota', \hat{\boldsymbol{\theta}}, \iota)$. If the couple is in agreement, this generates the pair of utility at the solution of (6). If the couple is in disagreement Φ maps these transactions into a pair of utilities $(\bar{u}_f, \bar{u}_m)$. The direct mechanism therefore defines, for a couple in disagreement, a game in the space of announcements, $\Theta^2 \times \mathcal{I}$. The reaction function for f in a couple $\boldsymbol{\theta}$ is

$$\chi^f_\Gamma(\hat{\boldsymbol{\theta}}_m, \hat{\iota}_m, \boldsymbol{\theta}) := \underset{(\hat{\boldsymbol{\theta}}, \hat{\iota}) \in \Theta^2 \times \mathcal{I}}{\operatorname{argmax}} \; \phi_f \left(\gamma^c_f(\hat{\boldsymbol{\theta}}, \hat{\iota}, \hat{\boldsymbol{\theta}}_m, \hat{\iota}_m), \gamma^c_m(\hat{\boldsymbol{\theta}}, \hat{\iota}, \hat{\boldsymbol{\theta}}_m, \hat{\iota}_m); \boldsymbol{\theta} \right), \tag{10}$$

with analogous definition for $\chi^m_\Gamma(\hat{\boldsymbol{\theta}}_f, \hat{\iota}_f, \boldsymbol{\theta})$. An equilibrium for this game,

$$(\boldsymbol{\theta}^*_f, \iota^*_f, \boldsymbol{\theta}^*_m, \iota^*_m) = \left(\chi^f_\Gamma(\hat{\boldsymbol{\theta}}^*_m, \iota^*_m, \boldsymbol{\theta}), \chi^m_\Gamma(\hat{\boldsymbol{\theta}}^*_f, \iota^*_f, \boldsymbol{\theta}) \right)$$

finally determines the threat points $\left(\bar{u}^f_\Gamma(\boldsymbol{\theta}), \bar{u}^m_\Gamma(\boldsymbol{\theta}) \right) \in \mathring{\mathcal{U}}_\Gamma(\boldsymbol{\theta})$.

Using these definitions, we characterize the set of incentive-feasible allocations through $z^i_\Gamma(\theta_i) = \gamma^s_i(\theta_i)$ for all θ_i, $i = f, m.$, $\boldsymbol{z}_\Gamma(\boldsymbol{\theta}) = \boldsymbol{\gamma}(\boldsymbol{\theta}, a, \boldsymbol{\theta}, a)$ for all $\boldsymbol{\theta}$, such that

i) For every single agent θ_i,

$$u_i(\gamma^s_i(\theta_i), \theta_i) \geq u_i(\gamma^s_i(\hat{\theta}_i), \theta_i) \; \forall \hat{\theta}_i, \; i = f, m.; \tag{11}$$

ii) For every couple $\boldsymbol{\theta}$, and all $(\hat{\boldsymbol{\theta}}, \hat{\hat{\boldsymbol{\theta}}}, \iota, \hat{\iota}) \in \Theta^2 \times \mathcal{I}^2$,

$$W(\boldsymbol{\gamma}(\boldsymbol{\theta}, a, \boldsymbol{\theta}, a), \boldsymbol{\theta}; \bar{u}_f, \bar{u}_m) \geq W(\boldsymbol{\gamma}(\hat{\boldsymbol{\theta}}, \iota, \hat{\hat{\boldsymbol{\theta}}}, \hat{\iota}), \boldsymbol{\theta}; \bar{u}_f, \bar{u}_m), \tag{12}$$

where,

$$\begin{pmatrix} \bar{u}_f \\ \bar{u}_m \end{pmatrix} = \begin{pmatrix} \phi_f \left(\boldsymbol{\gamma}(\boldsymbol{\theta}, d, \boldsymbol{\theta}, d); \boldsymbol{\theta} \right) \\ \phi_m \left(\boldsymbol{\gamma}(\boldsymbol{\theta}, d, \boldsymbol{\theta}, d); \boldsymbol{\theta} \right) \end{pmatrix}; \tag{13}$$

iii) For all $\boldsymbol{\theta}$,

$$(\boldsymbol{\theta}, d) \in \underset{\hat{\boldsymbol{\theta}} \in \Theta^2, \iota \in \mathcal{I}}{\operatorname{argmax}} \, \phi_f \left(\boldsymbol{\gamma}(\hat{\boldsymbol{\theta}}, \iota, \boldsymbol{\theta}, d); \boldsymbol{\theta} \right) \tag{14}$$

and

$$(\boldsymbol{\theta}, d) \in \underset{\hat{\boldsymbol{\theta}} \in \Theta^2, \iota \in \mathcal{I}}{\operatorname{argmax}} \, \phi_m \left(\boldsymbol{\gamma}(\boldsymbol{\theta}, d, \hat{\boldsymbol{\theta}}, \iota); \boldsymbol{\theta} \right) \tag{15}$$

iv)

$$G\left(\sum_{i=f,m}\left\{\sum_{\theta_i} z^i_\Gamma(\theta_i)\mu^i_{\rm s}(\theta_i)+\sum_{\boldsymbol{\theta}} z^{c,i}_\Gamma(\boldsymbol{\theta})\boldsymbol{\mu}(\boldsymbol{\theta})\right\}\right)\leq 0. \tag{16}$$

v) $\mu^i_{\rm s}$, $\boldsymbol{\mu}$, $\mathscr{H}^{\emptyset}_f(\xi(\emptyset)|\emptyset)$, $\mathscr{H}^{\theta_{-i}}_i(\xi(\theta_{-i})|\boldsymbol{\theta})$, $i,-i,=\ f,m.$, are determined from a stable matching of the marriage problem.

Item (i) is simply an incentive compatibility constraint for a single agent which implicitly assumes that agents cannot misreport their gender. Constraint (ii) guarantees that a type $\boldsymbol{\theta}$ couple in agreement does not misreport its type for threat points which are defined by the pair of utilities that arise when both decide to announce truthfully in case of disagreement. That truth-telling be an equilibrium announcement for spouses in disagreement is what (iii) guarantees. Constraint (iv) is the feasibility constraint. Constraint (v) guarantees that the distribution of marriages are consistent with a stable matching in the marriage problem.

A couple of things are worth noting. First, items (i) to (iii) are independent of marriage market outcomes. This is due to the separability between consumption utility and love and the invariability of love with respect to the state in which the couple finds itself. Second, we do not require feasibility under disagreement, but we do impose incentive-compatibility.

4.1 Constrained (In)Efficient Allocations

In the examples in Section 2 we have shown how the planner may improve upon an allocation by manipulating threat points. Now we ask how much latitude the planner has to proceed as in those examples.

Spouses share a type which they are asked to announce under the direct mechanism. Conflicting announcements means that a lie has been told. By harshly punishing couples who send conflicting messages, e.g., $\boldsymbol{\gamma}(\boldsymbol{\theta}',\iota',\boldsymbol{\theta}'',\iota'')=(0,0)$ for $(\boldsymbol{\theta}',\iota')\neq(\boldsymbol{\theta}'',\iota'')$, the mechanism guarantees that any pair of coinciding announcements is an equilibrium of the disagreement game. In particular, the truthful announcement, $(\boldsymbol{\theta},d,\boldsymbol{\theta},d)$, is an equilibrium announcement for a couple $\boldsymbol{\theta}$ in disagreement. The planner, therefore, enjoys quite some latitude in the choice of $\boldsymbol{\gamma}(\boldsymbol{\theta},d,\boldsymbol{\theta},d)$. It must only make sure that constraint (12) is satisfied for all $\boldsymbol{\theta}$: no couple $\boldsymbol{\theta}$ in agreement would rather announce to be in disagreement and coordinate on an announcement $(\boldsymbol{\theta}',d)$. That is, the only restriction that the planner faces in choosing $\bar{\boldsymbol{u}}$ is that the transaction $\boldsymbol{z}$ such that $\bar{\boldsymbol{u}}=\Phi(\boldsymbol{z},\boldsymbol{\theta})$, if it exists, is not strictly preferred by some couple in agreement. A sufficient condition is that $\boldsymbol{z}$ be in $\mathscr{Z}^c_{\mathscr{G}}$, for, in this case, $\boldsymbol{z}$ is already a feasible choice for all couples. This is the content of the next proposition.

Proposition 2. *For any $\mathcal{E}$, and any $\boldsymbol{\theta}$ let $\bar{U}_{\mathcal{E}}(\boldsymbol{\theta}) := \{\bar{\boldsymbol{u}} \in \mathbb{R}^2 | \exists \boldsymbol{z} \in \mathcal{Z}^c_{\mathcal{E}}; \bar{\boldsymbol{u}} = \Phi(\boldsymbol{z}, \boldsymbol{\theta})\}$. Then, for all $\bar{\boldsymbol{u}} \in \bar{U}_{\mathcal{E}}(\boldsymbol{\theta})$ there is a direct mechanism Γ, which implements $\bar{\boldsymbol{u}}$ without changing any other couple's transaction.*

Proof. See appendix A.1 □

Proposition 3, below, establishes an important restriction on the set of threat points which can be chosen without changes in the equilibrium allocation.

Proposition 3. *For a given $\bar{\boldsymbol{u}}$, assume that for all $\boldsymbol{z}$ for which $\bar{\boldsymbol{u}} = \Phi(\boldsymbol{z}, \boldsymbol{\theta})$ there is $\boldsymbol{\theta}'$ such that $\boldsymbol{z}_{\mathcal{E}}(\boldsymbol{\theta}') < \boldsymbol{z}$. Then, $\bar{\boldsymbol{u}}$ cannot be chosen without changing the equilibrium allocation.*

Proof. See appendix A.1 □

Without changing the equilibrium allocation, we cannot implement $\bar{\boldsymbol{u}}$ as disagreement utilities for $\boldsymbol{\theta}$-couples if it requires transactions $\boldsymbol{z}$ which are strictly larger than the agreement transactions, $\boldsymbol{z}'$, for some $\boldsymbol{\theta}'$-couple.

Propositions 2 and 3 refer to reforms that change threat points without changing the equilibrium transactions. Such reforms may improve the social objective in the case of transferable utilities, as the first example in Section 2 shows. Yet, in most cases, welfare improving reforms are associated with changes in equilibrium transactions. Because, the bounds defined by (12) depend on the equilibrium transactions, characterizing welfare improving reforms as a two step procedure for which one first changes threat points and only then derives the optimal allocations is not, in general, possible. A noteworthy exception is the second example of Section 2, which we formalize next.

We say that $\boldsymbol{z}_{\mathcal{E}}(\boldsymbol{\theta})$ is *distorted* if there is a feasible $\boldsymbol{z}'$ which expands the utility possibility set for household $\boldsymbol{\theta}$ without changing the transactions of all other agents. We say that $\boldsymbol{\theta}$ envies $\hat{\boldsymbol{\theta}}$'s allocation if $\boldsymbol{z}(\hat{\boldsymbol{\theta}}) \succsim_{\boldsymbol{\theta},\bar{\boldsymbol{u}}} \boldsymbol{z}(\boldsymbol{\theta})$. Using these definitions we can prove the following.

Proposition 4. *Given $\mathcal{E}$ define for any $\boldsymbol{\theta}$, and any $\bar{\boldsymbol{u}}$ in the interior of $\mathcal{U}_{\mathcal{E}}(\boldsymbol{\theta})$, the set*[27]

$$\Lambda_{\mathcal{E}}(\boldsymbol{\theta}, \bar{\boldsymbol{u}}) := \{(u_f, u_m) | \exists \alpha \in (0,1); (u_f, u_m) = \alpha \boldsymbol{u}_{\mathcal{E}}(\boldsymbol{\theta}) + (1-\alpha)\bar{\boldsymbol{u}}\}.$$

Assume that there is an alternative institutional environment $\mathcal{E}'$ such that $\boldsymbol{z}_{\mathcal{E}}(\boldsymbol{\theta}) = \boldsymbol{z}_{\mathcal{E}'}(\boldsymbol{\theta})$ for all $\boldsymbol{\theta}$, and $\bar{\boldsymbol{u}}_{\mathcal{E}'}(\hat{\boldsymbol{\theta}}) \in \Lambda_{\mathcal{E}}(\hat{\boldsymbol{\theta}}, \bar{\boldsymbol{u}}_{\mathcal{E}}(\hat{\boldsymbol{\theta}}))$ for some $\hat{\boldsymbol{\theta}}$. Let also, $\hat{H}_{\mathcal{E}} \subset \Theta^2$ be the set of all household types whose transactions are either envied by no one or are only envied by $\hat{\boldsymbol{\theta}}$.

Then, one of the following must be true: no transaction $\boldsymbol{z}(\boldsymbol{\theta})$ for $\boldsymbol{\theta} \in \hat{H}_{\mathcal{E}}$ is distorted, or; the allocation $(\boldsymbol{z}_{\mathcal{E}}(\boldsymbol{\theta}))_{\boldsymbol{\theta}}$ is constrained inefficient.

[27] $\Lambda_{\mathcal{E}}(\boldsymbol{\theta}, \bar{\boldsymbol{u}})$ is the line segment between $\bar{\boldsymbol{u}}_{\mathcal{E}}$ and $\boldsymbol{u}_{\mathcal{E}}$ in Figure 2

Proof. See Appendix A.1. □

If there are couples whose transactions are distorted then the allocation is not first best efficient. Inefficiencies are usually introduced to relax incentive constraints when a type envies the allocation intended to another type. What the proposition says is that a distorted allocation cannot be constrained efficient if there are feasible reforms that move threat points closer to equilibrium choices. Such reforms, if they existed, would allow the planner to relax the incentive constraints that motivated the distortions in the first place. Proposition 4 allows us to rule out constrained efficiency of an allocation implemented by an optimal tax schedule, for example, but it requires our being able to find such $\mathcal{E}'$. Corollary 1 provides a sufficient condition to identify when this is the case.

Corollary 1. *For a given $\mathcal{E}$, assume that for some $\boldsymbol{\theta}$ there is $\boldsymbol{z}' \in \mathcal{Z}^c_{\mathcal{E}}$, $\boldsymbol{z}' > \bar{\boldsymbol{z}}_{\mathcal{E}}(\boldsymbol{\theta})$ such that $\hat{\boldsymbol{z}}_{\mathcal{E}}(\boldsymbol{\theta}|\Phi(\boldsymbol{z}', \boldsymbol{\theta})) = \boldsymbol{z}_{\mathcal{E}}(\boldsymbol{\theta})$. Let also, $\hat{H}_{\mathcal{E}} \subset \Theta^2$ be the set of all households whose transactions are either envied by no one or are only envied by $\boldsymbol{\theta}$. Then, one of the following must be true: either i) no transaction $\boldsymbol{z}(\boldsymbol{\theta})$ for $\boldsymbol{\theta} \in \hat{H}_{\mathcal{E}}$ is distorted, or; ii) the allocation $(\boldsymbol{z}_{\mathcal{E}}(\boldsymbol{\theta}))_{\boldsymbol{\theta}}$ is constrained inefficient.*

That is, provided that there exists some transaction $\boldsymbol{z}'$ that is greater than the disagreement transactions, and such that the choice of this transaction induces the same household agreement choices, then the current allocation cannot be optimal.

5 Tax Systems

5.1 Simple Tax Systems

For a given $\boldsymbol{\psi}$, let

$$\chi^f_{\Psi}(z_m, \boldsymbol{\theta}) := \underset{z|\boldsymbol{\psi}(z,z_m)\leq 0}{\text{argmax}} \; \phi_f\left(z, z_m; \boldsymbol{\theta}\right) \tag{17}$$

define the reaction function for the wife with analogous definition, $\chi^m_{\Psi}(z_f, \boldsymbol{\theta})$, for the husband.

Assumption 5.1. *There exists a unique Nash equilibrium,*

$$(z^*_f, z^*_m) = \left(\chi^f_{\Psi}(z^*_m, \boldsymbol{\theta}), \chi^m_{\Psi}(z^*_f, \boldsymbol{\theta})\right),$$

for the disagreement game played under the tax system, Ψ.

The expression in Assumption 5.1 relates to a Pure Strategy Nash Equilibrium. This is for notation convenience only. If only mixed strategy equilibrium exists, then threat points are the expected utilities attained by each spouse. All results remain valid.

An important limitation of simple tax systems is that they do not use information regarding whether households are in agreement or disagreement. The next example exploits this limitation to show the sub-optimality of simple tax systems.

Example 5.1. *Let $(\boldsymbol{z}_\Gamma(\boldsymbol{\theta}))_{\boldsymbol{\theta}}$ be the allocation implemented by a direct mechanism. At a minimum, a candidate tax schedule that implements $(\boldsymbol{z}_\Gamma(\boldsymbol{\theta}))_{\boldsymbol{\theta}}$ must make sure that all transactions chosen by the couples are available: $\boldsymbol{\psi}(\boldsymbol{z}) \leq 0$ for all z such that there is $\boldsymbol{\theta}$ such that $\boldsymbol{z} < \boldsymbol{z}_\Gamma(\boldsymbol{\theta})$. For each couple $\bar{\boldsymbol{z}}_\Gamma(\boldsymbol{\theta})$ denotes the equilibrium choices for the disagreement game under Γ.*

If, for any $\boldsymbol{\theta}$, there exists $\boldsymbol{z}$ such that $\boldsymbol{\psi}\left(z, \bar{z}_\Gamma^{c,m}(\boldsymbol{\theta})\right) \leq 0$ and $\phi_f((z_f, \bar{z}_\Gamma^{c,m}(\boldsymbol{\theta})), \boldsymbol{\theta}) > \phi_f((\bar{z}_\Gamma^{c,f}(\boldsymbol{\theta}), \bar{z}_\Gamma^{c,m}(\boldsymbol{\theta})), \boldsymbol{\theta})$, or $\phi_m((\bar{z}_\Gamma^{c,f}(\boldsymbol{\theta}), z_m), \boldsymbol{\theta}) > \phi_m((\bar{z}_\Gamma^{c,f}(\boldsymbol{\theta}), \bar{z}_\Gamma^{c,m}(\boldsymbol{\theta})), \boldsymbol{\theta})$, then the allocation $\boldsymbol{z}_\Gamma(\boldsymbol{\theta})$ is not an equilibrium allocation for the tax system Ψ. Note that this may still be compatible with $\phi_f(\boldsymbol{\gamma}^c(\boldsymbol{\theta}_f^*, d, \boldsymbol{\theta}_m^*, d), \boldsymbol{\theta}) > \phi_f(\boldsymbol{\gamma}^c(\hat{\boldsymbol{\theta}}, \iota, \boldsymbol{\theta}_m^*, d), \boldsymbol{\theta}) \forall \hat{\boldsymbol{\theta}}, \iota \in \{a, d\}$ and $\phi_m(\boldsymbol{\gamma}^c(\boldsymbol{\theta}_f^*, d, \boldsymbol{\theta}_m^*, d), \boldsymbol{\theta}) > \phi_m(\boldsymbol{\gamma}^c(\boldsymbol{\theta}_f^*, d, \hat{\boldsymbol{\theta}}, d), \boldsymbol{\theta}) \forall \hat{\boldsymbol{\theta}}, , \iota \in \{a, d\}$.

5.2 Sophisticated Tax Systems

Although Simple Tax Systems disregard whether couples are in agreement or disagreement, a feature commonly found in actual tax systems can be used to address this limitation.

Define a Sophisticated Tax System as follows. For single agents tax schedules are as before, $\{z \in Z; \psi_i(z) \leq 0\}$. For couples two schedules are offered: *i)* a schedule which induces a budget set represented by the restriction $\boldsymbol{\psi}(\boldsymbol{z}) \leq 0$, and; *ii)* a fall-back schedule captured by the restriction $\bar{\boldsymbol{\psi}}(\boldsymbol{z}) \leq 0$, which must apply to both spouses if any of the spouses opts to use it. That is, $\boldsymbol{\psi}$ is used if both spouses agree, whereas if a spouse decides to use $\bar{\boldsymbol{\psi}}$ the other spouse must abide by his or her decision. For any $\boldsymbol{z} \in Z^2$ such that $\bar{\boldsymbol{\psi}}(\boldsymbol{z}) \leq 0$ we have $\boldsymbol{\psi}(\boldsymbol{z}) \leq 0$. We let $(\boldsymbol{\psi}, \bar{\boldsymbol{\psi}})$ denote such tax system.

Under the U.S.'s tax system spouses may opt to file jointly or individually. Moreover, as a first approximation, the budget set induced by individual filing under the U.S. tax code is a subset of that induced by joint filling, hence, never chosen by spouses in agreement. These filing options define exactly the type of tax systems we have in mind. [28]

Because $\boldsymbol{\psi}(\boldsymbol{z}) \leq 0$ for any $\boldsymbol{z} \in Z^2$ such that $\bar{\boldsymbol{\psi}}(\boldsymbol{z}) \leq 0$, choices conditional on threat points are unchanged, by the introduction of $\bar{\boldsymbol{\psi}}$. Yet, the strategy space defined

[28] da Costa and Diniz (2016) consider a tax system for which spouses are required to file jointly. If they are unable to reach an agreement, they choose their transactions non-cooperatively on the budget set generated by the joint tax schedule. da Costa and Diniz (2016) then explore the consequences of offering the possibility of individual filing, under the assumption that the household's budget set under individual filing is a subset of the budget set under joint filing.

by each institutional setting matters for determining which transactions arise in the equilibrium of the disagreement games. Under a Sophisticated Tax System, $(\boldsymbol{\psi}, \bar{\boldsymbol{\psi}})$, $\chi_f^{\bar{\psi}}(z_m, \boldsymbol{\theta}) := \operatorname{argmax}_{z|\bar{\boldsymbol{\psi}}(z,z_m)\leq 0}\, \phi_f\,(z, z_m; \boldsymbol{\theta})$ substitutes for (17) as the reaction function for the wife. An analogous expression defines the husband's reaction function, $\chi_m^{\bar{\psi}}(z_f, \boldsymbol{\theta})$. The sets of admissible deviations, implied by $\chi_f^{\bar{\psi}}(z_m, \boldsymbol{\theta})$ and $\chi_m^{\bar{\psi}}(z_f, \boldsymbol{\theta})$ are different from that of a simple tax system only in that $\bar{\boldsymbol{\psi}}$ substitutes for $\boldsymbol{\psi}$.

Now, two different instruments are given for the planner to pursue two different objectives: $\boldsymbol{\psi}$, to try to induce choices conditional on threat points, and $\bar{\boldsymbol{\psi}}$ to try to induce 'desirable' threat points.

The first question we want to answer is whether the set of allocations that is implementable by a Sophisticated Tax System is the set of implementable allocations (those implementable by a direct mechanism). Unfortunately, this is not the case. Although a Sophisticated Tax System will, in general, expand the set of allocations that may be implemented by a simple tax system, they cannot mimic all that may be accomplished by an arbitrary direct mechanism. Outcome functions that entail strong punishments for conflicting announcements are able to induce any transactions that respect incentive compatibility as equilibria for the disagreement games. Transactions under $\bar{\boldsymbol{\psi}}$ are equivalent to a single announcement (z^f, z^m, d) which does not allow lies to be detected![29]

Despite not being able to implement all incentive feasible allocations, Sophisticated Tax System do expand the set of allocations that may be implemented by an simple tax system. Again, interpreting transactions as announcements is useful. Under $\boldsymbol{\psi}$ the announcement is (z_f, z_m, a) and under $\bar{\boldsymbol{\psi}}$ the announcement is $(\tilde{z}_f, \tilde{z}_m, d)$, which is better than a single announcement, (z_f, z_m).

To formalize what can be implemented by a tax system, let $\tilde{\Psi}$ be a function of transactions $\boldsymbol{z}$ and extra signals $\varsigma \in \mathcal{S}$. In the simple tax system $\mathcal{S}$ is a singleton, whereas in the sophisticated tax system $\varsigma = \iota \in \{a, d\}$ indicates if the couple is in agreement or disagreement, $\psi(\cdot) = \tilde{\psi}(\cdot, a)$, and $\bar{\psi}(\cdot) = \tilde{\psi}(\cdot, d)$.

Proposition 5. *Let $\boldsymbol{z}_{\mathcal{M}}(\cdot)$ be an implementable allocation. $\boldsymbol{z}_{\mathcal{M}}(\cdot)$ can be implemented by $\tilde{\psi}$ if and only if*

i) There exists threat points $\bar{\boldsymbol{u}}_{\tilde{\psi}}(\cdot)$ such that $\tilde{\psi}$ implements $\boldsymbol{z}_{\mathcal{M}}(\cdot)$ conditional on them.[30]

ii) There exists disagreement transactions $\bar{\boldsymbol{z}}_{\tilde{\psi}}(\cdot) = (z^f_{\tilde{\psi}}(\cdot), z^m_{\tilde{\psi}}(\cdot))$ and announcements $\boldsymbol{\varsigma}(\cdot) = \left(\varsigma^f(\cdot), \varsigma^m(\cdot)\right)$ such that for every $\boldsymbol{\theta}$, $\bar{\boldsymbol{u}}_{\tilde{\psi}}(\boldsymbol{\theta}) = \Phi(\bar{\boldsymbol{z}}_{\tilde{\psi}}(\boldsymbol{\theta}), \boldsymbol{\theta})$, and, for $i, -i \in$

[29]Naturally, if we consider a tax system in which a tax schedule is offered for each different announcement, that including a 'type announcement' we are back to a direct mechanism. In this very broad sense, very complex tax systems, do implement any incentive feasible allocation.

[30]Notice that there is at least one such threat point by assumption, the one implemented by $\mathcal{M}$, $\bar{\boldsymbol{u}}_{\mathcal{M}}(\cdot)$, but it may not be the only one.

$\{f, m\}$, $(\bar{z}^i_{\tilde{\psi}}(\boldsymbol{\theta}), \varsigma^i(\boldsymbol{\theta})) \in \chi_i^{\tilde{\psi}}(\bar{z}^{-i}_{\tilde{\psi}}(\boldsymbol{\theta}), \varsigma^{-i}(\boldsymbol{\theta}), \boldsymbol{\theta})$.

Proof. See appendix A.1 $\square$

The last condition can be translated in words as follows. By knowing only one spouses' transactions and announcements the planner must be able to induce the other spouse to choose the desired transactions. Suppose we want to know what can be implemented by a Simple Tax System, in this case ς is just a constant. If we want to implement a particular disagreement transaction for couple $\boldsymbol{\theta}$, $\bar{\boldsymbol{z}}(\boldsymbol{\theta})$, the first idea would be to punish harshly all deviations. If $z_f = \bar{z}^f_{\mathcal{M}}(\boldsymbol{\theta})$ or $z_m = \bar{z}^m_{\mathcal{M}}(\boldsymbol{\theta})$, but not both, set $\Psi(z_f, z_m) = \infty$.

The problem is that the planner may want to implement a transaction (z_f, z_m) for another couple (in agreement or disagreement) such that $z_f = \bar{z}^f_{\mathcal{M}}(\boldsymbol{\theta})$ or $z_m = \bar{z}^m_{\mathcal{M}}(\boldsymbol{\theta})$, but not both, say $(\bar{z}^f_{\mathcal{M}}(\boldsymbol{\theta}), z_m)$. Now the planner cannot set $\Psi(\bar{z}^f_{\mathcal{M}}(\boldsymbol{\theta}), z_m) = \infty$, and the husband on couple $\boldsymbol{\theta}$ may prefer to lie and get transactions z_m instead of $\bar{z}^m_{\mathcal{M}}(\boldsymbol{\theta})$.

If we add an announcement of agreement or disagreement, as in the Sophisticated Tax System, then we restrict this interaction to couples in disagreement, which expands what can be implemented.

Naturally, if we make the set of announcements, $\mathcal{S}$, rich enough so that we can recover the type $\boldsymbol{\theta}$ from each spouse announcement individually, then we are back to what can be implemented by any mechanism. With this intuition we can provide sufficient conditions for an allocation to be implemented under the Sophisticated Tax System.

Proposition 6. *$\boldsymbol{z}_{\mathcal{M}}(\cdot)$ can be implemented by $(\psi, \bar{\psi})$ if the disagreement transactions induced by $\mathcal{M}$, $\bar{z}^f_{\mathcal{M}}(\cdot)$ and $\bar{z}^m_{\mathcal{M}}(\cdot)$ are both one-to-one, that is, if by looking at each spouse's transaction individually we can recover their type.*

Proof. See appendix A.1 $\square$

5.3 Optimal Taxation: The Variational Approach

Variational methods bring the promise of finally allowing for a complete characterization of optimal tax schedules. Toward this goal some issues must be addressed. First, as we have shown, by relying on tax schedules only we cannot be assured that we have implemented optimal allocations. In other words, there are missing policy instruments capable of improving upon these allocations through the manipulation of threat points even if we do not restrict the shape of tax schedules. The attempt to mimic these instruments may introduce some features or distortions in optimal tax schedules that would otherwise be absent. Second, traditional formulae derived under a variational approach rely on elasticities derived under the assumption that preferences are well defined and invariable with respect to policy. Yet, as we have seen, both the invariability of household preferences to changes in schedules as well as the very definition of those preferences may depend on the presence of policy tools capable of handling threat points. Both these issues must be born in mind if one is to read the findings in the literature.

We start with some definitions which allows us to approximate our notation to that used in Golosov et al. (2014). For couples in agreement only $c = c_f + c_m$ matters, not the specific way in which c is split between c_f and c_m. This allows us first to focus on a tax schedule of the form $\mathcal{T}(n_f, n_m) = T^f(n_f, n_m) + T^m(n_f, n_m)$.

Let $n^f_{\boldsymbol{\theta},\bar{\boldsymbol{u}}}(\mathcal{T})$ denote the absolute value of the second entry in vector $\hat{\boldsymbol{z}}_\Psi(\boldsymbol{\theta}|\bar{\boldsymbol{u}})$ and $n^m_{\boldsymbol{\theta},\bar{\boldsymbol{u}}}(\mathcal{T})$, the absolute value of the last entry.[31] That is, $\boldsymbol{n}_{\boldsymbol{\theta},\bar{\boldsymbol{u}}}(\mathcal{T}) = (n^f_{\boldsymbol{\theta},\bar{\boldsymbol{u}}}(\mathcal{T}), n^m_{\boldsymbol{\theta},\bar{\boldsymbol{u}}}(\mathcal{T}))$ denotes the vector of taxable income functions for a couple $\boldsymbol{\theta}$ when the threat points are $\bar{\boldsymbol{u}} = (\bar{u}^f, \bar{u}^m)$. It is a bidimensional version of a usual taxable income functions as defined in Saez (2001), for example; the relevant object to define a tax schedule.[32]

With these definitions we apply Golosov et al.'s (2014) approach to derive the effects of a perturbation in $\mathcal{T}$, starting with the behavioral effect. For a marginal retention rate $1 - \tau_i$ and non-labor income R let $n_i = \varsigma^i(1 - \tau_f, 1 - \tau_m, R)$ define the spouse i's taxable income supply function. This is a standard Marshallian taxable income supply function. Using $\tau_i(n_f, n_m) = \partial\mathcal{T}(n_f, n_m)/\partial n_i$ and $R(n_f, n_m) = \mathcal{T}(n_f, n_m) - \tau_f(n_f, n_m)n_f - \tau_m(n_f, n_m)n_m$, we linearize the tax schedule at the couple's optimal choice $\boldsymbol{n}_{\boldsymbol{\theta},\bar{\boldsymbol{u}}}(\mathcal{T})$.

While still holding $\bar{\boldsymbol{u}}$ fixed, we can then derive the behavioral effect of a perturbation of $\mathcal{T}$ in he direction $\mathcal{H}$. That is, we replace the tax schedule $\mathcal{T}$ by $\tilde{\mathcal{T}} = \mathcal{T} + a\mathcal{H}$ for an arbitrarily small a. Let $h_i(n_f, n_m) = \partial\mathcal{H}(n_f, n_m)/\partial n_i$ and $P(n_f, n_m) = \mathcal{H}(n_f, n_m) - h_f(n_f, n_m)n_f - h_m(n_f, n_m)n_m$. Using $\varsigma^i_j = \partial\varsigma^i/\partial(1 - \tau_j)$, $i, j = f, m.$, $\varsigma^i_3 = \partial\varsigma^i/\partial R$ and $\tau_{ij} = \partial\tau_i/\partial n_j$, we can show that

$$\begin{pmatrix} dn_f \\ dn_m \end{pmatrix} = -\left\{ I + \begin{bmatrix} \hat{\varsigma}^f_f & \hat{\varsigma}^f_m \\ \hat{\varsigma}^m_f & \hat{\varsigma}^m_m \end{bmatrix} \begin{bmatrix} \tau_{ff} & \tau_{fm} \\ \tau_{fm} & \tau_{mm} \end{bmatrix} \right\}^{-1} \begin{pmatrix} \varsigma^f_1 h_f + \varsigma^f_2 h_m - \varsigma^f_3 P \\ \varsigma^m_2 h_m + \varsigma^m_1 h_f - \varsigma^m_3 P \end{pmatrix} da, \tag{18}$$

where we have used $\hat{\varsigma}^j_i := \varsigma^j_i + \varsigma^j_3 n_i$ to define the compensated effects, and I to represent the identity matrix. This is a particular case of the general formulation in Golosov et al. (2014) derived under the assumption that threat points can be held fixed as we perturb $\mathcal{T}$. But can we hold $\bar{\boldsymbol{u}}$ fixed? If not, what are the consequences?

Using a simple tax system, the answer is no in general. Changes in the tax schedule

[31] Recall that $\boldsymbol{z}_\Psi(\boldsymbol{\theta}) := \hat{\boldsymbol{z}}_\Psi(\boldsymbol{\theta}|\bar{\boldsymbol{u}}_\Psi(\boldsymbol{\theta}))$, where $\hat{\boldsymbol{z}}_\Psi(\boldsymbol{\theta}|\bar{\boldsymbol{u}})$ are the transactions chosen by a type $\boldsymbol{\theta}$, if the threat points faced by $\boldsymbol{\theta}$ are $\bar{\boldsymbol{u}}$.

[32] Our notation is chosen to resemble the one used in Golosov et al. (2014).

are likely to change the outcome of the disagreement games played by spouses. Hence, the constancy of $\bar{\boldsymbol{u}}$ to a perturbation of $\mathcal{T}$ for all couples is a very unlikely outcome. Without this constancy, labor supply will not have the usual properties that characterize rational consumption behavior. The compensated matrix in (18) is, for example, replaced by a pseudo-Slutsky matrix in the sense of Browning and Chiappori (1998).

This is not all. When we perturb the tax system at the equilibrium choice of a couple $\boldsymbol{\theta}$, we may affect the threat points of a couple $\hat{\boldsymbol{\theta}}$, which may be very different from couple $\boldsymbol{\theta}$. These 'global' impacts of a perturbation may undermine the whole procedure if we do not have full knowledge of disagreement choices.

If, however, a Sophisticated Tax System – see Section 5.2 – or another mechanism for which the tax schedule $\mathcal{T}$ does not define the outcome functions (and/or strategy spaces) for the disagreement game is used, then holding $\bar{\boldsymbol{u}}$ fixed is a more reasonable assumption. But this also raises new possibilities.

The elasticity of taxable income (ETI) depends on the specific values for $\bar{\boldsymbol{u}}$ under which the choices were made. They can be manipulated if one can change the threat points.[33] An example of such manipulations of elasticities is the one described in Section 2, where by moving the original threat point to a new point $u_i^\alpha = \alpha u_i^* + (1-\alpha)\bar{u}_i$, where u_i^* is the solution for the Nash bargain for threat points $\bar{u}_i$, then the agreement utility doesn't change, but the 'Hicksian' elasticities of taxable income are reduced.[34]

As for the mechanical effect of a reform, one would be tempted to apply an envelope argument to (6), yet this assumes that household welfare functions are meaningful from a normative perspective and that there is no dissonance. Recent works by Immervoll et al. (2011); Cremer et al. (2016) assess the impact of dissonance for optimal taxation. For example, Cremer et al. (2016) consider a linear household welfare function, in the sense of Samuelson (1956), $\mathcal{U}(c, n_f, n_m) := \max_{\mathfrak{c}_f}\{\delta u_f(\mathfrak{c}_f, n_f) + u_m(c - \mathfrak{c}_f, n_m)\}$, and find that a corrective term which increases the labor wedge for the wife should be added to the Mirrlees' optimal tax formula. In fact, under a unitary approach there is a well defined household welfare function which is used to assess household welfare. However, as precisely put by Chiappori and Meghir (2014), "...what we are (or should be) ultimately interested in is individual welfare. Household welfare, if this notion has any sense, cannot be defined without considering the welfare of each member." Hence, even if well conditional household preferences are well defined, the difference between household preferences and social preferences lead to first order welfare effects due to behavioral

[33] Optimally choosing ETI's is also what Kopczuk and Slemrod (2002) explore, in a different context.
[34] In fact one can show that if we use $\hat{\mathfrak{s}}_\alpha$ to denote the Slutsky matrix in (18) for a given α, then, for any vector $a \in R^2$, $a \cdot \hat{\mathfrak{s}}_\alpha a \geq a \cdot \hat{\mathfrak{s}}_{\alpha'} a$ for $\alpha \geq \alpha'$.

responses.[35] It is this type of dissonance that explain Cremer et al.'s 2016 finding.[36]

Indeed, the welfare effect of perturbing the tax system in the direction $\mathscr{H}$ is given by

$$(1-\omega_f)d\mathscr{H} + \omega_f(\delta - 1)\left[d\mathfrak{c}_f - (1-\tau_m)dn_m - (1-\tau_f)dn_f\right], \tag{19}$$

where ω_f is the marginal social weight of this particular woman type. The term in brackets captures the effect of dissonance: the misalignment between the Utilitarian planner's objective and the household power structure. Behavioral responses which have no first order effects from the household perspective will usually do so from a social perspective. In the case of Cremer et al.'s reform, a correction term associated with an increase in the marginal tax rate on the wife is positive for $\delta > 1$ whenever the term in curly brackets is positive. A small increase in the wife's marginal tax accompanied by a small decrease in the husband's increases his labor supply and decreases hers, thus explaining their results.[37]

Although δ is exogenous in Cremer et al. (2016), we have shown how auxiliary policy instruments can affect δ and bring it closer to the planner's preferred value, 1. In this case, the importance of the corrective term in 19 may be reduced and, in some cases, be entirely eliminated. Note also that even if we hold $\bar{\boldsymbol{u}}$ fixed, δ will vary due to the curvature of W as defined in (6).

6 Extensions and Robustness

6.1 Generalizing Preferences and Technology

We have specified preferences and technology which are exactly that of Mirrlees (1971). This is for ease of communication only. Our framework can be extended to allow for preferences defined over more goods, multidimensional heterogeneity, etc. More importantly, for married agents we may allow for externalities, public goods and other forms of interdependence in utility.

Let $\mathfrak{x} = (\mathfrak{c}, \mathfrak{l})' \in \mathbb{R}^2_+$ denote the bundle consumed by an agent. All our results apply more generally to $\mathfrak{x} \in \mathbb{R}^n_+$, and $\nu_i(\mathfrak{x}_i, \mathfrak{x}_{-i}, \theta_i)$, $i, -i = f, m$. Again, consumption goods may be separated into assignable, $\mathfrak{x}_A$, and non-assignable, $\mathfrak{x}_{NA}$, with $\mathfrak{x} = (\mathfrak{x}_A, \mathfrak{x}_{NA})\ \mathbb{R}^{n_A+n_{NA}}_+$.

[35] For instance, consider the extreme case where all household choices are made dictatorially by one of the spouses.

[36] Cremer et al. (2016) use a Mirrlees (1971) approach. The same is true of Immervoll et al. (2011) who also explore the consequences of dissonance for the optimal taxation of couples.

[37] At least for separable preferences, and fixed Pareto weight, δ. Moreover, the perturbation may be complemented by an adjustment in the virtual income to handle the non-linearity of $\mathscr{T}$.

Similarly, Θ, could be any subset of $\mathbb{R}^m$. Finally, more sophisticated technologies mapping Z^2 with $Z \subset R^n$ into 2^X, $X \subset R^n_+$ may be considered.

Some of these generalizations are of great practical relevance. For instance, central to any discussion involving household choices are the investments made on children, usually thought of as a public good from spouses' perspective. We can easily accommodate the presence of public goods in our description. Let g be the amount of public good produced, and $\mathfrak{x}_i = (\mathfrak{c}_i, \mathfrak{g}_i, \mathfrak{l}_i)$. We can simply add to the description of $F^a(\boldsymbol{z}, \boldsymbol{\theta})$, the pair of constraints $\mathfrak{g}_f \leq g$, $\mathfrak{g}_m \leq g$, and modify the time constraints to $\mathfrak{l}_f \leq 1 - n_f - t_f$, and $\mathfrak{l}_m \leq 1 - n_m - t_m$, $g = f(t_f, t_m)$. Note that leisure is no longer assignable in this case.

The main question that one must face when these generalizations are pursued is which sensible properties one would want to impose on Φ. In many cases, it will be more practical to write the disagreement game explicitly.[38]

6.2 Divorce as Threat Point

Following, Bergstrom's (1996) insights, we have taken the flow utility attained by spouses when they fail to reach an agreement as the relevant threat points. Here, we consider allow external threat points to play a role. That is, there is still a non-cooperative game played by spouses, but we with the proviso that these flow utilities that agents attain cannot be lower than what spouses can obtain if they divorce. Whether threat points are internal or external is therefore an outcome, not an assumption.

That is we assume that at any moment and for any institutional setting either spouse of any couple can unilaterally end the marriage, in which case, he or she: i) incurs a utility loss of $\kappa > 0$, and; ii) obtains a flow utility from consumption identical to that of a single agent of the same type that never married attains.

If $\mathscr{a}$, a woman of type θ_f, is married to $\mathscr{b}$, a type θ_m man, then it must be the case that $u^{c,f}_{\mathscr{g}}(\theta_f, \theta_m) + \xi_{\mathscr{a}}(\theta_m) \geq u^f_{\mathscr{g}}(\theta_f) + \xi_{\mathscr{a}}(\emptyset) - \kappa$. The utility that she attains in marriage is strictly greater than the utility, $u^f_{\mathscr{g}}(\theta_f) + \xi_{\mathscr{a}}(\emptyset) - \kappa$ that she would obtain were she to divorce. The constraint introduced by allowing agents to divorce only establishes an institutionally dependent lower bound for the threat points of both spouses in any household.

For each $\boldsymbol{\theta}$, consider all couples, $(\mathscr{a}, \mathscr{b})$ which productivities are $\boldsymbol{\theta}$ that actually marry. Then, let $\Delta_f(\boldsymbol{\theta}) = \sup_{\mathscr{a}} \xi^f_{\mathscr{a}}(\emptyset) - \xi^f_{\mathscr{a}}(\theta_m)$, and $\Delta_m(\boldsymbol{\theta}) = \sup_{\mathscr{b}} \xi^m_{\mathscr{b}}(\emptyset) - \xi^m_{\mathscr{b}}(\theta_f)$, then

[38] The presence of public goods has, for example, been the main motivation for the use of non-cooperative games in household economics, e.g., Lundberg and Pollak (1993); Browing et al. (2010). A Nash equilibrium for these private contribution games has all the properties we have imposed on Φ.

vi) For any $\boldsymbol{\theta} = (\theta_f, \theta_m)$,

$$\bar{u}_{\Gamma}^{c,i}(\boldsymbol{\theta}) \geq u_{\Gamma}(\gamma_i^s(\hat{\theta}_i), \theta_i) + \Delta_i(\boldsymbol{\theta}) - \kappa \quad \forall \hat{\theta}_i, i = f, m. \tag{20}$$

must be added to the set of IC constraints.

Allowing agents to divorce at any stage implies that the planner is subject to another constraint in its attempt to manipulate threat points. Interestingly, when we compare Simple Tax Systems with general mechanism a potential new role for optimal policies arises. Assume that at the optimum schedule constraint vi) binds. This means that the tax schedules offered to singles is distorted due to the impact it has on the threat points faced by married agents. A direct mechanism may in this case be used to relax the constraints by, for example, moving the threat point along the line connecting the original threat point and the agreement utility pair, $\Lambda_\psi(\boldsymbol{\theta}, \bar{\boldsymbol{u}})$.

6.3 Randomization in Agreement

We have only considered deterministic choices by spouses in agreement. Yet randomization is sometimes allowed for in order to convexify utility possibility sets to apply Nash's (1950) axioms in their original form.[39] In an earlier version of the paper – da Costa and de Lima (2016) — we discuss this possibility, emphasizing the commitment assumption regarding each possible view of how randomization occurs in practice. For now, just note that if the most general form of randomization is possible, then the planner is restricted to using policies that induce convex utility possibility sets.

7 Conclusion

We have asked whether tax implementation is without loss when couples do not behave as a single individual, but instead make decisions through a bargain which outcome we assume to satisfy Nash's (1950) axioms.

We find that tax schedules are poor instruments to play the double role of inducing allocations conditional on households' objectives and defining these objectives through its impact on threat points. Tax induced allocations can be improved upon by more general mechanism through the inducement of convenient threat points that allow for: the alignment of household and the planner's objectives; the relaxation of incentive constraints.

[39]See Chiappori (2016) for an application in the context of household economics.

We have assumed that marriage market are characterized Bargain in Marriage – e.g., Lundberg and Pollak (2009); Pollak (2016) — with internal threat points as suggested by Bergstrom (1996). These assumptions allowed us to get a very sharp characterization of implementable allocations. Eventually, binding divorce threat points can be incorporated, as we have shown in Section 6.2. Departing from Bargain in Marriage by allowing some for of commitment or pre-marital transfers is bound to break the two stage representation that we obtain here. We believe that, provided that the distribution of power between spouses still depends on things that happen after marriage, the taxation principle will still fail and some of the forces identified here will have to be taken into account in the design of optimal policies.

References

Patricia Apps and Ray Rees. Taxation and the household. *Journal of Public Economics*, 35:355–369, 1988. 4, 5

Eduardo M. Azevedo and John W. Hatfield. Existence of equilibrium in large matching markets. mimeo. Wharton, PA, August 2015. 21

Gary S. Becker. Altruism in the family and selfishness in the market place. *Economica*, 48:1 – 15, 1981. 2

Gary S. Becker. *A Treatsie on the Family*. Harvard University Press, Cambridge, MA, 2nd edition, 1991. 4

Theodore. Bergstrom. A fresh look at the rotten kid theorem – and other household mysteries. *Journal of Political Economy*, 97(5):1138 – 1159, 1989. 2

Theodore. Bergstrom. Economics in a family way. *Journal of Economic Literature*, 34 (4):1138–1159, 1996. 9, 33, 35

Kenneth George Binmore. Bargaining and coalitions. In Alvin Roth, editor, *Game-theoretic models of bargaining*, pages 259 –304. Cambridge University Press, Cambridge, 1985. 9

Kenneth George Binmore, Avner Shaked, and John Sutton. An outside option experiment. *Quarterly Journal of Economics*, 104(4):753 – 770, 1989. 9

Aaron L. Bodoh-Creed. Risk and conflation in matching markets. Mimeo Haas-Berkeley, 2016. 21

Michael J. Boskin and Eytan Sheshinski. Optimal tax treatment of the family: Married couples. *Journal of Public Economics*, 20(281–297), 1983. 4

Martin Browing, Pierre-André Chiappori, and Valérie Lechene. Distributional effects in household models: separate spheres and income pooling. *The Economic Journal*, 120: 786 – 799, 2010. 33

Martin Browing, Pierre-André Chiappori, and Arthur Lewbel. Estimating consumption economies of scale, adult equivalence scales, and household bargaining power. *Review of Economic Studies*, 80:1267 – 1303, 2013. 2, 6

Martin Browning and Pierre-André Chiappori. Efficient intra-household allocation: a characterization and tests. *Econometrica*, 66(6):1241–1278, 1998. 2, 5, 31

Z. Chen and F. Wooley. A cournot-nash modelo of family decision making. *Economic Journal*, 111:722–748, 2001. 5

Raj Chetty. Behavioral economics and public policy: A pragmatic perspective. *American Economic Review Papers and Proceedings*, 105(5):1 – 33, 2015. 6

Pierre-André Chiappori. Randomized household labor supply. mimeo, 2016. 34

Pierre-André Chiappori and Maurizio Mazzocco. Static and intertemporal household decisions. *Journal of Economic Literature*, forthcoming, 2015. 5

Pierre-André Chiappori and Costas Meghir. Intra-household welfare. Technical Report 20189, NBER Working Paper, June 2014. 2, 31

Pierre-André Chiappori, Olivier Donni, and Ivana Komunjer. Learning from a piece of pie. *Review of Economic Studies*, 79(1):162 – 195, 2012. 5

Eugene Choo and Aloysius Siow. Who marries whom and why. *Journal of Political Economy*, 114(1):175 – 201, 2006.

Helmuth Cremer, Jean-Marie Lozachmeur, Dario Maldonado, and Kerstin Roeder. Household bargaining and the design of couples' income taxation. *European Economic Review*, 89:454–470, 2016. 6, 31, 32

Jacques Crémer and Richard McLean. Full extraction of the surplus in bayesian and dominant strategy auctions. *Econometrica*, 56(6):1247–1257, 1988. 5

Carlos E. da Costa and Lucas A. de Lima. Taxation of couples: a mirrleesian approach for non-unitary households. Ensaios Econômicos 781, EPGE-FGV, July 2016. 34

Carlos E. da Costa and Érica Diniz. Tax filing choices for the household. *Brazilian Review of Econometrics*, 36:63 – 96, 2016. 11, 13, 27

Bev Dahlby. Progressive taxation and the social marginal cost of public funds. *Journa of Public Economics*, 67:105 — 122, 1998. 4

Partha Dasgupta, Peter J. Hammond, and Eric Maskin. The implementation of social choice rules: some results on incentive compatibility. *Review of Economic Studies*, 46: 185–216, 1979. 3, 5

David Gale and Lloyd S. Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69(1):9 – 15, 1962. 4

George-Levi Gayle and Andrew Shephard. Optimal taxation, marriage, home production, and family labour supply. Mimeo Washington University in St. Louis, July 2016.

Mikhail Golosov, Aleh Tsyvinski, and Nicolas Werquin. A variational approach to the analysis of tax systems. Mimeo. Princeton University, December 2014. 5, 30

Roger Guesnerie. *A Contribution to the Pure Theory of Taxation*. Cambridge University Press, 1998. 3, 4

Peter J. Hammond. Straightforward individual incentive compatibility in large economies. *Review of Economic Studies*, 46(0):263–282, 1979. 1, 3, 4

Peter J Hammond. Markets as constraints: Multilateral incentive compatibility in continuum economies. *Review of Economic Studies*, 54(3):399–412, 1987. 2, 3, 4

Milton Harris and Robert M. Townsend. Resource allocation under asymmetric information. *Econometrica*, 49(1):33 – 64, 1981. 3, 5

Herwig Immervoll, Henrik Jacobsen Kleven, Claus Thustrup Kreiner, and Nicolaj Verdelin. Optimal tax and transfer programs for couples with extensive labor supply responses. *Journal of Public Economics*, 95(11-12):1485–1500, 2011. 6, 31, 32

Ravi Kanbur, Jukka Pirttilä, and Matti Tuomala. Non-welfarist optimal taxation and behavioral public economics. *Journal of Economic Surveys*, 20(5):849 – 868, 2006. 6

Wojciech Kopczuk and Joel Slemrod. The optimal elasticit of taxable income. *Journal of Public Economic*, 84(1):91 – 112, 2002. 5, 31

Shelly Lundberg and Robert A. Pollak. Separate spheres bargaining and the marriage market. *Journal of Political Economy*, 101(6):988–1010, 1993. 5, 8, 33

Shelly Lundberg and Robert A. Pollak. Marriage market equilibrium and bargaining. Presidential Address at the Society of Labor Economics Meeting - Boston, 2009. 8, 35

Shelly Lundberg, Robert A. Pollak, and Terence J. Wales. Do husbands and wives pool their resources? evidence from the uk child benefit. *Journal of Human Resources*, 32 (3):463 – 480, 1997. 2, 5, 9

Marilyn Manser and Murray Brown. Marriage and household decision making: a bargaining analysis. *International Economic Review*, 21(1):31–44, 1980. 5, 8

Marjorie B. McElroy and Mary Jane Horney. Nash-bargained decisions: Towards a generalization of the theory of demand. *International Economic Review*, 22(2):333–349, 1981. 5, 8

James A. Mirrlees. An exploration in the theory of optimal income taxation. *Review of Economic Studies*, 38:175–208, 1971. 2, 4, 6, 31, 32

Roger B. Myerson. Incentive-compatibility and the bargaining problem. *Econometrica*, 47(61 - 73), 1979. 3, 5

Roger B. Myerson. Multistage games with communication. *Econometrica*, 54(2):323–358, March 1986. 7

Roger B. Myerson. *Game Theory: Analysis of Conflict*. Harvard University Press, 1997.

18

John F. Nash. The bargaining problem. *Econometrica*, 18(2):pp. 155–162, 1950. 1, 2, 5, 8, 34

Thomas Piketty. La redistribuition fiscale face au chômage. *Revue Française d' Èconomie*, 12:157 – 201, 1997. 1, 4

Robert A. Pollak. Marriage market equilibrium. Working Paper 22309, NBER, June 2016. 4, 8, 21, 35

Frank Ramsey. A contribution to the theory of taxation. *Economic Journal*, 37:47–61, 1927. 1

Alvin E. Roth and Marilda A. Sotomayor. *Two-Sided Matching: A study in game-theoretic modeling and analysis*. Econometric Society Monographs. Cambridge University Press, Cambridge, 1990. 21

Emmanuel Saez. Using elasticities to derive optimal income tax rates. *Review of Economic Studies*, 68:205–229, 2001. 4, 30

Emmanuel Saez. The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. *Journal of Public Economics*, 83(2):217–230, 2002. 1

Paul Samuelson. Social indifference curves. *Quarterly Journal of Economics*, 70:1–22, 1956. 31

Eytan Sheshinski. The optimal linear income tax. *Review of Economic Studies*, 39:297 – 302, 1972. 1

Duncan Thomas. Intra-household resource allocation: An inferential approach. *The Journal of Human Resources*, 25(4):635 – 664, 1990. 2, 5

Eduardo Zambrano. 'vintage' nash bargaining without convexity. *Economics Letters*, 141:32 – 34, 2016. 8, 14

A Appendix

A.1 Proofs

Proof of Proposition 1. Consider a general mechanism as defined in (2). The original mechanism $\mathcal{M}$ is a set of message spaces for single females, Σ^f, single males, Σ^m, married females, $\Sigma^{c,f}$, and married males, $\Sigma^{c,m}$, and outcome functions for single females, $g^f : \Sigma^f \to \mathcal{Z}$, single males, $g^m : \Sigma^m \to \mathcal{Z}$, and couples $\boldsymbol{g}^c : \Sigma^{c,f} \times \Sigma^{c,m} \to \mathcal{Z}^2$. Let $\{\sigma_i^s(\theta_i)\}_{\theta_i, i=f,m}$, $\{\sigma^{c,i}(\boldsymbol{\theta}, a)\}_{\boldsymbol{\theta}, i=f,m}$, and $\{\sigma^{c,i}(\boldsymbol{\theta}, d)\}_{\boldsymbol{\theta}, i=f,m}$ denote, respectively, equilibrium announcements for singles, married agents in couples in agreement and married agents in couples in disagreement.

An equilibrium allocation for the mechanism is

$$\left\{(z^f_{\mathcal{M}}(\theta_f))_{\theta_f}, (z^m_{\mathcal{M}}(\theta_m))_{\theta_m} (\boldsymbol{z}_{\mathcal{M}}(\boldsymbol{\theta}))_{\boldsymbol{\theta}}\right\},$$

where

$$z^f_{\mathcal{M}}(\theta_f) = g^f(\sigma^f(\theta_f)),\ z^m_{\mathcal{M}}(\theta_m) = g^m(\sigma^m(\theta_m)),$$

and

$$\boldsymbol{z}_{\mathcal{M}}(\boldsymbol{\theta}) = \boldsymbol{g}^c\left(\sigma^{c,f}(\boldsymbol{\theta}, a), \sigma^{c,m}(\boldsymbol{\theta}, a)\right).$$

We consider the case in which the game induces a Pure Strategy Nash equilibrium. The case in which mixed strategy equilibrium are allowed follows the same step but requires the introduction of additional notation.

Here,

$$\sigma^f(\theta_f) = \operatorname*{argmax}_{m\in\Sigma^f} u(g^f(m), \theta_f), \tag{21}$$

$$\sigma^m(\theta_m) = \operatorname*{argmax}_{m\in\Sigma^f} u(g^m(m), \theta_m), \tag{22}$$

and

$$\left(\sigma^{c,f}(\boldsymbol{\theta}, a), \sigma^{c,m}(\boldsymbol{\theta}, a)\right) = \operatorname*{argmax}_{m\in\Sigma^f, m'\in\Sigma^m} W(\boldsymbol{g}^c(m, m'), \boldsymbol{\theta}, \bar{u}^f_{\mathcal{M}}(\boldsymbol{\theta}), \bar{u}^m_{\mathcal{M}}(\boldsymbol{\theta})). \tag{23}$$

The threat points in (23) are

$$\left(\bar{u}^f_{\mathcal{M}}(\boldsymbol{\theta}), \bar{u}^m_{\mathcal{M}}(\boldsymbol{\theta})\right) = \Phi\left(\sigma^{c,f}(\boldsymbol{\theta}, d), \sigma^{c,m}(\boldsymbol{\theta}, d), \boldsymbol{\theta}\right) \tag{24}$$

for

$$\sigma^{c,f}(\boldsymbol{\theta}, d) = \operatorname*{argmax}_{m\in\Sigma^f} \phi^f\left(\boldsymbol{g}^c\left(m, \sigma^{c,m}(\boldsymbol{\theta}, d)\right), \boldsymbol{\theta}\right), \tag{25}$$

and

$$\sigma^{c,m}(\boldsymbol{\theta}, d) = \operatorname*{argmax}_{m'\in\Sigma^m} \phi^m\left(\boldsymbol{g}^c\left(\sigma^{c,f}(\boldsymbol{\theta}, d), m'\right), \boldsymbol{\theta}\right). \tag{26}$$

Define the outcome for the direct mechanism as follows. For singles,

$$\gamma^i(\theta_i) = g^i(\sigma^i(\theta_i)), i = f, m..$$

For couples,

$$\boldsymbol{\gamma}^c(\boldsymbol{\theta}, \iota, \boldsymbol{\theta}', \iota') = \boldsymbol{g}^c\left(\sigma^{c,f}(\boldsymbol{\theta}, \iota), \sigma^{c,m}(\boldsymbol{\theta}', \iota')\right)$$

for all $\boldsymbol{\theta}$, $\boldsymbol{\theta}'$, $\iota, \iota' \in \{a, d\}$.

For singles the proof is standard and we omit for brevity. For couples, assume that threat points, $\bar{u}^i_{\mathcal{M}}(\boldsymbol{\theta}) = \bar{u}^i_{\Gamma}(\boldsymbol{\theta})\ \forall\boldsymbol{\theta},\ i = f, m$, are given by (24). Then, by (23) and

$\boldsymbol{\gamma}^c(\boldsymbol{\theta}, \iota, \boldsymbol{\theta}', \iota') = \boldsymbol{g}^c\left(\sigma^{c,f}(\boldsymbol{\theta}, \iota), \sigma^{c,m}(\boldsymbol{\theta}', \iota')\right)$ for all $\boldsymbol{\theta}$, $\boldsymbol{\theta}'$, $\iota, \iota' \in \{a, d\}$ for any $(\hat{\boldsymbol{\theta}}, \hat{\iota}, \hat{\boldsymbol{\theta}}', \hat{\iota}')$ it must be the case that

$$W(\boldsymbol{\gamma}^c(\hat{\boldsymbol{\theta}}, \hat{\iota}, \hat{\boldsymbol{\theta}}', \hat{\iota}'), \boldsymbol{\theta}, \bar{u}^f_{\mathcal{M}}(\boldsymbol{\theta}), \bar{u}^m_{\mathcal{M}}(\boldsymbol{\theta})) \leq W(\boldsymbol{\gamma}^c(\boldsymbol{\theta}, a, \boldsymbol{\theta}, a), \boldsymbol{\theta}, \bar{u}^f_{\mathcal{M}}(\boldsymbol{\theta}), \bar{u}^m_{\mathcal{M}}(\boldsymbol{\theta})).$$

So, we only need to show that $\bar{u}^i_{\mathcal{M}}(\boldsymbol{\theta}) = \bar{u}^i_{\Gamma}(\boldsymbol{\theta})$ $\forall \boldsymbol{\theta}$, $i = f, m$. We do it by showing that truth telling is an equilibrium for the disagreement game.

Assume that spouse f of couple $\boldsymbol{\theta}$ announces $(\boldsymbol{\theta}, d)$, and let $(\boldsymbol{\theta}', i)$ be an announcement that her husband finds better than $(\boldsymbol{\theta}, d)$, i.e., $\phi^m\left(\boldsymbol{\gamma}^c(\boldsymbol{\theta}, d, \boldsymbol{\theta}', i), \boldsymbol{\theta}\right) > \phi^m\left(\boldsymbol{\gamma}^c(\boldsymbol{\theta}, d, \boldsymbol{\theta}, d), \boldsymbol{\theta}\right)$. Then,

$$\begin{aligned}\phi^m\left(\boldsymbol{g}^c\left(\sigma^{c,f}(\boldsymbol{\theta}, d), \sigma^{c,m}(\boldsymbol{\theta}', i)\right), \boldsymbol{\theta}\right) &= \phi^m\left(\boldsymbol{\gamma}^c(\boldsymbol{\theta}, d, \boldsymbol{\theta}', i), \boldsymbol{\theta}\right) \\ &> \phi^m\left(\boldsymbol{\gamma}^c(\boldsymbol{\theta}, d, \boldsymbol{\theta}, d), \boldsymbol{\theta}\right) = \phi^m\left(\boldsymbol{g}^c\left(\sigma^{c,f}(\boldsymbol{\theta}, d), \sigma^{c,m}(\boldsymbol{\theta}, d)\right), \boldsymbol{\theta}\right).\end{aligned}$$

This is a contradiction with (26). □

Proof of Proposition 2. A direct mechanism for which $\mathbf{0} = \boldsymbol{\varphi}(\boldsymbol{\theta}, \iota, \boldsymbol{\theta}', \iota')$ whenever $(\boldsymbol{\theta}, \iota) \neq (\boldsymbol{\theta}', \iota')$ is such that any announcement $(\boldsymbol{\theta}, \iota, \boldsymbol{\theta}, \iota)$ is an equilibrium of the disagreement game. In particular, for every $\boldsymbol{\theta}$, $(\boldsymbol{\theta}, d, \boldsymbol{\theta}, d)$ is an equilibrium.

All the planner must worry is not to violate (12). If no new transaction is introduced in $\mathcal{Z}^c_{\mathcal{E}}(\boldsymbol{\theta})$, then this is guaranteed. □

Proof of Proposition 3. If any transaction, $\boldsymbol{z}$, is it to exist, such that $\bar{\boldsymbol{u}} = \Phi(\boldsymbol{z}, \boldsymbol{\theta})$ is such that there is $\boldsymbol{\theta}'$ such that $\boldsymbol{z} > \boldsymbol{z}_{\mathcal{E}}(\boldsymbol{\theta}')$, then the couple strictly prefers this allocation since $\mathcal{U}_{\boldsymbol{z}}(\boldsymbol{\theta}') \supsetneq \mathcal{U}_{\boldsymbol{z}_{\mathcal{E}}(\boldsymbol{\theta}')}(\boldsymbol{\theta}')$ □

Proof of Proposition 4. Immediate from Lemma 1. □

A.2 Lemmata

Lemma 1. *Let*

$$W(\boldsymbol{z}_{\mathcal{E}}(\boldsymbol{\theta}), \boldsymbol{\theta}, \bar{u}^f_{\mathcal{E}}(\boldsymbol{\theta}), \bar{u}^m_{\mathcal{E}}(\boldsymbol{\theta})) = W(\boldsymbol{z}_{\mathcal{E}}(\boldsymbol{\theta}'), \boldsymbol{\theta}, \bar{u}^f_{\mathcal{E}}(\boldsymbol{\theta}), \bar{u}^m_{\mathcal{E}}(\boldsymbol{\theta}))$$

for some $\boldsymbol{\theta}' \neq \boldsymbol{\theta}$.

Assume that there is $\mathcal{E}'$ *such that*

$$(\bar{u}^f_{\mathcal{E}'}(\boldsymbol{\theta}), \bar{u}^m_{\mathcal{E}'}(\boldsymbol{\theta})) = \alpha(\bar{u}^f_{\mathcal{E}}(\boldsymbol{\theta}), \bar{u}^m_{\mathcal{E}}(\boldsymbol{\theta})) + (1-\alpha)(u^f_{\mathcal{E}}(\boldsymbol{\theta}), u^m_{\mathcal{E}}(\boldsymbol{\theta}))$$

for $\alpha \in (0, 1)$.

Then,

$$W(\boldsymbol{z}_{\mathcal{E}}(\boldsymbol{\theta}), \boldsymbol{\theta}, \bar{u}^f_{\mathcal{E}'}(\boldsymbol{\theta}), \bar{u}^m_{\mathcal{E}'}(\boldsymbol{\theta})) > W(\boldsymbol{z}_{\mathcal{E}}(\boldsymbol{\theta}'), \boldsymbol{\theta}, \bar{u}^f_{\mathcal{E}'}(\boldsymbol{\theta}), \bar{u}^m_{\mathcal{E}'}(\boldsymbol{\theta})).$$

Proof. Without loss, let the spouses utilities at the original threat point be $(0, 0)$. Let (u_f, u_m) be the corresponding agreement solution. We know, in this case, that $(u_f, u_m) \geq (u'_f, u'_m)$ for any feasible utility pair (u'_f, u'_m).

Suppose that there is $\alpha \in (0, \frac{1}{2})$ such that the planner can induce disagreement utilities $(\alpha u_f, \alpha u_m)$. In this case, we have

$$\begin{aligned} u'_f u'_m \leq (u'_f u'_m)^{\alpha}(u_f u_m)^{1-\alpha} = (u'_f u_m)^{\alpha}(u'_f u'_m)^{\alpha}(u_f u_m)^{1-2\alpha} \\ < \alpha u'_f u_m + \alpha u_f u'_m + (1-2\alpha) u_f u_m \end{aligned}$$

Adding $\alpha^2 u_f u_m$ in both sides and rearranging we get

$$u'_f u'_m - \alpha u'_f u_m + \alpha u_f u'_m + \alpha^2 u_f u_m < (1 - 2\alpha + \alpha^2) u_f u_m$$

Which we can rewrite as

$$(u'_f - \alpha u_f)(u'_m - \alpha u_m) < (u_f - \alpha u_f)(u_m - \alpha u_m)$$

If we can only find $\alpha > \frac{1}{2}$, then we just go to $\frac{1}{2}$, then to $\frac{3}{4}$, and so forth until we reach α and apply the same argument as above. □

Lemma 2. *Assume that utilities that represent agents preferences are of the form $u_i(\mathfrak{c}, \mathfrak{l}) = \mathfrak{c} + h(\mathfrak{l})$ for $h(\cdot)$ strictly increasing and concave. Then, optimal transactions are independent of threat points.*

Proof. Let $\mathfrak{h}^i = h(n^i/\theta^i)$, and $\hat{\mathfrak{h}}^i = h(\hat{n}^i/\theta^i)$, $i = f, m$. The household prefers transactions $(c^f, -n^f, c^m, -n^m)$ to $(\hat{c}^f, -\hat{n}^f, \hat{c}^m, -\hat{n}^m)$ if

$$\begin{aligned} \max_{\mathfrak{c}^f} \left[\mathfrak{c}^f - \mathfrak{h}^f - \bar{u}^f\right] \left[x - \mathfrak{c}^f - \mathfrak{h}^m - \bar{u}^m\right] \geq \\ \max_{\mathfrak{c}^f} \left[\mathfrak{c}^f - \hat{\mathfrak{h}}^f - \bar{u}^f\right] \left[\hat{x} - \mathfrak{c}^f - \hat{\mathfrak{h}}^m - \bar{u}^m\right], \quad (27) \end{aligned}$$

for $x = c^f + c^m$, $\hat{x} = \hat{c}^f + \hat{c}^m$.

At the optimum, for the maximization problems above we have

$$\left[x - \mathfrak{c}^f + \mathfrak{h}^m - \bar{u}^m\right] = \left[\mathfrak{c}^f + \mathfrak{h}^f - \bar{u}^f,\right]$$

hence,

$$\frac{x + (\upsilon^f + \bar{u}^f) - (\upsilon^m + \bar{u}^m)}{2} = \mathfrak{c}^f$$

with analogous expressions for $(\hat{c}^f, -\hat{n}^f, \hat{c}^m, -\hat{n}^m)$.

The value of the program, for a given $\boldsymbol{z} = (c^f, -n^f, c^m, -n^m)$ is, therefore,

$$\left[\frac{x + (\mathfrak{h}^f + \bar{u}^f) - (\mathfrak{h}^m + \bar{u}^m)}{2} + \mathfrak{h}^f - \bar{u}^f\right]\left[x - \frac{x + (\mathfrak{h}^f + \bar{u}^f) + (\mathfrak{h}^m - \bar{u}^m)}{2} + \mathfrak{h}^m - \bar{u}^m\right],$$

or

$$\frac{1}{4}\left[x + (\mathfrak{h}^f - \bar{u}^f) + (\mathfrak{h}^m - \bar{u}^m)\right]^2.$$

Equation (27) is therefore equivalent to

$$\left[\underbrace{x - (\mathfrak{h}^f + \mathfrak{h}^m)}_{y} - \underbrace{(\bar{u}^f + \bar{u}^m)}_{\bar{\upsilon}}\right]^2 \geq \left[\underbrace{\hat{x} - (\hat{\mathfrak{h}}^f + \hat{\mathfrak{h}}^m)}_{\hat{y}} - \underbrace{(\bar{u}^f + \bar{u}^m)}_{\bar{\upsilon}}\right]^2$$

Finally, using $y - \bar{\upsilon} > 0$, and $\hat{y} - \bar{\upsilon} > 0$, (27) is shown to be equivalent to

$$y - \bar{\upsilon} \geq \hat{y} - \bar{\upsilon} \Leftrightarrow y \geq \hat{y}.$$

□

Lemma 3. *For a given $\mathcal{E}$, assume that there is a subset $H_{\mathcal{E}} \subset \Theta^2$ of positive measure such that for all $\boldsymbol{\theta}$ in $H_{\mathcal{E}}$ there is $\bar{\boldsymbol{z}}(\boldsymbol{\theta}) \in \mathcal{Z}^c_{\mathcal{E}}$ such that $\Phi(\bar{\boldsymbol{z}}(\boldsymbol{\theta}), \boldsymbol{\theta}) \in \Lambda_{\mathcal{E}}(\boldsymbol{\theta}, \bar{\boldsymbol{u}}_{\mathcal{E}}(\boldsymbol{\theta}))$. Let also, $\hat{H}_{\mathcal{E}} \subset \Theta^2$ be the set of all households whose transactions are either envied by no one or are only envied by households in the set $H_{\mathcal{E}}$. Then, one of the following must be true: no transaction $\boldsymbol{z}(\boldsymbol{\theta})$ for $\boldsymbol{\theta} \in \hat{H}_{\mathcal{E}}$ is distorted, or; the allocation $(\boldsymbol{z}_{\mathcal{E}}(\boldsymbol{\theta}))_{\boldsymbol{\theta}}$ is constrained inefficient.*

Proof. Just note that any $\bar{\boldsymbol{z}}(\boldsymbol{\theta}) \in \mathcal{Z}^c_{\mathcal{E}}$ respects incentive compatibility. □

Lemma 4. *Assume that there is $\boldsymbol{z}' \in \mathcal{Z}^c_{\mathcal{E}}$ such that $\boldsymbol{z}' > \bar{\boldsymbol{z}}_{\mathcal{E}}(\boldsymbol{\theta})$. Further assume that $\hat{\boldsymbol{z}}_{\mathcal{E}}(\boldsymbol{\theta}, \Phi(\boldsymbol{z}', \boldsymbol{\theta})) = \boldsymbol{z}_{\mathcal{E}}(\boldsymbol{\theta})$. Then, $\Phi(\boldsymbol{z}', \boldsymbol{\theta}) = \Lambda_{\mathcal{E}}(\boldsymbol{\theta}, \bar{\boldsymbol{u}}_{\mathcal{E}}(\boldsymbol{\theta}))$.*

Proof. If $\boldsymbol{z}' \in \mathcal{Z}^c_{\mathcal{E}}$, property (iv) of Φ guarantees that at least one spouse is better-off. But if $\hat{\boldsymbol{z}}_{\mathcal{E}}(\boldsymbol{\theta}, \Phi(\boldsymbol{z}', \boldsymbol{\theta})) = \boldsymbol{z}_{\mathcal{E}}(\boldsymbol{\theta})$ then it can only be because the $\Phi(\boldsymbol{z}', \boldsymbol{\theta}) = \Lambda_{\mathcal{E}}(\boldsymbol{\theta}, \bar{\boldsymbol{u}}_{\mathcal{E}}(\boldsymbol{\theta}))$. □