The Optimal Taxation of Couples*

Mikhail Golosov
University of Chicago

Ilia Krasikov
Arizona State University

August 2024

Abstract

We study optimal nonlinear taxation of single and married households. Taxes on couples depend on the earnings of both spouses and are an example of multi-dimensional tax schedules. We develop novel analytical techniques to study properties of such taxes. We show that the optimal marginal taxes for married individuals are generally lower than for single individuals because resource-sharing within couples provides socially valuable redistribution. Under realistic assumptions, the optimal tax rates for married individuals increase with the correlation of spousal earnings, the marginal tax rates for one spouse increase (decrease) in the earnings of the other if both spouses have low (high) earnings, and the primary earner faces lower marginal taxes than the secondary earner. We extend our approach to consider normative tax implications of within-family public goods, home production, extensive margin in labor supply, selection into marriage, bargaining over marital surplus, and gender differences.

*We thank Ben Brooks, Mariacristina De Nardi, Marina Halac, Michael Kremer, Elliot Lipnowski, Phil Reny, Emmanuel Saez, Aleh Tsyvinski, Alessandra Voena, Nicolas Werquin and the seminar participants at the EIEF, Federal Reserve Bank of Chicago, Federal Reserve Bank of St Louis, 2023 Hydra conference, Minnesota Macro, NBER Summer Institute, Paris-London Public Economics conference, SED, SEA, SAET, the University of Chicago for comments. We thank Karan Jain for excellent research assistance. Golosov thanks the NSF for support (grant #36354.00.00.00).

1 Introduction

In developed countries, a significant portion of income taxes is paid by households comprising two adult members. For example, married couples in the U.S. contribute to over 70 percent of federal income taxes. Despite this, the theory of optimal taxation of family income remains poorly understood. What economic factors determine the shape and magnitude of the optimal tax schedule for two-earner households? How should the taxes of one member be influenced by the earnings of the other? Is it ever beneficial to tax each individual in a couple separately, or should total family income be used as the sole basis for taxation? How should taxes on married households compare to those on single individuals?

In this paper, we take a step towards answering these questions. We study optimal taxation in a simple model of couple formation. In our model, all individuals are identical ex-ante, share the same preferences (which, for tractability, we assume to be quasi-linear with a constant elasticity of labor supply), and draw productivities from the same distribution. Each individual decides whether to stay single or to get married based on both the pecuniary utility of consumption and leisure and non-pecuniary preference shocks associated with marriage. Within marriage, spouses divide their marital surplus equally. Our model is flexible in that it can accommodate any level of assortativity in the marriage market, allowing for any joint distribution of spouses' productivities.

We study optimal taxation of single and married households in this economy. In line with a longstanding tradition in public finance, we posit that the social planner sets taxes with the aim of redistributing wealth from more productive to less productive individuals. This objective is encapsulated by Pareto weights, which decrease as an individual's productivity increases. The social planner chooses tax functions $T^s(y)$ for single households with earnings y and $T^m(y_1, y_2)$ for married spouses who have earnings y_1 and y_2. Beyond the requirement of budget feasibility, we do not impose any additional constraints on these functions.

The optimal tax problem can be formulated as a problem of a fictitious mechanism designer who chooses allocations based on individuals' reports about their productivities subject to incentive compatibility and feasibility constraints. Assuming that only local incentive constraints bind at the optimum – known as the first-order approach (FOA) – we demonstrate that the optimal tax distortions[1] of single individuals are characterized by a linear ordinary differential equation. This equation is straightforward to solve analytically, and it has been

[1] By "distortion" we mean the monotone transformation of the tax rate, $\frac{\frac{\partial}{\partial y}T^s}{1-\frac{\partial}{\partial y}T^s}$ for a single household, and $\frac{\frac{\partial}{\partial y_i}T^m}{1-\frac{\partial}{\partial y_i}T^m}$ for each spouse i in a married household.

extensively studied in the literature since the seminal work of Mirrlees (1971). In contrast, the optimal tax distortions for married individuals are described by a system of nonlinear partial differential equations. These equations are significantly more challenging, and no general analytical techniques exist to solve them. We show that this challenge can be partially addressed by employing two mathematical results from multivariate calculus: the divergence theorem and the coarea formula. By combining these results, we can analytically characterize a large family of conditional averages that the optimal distortions for married individuals must satisfy.

The optimal distortions for single individuals are described by the classical "ABC" formula developed by Diamond (1998). This formula indicates that the optimal distortion for a single individual with a given productivity level is determined by the ratio of benefits from redistribution — capturing welfare gains from transferring resources from singles with higher productivity levels to an average single individual — to the costs of tax distortions, which depend on the elasticity of labor supply and the elasticity of the productivity distribution with respect to the productivity level. The optimal distortions for married individuals follow a broadly similar form, but benefits and costs of taxation are balanced across all possible slices of the two-dimensional productivity distribution of couples. Each slice provides insights into a certain conditional moment of marginal tax rates for married spouses, and we use different moments to study various properties of the optimal tax schedule.

Using our formulas, we obtain three sets of economic insights about the shape and the magnitude of optimal tax schedule. First, we establish that the optimal distortions for married individuals are, on average, less than those for single individuals with equivalent productivity. This is due to the intra-family redistribution that occurs through resource sharing within couples, a mechanism valued by the social planner. Replacing this intra-family redistribution with redistribution through the tax system is costly, leading the planner to choose lower tax rates for couples. One implication of this finding is that the social planner subsidizes marriages, resulting in a higher marriage rate in the optimum compared to the laissez-faire economy.

We also show that the distortions for married individuals increase with the degree of assortativity in marriages and coincide with the distortions for single individuals under perfect assortative matching. Additionally, we exhibit a natural order for ranking the redistributiveness of the social objective function, and demonstrate that Pareto weights that are more redistributive in this order lead to higher distortions for both single and married households.

Second, we study how the optimal marginal tax rates on one spouse depend on the earnings of their partner, which is a crucial aspect of multi-dimensional taxes. Positive jointness – where the marginal taxes of one spouse increase with the earnings of the other – allows the planner to better target taxes toward the richest couples. Negative jointness, on the other hand, enables the planner to better direct transfers to the poorest couples. Both forms of targeting are socially valuable, and optimal jointness depends on their relative costs and benefits. Under random matching, these effects cancel each other out, resulting in the optimal taxes on married spouses that are additively separable in their earnings. We show that, under realistic models of positive assortative matching, optimal jointness is negative for high-earning couples and positive for low-earning ones.

Third, we examine how distortions faced by two spouses in the same couple compare to each other. We demonstrate that, under mild restrictions on social weights, spouses who earn a significantly smaller fraction of family income face, on average, higher distortions than their partners. This allows the planner to efficiently target transfers to the poorest married individuals.

We also use our formulas to analytically investigate the conditions under which the first-order approach (FOA) is valid for both single and married households. When matching is random, it is possible to derive necessary and sufficient conditions under which the FOA holds for each type of household. It turns out that the conditions for the validity of the FOA for married households are strictly less stringent than those for single households. This stands in marked contrast to the results in the industrial organization literature, which concludes that the FOA generally fails in multi-dimensional monopolist pricing models (e.g., Rochet and Chone (1998)).

Our approach to characterizing optimal joint taxation can be easily adapted to richer environments. In this paper, we consider several extensions that incorporate various economic mechanisms emphasized by the family economics literature.

First, we show that within-marriage public goods and consumption economies of scale make marriages more economically efficient but also amplify consumption inequality among married individuals. As a result, the optimal distortions increase for married individuals and decrease for single individuals compared to the economy without those features. We also incorporate home production and intra-household division of labor and describe how the optimal tax formulas account for the fact that that the market labor supply of a married person is affected, via intra-family specialization, by the marginal tax rates of their partner.

Second, we investigate how optimal taxation is affected when spouses bargain over marital surplus rather than share it equally. Under classical Nash bargaining, where spouses use their outside options of singledom to determine the split of their marital surplus, the optimal taxes for married individuals remain the same as under equal consumption sharing, but the optimal taxes for single individuals are higher. The higher optimal taxes for single individuals arise because, by reducing inequality among them, the planner facilitates a more equitable division of surplus within marriages. The optimal marginal taxes for married individuals remain unaffected because, under classical Nash bargaining, each marginal dollar of family surplus is equally divided between spouses, even though the overall division of surplus is unequal. The optimal tax rates are determined by how couples share the marginal dollar of after-tax joint income, which aligns with the distribution under equal consumption sharing. Additionally, we show that generalized bargaining, wherein more productive spouses possess both greater bargaining power and superior outside options, has an ambiguous effect on the magnitude of optimal distortions for married individuals. This ambiguity arises from the conflict in the planner's incentives to redistribute within couples versus between married and single households.

Third, we extend the model by allowing individuals to adjust their labor supply along both the intensive and extensive margins. The analysis of optimal taxation of couples becomes considerably more complicated. If matching is random, the formulas that describe optimal distortions of single and married individuals are similar to our baseline case but include an additional cost of distortion due to the extensive margin adjustment. Under certain conditions, we can compare them and show that married individuals still face lower distortions than singles, but the optimal marriage rate, and hence the total marriage tax subsidy, decreases with the strength of the extensive margin response.

Forth, we extend our approach to economies with observable heterogeneity, such as gender differences. We characterize both gender-specific and gender-neutral optimal taxes and demonstrate a very close relationship between the two. The optimal distortions under gender-neutral taxation are equal to the weighted average of the optimal distortions under gender-specific taxation, with weights determined by relative fractions of people of each gender at any given level of productivity. This same insight holds more broadly. We derive expressions for the optimal taxes on married individuals that are restricted to be disjointed, and that are restricted to depend only on their total family earnings. In both cases, the optimal distortions under such restricted tax systems are equal to appropriate weighted averages of the optimal distortions under the unrestricted tax system.

In the final theoretical section of the paper, we investigate conditions on primitives under which the optimal taxes on married individuals are based solely on their total family earnings. We demonstrate that family-earnings based taxation is optimal if the planner uses Pareto weights that explicitly favor such taxation (capturing the idea of the "horizontal equity" concern) and, in addition, a certain average of productivities of the two spouses is independent of their relative productivities. We also characterize how departing from pure family-earnings based taxation improves welfare of such planner when this condition is not satisfied.

Finally, we we quantitatively investigate optimal taxation within the baseline model. We use data on the earnings of married households and the U.S. tax schedule to obtain the joint distribution of productivities. We show that the Gaussian copula with Pareto-lognormal marginal distributions can well approximate this distribution. We find that our analytical formulas provide excellent guidance to numerical properties of the optimal tax schedule. In the U.S. data, spousal productivities are positively but not perfectly correlated, so the optimal taxes on married individuals are higher than in the economy with random matching but lower than in the uni-dimensional models such as Diamond (1998). Consistent with our theoretical results, we find that optimal jointness is positive at the bottom and negative at the top, and that secondary earners face higher marginal tax rates. Quantitatively, optimal jointness is small and the optimal taxes for couples are well approximated by a disjointed tax schedule. In contrast, taxation based only on total family earnings is generally quite far from the unrestricted optimum.

Our paper is related to several strands of literature. Mirrlees (1976, 1986) derived optimality conditions for multi-dimensional tax problems under the FOA and pointed out that they are much more challenging than their uni-dimensional analogs. Subsequent literature typically imposed additional simplifying assumptions to avoid confronting these partial differential equations. For example, Kleven et al. (2009) studied taxation of couples but restricted one spouse to make only binary labor supply choices. Frankel (2014) considered the case in which a binary distribution describes spouses' productivities. Ales and Sleet (2022) studied couples taxation in a discrete choice environment. Moser and de Souza e Silva (2019) analyzed paternalistic savings policies in a model with two-dimensional discrete heterogeneity. Alves et al. (2021) considered the optimal tax problem of couples but imposed enough structure to collapse it into a uni-dimensional problem. Golosov et al. (2013) and Lockwood and Weinzierl (2015) pursued a similar approach in labor and commodity taxation with preference heterogeneity. Hellwig and Werquin (2022) discussed a generalization of their ideas of redistributional arbitrage to multi-dimensional type spaces. In a series of papers, Rothschild and Scheuer (2013; 2014; 2016) developed a mechanism design approach to study optimal taxation in models with multi-dimensional private information but with uni-dimensional tax instruments. In contrast to these papers, we develop an approach that allows us to analytically characterize optimal taxation in a fairly unrestricted multi-dimensional environment and shed light on economic forces that are hard to see in more specialized settings.[2]

The closest work related to our study is the unpublished Section 3 of the working paper by Kleven et al. (2007). Those authors derived the expression for optimal average distortions of a married individual, and characterized the sign of optimal jointness under the assumption of random matching into couples. The family of moments that our approach allows us to characterize is significantly richer. Our techniques to study jointness under assortative matchings are new and they offer a broader perspective on the trade-offs that determine optimal jointness. The comparison of conditions for validity of the FOA in uni- and bi-dimensional settings, comparative statics results, and the numerous extensions we consider are all novel to our work.

Several authors, such as Golosov et al. (2014), Spiritus et al. (2022), Ferey et al. (2022) study optimal multidimensional taxation using an alternative, variational approach. They consider perturbations of tax schedules and derive expressions for optimal rates in terms of sufficient statistics. While their approach has many appealing features, its key limitation is that the optimal tax rates are expressed in terms of endogenous objects that are themselves functions of the optimal tax schedule. This makes it difficult to use those expressions to understand how the model structural parameters affect optimal taxes. In contrast, our formulas are derived in terms of the exogenous primitives, which allows us to prove sharp theoretical results.

Gayle and Shephard (2019) and Spiritus et al. (2022) use numerical methods to study the optimal joint taxation of couples. Boerma et al. (2022) developed techniques to tackle multi-dimensional mechanism design problems when the FOA fails. Our work is complementary to theirs. Our analytical results provide insights about the forces determining the optimal taxes that are often hard to see with numerical approaches.

The rest of the paper is organized as follows. In Section 2, we present our benchmark economy. In Section 3, we describe the mechanism design approach and characterize optimal taxes in that benchmark economy. Section 5 considers various extensions. Section 6 provides calibration and quantitative analysis. Section 7 concludes.

[2] In addition to these papers, our work is also related to the New Dynamic Public Finance literature (see, e.g., Golosov et al. (2003), Albanesi and Sleet (2006), Farhi and Werning (2013), Golosov et al. (2016), Stantcheva (2017), Ndiaye (2018)) that studies optimal nonlinear taxes in dynamic environments in which information is revealed over time. In those models, optimal taxes in a given period are a nonlinear function of earnings in previous periods, but the dynamic nature of information revelation allows collapsing the mechanism design problem to a sequence of problems with uni-dimensional incentive constraints. Also related is the recent work by Kushnir and Shourideh (2022) who explore alternative ways to relax multidimensional mechanism design problems.

2 The environment

Consider an economy comprised of a unit measure of ex-ante identical individuals, each with a utility function given by $c - \gamma l^{1/\gamma}$, where c and l denote consumption and labor, and $\gamma \in (0,1)$ is the parameter capturing the elasticity of labor supply.[3] Each person decides whether to stay single or get married, and how much to work and consume. These decisions occur in three stages.

Stage 1. Each person draws a shock ε that captures idiosyncratic, non-pecuniary benefits of singlehood and decides whether to go to the marriage market. Let $\mathbb{E}U^s$ and $\mathbb{E}U^m$ be the expected pecuniary benefits of singlehood and marriage, respectively, which we will define formally in Stage 3. A person with a shock ε goes to the marriage market if $\mathbb{E}U^m > \mathbb{E}U^s + \varepsilon$ and remains single if $\mathbb{E}U^m \leq \mathbb{E}U^s + \varepsilon$.

Stage 2. Each person on the marriage market draws a publicly observable signal q, which is potentially correlated with their productivity that will realize in Stage 3, and marry another person with the same value of the signal. Two married spouses agree to share their marital surplus equally.

Stage 3. Each person draws a productivity w from a cumulative probability distribution G and decides how much to work and consume. A person with productivity w who supplies l units of labor earns pre-tax income $y = wl$. Let $T^s(y)$ and $T^m(y_1, y_2)$ be the taxes on single and married households, respectively. The decision problem of a single household is

$$v^s(w) := \max_{c,y} c - \gamma \left(\frac{y}{w}\right)^{1/\gamma} \text{ s.t. } c \leq y - T^s(y) \text{ and } y \geq 0,$$

which pins down utility from singlehood to be $U^s = v^s$. Married spouses act collectively and solve

$$v^m(w_1, w_2) := \max_{c,\{y_i\}_{i=1}^2} c - \sum_{i=1}^2 \gamma \left(\frac{y_i}{w_i}\right)^{1/\gamma} \text{ s.t. } c \leq \sum_{i=1}^2 y_i - T^m(y_1, y_2) \text{ and } y_1, y_2 \geq 0.$$

The decision problem of married spouses pins down their joint consumption c that is allocated so that each spouse receives equal utility from marriage, $U^m = \frac{1}{2}v^m$.[4] Given these definitions, expected pecuniary utilities in Stage 1 are $\mathbb{E}U^m = \frac{1}{2}\mathbb{E}v^m$ and $\mathbb{E}U^s = \mathbb{E}v^s$, respectively.

Our economy provides a simple way to model many realistic features of marriage, labor supply and resource allocations. Shocks ε capture the innate personal inclination for marriage or singlehood. These shocks also imply that marriage decisions are not influenced elusively

[3] The relationship between γ and the elasticity of labor supply e is $1/\gamma = 1 + 1/e$.

[4] Individual consumptions of two spouses (c_1, c_2) splits their joint consumption $c = c_1 + c_2$ so that each spouse obtains the same utility from marriage that equals to the half of their marital surplus.

by pecuniary benefits and keep the elasticity of marriage rates to a marriage tax penalty or bonus finite. Shocks q capture the fact that married individuals often share similar socio-economic characteristics that are predictive of their future earnings. For our approach, we can be agnostic about how assortative the marriage market is. Our model is flexible in that it allows for random matching on productivities (if q is independent of w), positive assortative matching (if q is positively correlated with w) and even negative assortative matching (if q is negatively correlated with w). Until Section 5.6, we do not need to model explicitly the relationship between q and w and simply use F to denote the joint distribution of productivities of married persons that emerges as the outcome of the matching process in Stage 2.

Our model isolates some key economic mechanisms that shape optimal taxation of single and married households without introducing additional complexities. Since all individuals are ex-ante identical and select into marriage before any information about their future productivity is revealed, married individuals are statistically identical to single individuals as they all draw their productivities from the same distribution and have the same elasticity of labor supply. We abstract for now from home production, extensive margin responses in labor supplies, more sophisticated intra-household bargaining. We discuss implications of these features in Section 5.

We impose mild regularity conditions to simplify exposition. It is without loss of generality to take F to be symmetric, with both marginals equal to G; furthermore, we assume that F admits a continuously differentiable density f that is strictly positive on $\mathbb{R}^2_+$ and denote the density of G by g. To streamline our exposition, we assume that the distribution of productivities F satisfies $\int \max\limits_{i=1,2} w_i^{1/1-\gamma} dF < \infty$. Shocks ε are drawn from an absolutely continuous probability distribution supported on $\mathbb{R}$; we use Φ to denote the inverse of this probability distribution. We use μ to denote the marriage rate.

It is useful to first describe the equilibrium in the absence of taxes. In the laissez-faire economy, the expected amount of resources available to each person is independent of that person's marital status. Randomness in productivity realizations and the matching process introduce uncertainty about ex-post resource allocations, but this uncertainty does not affect individuals' ex-ante utilities since they are risk-neutral. This implies that the pecuniary benefits of marriage and singlehood are the same in the absence of taxation. We record this observation in the lemma that follows, where we use superscripts "LF" to denote the laissez-faire allocations.

Lemma 1. *In the laissez-faire economy,* $\mathbb{E}U^{m,\mathrm{LF}} = \mathbb{E}U^{s,\mathrm{LF}}$.

We now turn to the problem of optimal taxation. Following a long tradition in public finance going back to the work of Mirrlees (1971), we assume that the social planner chooses taxes to redistribute resources from more productive to less productive individuals. Let $\mathbb{E}[U|w]$ be the expected utility of a person with productivity w. This utility can be written as

$$\mathbb{E}[U|w] = \mu\mathbb{E}[U^m|w] + (1-\mu)\mathbb{E}[U^s|w] + \int_\mu^1 \Phi(\varepsilon)\, d\varepsilon, \tag{1}$$

where the first two terms capture their pecuniary utilities from marriage and singlehood, respectively, and the last term corresponds to the non-pecuniary benefit of singlehood. We take social welfare to be given by the Pareto-weighted sum of these utilities, that is $\mathcal{W} := \int \alpha(w)\mathbb{E}[U|w]\, dG$, where α is a non-negative, strictly decreasing, bounded, continuous function normalized so that $\int \alpha dG = 1$. This definition of Pareto weights is the natural benchmark. Under these weights, the social planner values an extra dollar of consumption of a person with productivity w with the same weight $\alpha(w)$ irrespective of whether the person is single or married, which implies that the planner has no inherent preference for or against marriage.

The social planner chooses T^s and T^m to maximize social welfare. For now, we impose no restrictions on the form of these tax functions other than that total tax revenues in equilibrium must be non-negative. In Section 5.9, we consider the problem of a planner who faces additional ad-hoc restrictions on the form of taxes that she can use and show that there is a close relationship between the optimal taxes with and without ad-hoc restrictions.

3 Optimal taxation as a mechanism design problem

We use the mechanism design approach to study optimal taxation. Since this approach is well known, we present it heuristically, leaving technical details for the appendix.

Let $\mathbf{w} = (w_1, w_2)$ be a pair of productivities of a given couple. Using the taxation principle (see, e.g., Hammond (1979)), one can show that T^s, T^m are budget feasible if and only if there exists a $\mu \in [0,1]$ and tuples (v^s, c^s, y^s) and (v^m, c^m, y_1^m, y_2^m) that satisfy

$$v^s(w) = c^s(w) - \gamma\left(\frac{y^s(w)}{w}\right)^{1/\gamma} \forall w, \quad v^m(\mathbf{w}) = c^m(\mathbf{w}) - \sum_{i=1}^{2}\gamma\left(\frac{y_i^m(\mathbf{w})}{w_i}\right)^{1/\gamma} \forall \mathbf{w}, \tag{2}$$

$$v^s(w) \geq c^s(\widehat{w}) - \gamma\left(\frac{y^s(\widehat{w})}{w}\right)^{1/\gamma} \forall w, \widehat{w}, \quad v^m(\mathbf{w}) \geq c^m(\widehat{\mathbf{w}}) - \sum_{i=1}^{2}\gamma\left(\frac{y_i^m(\widehat{\mathbf{w}})}{w_i}\right)^{1/\gamma} \forall \mathbf{w}, \widehat{\mathbf{w}}, \tag{3}$$

$$\frac{\mu}{2}\int\left(\sum_{i=1}^{2} y_i^m - c^m\right)dF + (1-\mu)\int (y^s - c^s)\, dG \geq 0, \tag{4}$$

$$\Phi(\mu) = \frac{1}{2}\int v^m dF - \int v^s dG. \tag{5}$$

Equations (2) define utilities of single and married households and (3) are their incentive constraints. Equation (4) is the resource constraint that ensures that total tax revenues are non-negative. Equation (5) determines the cut-off of the preference shock $\bar{\varepsilon}$ at which a person is indifferent between getting married and remaining single. This indifference cut-off is given by $\bar{\varepsilon} = \frac{1}{2}\mathbb{E}v^m - \mathbb{E}v^s$, which is the same equation as (5) because the marriage rate μ satisfies $\Phi(\mu) = \bar{\varepsilon}$. The planner's objective function $\mathcal{W}$ can be written in terms of v^m and v^s as

$$\mathcal{W} = \frac{\mu}{2}\int \alpha^m v^m dF + (1-\mu)\int \alpha v^s dG + \int_\mu^1 \Phi d\varepsilon, \tag{6}$$

where

$$\alpha^m(w_1, w_2) := \frac{1}{2}\alpha(w_1) + \frac{1}{2}\alpha(w_2). \tag{7}$$

Thus, the mechanism design problem can be stated as finding μ, (v^s, c^s, y^s), (v^m, c^m, y_1^m, y_2^m) that maximize (6) subject to (2) – (5).

It will be convenient to simplify this maximization problem by getting rid of redundant variables. Due to the envelope theorem, Equation (3) implies

$$\frac{\partial v^s}{\partial w} = \frac{(y^s)^{1/\gamma}}{w^{1+1/\gamma}}, \quad \frac{\partial v^m}{\partial w_i} = \frac{(y_i^m)^{1/\gamma}}{w_i^{1+1/\gamma}}. \tag{8}$$

These equations can be thought of as local incentive constraints, as they ensure that no person can gain from small misreporting of their type.

One challenge of studying mechanism design problems is the very large number of incentive constraints (3). The standard approach to overcome this problem, known as *the first-order approach (FOA),* is to consider a relaxed problem in which the global incentive constraints (3) are replaced by the local constraints, (8). This approach is known to hold in many realistic uni-dimensional tax models. For now, we follow it and characterize the optimal tax schedule under the assumption that the FOA is valid. In Section 4 we theoretically examine conditions for the validity of the FOA and show that the FOA is more likely to hold for multi-dimensional tax problem of couples in an important special case of our environment. In Section 6 we verify numerically that the FOA is valid in the calibrated economy.

Using Equations (2) and (8), we can substitute out for consumption and labor earnings. The relaxed problem is to choose v^s, v^m and μ to maximize welfare (6) subject to the marriage indifference condition (5) and the feasibility constraint,

$$\frac{\mu}{2}\int \sum_{i=1}^{2}\left(w_i^{1+\gamma}\left(\frac{\partial v^m}{\partial w_i}\right)^\gamma - \gamma w_i \frac{\partial v^m}{\partial w_i}\right) dF + (1-\mu)\int \left(w^{1+\gamma}\left(\frac{\partial v^s}{\partial w}\right)^\gamma - \gamma w\frac{\partial v^s}{\partial w}\right) dG \geq$$

$$\geq \frac{\mu}{2}\int v^m dF + (1-\mu)\int v^s dG. \tag{9}$$

3.1 Optimality conditions to the relaxed problem

Let $v^{s,*}$, $v^{m,*}$, and μ^* denote the solution to the relaxed mechanism design problem. This solution determines the optimal earnings $y^{s,*}$, $y^{m,*}$ via (8) and the optimal marginal tax rates. It will be convenient to work with monotone transformations of the optimal marginal tax rates. We define these transformations as

$$\lambda^{s,*}(w) := \frac{\frac{\partial}{\partial y}T^{s,*}\left(y^{s,*}(w)\right)}{1-\frac{\partial}{\partial y}T^{s,*}\left(y^{s,*}(w)\right)}, \quad \lambda_i^{m,*}(\mathbf{w}) := \frac{\frac{\partial}{\partial y_i}T^{m,*}\left(\mathbf{y}^{m,*}(\mathbf{w})\right)}{1-\frac{\partial}{\partial y_i}T^{m,*}\left(\mathbf{y}^{m,*}(\mathbf{w})\right)}. \tag{10}$$

and refer to them as *optimal distortions.*

We now derive the set of optimality conditions for the relaxed mechanism design problem. Using $\int \alpha dG = 1$, it is easy to show that the Lagrange multipliers on (5) and (9) must be zero and one, respectively. These results have simple economic interpretations. Since the planner is inherently indifferent about each person's marriage status, the constraint that determines which agents select into marriage, Equation (5), is slack. The Lagrange multiplier on (9) captures how much the social planner values an extra unit of consumption in the hands of an average person in the economy, which is equal to the average value of Pareto weights that we normalized to one.

Using our observation about the values of Lagrange multipliers, the optimality conditions for $v^{s,*}$ can be written as

$$\frac{\partial}{\partial w}\left(\lambda^{s,*}\gamma w g\right) = (\alpha-1)\,g, \quad \lim_{w\to 0,\infty}\lambda^{s,*}(w)\,wg(w) = 0. \tag{11}$$

This is a linear ODE that is easy to solve. Integrate it from t to ∞ and use the boundary condition to obtain $\lambda^{s,*}(t) = \frac{\int_t^\infty (1-\alpha)dG}{\gamma t g(t)}$. It will be slightly more convenient to normalize both the numerator and the denominator by $1-G(t)$ so that we can interpret integrals as conditional expectations. Using this normalization, the optimal distortions for single individuals can be written as

$$\lambda^{s,*}(t) = \frac{1-\mathbb{E}\left[\alpha|w\geq t\right]}{\gamma\theta(t)}, \tag{12}$$

where $\theta(t)$ is the tail statistics of distribution G defined by

$$\theta(t) := \frac{tg(t)}{1-G(t)} = \frac{-d\ln\Pr(w\geq t)}{d\ln t}. \tag{13}$$

Equation (12) expresses $\lambda^{s,*}$ for each productivity level t in terms of the primitives of our environment and thus fully characterizes optimal marginal taxes. The optimal marginal taxes for single households are independent of the marriage rate and coincide with the optimal taxes in a version of our model with only single households. Equation (12) is a version of famous Diamond's "ABC formula" generalized to our economy in which single and married households co-exist.[5]

It is useful to highlight the intuition behind Equation (12). Consider a thought experiment of increasing the marginal tax rate on single persons with productivity $w = t$. This increases the level of taxes for all single households whose productivities $w > t$. To balance the government budget, we adjust the intercepts of tax functions, $T^{s,*}(0)$ and $T^{m,*}(0,0)$, so that the marriage rate μ^* is unaffected. One can think of the welfare effect of this perturbation by separately considering the mechanical effect, which occurs if agents do not change their behavior, and the behavioral effect, which captures how government tax revenues are effected by agents re-optimazing their choices in response to this tax change.

Let start with the mechanical effect. This tax perturbation takes an extra dollar from each single household with $w > t$ and gives it to "average" single and married persons. The social value of a dollar in the hands of a person with productivity w is $\alpha(w)$, while the social value of a dollar in the hands of an average person, single or married, is one. There are $(1-\mu^*)(1-G(t))$ single households with productivities $w > t$, so that the total change of social welfare due to the mechanical effect is $(1-\mu^*)(1-G(t)) \times (1 - \mathbb{E}[\alpha|w \geq t])$. Since our perturbation does not affect the marriage rate, the behavioral effect arises only because single individuals with productivity t reduce their labor supply. The reduction of tax revenues is a product of the tax distortion for single $\lambda^{s,*}(t)$, the elasticity parameter of labor supply γ, their productivity t, and the mass of single households affected by this perturbation, $(1-\mu^*)g(t)$. In the optimum, the sum of the mechanical and behavioral effects must be zero, which gives Equation (12). This thought experiment shows that the optimal distortions are given by the ratio of the benefits of redistribution (the numerator on the right hand side of (12)) to the costs of tax distortions (the denominator on the right hand side of (12)).

We now turn to describing distortions for married households. The same variational techniques show that $\boldsymbol{\lambda}^{m,*}$ satisfies the following system of equations:

$$\sum_{i=1}^{2} \frac{\partial}{\partial w_i}\left(\lambda_i^{m,*}\gamma w_i f\right) = \left(\alpha^m - 1\right) f, \qquad \lim_{w_i \to 0,\infty} \lambda_i^{m,*}\left(\mathbf{w}\right) w_i f\left(\mathbf{w}\right) = 0 \text{ for all } w_{-i} \tag{14}$$

and

$$\frac{\partial}{\partial \ln w_2}\left(\frac{w_1}{1+\lambda_1^{m,*}}\right)^{1/(1-\gamma)} = \frac{\partial}{\partial \ln w_1}\left(\frac{w_2}{1+\lambda_2^{m,*}}\right)^{1/(1-\gamma)}. \tag{15}$$

Equation (14) is very similar to (11) except now it has a sum over spouse-specific distortions. Unlike the uni-dimensional case, Equation (14) is not sufficient for optimality: there are many

[5] In Diamond's ABC terminology, $\frac{1}{\gamma}$ is "term A", $\frac{1}{\theta(t)}$ is "term B", and $1 - \mathbb{E}[\alpha|w \geq t]$ is "term C".

functions $\{\lambda_i^{m,*}\}_{i=1}^2$ that satisfy this equation but these functions cannot be chosen in isolation from each other as they are transformations of derivatives of the same tax function $T^{m,*}$. The cross-partial derivatives of $T^{m,*}$ must agree, $\frac{\partial}{\partial y_2}\frac{\partial}{\partial y_1}T^{m,*} = \frac{\partial}{\partial y_1}\frac{\partial}{\partial y_2}T^{m,*}$, which imposes an additional restriction on admissible $\{\lambda_i^{m,*}\}_{i=1}^2$ that is given by Equation (15). Under mild regularity conditions, Equations (14) and (15) are not only necessary but also sufficient for the optimum.

In contrast to (11), it is very hard to use (14) and (15) to find $\boldsymbol{\lambda}^{m,*}$ explicitly. These equations form a system of non-linear PDEs, and there are no readily available techniques to solve them analytically.[6] Our approach is to sidestep the difficult task of characterizing $\boldsymbol{\lambda}^{m,*}$ analytically at every point $\mathbf{w}$. Instead, we exploit the fact that Equation (14) is a relatively tractable linear differential equation in $\boldsymbol{\lambda}^{m,*}$ and use it to derive a rich family of conditional averages that the optimal distortions satisfy.

Consider any continuous function $Q : \mathbb{R}^2_{++} \to \mathbb{R}_{++}$. For any $t > 0$, the set of couples $\mathbf{w}$ for whom $Q(\mathbf{w}) = t$, or $\{Q = t\}$ for brevity, divides the space of all couples into two regions: couples for whom $Q > t$ and couples for whom $Q < t$ (see Figure 1 for an illustration). Our key result of this section is that the optimality condition (14) can be integrated over regions $\{Q > t\}$ or $\{Q < t\}$ and, by using two mathematical results from multi-variable calculus – the divergence theorem and the coarea formula – these integrals can be expressed as a conditional average of optimal distortions along the boundary $\{Q = t\}$.

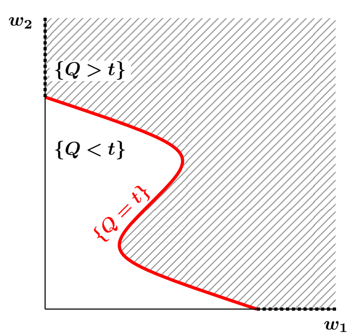


Figure 1: Using a function Q to define areas in the type space. The thick line is the set $\{Q = t\}$, the hashed area is the set $\{Q > t\}$, and white region is the set $\{Q < t\}$.

Theorem 1. *Let $Q : \mathbb{R}^2_{++} \to \mathbb{R}_+$ be an onto locally Lipshitz function that satisfies mild*

[6] Renes and Zoutman (2017) describe how (their equivalent of our) Equation (14) can be exploited to find $\boldsymbol{\lambda}^{m,*}$ using so-called Green functions if one assumes that $\boldsymbol{\lambda}^{m,*}$ is a conservative vector field, i.e., it satisfies $\frac{\partial}{\partial w_2}\lambda_1^{m,*} = \frac{\partial}{\partial w_1}\lambda_2^{m,*}$. Unfortunately, $\boldsymbol{\lambda}^{m,*}$ does not need to form a conservative vector field but instead has to satisfy Equation (15). This condition is non-linear, which makes analysis much more difficult.

regularity conditions (Q1)-(Q4) stated in the appendix. Then, the optimal distortions satisfy

$$\mathbb{E}\left[\sum_{i=1}^{2}\lambda_i^{m,*}\frac{\partial\ln Q}{\partial\ln w_i}\Big|Q=t\right]=\frac{1-\mathbb{E}\left[\alpha^m|Q\geq t\right]}{\gamma\frac{-d\ln\Pr(Q\geq t)}{d\ln t}}. \tag{16}$$

Equation (16) shares a lot of similarities with (12). The numerator on the right-hand side of (16) captures the average Pareto weight of persons in the $\{Q > t\}$ region. The denominator is a product of the elasticity parameter γ and the density of households on the boundary of $\{Q = t\}$ relative to the mass of $\{Q > t\}$. This ratio generalizes the tail statistics $\theta(t)$ that appeared in Equation (12). Since such a function Q and its value t are arbitrary, Theorem 1 allows us to characterize a large set of moments that the optimal distortions of married persons must satisfy. We unpack some of the implications of these moments in the next section.

Similarly to Equation (12), the right-hand side of (16) is a ratio of the benefits of redistribution from couples in $\{Q > t\}$ relative to the costs of distortion that this redistribution entails. This equation can be derived heuristically similarly to (12) by considering a tax perturbation that increases the level of taxes of all $\{Q > t\}$ by the same infinitesimal amount, with the intercepts, $T^{s,*}(0)$ and $T^{m,*}(0,0)$, being adjusted in a way that keeps the marriage rate constant. The numerator on the right-hand side of (16) captures the mechanical effect of this perturbation and its redistributive gains. The denominator on the right-hand side of (16) corresponds to the number of agents affected by this perturbation, adjusted by their productivities and labor supply elasticities. The gradient $\{\frac{\partial\ln Q}{\partial\ln w_i}\}_{i=1}^2$ that appears on the left hand side of (16) reflects how much marginal tax rates on each spouse need to be adjusted to engineer a uniform increase in tax levels for all $\{Q > t\}$.

A remarkable feature of Equations (12) and (16) is that the optimal marginal tax rates for both single and married households are independent of the marriage rate or the elasticity of the marriage rate to taxes. The optimal marriage rate μ^* is determined by the following first-order condition:

$$\frac{1-\gamma}{2}\int\sum_{i=1}^{2}w_i\left(\frac{w_i}{1+\lambda_i^{m,*}}\right)^{\gamma/(1-\gamma)}dF-(1-\gamma)\int w\left(\frac{w}{1+\lambda^{s,*}}\right)^{\gamma/(1-\gamma)}dG=\Phi(\mu^*). \tag{17}$$

The optimal marriage rate depends not only on the optimal distortions but also on the elasticity of the marriage rates to taxes, captured by the distribution Φ of non-pecuniary "love" shocks ε. As we demonstrate in the subsequent section, this equation yields sharp qualitative predictions about the optimal marriage rate μ^*.

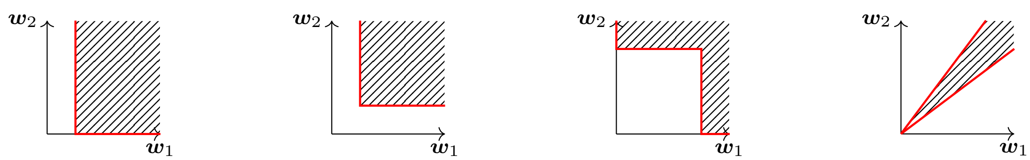
The discussion above describes the procedure to characterize the optimal marginal tax rates, $\frac{\partial}{\partial y}T^{s,*}$ and $\{\frac{\partial}{\partial y_i}T^{m,*}\}_{i=1}^2$, and the optimal marriage rate μ^*. The only remaining moments of the optimal tax system are the intercepts, $T^{s,*}(0)$ and $T^{m,*}(0,0)$. These intercepts are pinned by the marriage indifference condition (5) and by the feasibility condition (4) holding with equality.

3.2 Properties of the optimal joint taxes

In this section, we use Theorem 1 to shed light on four specific moments of the optimal distortions for couples:

$$\mathbb{E}\left[\lambda_i^{m,*}|w_i = t\right], \mathbb{E}\left[\lambda_i^{m,*}|w_i = t, w_{-i} \geq t\right], \mathbb{E}\left[\lambda_i^{m,*}|w_i = t, w_{-i} \leq t\right], \mathbb{E}\Big[\lambda_i^{m,*} - \lambda_{-i}^{m,*}|w_i = \iota w_{-i}\Big].$$

These four moments highlights different aspects of the optimal tax function $T^{m,*}$. The first moment encapsulates the mean value of the optimal distortion for a married person with productivity t. This moment is the most direct counterpart of $\lambda^{s,*}(t)$ for a single person. The second and third moments quantify how the optimal distortion of a married person depends on their spouse's productivity capturing the mean value of optimal distortion given that the spouse is more and less productive, respectively. We refer to the ratio of these two moments as jointness. The forth moment compares the optimal distortions of two spouses within the same couple when one of the spouses is ι times more productive than the other. These last moment can be obtained from Theorem 1 by considering Q functions of the form $Q = w_i$, $Q = \min \boldsymbol{w}$, $Q = \max \boldsymbol{w}$, and $Q = \frac{\min \boldsymbol{w}}{\max \boldsymbol{w}}$, see Figure 2 for an illustration.



(a) $w_1 > t$. (b) $\min \boldsymbol{w} > t$. (c) $\max \boldsymbol{w} > t$. (d) $\frac{\min \boldsymbol{w}}{\max \boldsymbol{w}} > \iota$.

Figure 2: Q functions that are used to study various properties of the optimal tax schedule. The thick line is the set $\{Q = t\}$, the hashed area is the set $\{Q > t\}$, and white region is the set $\{Q < t\}$.

3.2.1 Average optimal distortions

By setting $Q = w_i$ in Formula (16), we obtain

$$\mathbb{E}\left[\lambda_i^{m,*}|w_i = t\right] = \frac{1 - \mathbb{E}\left[\alpha^m|w_i \geq t\right]}{\gamma\theta(t)}. \tag{18}$$

It is instructive to compare this equation to its analogue for a single person, Equation (12). Both formulas have the same denominator, $\gamma\theta(t)$, but different numerators. The optimal distortion of a single person depends on $\mathbb{E}\left[\alpha|w_i \geq t\right]$, which is the average Pareto weight of a person who has productivity higher than t. In contrast, the average optimal distortion of a married person depends on $\mathbb{E}[\alpha^m | w_i \geq t]$, which is the average Pareto weight of all people in couples in which one of the spouses has productivity higher than t. The difference in these formulas arises because an increase in the marginal taxes on a married person with productivity $w_i = t$ raises average taxes for all couples with a $w_i > t$ earner, but the burden of these higher taxes are shared by both spouses in those couples.

An important implication of Equation (18) is that the degree to which people marry partners with similar productivities plays a crucial role in optimal taxation. To build the intuition for a general result, we first consider two special cases. If the marriage market is perfectly assortative, so that individuals always marry spouses with the same productivity, then $\mathbb{E}[\alpha^m | w_i \geq t] = \mathbb{E}[\alpha | w_i \geq t]$ for all t and the optimal distortions for single and married persons coincide. In contrast, if the marriage market is random, so that productivities of spouses are statistically independent from each other, then $\mathbb{E}[\alpha^m | w_i \geq t] = \frac{1}{2} + \frac{1}{2}\mathbb{E}[\alpha | w_i \geq t]$ and Equation (18) implies that the labor distortion of a married person is on average one half of the distortion of a single person with the same productivity. Random matching cuts the redistributory benefits of taxation in half since a person with any productivity t shares, in expectation, half of their tax burden with a spouse of average productivity.

This discussion suggests that married individuals should face lower distortions than single individuals and that the magnitude of this "marriage subsidy" should depend on the assortativity of marriages. To show that this is indeed the case, we need to introduce a notion of assortativity that can be applied to general, non-parametric settings. Consider two joint symmetric distributions F^a and F^b with the same marginals G. We say that F^b is *independent* if $F^b(w_1, w_2) = G(w_1)G(w_2)$ for all w_1, w_2, or $F^b = G^2$ for short, *positively dependent* if $F^b \geq G^2$, and *more dependent* than F^a if $F^b \geq F^a$. We denote this latter relationship by $F^b \geq_{PQD} F^a$. Any F satisfies bounds $\overline{F} \geq_{PQD} F \geq_{PQD} \underline{F}$, where $\overline{F}$ and $\underline{F}$ are distributions under perfect positive and negative assortative matchings, respectively. A reader familiar with literature on multi-variate stochastic orders will recognize this notion as the positive quadrant dependence partial order (e.g., see Shaked and Shanthikumar (2007) or Nelsen (2006)). It is widely used in statistical literature and is equivalent to the condition that $\mathbb{C}ov(\phi_1(w_1), \phi_2(w_2)) \geq 0$ for any two increasing functions ϕ_1 and ϕ_2. To see an illustration of this definition, consider two families of joint distributions, given by the Gaussian and FGM copulas with some correlation parameter ρ.[7] For both copulas, independence and positive dependence of F^b is equivalent to

[7] Copulas provide a convenient with to construct joint distributions using arbitrary marginal distributions. In our setting, a copula C is a mapping $C : [0,1]^2 \rightarrow [0,1]$, where $C(u_1, u_2)$ is the joint probability that the productivity of spouse 1 is in the u_1^{th} quantile of their marginal distribution and the pro-

$\rho = 0$ and $\rho \geq 0$, respectively, and $F^b \geq_{PQD} F^a$ is equivalent to $\rho^b \geq \rho^a$.

Using our notion of dependence, we can derive a sharp comparative statics characterization of the average value of optimal distortions for married individuals.

Lemma 2. *Consider two economies, a and b, that are identical in all respects except the joint distribution of productivities, and assume that $F^a \leq_{PQD} F^b$. The relationship between optimal distortions in the two economies is*

$$\mathbb{E}^a \left[\lambda_i^{m,a,*} | w_i = t\right] \leq \mathbb{E}^b \left[\lambda_i^{m,b,*} | w_i = t\right] < \lambda^{s,b,*}(t) = \lambda^{s,a,*}(t) \text{ for all } t.$$

The second inequality would be equality if we relax our assumption on F to include distributions without densities and set $F^b = \overline{F}$.

Furthermore, $\mathbb{E}^a \left[\lambda_i^{m,a,*} | w_i = t\right] > 0$ for all t if F^a is positively dependent.

To understand the intuition for this result, recall that the social planner uses distortionary taxes to provide redistribution among individuals. When two individuals form a couple, they pool their resources together, providing an alternative, intra-family, channel of redistribution. This intra-family redistribution is valued by the planner. Consequently, the planner responds by setting lower distortions for married individuals than for singles. The more mixing of productivities there is in the marriage market, the more redistribution families provide, and the higher the optimal marriage tax bonus.

A direct implication of Lemma 2 is that it is optimal to subsidize marriage.

Corollary 1. *The pecuniary gains from marriage and the marriage rate are higher in the optimum than in the laissez-faire: $\mathbb{E}U^{m,*} - \mathbb{E}U^{s,*} > \mathbb{E}U^{m,\mathrm{LF}} - \mathbb{E}U^{s,\mathrm{LF}}$ and $\mu^* > \mu^{\mathrm{LF}}$.*

Proof. Recall from Lemma 1 that in the absence of taxes $\frac{1}{2}\mathbb{E}v^{m,\mathrm{LF}} - \mathbb{E}v^{s,\mathrm{LF}} = \mathbb{E}U^{m,\mathrm{LF}} - \mathbb{E}U^{s,\mathrm{LF}} = 0$ and, therefore, the laissez-faire marriage rate satisfies $\Phi(\mu^{\mathrm{LF}}) = 0$. The optimal marriage rate μ^* is pinned down by Equation (17). Since $x \mapsto (1+x)^{\gamma/(\gamma-1)}$ is decreasing and convex, we have

$$\mathbb{E}\left[\left(1+\lambda_i^{m,*}\right)^{\gamma/(\gamma-1)} | w_i = t\right] \geq \left(1 + \mathbb{E}\left[\lambda_i^{m,*} | w_i = t\right]\right)^{\gamma/(\gamma-1)} > \left(1+\lambda^{s,*}(t)\right)^{\gamma/(\gamma-1)},$$

where the last inequality follows from Lemma 2. Substitute this into Equation (17) to show that $\Phi(\mu^*) > 0$ and, therefore, $\mu^* > \mu^{\mathrm{LF}}$. Equation (5) then implies that $\mathbb{E}U^{m,*} - \mathbb{E}U^{s,*} > 0$. □

ductivity of spouse 2 is in the u_2^{th} quantile. Copulas allow one to isolate dependence properties of F from properties of its marginal distributions G_1, G_2 in general settings. The Gaussian copula is defined as $C(u_1, u_2) \propto \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \exp\left[-\frac{\left(s_1^2 - 2\rho s_1 s_2 + s_2^2\right)}{2\left(1-\rho^2\right)}\right] ds_1 ds_2$, where Φ is the distribution of a standard normal random variable, and it generalizes dependence properties of the bi-variate normal to arbitrary marginals, e.g., a joint log-normal distribution is characterized by the Gaussian copula. The FGM copula is defined by $C(u_1, u_2) = u_1 u_2 \left[1 + \rho(1-u_1)(1-u_2)\right]$. In both cases, the parameter $\rho \in (-1, 1)$ captures the degree of dependence. See Nelsen (2006) for an introduction to copulas.

Our discussion suggests that a more "redistributive" social planner should set higher marginal taxes. A natural way to compare the redistributiveness of two Pareto weights, α^a and α^b, is to consider their likelihood ratio α^b/α^a. If this ratio is decreasing in w, then the planner places a uniformly higher value on less productive individuals under α^b than under α^a.

Lemma 3. *Consider two economies, a and b, that are identical in all respects except the Pareto weights, and assume that α^b/α^a is decreasing in w. Then, $\lambda^{s,a,*}(t) \leq \lambda^{s,b,*}(t)$ for all t. In addition, if f is log-supermodular, then $\mathbb{E}\left[\lambda_i^{m,a,*}|w_i = t\right] \leq \mathbb{E}\left[\lambda_i^{m,b,*}|w_i = t\right]$ for all t.*

Consistent with intuition, a more redistributive social planner uses higher labor distortions for single households. The same result extends to married households provided that f is log-supermodular. Log-supermodularity captures a form of positive dependence among variables and is satisfied by many commonly used joint distributions, such as those constructed by the Gaussian and FGM copulas with $\rho \geq 0$.

3.2.2 Optimal average jointness

We now turn to the discussion of how distortions of a married person depend on productivity of their spouse.[8] We consider the following measure of jointness:

$$J(t) = \frac{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t \leq w_{-i}\right]}{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t \geq w_{-i}\right]} - 1. \tag{19}$$

This term is positive (negative) if a person's distortion is higher (lower), on average, if they marry a more productive spouse than themselves. By setting $Q = \min \mathbf{w}$ in Formula (16) and using symmetry of F, we obtain

$$\mathbb{E}\left[\lambda_i^{m,*}|w_i = t \leq w_{-i}\right] = \frac{1 - \mathbb{E}\left[\alpha^m | \min \mathbf{w} \geq t\right]}{\gamma \frac{-d\ln \Pr(\min \mathbf{w} \geq t)}{d \ln t}} = \frac{1}{\gamma \frac{-d \ln \Pr(w_i \geq t)}{d \ln t}} \frac{1 - \mathbb{E}\left[\alpha^m | \min \mathbf{w} \geq t\right]}{\frac{d \ln \Pr(\min \mathbf{w} \geq t)}{d \ln \Pr(w_i \geq t)}}. \tag{20}$$

The other moment, $\mathbb{E}\left[\lambda_i^{m,*}|w_i = t \leq w_{-i}\right]$, can be unpacked analogously by setting $Q = \max \mathbf{w}$ in Formula (16):

$$\mathbb{E}\left[\lambda_i^{m,*}|w_i = t \geq w_{-i}\right] = \frac{1 - \mathbb{E}\left[\alpha^m | \max \mathbf{w} \geq t\right]}{\gamma \frac{-d\ln \Pr(\max \mathbf{w} \geq t)}{d \ln t}} = \frac{1}{\gamma \frac{-d \ln \Pr(w_i \geq t)}{d \ln t}} \frac{1 - \mathbb{E}\left[\alpha^m | \max \mathbf{w} \geq t\right]}{\frac{d \ln \Pr(\max \mathbf{w} \geq t)}{d \ln \Pr(w_i \geq t)}}. \tag{21}$$

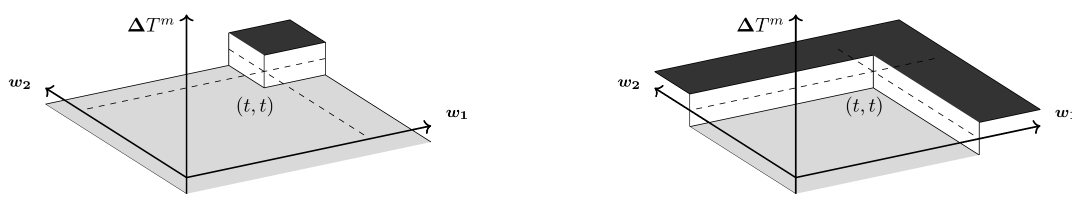
As a result, we can write J as

$$J(t) = \frac{1 - \mathbb{E}\left[\alpha^m | \min \mathbf{w} \geq t\right]}{\frac{d \ln \Pr(\min \mathbf{w} \geq t)}{d \ln \Pr(w_i \geq t)}} \div \frac{1 - \mathbb{E}\left[\alpha^m | \max \mathbf{w} \geq t\right]}{\frac{d \ln \Pr(\max \mathbf{w} \geq t)}{d \ln \Pr(w_i \geq t)}} - 1. \tag{22}$$

[8] Kleven et al. (2007) discuss jointness in their model of couples taxation. Jointness in their setup is driven by a different mechanism from the one studied in this section. We discuss the connection with their results in Section 5.1.

Before discussing implications of this equation, it will be insightful to consider the economics of tax jointness.

Suppose we start with a separable tax schedule $T^m(y_1, y_2) = \widetilde{T}^m(y_1) + \widetilde{T}^m(y_2)$. Pick some level of productivity t and consider two different reforms of this tax schedule, which we call Reform I and Reform II. Under Reform I, the planner increases average taxes by the same infinitesimal amount for all couples in which both spouses are more productive than t. The set of those couples is $\{\min \mathbf{w} > t\}$. Under Reform II the planner increases average taxes for all couples in which at least one spouse is more productive than t, with the set of those couples is $\{\max \mathbf{w} > t\}$. In both cases, the marginal taxes are increased on the boundaries of these sets, $\{\min \mathbf{w} = t\}$ and $\{\max \mathbf{w} = t\}$, to attain this increase in average taxes. Tax revenues are redistributed uniformly to all households in a way that keeps the marriage rate unchanged. See Figure 3 for a graphical description of Reforms I and II.



(a) Reform I

(b) Reform II

Figure 3: The change of the tax schedule on married ΔT^m due to Reforms I and II.

It is easy to pick visually that redistributory gains of these reforms are $1 - \mathbb{E}\left[\alpha^m \middle| \min \mathbf{w} \geq t\right]$ and $1 - \mathbb{E}\left[\alpha^m \middle| \max \mathbf{w} \geq t\right]$, and the densities of agents who face higher distortions are $\frac{-d \ln \Pr(\min \mathbf{w} \geq t)}{dt}$ and $\frac{-d \ln \Pr(\max \mathbf{w} \geq t)}{dt}$, respectively. Thus, J is equal to the ratio of redistributory benefits of the two reforms normalized by the number of distorted individuals.

There are several insightful observations about these reforms that illuminate the economics behind optimal joint taxation. Reform I increases taxes on the richest couples, $\{\min \mathbf{w} > t\}$, that are then redistributed by increasing transfers to $\{\min \mathbf{w} < t\}$. Reform II increases transfers to the poorest couples, $\{\max \mathbf{w} < t\}$ by raising taxes on $\{\max \mathbf{w} > t\}$. Both reforms enhance the planner's ability to target the tax and transfer system more effectively. Reform I improves the targeting of taxes on the richest couples, while Reform II better targets transfers to the poorest couples. Despite these improvements in targeting, the reforms have opposite implications for jointness: Reform I introduces positive jointness into the tax schedule, whereas Reform II introduces negative jointness. This creates a trade-off for the planner. Positive jointness allows for better targeting of taxes but results in less effective transfer targeting, and vice versa. The optimal degree of jointness is determined by weighing the cost-adjusted benefits of the two reforms.

The benefits of transfer targeting can be made more salient in the formula for optimal jointness. Since Pareto weights α^m and probabilities $\mathbf{w}$ integrate to one, we can re-write Equation (22) as

$$J(t) = \frac{\mathbb{E}\left[\alpha^m \middle| \min \mathbf{w} \le t\right] - 1}{\frac{d \ln \Pr(\min \mathbf{w} \le t)}{d \ln \Pr(w_i \le t)}} \div \frac{\mathbb{E}\left[\alpha^m \middle| \max \mathbf{w} \le t\right] - 1}{\frac{d \ln \Pr(\max \mathbf{w} \le t)}{d \ln \Pr(w_i \le t)}} - 1. \tag{23}$$

Mathematically, this equation is identical to (22), but it is written in a way that emphasizes redistributory benefits of cost-adjusted transfers.

An examination of the cost adjustments that appear in Equations (22) and (23) reveals that they are determined by assortativity in matching. Thus, one should expect the correlation of spousal productivities to play a significant role in determining the optimal degree of jointness.

First, consider two special cases of assortativity: perfectly assortative matching and random matching. Under perfectly assortative matching, both $\mathbb{E}\left[\alpha^m \middle| \min \mathbf{w} \ge t\right]$ and $\mathbb{E}\left[\alpha^m \middle| \max \mathbf{w} \ge t\right]$ are equal to $\mathbb{E}\left[\alpha | w \ge t\right]$, and $\frac{d \ln \Pr(\min \mathbf{w} \ge t)}{d \ln \Pr(w_i \ge t)}$ and $\frac{d \ln \Pr(\min \mathbf{w} \ge t)}{d \ln \Pr(w_i \ge t)}$ are equal to one. Therefore, Equation (22) implies that $J(t) = 0$ for all t. Under random matching, both $\mathbb{E}\left[\alpha^m \middle| \min \mathbf{w} \ge t\right]$ and $\mathbb{E}\left[\alpha^m \middle| \max \mathbf{w} \ge t\right]$ are equal to $\frac{1}{2} + \frac{1}{2}\mathbb{E}\left[\alpha | w \ge t\right]$, and $\frac{d \ln \Pr(\min \mathbf{w} \ge t)}{d \ln \Pr(w_i \ge t)}$ and $\frac{d \ln \Pr(\min \mathbf{w} \ge t)}{d \ln \Pr(w_i \ge t)}$ are equal to two. This implies that under random matching $J(t) = 0$ for all t, but for a different reason from perfectly assortative matching. In the case of perfectly assortative matching, jointness gives no additional benefits of tax and transfer targeting but it also has no additional costs. In contrast, under random matching, both tax and transfer targeting improves welfare, but their relative costs exactly offset each other. In both cases, a separable tax system can implement the optimal allocations. We summarize this discussion in the following lemma.

Lemma 4. *If matching is either random or perfectly assortative, then the optimal taxes $T^{m,*}$ are separable and the optimal distortions $\lambda_i^{m,*}$ are independent of w_{-i}. These distortions satisfy, respectively, $\lambda_i^{m,*}(t, w_{-i}) = \frac{1}{2}\lambda^{s,*}(t)$ and $\lambda_i^{m,*}(t, w_{-i}) = \lambda^{s,*}(t)$ for all t, w_{-i}.*

To gain insights into the empirically relevant case of positive but not perfectly assortative matching, suppose that F is given by the Gaussian copula with the correlation parameter $\rho > 0$. We first consider properties of J for high and low values of t. Let $\alpha(\infty) = \lim_{t \to \infty} \alpha(t)$ be the asymptotic weight on the richest individual.

Since productivities are positively correlated, under the Gaussian copula, we have

$$\lim_{t \to \infty} \mathbb{E}\left[\alpha^m \middle| \min \mathbf{w} \ge t\right] = \lim_{t \to \infty} \mathbb{E}\left[\alpha^m \middle| \max \mathbf{w} \ge t\right] = \alpha(\infty), \tag{24}$$

which shows that the redistributory benefits of tax and transfer targeting coincide for the richest couples. At the same time, one can show that

$$\lim_{t\to\infty}\frac{d\ln\Pr\left(\min\mathbf{w}\geq t\right)}{d\ln\Pr\left(w_i\geq t\right)}=\frac{2}{1+\rho},\quad \lim_{t\to\infty}\frac{d\ln\Pr\left(\max\mathbf{w}\geq t\right)}{d\ln\Pr\left(w_i\geq t\right)}=1, \tag{25}$$

so that the costs of the two targeting schemes differ. Since $\rho<1$, the cost of tax targeting is higher than the cost of transfer targeting for high earners. Therefore, Equation (22) implies that $J(t)<0$ for all sufficiently large t, indicating that optimal jointness for high earners is negative.

The analysis of optimal jointness in the left tail is similar, but the results are easier to see using Equation (23). According to this equation, as $t\to 0$, the redistributory benefits of tax and transfer targeting converge to $\alpha(0)$, but their costs $\frac{d\ln\Pr(\min\mathbf{w}\leq t)}{d\ln\Pr(w_i\leq t)}$ and $\frac{d\ln\Pr(\max\mathbf{w}\leq t)}{d\ln\Pr(w_i\leq t)}$ converge to 1 and $\frac{2}{1+\rho}$ respectively. This implies that $J(t)>0$ for all sufficiently small t, indicating that optimal jointnessss for low earners is positive.

It is hard to analyze analytically the sign of optimal jointness for intermediate values of t in our general, non-parametric settings. If we simplify the structure of the joint distribution of productivities by assuming that it is given by the analytically tractable FGM copula, we can extend our conclusions to all t. We summarize our discussion in the following lemma.

Lemma 5. *Suppose that F is given either by the Gaussian or the FGM copula with $\rho>0$. In both cases, $J(t)<0$ for all t sufficiently large, and $J(t)>0$ for all t sufficiently small. Moreover, in the case of the FGM copula, there exists a threshold $\bar{t}>0$ such that $J(t)>0$ for all $t<\bar{t}$ and $J(t)<0$ for all $t>\bar{t}$.*

3.2.3 Distortions for primary and secondary earners

We now turn to comparison of the optimal distortions within family. We refer to a spouse with higher productivity as the *primary earner* and to the other spouse as the *secondary earner*, and we shall use $\lambda_{pr}^{m,*}(\mathbf{w})$ and $\lambda_{sec}^{m,*}(\mathbf{w})$ to denote their optimal distortions, respectively.[9] Let $I(\mathbf{w})=\frac{\min\boldsymbol{w}}{\max\boldsymbol{w}}$ be the productivity of the secondary earner relative to that of the primary earner. By setting $Q=I$ in Formula (16), we obtain

$$\mathbb{E}\left[\lambda_{sec}^{m,*}-\lambda_{pr}^{m,*}|I=\iota\right]=\frac{1-\mathbb{E}\left[\alpha^m|I\geq\iota\right]}{\gamma\theta_\iota(\iota)}, \tag{26}$$

where $\theta_\iota(\iota)=\frac{-d\ln\Pr(I\geq\iota)}{d\ln\iota}$ is the tail statistics of the distribution of relative productivities of two spouses. In this formula, $\mathbb{E}\left[\lambda_{sec}^{m,*}-\lambda_{pr}^{m,*}|I=\iota\right]$ is the average difference in labor distortions

[9] By definition, these optimal distortion are given by $\lambda_{pr}^{m,*}(\mathbf{w})=\lambda_i^{m,*}$ and $\lambda_{sec}^{m,*}=\lambda_{-i}^{m,*}$ when $w_i\geq w_{-i}$.

between the secondary and primary earners in couples in which the primary earner is $1/\iota$ times more productive. When spousal productivities are equal, $\iota = 1$, their optimal distortions are the same, $\lambda_{sec}^{m,*}(w,w) = \lambda_{pr}^{m,*}(w,w)$ for all w. On the other hand, if spousal productivities are very unequal, then the optimal distortions for the secondary earner are on average higher than for the primary earner.

Lemma 6. *Suppose that $\alpha(0) > 2$ and $\lim_{\iota \to 0} \mathbb{E}\left[w_i | w_{-i} \leq \iota w_i\right] < \infty$. Then,*

$$\mathbb{E}\Big[\lambda_{sec}^{m,*} - \lambda_{pr}^{m,*} \big| I = \iota\Big] > 0$$

for all sufficiently small ι.

To understand the sufficient condition $\alpha(0) > 2$ and the intuition behind the result, first recall that the weights α are monotonically decreasing and integrate to one. Therefore, this condition means that the social planner values a dollar in the hands of the least productive person at least twice as much as in the hands of an average person in the economy. Under this condition, the planner values redistribution to an unproductive married spouse i irrespective of whom that person is matched to, i.e., $\alpha^m(\boldsymbol{w}) = \frac{\alpha(w_i)}{2} + \frac{\alpha(w_{-i})}{2} \geq 1$ for all w_{-i} when w_i is sufficiently small. As a result, the planner wants to transfer some resources to all couples with unproductive secondary earners. Such transfers are phased out as the earnings of the secondary earner grow, leading to high implicit marginal taxes on the secondary earner.

4 On the validity of the first order approach

In our characterization of optimal taxes, we followed the common approach of simplifying the mechanism design problem by replacing the global incentives constraints, Equations (3), with their local analogs, Equations (8). While this method is widely used for uni-dimensional problems, there is a common perception in the literature that it may fail in multi-dimensional settings. For example, in their classic study of the optimal taxation of couples, Kleven et al. (2009) note (p. 538), "very few studies in the optimal tax literature have attempted to deal with multidimensional screening problems. The nonlinear pricing literature in industrial organization has analyzed such problems extensively. A central complication of multidimensional screening problems is that first order conditions are often not sufficient to characterize the optimal solution. The reason is that solutions usually display "bunching" at the bottom (Armstrong (1996), Rochet and Chone (1998)), whereby agents with different types are making the same choices." To sidestep this perceived difficulty, Kleven et al. (2009) further restrict agents' choices by allowing one of the spouses to make only binary labor supply decisions. They explain (p. 538), "Our framework with a binary labor supply outcome for the secondary earner along with continuous earnings for the primary earner avoids the bunching complexities and offers a simple understanding of the shape of optimal taxes based on graphical exposition."

In this section, we theoretically examine the validity of the FOA. We focus on the case of random matching, where we can fully describe conditions under which the FOA is valid for characterizing optimal taxes for single and for characterizing optimal taxes for married households, and compare these conditions.

Recall that the relaxed problem is to choose v^s, v^m and μ to maximize welfare (6) subject to (5) and (9). We can write the omitted global constraints in terms of v^s and v^m as follows:

$$v^s(w) \geqslant v^s(\widehat{w}) + \gamma \widehat{w} \frac{\partial v^s(\widehat{w})}{\partial w} \left(\left(\frac{\widehat{w}}{w} \right)^{1/\gamma} - 1 \right) \quad \forall w, \widehat{w}, \tag{27}$$

$$v^m(\mathbf{w}) \geqslant v^m(\widehat{\mathbf{w}}) + \sum_{i=1}^{2} \gamma \widehat{w}_i \frac{\partial v^m(\widehat{\mathbf{w}})}{\partial w_i} \left(\left(\frac{\widehat{w}_i}{w_i} \right)^{1/\gamma} - 1 \right) \quad \forall \mathbf{w}, \widehat{\mathbf{w}}. \tag{28}$$

The FOA is valid for single households if $v^{s,*}$ satisfies (27) and valid for married households if $v^{m,*}$ satisfies (28).

Equations (27) and (28) are complicated, and using them directly to verify the validity of the FOA is difficult. The analysis becomes significantly easier once we observe that these equations can be simplified if we use the transformation $x \leftrightarrow w^{-1/\gamma}$. Let $v^{m,x}$ and $v^{s,x}$ be utilities of married and single households in the transformed type variables, that is $v^{m,x}(x_1, x_2) = v^m\left(w_1^{-\gamma}, w_2^{-\gamma}\right)$ and $v^{s,x}(x) = v^s(w^{-\gamma})$. When written in the x-space, Equations (27) and (28) become

$$v^{s,x}(x) \geqslant v^{s,x}(\widehat{x}) + \frac{\partial v^{s,x}(\widehat{x})}{\partial x}(x - \widehat{x}) \quad \forall x, \widehat{x},$$

$$v^{m,x}(\mathbf{x}) \geqslant v^{m,x}(\widehat{\mathbf{x}}) + \sum_{i=1}^{2} \frac{\partial v^{m,x}(\widehat{\mathbf{x}})}{\partial x_i}(x_i - \widehat{x}_i) \quad \forall \mathbf{x}, \widehat{\mathbf{x}}.$$

These equations are equivalent to the requirement that v^s and v^m are convex in the x-space, i.e., $v^{s,x}$ and $v^{m,x}$ are both convex functions. Thus, the FOA is valid if the solution to the relaxed problem is convex in the x-space.

We characterized $v^{s,*}$ and $v^{m,*}$ in the case of random matching in the w-space explicitly in Section 3. By transforming those solutions into the x-space and using routine algebra, we can establish the following result.

Proposition 1. *Suppose that the matching is random and g, α are such that $\lambda^{s,*}$ defined in Equation (12) is bounded, continuously differentiable, and has bounded derivatives.*

The FOA for single households is valid if and only if

$$x \cdot \left(1 + \lambda^{s,*}\left(x^{-\gamma}\right)\right) \quad \textit{is increasing in } x. \tag{29}$$

The FOA for married households is valid if and only if

$$x \cdot \left(1 + \frac{1}{2}\lambda^{s,*}\left(x^{-\gamma}\right)\right) \text{ is increasing in } x. \tag{30}$$

In particular, (30) holds whenever (29) holds.

The remarkable conclusion of this proposition is that the FOA is *more likely* to hold to study the optimal taxation of married than of single households in this economy. To understand the intuition behind this result, it is insightful to consider the economic interpretation of Equations (29) and (30).

Equation (29), which characterizes conditions for validity of the FOA for single households, can equivalently be written as

$$\left[1 + \left(1 + \gamma\frac{\partial \ln(wg)}{\partial \ln w}\right)\lambda^{s,*}\right] + [1-\alpha] \geq 0 \text{ for all } w. \tag{31}$$

The term in the first square bracket is typically positive, whereas the term in the second square bracket is negative for low w and positive for high w. Thus, (31) is violated if the second term is sufficiently negative relative to the first, which occurs if the planner puts sufficiently high Pareto weights on some low types. In other words, the FOA holds for single households if the planner is not "too redistributive" in the precise sense given by Equation (31). If the planner becomes "too redistributive" for some w, she would want to increase redistribution around those types. The relaxed problem would call for a sharply increasing marginal taxes for those households, which would make their implied after-tax budget constraint highly non-convex.

Equation (30), which characterizes conditions for validity of the FOA for married households, is very similar to (31) since it captures the same economic mechanism. However, the set of primitives under which (30) holds is strictly larger. As we showed in Section 3.2, the social planner endogenously chooses to redistribute less among married than among single because of the intra-family redistribution within couples. This implies that there are fewer cases in which the relaxed problem would choose highly non-convex tax functions.

The conclusion of Proposition 1 is a special case of a more general insight that multi-dimensional mechanism design problems in public finance are fundamentally different from multi-dimensional pricing problems studied by Armstrong (1996) and Rochet and Chone (1998). The mechanism designer in pricing problems aims to extract maximum surplus from agents. In contrast, the mechanism designer in public finance settings aims to redistribute resources and the FOA holds for a wide class of Pareto weights.

5 Extensions

In this section, we consider several extensions of our economy. We discuss the implications of different welfare criteria, intra-household public goods, labor specialization, bargaining, the extensive margin in labor supply decisions, selection into marriage, and gender differences for optimal taxation. Additionally, we explain how to apply our techniques to analyze the optimal taxes that are constrained to certain forms, such as being separable. Finally, we establish a close connection between the optimal unrestricted taxes, which are the main focus of our analysis, and the optimal taxes with exogenous restrictions. Each of these extensions is discussed in a separate subsection. To maintain a clear and concise presentation, we focus on highlighting a few key takeaways and relegate all the proofs to the appendix.

5.1 The role of Pareto weights

In previous sections, we assumed that the social planner uses the same weight $\alpha(w)$ for any person with productivity w irrespective of their marital status or the productivity of their spouse. We now discuss implications of alternative assumptions about Pareto weights that the social planner could use to evaluate welfare. Suppose that a person with productivity w is valued by the planner with weight $\alpha^s(w)$ if single and with weight $\beta(w|w_{-i})$ if married to a spouse with productivity w_{-i}. Let $\alpha^m(\boldsymbol{w}) := \frac{1}{2}\beta(w_i|w_{-i}) + \frac{1}{2}\beta(w_{-i}|w_i)$ be the average Pareto weight within a couple. We allow $\mathbb{E}\alpha^s = \int \alpha^s dG$ and $\mathbb{E}\alpha^m = \int \beta dF$ to be arbitrary positive numbers.

Following the same steps as in Section 3, one can show that the optimal distortions for single and married persons satisfy

$$\lambda^{s,*}(t) = \frac{1 - \mathbb{E}[\frac{\alpha^s}{\mathbb{E}\alpha^s}|w \geq t]}{\gamma\theta(t)} \times \frac{\mathbb{E}\alpha^s}{(1-\mu^*)\mathbb{E}\alpha^s + \mu^*\mathbb{E}\alpha^m}, \tag{32}$$

$$\mathbb{E}\left[\sum_{i=1}^{2} \lambda_i^{m,*} \frac{\partial \ln Q}{\partial \ln w_i}\Big|Q = t\right] = \frac{1 - \mathbb{E}\left[\frac{\alpha^m}{\mathbb{E}\alpha^m}|Q \geq t\right]}{\gamma\frac{-d\ln \Pr(Q \geq t)}{d\ln t}} \times \frac{\mathbb{E}\alpha^m}{(1-\mu^*)\mathbb{E}\alpha^s + \mu^*\mathbb{E}\alpha^m}. \tag{33}$$

The first terms on the right-hand side of these equations capture the benefits of redistribution within single and within married households, respectively. These terms have the same interpretations as terms that appear in Equations (12) and (16). For example, $\frac{\alpha^s(w)}{\mathbb{E}\alpha^s}$ is the social value of a dollar in the hands of a single person with productivity w relative to that of an average single person, and $\frac{\alpha^m(\mathbf{w})}{\mathbb{E}\alpha^m}$ is the couple's analogue of that ratio. The second terms are new. To see their implications, first suppose that the planner is indifferent about a person's marital status, in the sense that $\mathbb{E}\alpha^s = \mathbb{E}\alpha^m$. In this case, they are both equal to one and the optimal tax formulas collapse to Equations (12) and (16) except that α^m is now the average of β rather than α (recall that in Equations (12) and (16) we normalized $\mathbb{E}\alpha$ to one). The economic trade-offs are exactly the same as in Section 3. If the planner values married people more highly, $\mathbb{E}\alpha^m > \mathbb{E}\alpha^s$, then there is a new force that calls for higher distortions for married and lower distortions for single. The results are the opposite when $\mathbb{E}\alpha^m < \mathbb{E}\alpha^s$.

Why does a preference for marriage incentivize the planner to increase distortions for married persons? This happens because the planner cares more about both the level of their expected utility and the inequality among them. Thus, the planner finds it optimal to increase redistribution among married households financed by using lower distortions for single households.

We now look at the implications of the dependence of each spouse i's social weight on their parner's productivity w_{-i} for the optimal taxes on married. If $\beta(w_i|w_{-i})$ is independent of w_{-i}, so that the planner values a dollar in the hands of a married person in the same way irrespective of whom they marry, then most of the results in Section 3.2 are unchanged. In particular, Lemmas 3, 4, 5, and 6 still hold except α is replaced by β in the statements of these results. Similarly, the comparative statics in Lemma 2 also holds provided that $\mathbb{E}\alpha^s = \mathbb{E}\alpha^m$.

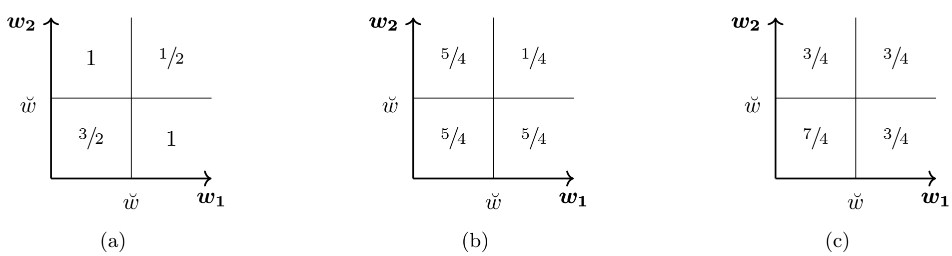


Figure 4: Illustrating Pareto weights α^m that are (a) separable, (b) submodular (b), and (c) supermodular.

If $\beta(w_i|w_{-i})$ is a non-trivial function of w_{-i}, then the planner's valuation of a dollar in the hands of a married person depends on a productivity of that person's spouse. In particular, if β is submodular, then the planner has a higher value for that dollar when it comes from a spouse who is different in terms of their productivity; if β is supermodular the planner has a lower value for it. Sub- and super-modular weights capture planner's preference for hetero- and homophily, respectively. If β is submodular or supermodular, so is α^m.

It would be helpful to visualize the implications of such weights using a simple example. Consider an economy with random matching and suppose first that $\beta(w_i|w_{-i})$ is independent of w_{-i} and takes the form $\beta(w_i) = \frac{3}{2}$ if w_i is below the median of G denoted by $\breve{w}$, $\beta(w_i) = \frac{1}{2}$ if w_i is above $\breve{w}$. The implied weights α^m are shown in Panel (a) of Figure 4. Panel (b) and

(c) construct sub- and super-modular weights from this β that preserve the property that the expected weight of a married person with any productivity w is the same as in Panel (a). It is easy to see visually that supermodularity "moves" Pareto weights closer to the 45 degree line, whereas submodularity moves them away from that line. Comparing this figure to Figure 3, which describes the economics of tax jointness, it is easy to see that submodularity amplifies benefits of targeting taxes to the richest couples (Reform I in Figure 3) while supermodularity amplifies benefits of targeting transfers to the poorest couples (Reform II in Figure 3).

Recall that in the random matching economy with separable weights the costs and benefits of the two reforms exactly cancel out, resulting in the disjointed optimal taxes (Lemma 4). Submodularity tilts the balance in favor of Reform I, while supermodularity tilts the balance in favor of Reform II. Thus, under random matching, optimal jointness is positive (negative) if α^m is supermodular (submodular). Even a small amount of correlation can break this result. In particular, observe that our discussion in Section 3.2.2 of optimal jointness under a Gaussian copula remains virtually unchanged, except $\alpha(\infty)$ is replaced by $\beta(\infty|\infty)$ in Equation (24). As the result, optimal jointess is still negative at the top and positive at the bottom for all $\rho > 0$. We summarize this result in the following corollary.

Corollary 2. *Suppose that $\alpha^m : \mathbb{R}_+^2 \to \mathbb{R}_+$ is symmetric and strictly decreasing in each argument.*

(a) Under random matching, $J(t) \leq 0$ for all t when α^m is supermodular, and $J(t) \geq 0$ for all t when α^m is submodular;

(b) If F is given by the Gaussian copula with $\rho > 0$, then $J(t) < 0$ for all t sufficiently large, and $J(t) > 0$ for all t sufficiently small.

Part (a) of this corollary is closely related to the findings in the working paper by Kleven et al. (2007). Those authors study jointness in the economy in which all households are married, their productivities are independent (i.e., people are matched randomly) and social welfare is given by $\int W(v^m)\, dF$, where W is a strictly concave social welfare function. They show that optimal jointess is negative (positive) when W''' is strictly positive (negative). One can show that there is one-to-one relationship between the sign of W''' and the modularity of the implied weights, and so part (a) of Corollary 2 is a version of their result. We provide details in the appendix.

5.2 Public goods and economies of scale

Our benchmark economy presents a somewhat idealized view of marriage. The decision to form couples in the absence of taxation is driven purely by a non-pecuniary shock ε, aka "falling in love", and married individuals commit to share their resources equally, "in sickness and in health", irrespective of their subsequent productivities and earnings. A sizable part of the family economics literature argues that this view of marriages is oversimplified. There are many economic incentives to form couples. For example, married individuals can attain economies of scale in consumption and allocate their labor supplies more efficiently between home and market activities. A division of their marital surplus can also be affected by various economic forces. Spouses with better outside-of-marriage economic prospects often attain a more favorable division of resources.

In the next several sections, we explore the implications of these mechanisms for optimal taxation. We start with one of the most obvious economic rationales for forming couples. Many goods and services consumed by households, from housing to child-rearing to Netflix subscriptions, have a public goods aspect as they provide non-rival benefits to both spouses in the household. This feature has long been recognized by the empirical labor literature from an early work by Lazear and Michael (1980) to the modern state-of-the-art estimations by Browning et al. (2013). In this section, we incorporate public goods into our model along the lines of Chapter 2.1 in Browning et al. (2014). We assume that each person's utility is $\phi(c^{pr}, c^{pub}) - \frac{1}{\gamma} l^{\gamma}$, where ϕ is a strictly increasing function with constant returns to scale, and c^{pr} and c^{pub} are consumption levels of private and within-household public goods, respectively. Without loss of generality, we set the prices of both goods to be equal to one. Single and married households solve

$$\max_{c^{pr}, c^{pub}, y} \phi\left(c^{pr}, c^{pub}\right) - \frac{1}{\gamma}\left(\frac{y}{w}\right)^{\gamma} \text{ s.t. } c^{pr} + c^{pub} \leq y - T^{s}(y) \text{ and } y \geq 0,$$

and

$$\max_{c^{pub}, \{c_i^{pr}, y_i\}_{i=1}^{2}} \sum_{i=1}^{2}\left(\phi\left(c_i^{pr}, c^{pub}\right) - \frac{1}{\gamma}\left(\frac{y_i}{w_i}\right)^{\gamma}\right) \text{ s.t. } \sum_{i}^{2} c_i^{pr} + c^{pub} \leq \sum_{i=1}^{2} y_i - T^{m}(y_1, y_2) \text{ and } y_1, y_2 \geq 0,$$

respectively. The rest of the model is as in Section 2.

Consumption expenditures can be conveniently aggregated. These maximization problems can be written as

$$\max_{c,y} k^{s} c - \frac{1}{\gamma}\left(\frac{y}{w}\right)^{\gamma} \text{ s.t. } c \leq y - T^{s}(y) \text{ and } y \geq 0;$$

$$\max_{c, \{y_i\}_{i=1}^{2}} k^{m} c - \frac{1}{\gamma}\sum_{i=1}^{2}\left(\frac{y_i}{w_i}\right)^{\gamma} \text{ s.t. } c \leq \sum_{i=1}^{2} y_i - T^{m}(y_1, y_2) \text{ and } y_1, y_2 \geq 0,$$

where

$$k^{s} := \max_{c^{pr} + c^{pub} = 1} \phi(c^{pr}, c^{pub}), \qquad k^{m} := \max_{c^{pr} + c^{pub} = 1} \phi(c^{pr}, 2c^{pub}).$$

It is easy to see that $k^m > k^s$, which captures the efficiency gains in consumption that marriage offers.

The key observation from these equations is that this model is isomorphic to a model in which an individual with a productivity w has a Pareto weight $k^s\alpha(w)$ if single and $k^m\alpha(w)$ if married. Thus, the insights of Section 5.1 apply directly. The optimal distortions satisfy

$$\lambda^{s,*}(t) = \frac{1-\mathbb{E}[\alpha|w\geq t]}{\gamma\theta(t)}\left(1-\mu^*+\mu^*\frac{k^s}{k^m}\right), \tag{34}$$

$$\mathbb{E}\left[\lambda_i^{m,*}\middle|w_i=t\right] = \frac{1-\mathbb{E}[\alpha^m|w_i\geq t]}{\gamma\theta(t)}\left((1-\mu^*)\frac{k^m}{k^s}+\mu^*\right). \tag{35}$$

Public goods make married households more economically efficient but, as consumption of public goods scales with income, also increase inequality. As a result, the social planner, while being inherently indifferent about a person's marriage status, acts as if she assigns a higher weight to married households in the social objective.

5.3 Home production and division of labor within families

Household consumption includes not only goods and services purchased in the marketplace but also those produced at home. Economists, since the early work of Reid (1934) and the influential contributions of Becker (1973, 1981), have argued that home production and labor specialization provide important economic rationales for forming couples. In this section, we incorporate these mechanisms into our model.

Let d be consumption of the home good and x be the effort required to produce it. We assume that preferences of each individual are given by $c+\frac{1}{1-\sigma}d^{1-\sigma}-\gamma\left(l^p+x^p\right)^{1/(\gamma p)}$. For single households, the production technology for the home good is $D^s(x)=x$, and their maximization problem is

$$\max_{c,y,x}\frac{x^{1-\sigma}}{1-\sigma}+c-\gamma\left(\left(\frac{y}{w}\right)^p+x^p\right)^{1/(\gamma p)} \text{ s.t. } c\leq y-T^s(y) \text{ and } y,x\geq 0.$$

For married households, the home production technology is $D^m(\mathbf{x})=\left(x_1^{1/q}+x_2^{1/q}\right)^q$, and their maximization problem is

$$\max_{c,\{y_i,x_i\}_{i=1}^2} 2\frac{\left(x_1^{1/q}+x_2^{1/q}\right)^{q(1-\sigma)}}{1-\sigma}+c-\sum_{i=1}^2\gamma\left(\left(\frac{y_i}{w_i}\right)^p+x_i^p\right)^{1/(\gamma p)} \text{ s.t. } c\leq\sum_{i=1}^2 y_i-T^m(\mathbf{y})$$

and $y_1,x_1,y_2,x_2\geq 0$. The parameter γ captures the elasticity of total (i.e., at home and at the market) labor supply, while the parameter $p\in(1,1/\gamma)$ captures the elasticity of substitution between hours at home and at work. The parameter $\sigma\in[0,1)$ captures the degree of returns in utility obtained from consumption of the home produced good. Finally, the parameter $q \in [1, 1/(1-\sigma)]$ allows home produced goods by the two married spouses to be imperfect substitutes. Our restrictions on these parameters ensure that all choices are interior, and hence we can abstract from corner solutions.

We analyze this problem in two steps. First, we fix market labor supply of single and married households, and characterize their optimal production and consumption of home goods. The optimal allocation of resources for home goods by single and married households, for any given levels of market labor supplies l and (l_1, l_2), solve the following two problems:

$$
\begin{aligned}
N^s(l) &:= \min_x -\frac{x^{1-\sigma}}{1-\sigma} + \gamma\left(l^p + x^p\right)^{1/(\gamma p)} \text{ s.t. } x \geq 0, \\
N^m(l_1, l_2) &:= \min_{\{x_i\}_{i=1}^2} -2\frac{\left(\sum_{i=1}^2 x_i^{1/q}\right)^{q(1-\sigma)}}{1-\sigma} + \gamma \sum_{i=1}^2 \left(l_i^p + x_i^p\right)^{1/(\gamma p)} \text{ s.t. } x_1, x_2 \geq 0.
\end{aligned}
$$

Second, we use N^s and N^m to characterize the optimal market labor supply:

$$
\max_{c,y} c - N^s\left(\frac{y}{w}\right) \text{ s.t. } c \leq y - T^s(y) \text{ and } y \geq 0,
$$

$$
\max_{c,\{y_i\}_{i=1}^2} \sum_{i=1}^2 c_i - N^m\left(\frac{y_1}{w_1}, \frac{y_2}{w_2}\right) \text{ s.t. } \sum_{i=1}^2 c_i \leq \sum_{i=1}^2 y_i - T^m(y_1, y_2) \text{ and } y_1, y_2 \geq 0.
$$

Note that this problem is isomorphic to the problem without home production but in which the disutility of market labor supply for single and married is given by the reduced-form functions N^s and N^m, respectively. These reduced-form functions introduce two changes relative to our benchmark specification: the elasticity of labor supply is no longer constant, and it depends both on a person's labor supply and, in the case of married individuals, on the labor supply of their spouse.

Despite these difference, the analysis of this economy is very similar to Section 3. Before stating the formulas for the optimal distortions that arise in this economy, it would be useful to define several economically meaningful objects. Let $\widetilde{e}^s$ be the elasticity of the market labor supply of a single individual, and $\widetilde{e}^m$ be the 2×2 cross-elasticity matrix of market labor supplies of married individuals.[10] Let $\widetilde{\gamma}^s$ and $\widetilde{\gamma}^m$ be their transformations defined as

$$
\widetilde{\gamma}^s := \left(1 + \left(\widetilde{e}^s\right)^{-1}\right)^{-1}, \quad \widetilde{\gamma}^m := \left(1 + \left(\widetilde{e}^m\right)^{-1}\right)^{-1}.
$$

Functions $\widetilde{\gamma}^s$ and $\widetilde{\gamma}^m$, just like elasticities $\widetilde{e}^s$ and $\widetilde{e}^m$, may in general depend on market labor supplies. In the benchmark economy, they were constant and given by $\widetilde{\gamma}^s = \gamma$ and $\widetilde{\gamma}^m = \gamma I$,

[10] The elasticity $\widetilde{e}^s$ can be written as $\widetilde{e}^s := \frac{\partial \ln l}{\partial \ln(\partial N^s/\partial l)}$. The matrix $\widetilde{e}^m$ is defined analogously, see the appendix for details.

where I is the 2×2 identity matrix. We use the notation $\widetilde{\gamma}^{s,*}(w)$ and $\widetilde{\gamma}^{m,*}(\mathbf{w})$ to denote values of these elasticity measures at the optimum, for individuals and couples with productivities w and $\mathbf{w}$, respectively. Let $\widetilde{\gamma}^{m,*}_{i,j}$ denote the ij^{th}-element of the matrix $\widetilde{\gamma}^{m,*}$.

With this notation in place, it is easy to derive the optimality conditions. In particular, we have

$$\widetilde{\gamma}^{s,*}(t)\lambda^{s,*}(t) = \frac{1 - \mathbb{E}\left[\alpha | w \geq t\right]}{\theta\left(t\right)}, \tag{36}$$

$$\mathbb{E}\left[\widetilde{\gamma}^{m,*}_{i,i}\lambda^{m,*}_{i} + \widetilde{\gamma}^{m,*}_{i,-i}\lambda^{m,*}_{-i}|w_i = t\right] = \frac{1 - \mathbb{E}\left[\alpha^m | w_i \geq t\right]}{\theta\left(t\right)}. \tag{37}$$

Equation (36) shows that the optimal tax for a single individual is still given by Diamond's ABC formula. Equation (37) reveals that the optimal tax formula for married households now captures a weighted sum of distortions of two spouses. The redistributive benefits of marginal taxes on married individuals are the same as in the benchmark economy, but the marginal tax on spouse i also affects the labor supply of their spouse. These labor supply responses are captured by $\widetilde{\gamma}^{m,*}_{i,i}$ and $\widetilde{\gamma}^{m,*}_{i,-i}$, respectively.

5.4 Bargaining and the allocation of resources within couples

In Section 2, we assumed that spouses share their marital surplus equally. A substantial empirical literature (see, e.g., Voena (2015) or the handbook chapter by Almas et al. (2023)) has documented that relative economic opportunities of spouses and their post-divorce outside options affect how resources are allocated within families. In this section, we study normative tax implications of these mechanisms.

We use Nash bargaining to model the allocation of resources within couples. We modify descriptions of Stages 2 and 3 of our model and assume that spouses bargain over surplus division after their productivities $\mathbf{w}$ are realized but before they supply labor to the market. When bargaining, each spouse uses the threat of divorce at a personal cost $\varrho > 0$. If divorce occurs, both spouses become single and do not remarry. Thus, the outside option of spouse i is $v^s(w_i) - \varrho$. The cost parameter ϱ is assumed to be sufficiently high so that it is not socially efficient for couples to get divorced. The couples' joint surplus v^m is given as in Section 2, and its allocation between spouses is obtained by solving

$$\max_{U_1^m, U_2^m} \left[U_1^m - (v^s(w_1) - \varrho)\right]^{1/2} \left[U_2^m - (v^s(w_2) - \varrho)\right]^{1/2} \text{ s.t. } U_1^m + U_2^m = v^m, \tag{38}$$

which pins down the pecuniary utility of spouse i to be

$$U^m(w_i | w_{-i}) = \frac{1}{2} v^m(\boldsymbol{w}) + \frac{v^s(w_i) - v^s(w_{-i})}{2}. \tag{39}$$

As can been from this equation, the primary earner obtains a larger share of the marital surplus because of a better outside option.

Proceeding as in Section 3, we can show that the optimal distortions for a single person are given by

$$\lambda^{s,*}(t) = \frac{1 - \mathbb{E}\left[\alpha|w \geq t\right]}{\gamma\theta\left(t\right)} + \frac{\mu^*}{1-\mu^*}\frac{\mathbb{E}\left[\xi|w \geq t\right]}{\gamma\theta\left(t\right)}, \tag{40}$$

where the function ξ is defined by $\xi\left(w\right) := \frac{1}{2}\mathbb{E}\left[\alpha(w_{-i}) - \alpha(w_i)|w_i = w\right]$. Under positive assortative matching, we have $\mathbb{E}\left[\xi|w \geq t\right] > 0$; thus, Equation (40) implies that the optimal marginal taxes on singles are higher under bargaining than under equal consumption sharing. The intuition for this result is as follows. Spouses use the threat of becoming single to secure their share of marital surplus. By making single households more equal via distortionary taxation, the planner compresses outside options of married spouses and endogenously facilitates consumption sharing within couples.

The optimal distortions for married persons $\boldsymbol{\lambda}^{m,*}$ are the same as in Section 3 and are characterized by Equation (16). This may appear surprising. After all, the optimal taxes for married individuals are lower than for single individuals if surplus is shared equally by spouses. Bargaining leads to unequal surplus division, so why is it that the optimal taxes are the same in these two cases? To answer this question, observe that according to Equation (39), while a more productive spouse receives a bigger share of marital surplus, they split any marginal change in v^m equally with their partner. The optimal labor distortions are pinned down by how couples share a marginal dollar, which is the same under equal consumption sharing and under Nash bargaining.

One might conjecture that if a more productive spouse has higher bargaining power, so that they obtain a larger fraction of the marginal dollar, then the optimal distortions for married individuals would be higher than under equal consumption sharing. This conjecture turns out to be false and the tax implications of such bargaining are more nuanced. To be concrete, consider the generalized bargaining solution in which the marital surplus is shared according to

$$\max_{U_1^m, U_2^m} \left[U_1^m - (v^s(w_1) - \varrho)\right]^{\eta(w_1|w_2)} \left[U_2^m - (v^s(w_2) - \varrho)\right]^{\eta(w_2|w_1)} \text{ s.t. } U_1^m + U_2^m = v^m.$$

Here, η captures the relative bargaining powers of spouses. We assume that $\eta(w_i|w_{-i}) \geq 0$, $\eta(w_1|w_2) + \eta(w_2|w_1) = 1$, and that $\eta(\cdot|w_{-i})$ is increasing. This implies that a more productive partner has larger bargaining power since $\eta(w_i|w_{-i}) \geq 1/2$ when $w_i \geq w_{-i}$. It is easy to see that, under generalized bargaining, the pecuniary utility of spouse i is

$$U^m(w_i|w_{-i}) = \eta(w_i|w_{-i})v^m(\boldsymbol{w}) + \left[\eta(w_{-i}|w_i)\left(v^s(w_i) - \varrho\right) - \eta(w_i|w_{-i})\left(v^s(w_{-i}) - \varrho\right)\right], \tag{41}$$

so that a more productive spouse receives a higher share of both the total family surplus and of each marginal dollar.

Using (41) and the same arguments as in Section 3, we can show that the optimal distortion for single persons $\lambda^{s,*}$ takes a form similar to (40). In contrast, the average optimal distortion for married persons satisfies

$$\mathbb{E}\left[\lambda_i^{m,*}|w_i = t\right] = \frac{1 - \mathbb{E}\left[\frac{\overline{\alpha}^m}{\mathbb{E}\overline{\alpha}^m}|w \geq t\right]}{\gamma\theta\left(t\right)} \times \mathbb{E}\overline{\alpha}^m, \tag{42}$$

where $\overline{\alpha}^m(\boldsymbol{w}) := \eta\left(w_1|w_2\right)\alpha\left(w_1\right) + \eta\left(w_2|w_1\right)\alpha\left(w_2\right)$. There is a close parallel between this equation and our discussion of optimal distortions under arbitrary Pareto weights in Section 5.1. As in that section, the average optimal distortion is a product of two terms. The first term captures the planner's preference for redistribution within married households. This preference is captured by the quasi-weight $\overline{\alpha}^m$ that takes into account how marginal dollars are allocated by spouses through bargaining. This term generally calls for higher optimal distortions than under equal consumption sharing. The second term in Equation (42), $\mathbb{E}\overline{\alpha}^m$, captures how generalized bargaining affects the planner's attitude towards marriage. Under positive assortative matching, this term is less than one so the planner endogenously cares less about married households. As we discussed in Section 5.1, this effect calls for lower marginal taxes on married households. Thus, the net impact of generalized bargaining on optimal distortions for married households is ambiguous.

To understand the intuition for why marriage is less desirable under generalized bargaining, consider a thought experiment of giving an additional \$2 to all couples. Under both equal consumption sharing and classical Nash bargaining, spouses share this transfer equally so that consumption of every married individual increases by \$1. In contrast, under generalized bargaining a more productive spouse gets a larger fraction of this transfer and so the uniform \$2 transfer to married couples gets endogenously allocated to more productive persons, who have a lower weight in the planner's objective. This makes the planner want to shift resources from married to single households.

To see concrete implications of (42), suppose that $\mathbb{E}\left[\eta(w_i|w_{-i})|w_i = t\right] \to 1$ as $t \to \infty$, which means that very productive primary earners grab the whole marginal dollar. Taking the limit of (42) as $t \to \infty$, we obtain

$$\lim_{t\to\infty} \mathbb{E}\left[\lambda_i^{m,*}|w_i = t\right] = \frac{1}{\gamma \lim_{t\to\infty} \theta\left(t\right)} \times \left(\mathbb{E}\overline{\alpha}^m - \alpha(\infty)\right).$$

As shown in the previous section, under equal consumption sharing, the optimal average tax rate satisfies

$$\lim_{t\to\infty}\mathbb{E}\left[\lambda_i^{m,*}|w_i=t\right]=\frac{1}{\gamma\lim_{t\to\infty}\theta\left(t\right)}\times\left(1-\frac{1}{2}\alpha(\infty)-\lim_{t\to\infty}\frac{1}{2}\mathbb{E}[\alpha(w_{-i})|w_i=t]\right).$$

In general, there is no clear ranking of these limiting distortions. For example, if the joint distribution is given by the Gaussian copula with $\rho>0$, then $\lim_{t\to\infty}\frac{1}{2}\mathbb{E}[\alpha(w_{-i})|w_i=t]=\alpha(\infty)$, and the optimal distortions are smaller under generalized bargaining due to $\mathbb{E}\overline{\alpha}^m<1$. The opposite may be true for some Pareto weights under alternative assumptions about the distribution of types.

Finally, it is worthwhile to remark that generalized bargaining often endogenously imposes supermodularity on the planner's objective function. For example, if primary earners always receive the whole marginal dollar, then $\overline{\alpha}^m(\boldsymbol{w})=\min\{\alpha(w_1),\alpha(w_2)\}$, which is a supermodular function. So, the insights of Corollary 2 apply directly to optimal taxation with differential bargaining powers.

5.5 Extensive margin of labor supply

In our benchmark economy, all labor supply adjustments are done along the intensive margin. We now discuss the implications of adding the extensive margin.

We model the extensive margin along the lines of Jacquet et al. (2013). Each individual has an idiosyncratic fixed disutility cost κ to participate in the labor market, so that the utility of each person is given by $c-\gamma l^{1/\gamma}-e\kappa$, where $e\in\{0,1\}$ is the labor market participation decision. We assume that κ is distributed independently from other variables according to some distribution H supported on $[\underline{\kappa},\overline{\kappa}]$, where $0\le\underline{\kappa}\le\overline{\kappa}<\infty$, that admits a continuously differentiable density h if $\underline{\kappa}<\overline{\kappa}$. Following Jacquet et al. (2013), we assume that e is observable and taxes may directly depend on it. Each person learns their realization of shock κ in Stage 3 before deciding on their labor supply. The rest of the assumptions are as in Section 2.

The analysis of the model with the extensive margin requires several adjustments. We first describe them for single individuals. Our tax system consists of a uniform lump-sum transfer b^s to all single persons, and a tax schedule $T^s(y)$ for working singles, i.e., those with $e=1$. Let v^s be the utility of a single working person defined analogously to the definition of v^s in Section 2. The local incentive constraints, that capture the intensive margin labor response, are given by the same envelope condition (8). Taking into account the decision whether to participate in the labor market, the pecuniary utility of a single person can be written as

$$\max\{v^s(w)-\kappa,0\}+b^s.$$

Clearly, a single person with productivity w participates in the labor market if κ is below the threshold value of $v^s(w)$ that happens with probability $H(v^s(w))$. The necessary conditions for optimality in our model with extensive margin can be written as

$$\frac{\partial}{\partial w}\left(\lambda^{s,*}\gamma w H(v^{s,*})g\right) = (\alpha - 1)\, H(v^{s,*})g + T^{s,*}h(v^{s,*})g, \tag{43}$$

where $T^{s,*}$ are the taxes paid by a working single person. These taxes are determined by $v^{s,*}$ via $T^{s,*}(y^{s,*}(w)) = w^{1+\gamma}\left(\frac{\partial v^{s,*}(w)}{\partial w}\right)^{\gamma} - \gamma w \frac{\partial v^{s,*}(w)}{\partial w} - v^{s,*}(w)$.

Comparing this equation with (11), we see that the extensive margin introduces two changes to the optimality condition. First, the density of types $g(w)$ is multiplied by $H^s(w) := H(v^{s,*}(w))$, which is the density of individuals who participate in the labor market. Second, there is an additional term, $T^{s,*}h^s g$, that captures the extensive margin response.

To state the expression for the optimal distortions, let $g^s := H^s g$ be the density of productivities of single agents who participate in the labor market, $\mathbb{E}^s$ be the expectation with respect to this distribution, θ^s be its tail statistics defined analogously to (13). Integrating (43) we obtain[11]

$$\lambda^{s,*}(t) = \frac{1 - \mathbb{E}^s[\alpha | w \geq t]}{\gamma \theta^s(t)} - \frac{\mathbb{E}^s\left[T^{s,*}\frac{h^s}{H^s} | w \geq t\right]}{\gamma \theta^s(t)}. \tag{44}$$

This equation is a version of optimal tax formulas derived by Jacquet et al. (2013), who studied optimal uni-dimensional taxation in models with both intensive and extensive margins. The first term on the right-hand side of (44) has the same interpretation as Equation (12). The second term on the right-hand side of (44) captures the behavioral costs of labor force participation decision. To understand the intuition for this term, observe that higher marginal taxes on single persons with productivity $w = t$ increase average taxes for individuals with productivity $w > t$. They respond along the extensive margin, and their extensive margin elasticity is captured by $\frac{h^s}{H^s}$. Thus, $\mathbb{E}^s\left[T^{s,*}\frac{h^s}{H^s} | w \geq t\right]$ measures tax revenues lost by the government due to this behavioral response.

The problem of married households is more difficult. Each couple has four labor market participation choices: (i) both spouses work, (ii) and (iii) only one of the spouses works, and (iv) both exit the labor force. Let $v^m(\mathbf{w})$ be the marital surplus of couples with productivities $\mathbf{w}$ when both spouses work and $\widetilde{v}^m(w_i)$ be their surplus when only spouse i works. The labor force participation decision of a couple with productivities (w_1, w_2) and participation costs

[11] In writing this equation, we implicitly assumed that $\underline{\kappa} = 0$ so that there are persons of all productivities who participate in the labor market. When $\underline{\kappa} > 0$, there is a productivity cut-off so that persons with productivity below that cut-off do not participate in the market. Equation (44) applies to all productivities above that cut-off. The same is true for (45).

(κ_1, κ_2) can be written as

$$\max\{v^m(w_1, w_2) - \kappa_1 - \kappa_2, \widetilde{v}^m(w_1) - \kappa_1, \widetilde{v}^m(w_2) - \kappa_2, 0\} + b^s.$$

Thus, the optimal tax problem involves choosing two functions for couples, $v^{m,*}$ and $\widetilde{v}^{m,*}$, corresponding to cases when both spouses work and only one spouse works. The optimality conditions for those functions, which we state in the appendix, are inter-related, as, for example, a perturbation of $\widetilde{v}^m$ affects participation decisions of both one- and two-earner couples.

To streamline our exposition, we focus on the random matching economy. In this economy, one can guess that the relationship between $v^{m,*}$ and $\widetilde{v}^{m,*}$ is given by $v^{m,*}(w_1, w_2) = \widetilde{v}^{m,*}(w_1) + \widetilde{v}^{m,*}(w_2)$, solve explicitly for $\widetilde{v}^{m,*}$, and verify that this function is consistent with the remaining optimality conditions. This approach implies that the tax schedule on married household is separable, $T^{m,*}(y_1, y_2) = \widetilde{T}^{m,*}(y_1) + \widetilde{T}^{m,*}(y_2)$, and the optimal distortions satisfy

$$\lambda^{m,*}(t) = \frac{1}{2}\frac{1 - \mathbb{E}^m[\alpha | w \geq t]}{\gamma\theta^m(t)} - \frac{\mathbb{E}^m\left[\widetilde{T}^{m,*}\frac{h^m}{H^m} | w \geq t\right]}{\gamma\theta^m(t)}, \tag{45}$$

where θ^m, $\mathbb{E}^m$, H^m are defined analogously to θ^s, $\mathbb{E}^s$, H^s for single agents (see the appendix for the details).

Equation (45) combines the economic mechanisms that we discussed in Section 3 and in Equation (44). Consumption sharing implies that the redistributory benefits of taxation are cut in half for married persons, which is the reason for why $\frac{1}{2}$ appears in the front of the first term on the right hand side of (44). At the same time, an increase in average taxes on persons with productivity $w > t$ triggers their extensive margin response, which is captured by the second term in Equation (44).

While Equations (44) and (45) provide a lot of insights about optimal taxation of single and married individuals, they have two limitations. First, both distributions, H^s and H^m, and tax functions, $T^{s,*}$ and $\widetilde{T}^{m,*}$, that appear on their right-hand sides are endogenous objects. This makes it difficult to use these equations to derive sharp comparisons of taxes for single and married or to conduct comparative statics analysis. Second, the mechanism design problem with the extensive margin is no longer convex. While the guess and verify technique ensures that solution (44) satisfies the necessary conditions for optimality, it cannot rule out a possibility that there is another, more complicated solution that satisfies the same optimality conditions and yields higher welfare. We address both of these limitations by considering the special case of our economy in which all persons have the same cost of labor market participation.

Lemma 7. *Consider the model with random matching and* $\underline{\kappa} = \overline{\kappa} > 0$.

(a) The optimal tax for married couples is separable, and marginal taxes for single and married individuals coincide with those derived in Section 4.

(b). There are productivity cut-offs $\underline{w}^s, \underline{w}^m$ *for single and married, with* $\underline{w}^s > \underline{w}^m$*, so that single and married individuals participate in the labor market if and only if their productivity is above those respective cut-offs.*

(c). The optimal marriage rate μ^* *and the marriage subsidy* $\mathbb{E}U^{m,*} - \mathbb{E}U^{s,*}$ *are decreasing in* $\overline{\kappa}$ *and converge to* μ^{LF} *and 0, respectively, as* $\overline{\kappa} \to \infty$*.*

The hardest part of proving this lemma lies in verifying that our candidate solution $v^{m,*}$ is globally optimal. In order to do it, we consider an extended mechanism design problem, in which persons choose probabilities of labor force participation and the mechanism designer can choose allocations that depend on those probabilities. That extended mechanism design problem is convex and, using the guess and verify technique, we can characterize its solution explicitly. We then verify that this solution also satisfies the optimality conditions of the more restrictive original mechanism design problem, which ensures that we found the global optimum.

The economy described in Lemma 7 has a simple yet very insightful structure. If labor market participation costs are the same for all individuals, then they participate in the labor market if and only if their productivity is above certain cut-offs. The individuals who are exactly at their cut-offs respond to tax changes along the extensive margin, the individuals above those cut-offs respond along the intensive margin. Lemma 7 gives several additional insights about optimal taxation of single and married individuals with extensive margin responses. Part (b) shows that the productivity cut-off in the optimum is higher for single persons so that they are less likely to participate in the labor market than married. The intuition for this result is that single individuals face higher optimal taxes, which increases their incentives to drop out of the labor force. Part (c) shows that the marriage rate and the marriage subsidy decrease in the participation costs. Larger participation costs mean that fewer married individuals work, resulting in less surplus-sharing within couples. This decreases redistributory benefits of marriage and reduces the size of the optimal marriage subsidy.

5.6 Selection into marriage

In our economy described in Section 2, the probability of getting married was the same for all persons. However, in the data, marriage rates are correlated with various indicators of a person's socio-economic status and their earnings. In this section, we extend our model to incorporate such heterogeneous selection into marriage.

The simplest way to introduce the correlation of marriage rates and productivities is to switch the order of Stages 1 and 2, i.e., to assume that individuals first observe signals q about their individual productivities and then decide whether to enter the marriage market. The rest of the setup is as in Section 2. To highlight the key differences that this setup introduces, we assume that there are two possible signals, $q = h$ and $q = l$, that occur with equal probabilities. Let H_q be the distribution of productivities of individuals who receive a signal q, and μ_q be their marriage rate. The average marriage rate μ in the economy is given by $\mu = \frac{1}{2}(\mu_h + \mu_l)$.

There are several parallels with our discussion in the previous section. First of all, taxes affect skill composition of single and married persons. Let G^s and F^m denote the skill distributions of single and married persons in the optimum. Let $\mathbb{E}^s$ and $\mathbb{E}^m$ be expectations under these distributions, and θ^s and θ^m be their tail statistics. Setting up the mechanism design problem, one can show that the optimal distortions satisfy

$$\begin{aligned} \lambda^{s,*}(t) &= \frac{1 - \mathbb{E}^s[\alpha|w \geq t]}{\gamma\theta^s(t)} + \frac{1}{1-\mu^*}\frac{\delta_h(1 - H_h(t)) + \delta_l(1 - H_l(t))}{\gamma\theta^s(t)}, \\ \mathbb{E}[\lambda_i^{m,*}|w_i = t] &= \frac{1 - \mathbb{E}^m[\alpha^m|w_i \geq t]}{\gamma\theta^m(t)} - \frac{1}{\mu^*}\frac{\delta_h(1 - H_h(t)) + \delta_l(1 - H_l(t))}{\gamma\theta^m(t)}, \end{aligned}$$

where δ_h and δ_l are the Lagrange multipliers on the two analogs of Equation (5), for q_h and q_l, respectively. Furthemore, it can be shown that these multipliers must satisfy $\mu^*\delta_h + (1-\mu^*)\delta_l = 0$.

Examination of these equations reveals that both optimal tax formulas consist of two terms. The first terms on the right-hand sides are the same as in Equations (12) and (18), except now they take into account that the tail statistics θ^s and θ^m summarizing the marginal distributions of productivities may differ across single and married individuals. The second terms on the right-hand side of these equations capture the additional effect from behavioral responses to changes in pecuniary benefits from singlehood and marriage on sorting. The planner recognizes that taxes affect sorting into marriage, influencing the distributions of productivities among single and married individuals, and takes these behavioral responses into account when choosing tax rates. These terms take opposite signs for single and married persons, and their economic implications are similar to those of exogenously different Pareto weights in Section 5.1.

5.7 Optimality of taxation of family earnings

Several countries, e.g., the U.S., tax married couples based on their total earnings. Under such *family-earnings-based taxation*, a tax function T^m can be written as $T^m(y_1, y_2) = \widetilde{T}^{fam}(y_1 + y_2)$. In this section, we explore conditions under which such taxation is optimal.

Let $Y = y_1 + y_2$. Under family-earning-based taxation, both spouses face identical marginal tax rates, and their labor optimality conditions are

$$y_i = w_i^{1/(1-\gamma)} \left(1 - \frac{\partial}{\partial Y} \widetilde{T}^{fam}(Y)\right)^{\gamma/(1-\gamma)} \text{ for } i = 1, 2.$$

This equation shows that total earnings of a couple $\mathbf{w}$ under such tax system are proportional to $w_1^{1/(1-\gamma)} + w_2^{1/(1-\gamma)}$. Therefore, finding conditions for optimality of family-earnings based taxation is equivalent to characterizing conditions under which the optimal allocations for each couple depend only on a uni-dimensional summary statistics given by

$$R(\mathbf{w}) := \left(w_1^{1/(1-\gamma)} + w_2^{1/(1-\gamma)}\right)^{1-\gamma},$$

which can be interpreted as the measure of the average family productivity. Recall the function $I(\mathbf{w}) = \frac{\min \boldsymbol{w}}{\max \boldsymbol{w}}$ that we defined in Section 3.2. Taken together, R and I conveniently redefine the coordinate system $\mathbf{w} \leftrightarrow (r, \iota)$, where $r = R(\mathbf{w})$ is the average family productivity and $\iota = I(\mathbf{w})$ is the relative spousal productivity. Family-earnings-based taxation is optimal if the optimal distortions depend only on r and not on ι.

Set $Q = R$ in Theorem 1 to obtain that the optimal distortions satisfy

$$\mathbb{E}\left[\sum_{i=1}^{2} \frac{w_i^{1/(1-\gamma)}}{w_1^{1/(1-\gamma)} + w_2^{1/(1-\gamma)}} \lambda_i^{m,*} \middle| R = r\right] = \frac{1 - \mathbb{E}\left[\alpha^m | R \geq r\right]}{\gamma \theta_r(r)}, \tag{46}$$

where $\theta_r(r) := \frac{-d \ln \Pr(R \geq r)}{d \ln r}$. If family-earnings-based taxation is optimal, then $\lambda_1^{m,*} = \lambda_2^{m,*}$ and this equation simplifies to

$$\lambda_1^{m,*}(\mathbf{w}) = \lambda_2^{m,*}(\mathbf{w}) = \frac{1 - \mathbb{E}[\alpha^m | R \geq r]}{\gamma \theta_r(r)} \text{ for all } \mathbf{w} \text{ s.t. } R(\mathbf{w}) = r.$$

Conversely, family-earnings-based taxation is optimal if these $(\lambda_1^{m,*}, \lambda_2^{m,*})$ satisfy (14) and (15). The cross-partial condition (15) is satisfied automatically since both spouses face the same marginal tax rates. The next lemma describe conditions under which (14) holds for some symmetric α^m with $\mathbb{E}\alpha^m = 1$ and the distribution of (r, ι) as primitives of the model.

Lemma 8. *The optimal tax for couples is family-earnings-based if and only if*

$$\frac{1 - \mathbb{E}\left[\alpha^m | R \geq r, I = \iota\right]}{\frac{-d \ln \Pr(R \geq r | I = \iota)}{d \ln r}}$$

is independent of ι *for all* r. *In particular, if Pareto weights for couples* α^m *take the form* $\alpha^m(\mathbf{w}) = \widetilde{\alpha}^m(R(\mathbf{w}))$, *then such taxation is optimal if and only if the distribution* F *is such that* r *and* ι *are independent.*

The idea that families with the same total income should be treated and valued identically by the planner is known as "horizontal equity".[12] One interpretation of horizontal equity in our context is that taxes are family-earnings-based and for each level of total earnings Y, the planner assigns the same implicit Pareto weight to all couples $\boldsymbol{w}$ with that level of total earnings, i.e., $y_1^m(\boldsymbol{w}) + y_2^m(\boldsymbol{w}) = Y$. Under family-earnings-based taxation, the latter property is captured by Pareto weights α^m that treat identically all couples with the same average productivity. According to Lemma 8, when the planner has an inherent preference for horizontal equity, the necessary and sufficient condition for the optimality of family-earnings-based taxation is independence of the average family productivity r from the relative spousal productivity ι. When r and ι are independent, there is no trade-off between pursuing horizontal equity and redistribution between rich and poor couples, which is also known in the literature as "vertical equity". However, if ι contains some information about r, it may be optimal to sacrifice some of horizontal equity in favor of more efficient redistribution.[13]

One can analyze implications of correlation between r and ι along the lines of our analysis in Section 3.2. For example, one can show that higher dependence of r and ι in the PQD sense implies that the secondary earner should, on average, face higher distortions.

5.8 Gender differences

The family economics literature has documented systematic differences in earnings between males and females. In this section, we incorporate this heterogeneity in our model. We assume that there are two fixed genders, $o = 1$ and $o = 2$, each of equal measure, and that married couples consist of one spouse of each gender. Persons 1 and 2 draw their productivities from distributions $\hat{G}_1$ and $\hat{G}_2$ with densities $\hat{g}_1$ and $\hat{g}_2$. The description of the matching process is as in Section 2 with one modification. Since individuals of two genders are ex-ante different, there is no reason to expect that the same number of males and females will arrive on the marriage market. To clear this market, we introduce rationing and return individuals of the "surplus" gender with highest values of ε back to the singlehood. We assume that social welfare is given by

$$\mathcal{W} := \frac{1}{2}\int \hat{\alpha}_1(w_1)\mathbb{E}\left[U_1|w_1\right] d\hat{G}_1 + \frac{1}{2}\int \hat{\alpha}_2(w_2)\mathbb{E}\left[U_2|w_2\right] d\hat{G}_2,$$

where weights $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are not necessarily the same, and set $\int \hat{\alpha}_o d\hat{G}_o = 1$ for $o = 1, 2$.

Let $\hat{F}$ be the joint distribution of productivites of married couples, where its two coordinates

[12] See, e.g., Liebman and Ramsey (2019).

[13] Other authors have criticized the U.S. family-earnings-based tax policy as well. For example, Borella et al. (2023) estimate a dynamic life-cycle model of labor supply and argue that the U.S. tax system is suboptimal because it discourages secondary earners (often women) from working.

correspond to productivities of persons of genders 1 and 2. $\hat{F}$ does not need to be symmetric, let $\hat{G}_1$ and $\hat{G}_2$ be its marginals with corresponding tail statistics $\hat{\theta}_1$ and $\hat{\theta}_2$. Let $\hat{\alpha}^m$ be defined by $\hat{\alpha}^m(w_1, w_2) := \frac{1}{2}\hat{\alpha}_1(w_1) + \frac{1}{2}\hat{\alpha}_2(w_2)$.

If gender is observable, the planner can, in principle, use taxes that depend both on the marital status and gender. We consider both gender-specific taxes, in which the planner uses gender information, and gender-neutral taxes, in which the planner cannot use this information. It turns out that there is a close connection between optimality conditions for these two tax systems.

We start by considering gender-specific taxation. Such taxes can be studied using essentially the same mechanism design problem as described in Section 3, except that now there are two asymmetric groups of agents, a and b. The optimal gender-specific distortions, $\{\hat{\lambda}_o^{s,*}, \hat{\lambda}_o^{m,*}\}_{o\in\{1,2\}}$, satisfy

$$\hat{\lambda}_o^{s,*}(t) = \frac{1 - \hat{\mathbb{E}}_o[\hat{\alpha}_o|w_o \geq t]}{\gamma\hat{\theta}_o(t)}, \quad \hat{\mathbb{E}}[\hat{\lambda}_o^{m,*}|w_o = t] = \frac{1 - \hat{\mathbb{E}}\left[\hat{\alpha}^m|w_o \geq t\right]}{\gamma\hat{\theta}_o(t)} \text{ for } o = 1, 2, \tag{47}$$

where $\hat{\mathbb{E}}_o$ and $\hat{\mathbb{E}}$ denote expectations with respect to the probability distributions $\hat{G}_o$ and $\hat{F}$, respectively. These equations are analogs of (12) and (18) in the benchmark economy.

The characterization of optimal gender-neutral taxes is even simpler as it reduces to a special case of the mechanism design problem that we considered in Section 5.1. Under gender-neutral taxes, any person with productivity w must receive the same utility irrespective of their gender. Consequently, to set taxes for single individuals, the planner can use information only about their unconditional (i.e., not gender-specific) distribution of productivities, $G = \frac{1}{2}(\hat{G}_1 + \hat{G}_2)$. Similarly, taxes for married individuals cannot be based on gender-specific information about the joint distribution of productivities. The joint distribution with "scrambled" gender information is $F(\mathbf{w}) = \frac{1}{2}(\hat{F}(w_1, w_2) + \hat{F}(w_2, w_1))$. Finally, since neither v^s nor v^m contain gender information, the Pareto weights that appear in the planner's objective function can be written as $\alpha = \frac{1}{2g}(\hat{\alpha}_1\hat{g}_1 + \hat{\alpha}_2\hat{g}_2)$ and $\alpha^m(\mathbf{w}) = \frac{1}{2f}(\hat{\alpha}^m(w_1, w_2)\hat{f}(w_1, w_2) + \hat{\alpha}^m(w_2, w_1)\hat{f}(w_2, w_1))$, where g and f are densities of G and F. Taking all pieces together, we can show that the optimal distortions under gender-neutral taxation are given by

$$\lambda^{s,*}(t) = \frac{1 - \mathbb{E}[\alpha|w_i \geq t]}{\gamma\theta(t)}, \quad \mathbb{E}[\lambda_i^{m,*}|w_i = t] = \frac{1 - \mathbb{E}\left[\alpha^m|w_i \geq t\right]}{\gamma\theta(t)}. \tag{48}$$

There is a close relationship between the optimal distortions under gender-specific and gender-neutral taxation. In particular, it is easy to verify that these optimal distortions described in (47) and (48) satisfy

$$\lambda^{s,*}(t) = \sum_{o=1,2} \omega_o(t)\hat{\lambda}_o^{s,*}(t), \quad \mathbb{E}[\lambda_i^{m,*}|w_i = t] = \sum_{o=1,2} \omega_o(t)\hat{\mathbb{E}}[\hat{\lambda}_o^{m,*}|w_o = t], \tag{49}$$

where $\omega_o(t) := \frac{g_o(t)}{g_1(t)+g_2(t)}$. Thus, the optimal distortions under gender-neutral taxation are equal to the weighted average of the optimal distortions under gender-specific taxation. The total "amount" of distortions is the same under both tax systems but gender-specific taxation allows the planner to improve welfare by allocating them more efficiently between genders.

5.9 Optimal restricted taxation

The close relationship between optimal unrestricted and restricted taxation uncovered by Equation (49) holds in other contexts as well. For example, suppose we exogenously restrict taxes for married households to take the form $T^m(y_1, y_2) = \widetilde{T}^m(y_1) + \widetilde{T}^m(y_2)$. It is easy to verify that under such taxes, the marital surplus v^m must be additively separable in w_1 and w_2, i.e., $v^m(w_1, w_2) = \widetilde{v}^m(w_1) + \widetilde{v}^m(w_2)$. Thus, studying separable taxation is equivalent to imposing an additional restriction on the mechanism design problem that v^m belongs to the class of additively separable functions. Since $\widetilde{v}^m$ is a uni-dimensional function, it is easy to analyze using standard uni-dimensional techniques. In particular, one can derive the optimal distortions from the optimal separable taxes, $\widetilde{\lambda}^{m,*}$, and show that they are equal to the average optimal distortions in the unrestricted tax system:

$$\widetilde{\lambda}^{m,*}(t) = \mathbb{E}[\lambda_i^{m,*} | w_i = t].$$

Similarly, suppose that we require taxes for married households to be family-earnings-based, $T^m(y_1, y_2) = \widetilde{T}^{fam}(y_1 + y_2)$. Under such system, utility of married couples must take the form $v^m(\boldsymbol{w}) = \widetilde{v}^m(R(\mathbf{w}))$, i.e., it depends only on the total family productivity. Since such $\widetilde{v}^m$ is uni-dimensional, it is also easy to characterize and derive the optimal distortions $\lambda^{m,fam,*}$ implied by this restricted tax system. Once again, these distortions are just the average of distortions under the unrestricted tax system, where averages are now taking across all couples $\mathbf{w}$ with a given total productivity r:

$$\lambda^{m,fam,*}(r) = \mathbb{E}\left[\sum_{i=1}^{2} \frac{w_i^{1/(1-\gamma)}}{w_1^{1/(1-\gamma)} + w_2^{1/(1-\gamma)}} \lambda_i^{m,*} \middle| R = r \right].$$

Observations in this and in the previous section show a remarkable close connection between optimal restricted and unrestricted taxation. In all three cases, the optimal distortions under restricted taxes are equal, on average, to the optimal distortions under the unrestricted tax system. Unrestricted and restricted optimal taxes are chosen to balance-off exactly the same redistributory benefits and distortionary costs. The unrestricted tax system allows the planner to allocate those benefits and costs more efficiently among different households.[14]

[14]See also Golosov et al. (2014) who make a similar observation in a different context.

6 Quantitative analysis

In this section, we illustrate the theoretical implications of our analysis using a quantitative model. We focus on our baseline model, which we view as a natural benchmark. We calibrate this model, compute the optimal taxes, and compare the properties of these taxes to the theoretical predictions derived using our analytical techniques. Additionally, we numerically check the validity of the FOA that was implicitly invoked in those derivations. Due to space constraints, we defer the comprehensive quantitative analysis of various extensions that we considered in Section 5 to future work.

Our baseline model has three types of primitives: the elasticity parameter γ, the distribution of productivities G and the joint distribution F, and the set of Pareto weights α. We set the elasticity parameter γ to 1/4, so that the implied labor supply elasticity $\frac{\gamma}{1-\gamma}$ is equal to 1/3, which is the mid-range of values considered by Diamond (1998).

We use observed taxes and the distribution of earnings to invert it to obtain the distributions G and F. Following Guner et al. (2014) and Heathcote et al. (2017), who argued that the U.S. tax schedule is such that family post-tax earnings are approximately a log-linear function of family pre-tax earnings, we assume that households in the data face taxes of the form $T(y_1, y_2) = (y_1 + y_2) - \nu (y_1 + y_2)^{1-\tau}$, where τ and ν are parameters. We refer to this functional form as the HSV tax schedule. Under this tax system, the relationship between an observed vector of earnings $\mathbf{y}$ and an unobserved vector of productivities $\mathbf{w}$ is given by

$$w_i^{1/\gamma} = \frac{1}{(1-\tau)\nu} y_i^{1/\gamma - 1} (y_1 + y_2)^{\tau}. \tag{50}$$

To invert this mapping, we use the parameter values of (τ, ν) that Guner et al. (2014) estimated for U.S. married couples. To obtain the joint distribution of earnings we construct the dataset from the 2020 CPS survey, restricting attention to couples in which both individuals are between 25 and 65 years old and worked at least 20 weeks in 2019. To ensure that two married spouses are statistically identical, we create a copy of each couple with permuted spousal earnings. This provides us with the empirical joint distribution of spousal earnings, which we then invert to obtain the empirical joint distribution of productivities. To isolate dependence properties from properties of its marginals, we decompose this empirical joint distribution into its empirical marginal G^e and copula C^e.[15] We calibrate these components separately.

We choose a parsimonious representation of the marginal distribution of $\mathbf{w}$. Consistent with earlier literature (e.g., Badel et al. (2020) or Golosov et al. (2016)), we find that the

[15] $G^e(w)$ corresponds to the sample fraction of individuals with a productivity less or equal to w and $C^e(u_1, u_2)$ is the sample fraction of couples in which a pair of spousal empirical percentiles are less or equal to (u_1, u_2).

empirical marginal distribution can be well approximated by G that belongs to the parametric family of Pareto lognormal (PLN) distributions.[16] We choose the three parameters (η, σ, a) of the PLN distribution to match the mean level of productivity, the Gini coefficient, and the tail parameter. These three moments can be expressed analytically in terms of (η, σ, a), allowing us to obtain these parameters by a simple inversion of the respective moment conditions (see appendix for the details). Panel (a) of Figure 5 shows the empirical and calibrated marginal distributions of productivities.

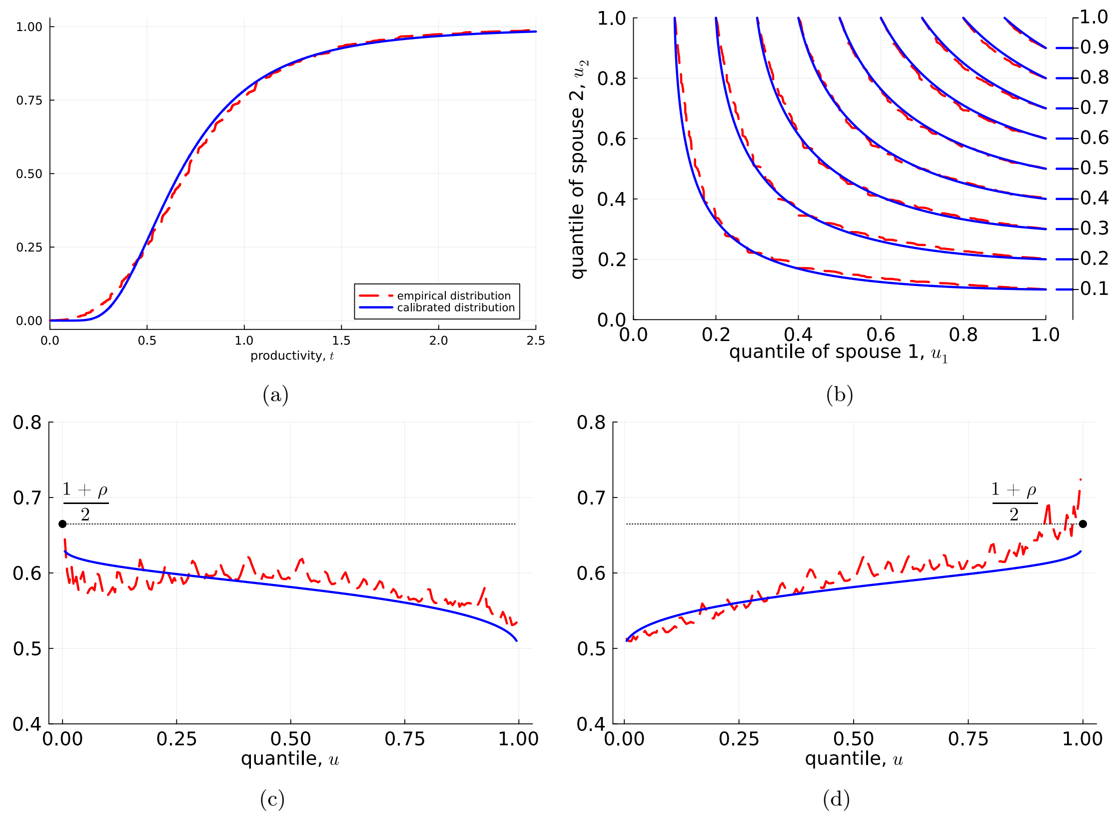


Figure 5: Empirical (dashed red lines) and calibrated (solid blue lines) distributions of productivities. Panel (a) plots their marginals; Panel (b) visualizes the isoquants of their copulas; and Panels (c) and (d) show the left-tail and right-tail statistics defined by $\frac{\ln u}{\ln C(u,u)}$ and $\frac{\ln(1-u)}{\ln(1+C(u,u)-2u)}$, respectively.

Recall that F can be uniquely identified with the bi-dimensional distribution C with uniform marginals called copula via $F(w_1, w_2) = C(G(w_1), G(w_2))$, e.g., see Nelsen (2006). We use this representation and calibrate the copula of F to the empirical copula C^e. We find that the Gaussian copula fits the data very well. It has only one parameter, ρ, that we set

[16]The PLN family was introduced in Colombi (1990) as a model of the income distribution, and since then, it has been used extensively in various studies. It is defined as $G(t) = \Phi\left(\frac{\ln t - \eta}{\sigma}\right) - t^{-a} \exp\left(a\eta + a^2\sigma^2/2\right) \Phi\left(\frac{\ln t - \eta - a\sigma^2}{\sigma}\right)$, where Φ is the standard normal distribution.

Parameter	Value	Definition	Target
γ	0.25	Measure of the labor supply elasticity	Elasticity of labor supply, 0.33
a	2.95	Pareto tail of the PLN distribution	Sample Pareto statistics at 99% of individual productivities, 2.95
η	-0.71	Location parameter of the PLN distribution	Sample mean individual productivity, 0.81
σ	0.40	Shape parameter of the PLN distribution	Sample Gini of individual productivities, 0.31
ρ	0.33	Correlation parameter of the Gaussian copula	Sample Kendell's tau of spousal productivities, 0.21

Table 1: Calibrated parameters

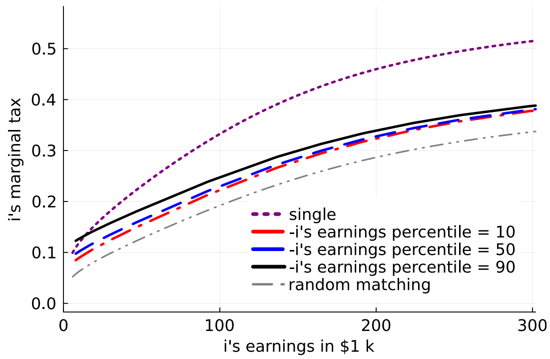
to match the Kendell's tau dependence coefficient[17] of the productivity distribution. This parsimonious specification matches well a number of other moments of as well. For example, in Panel (b) of Figure 5, we plot the isoquants of both the empirical copula and the calibrated Gaussian copula, showing a very close fit. The Gaussian copula also has a property that $\lim_{u\to 0}\frac{\ln u}{\ln C(u,u)} = \lim_{u\to 1}\frac{\ln(1-u)}{\ln(1+C(u,u)-2u)} = \frac{1+\rho}{2}$, the fact that we used in proving Lemma 5. This is consistent with the properties of the empirical copula as shown in Panels (c) and (d) of Figure 5. Table 1 summarizes all our parameters and their empirical counterparts.

We use the Pareto weights α that are given by $\alpha(w) = const \times e^{-mw^{1/(1-\gamma)}}$, where $const$ is chosen so that they integrate to one. We set the parameter m to 0.35 to match the average marginal tax rates of married persons in the data and in the optimum. This way, the total amount of redistribution is similar in our model and the data. To compute the optimal taxes, we first solve the relaxed problem and then verify that the solution satisfies global incentive constraints. In all cases that we considered we found that the FOA was valid.

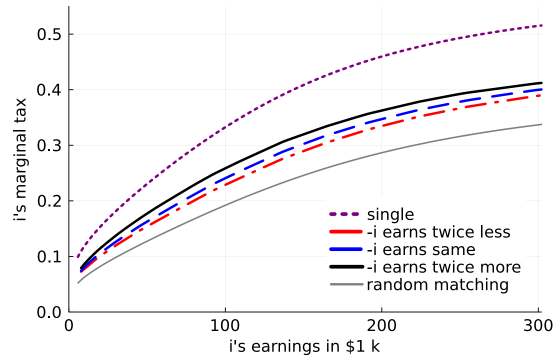
Figure 6 visualizes the optimal marginal taxes on single and married households and compares these schedules to the U.S. tax rates that implied by the estimated HSV function, which is reported in Guner et al. (2014).[18] We report $\frac{\partial}{\partial y_i}T^{m,*}(y_i, y_{-i})$ and their empirical analog, $\frac{\partial}{\partial y_i}T^{US}(y_i, y_{-i})$, in two ways: when earnings y_{-i} are fixed at different percentiles of earnings distribution, and when earnings y_{-i} are fixed multiples of y_i. The former way is more informa-

[17]The Kendell's tau is a measure of rank correlation of the joint distribution, thus it is independent of its marginals. The relationship between Kendell's tau and parameter ρ of the Gaussian copula is given by Kendell's tau $= 2\frac{\arcsin\rho}{\pi}$.

[18]We use the second column of Table 10 for married and the first column of Table 11 for single. These authors record earnings in multipliers of 53K, whereas we use multipliers of \$100K. To ensure that total tax liabilities are identical in dollars terms, we re-normalize their estimates of ν by $\left(\frac{53}{100}\right)^{\tau}$.

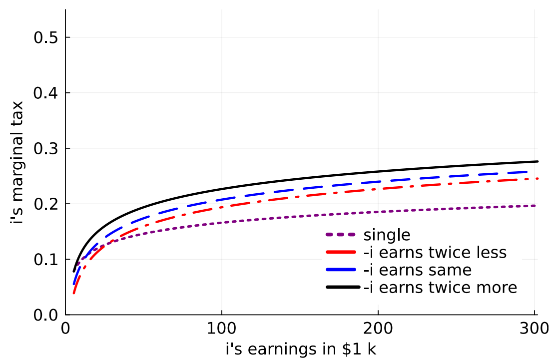


(a) Optimum



(b) Optimum

(c) U.S.



(d) U.S.

Figure 6: Optimal taxes (Panels (a) and (b)) and the U.S. taxes implied by the estimated HSV schedule (Panels (c) and (d)).

tive about jointness, the latter about the comparison of marginal taxes between the primary and the secondary earners. As a reference, we also report the optimal marginal taxes in the economy with random matching (in which case $\frac{\partial}{\partial y_i} T^{m,*}(y_i, y_{-i})$ is independent of y_{-i} as we showed in Lemma 4).

The optimal tax schedules share properties implied by our analytical formulas, both qualitatively and quantitatively. First of all, the optimal tax rates on married individuals under calibrated assortativity lie between the optimal tax rates on singles and the optimal tax rates on married individuals under random matching, consistent with our Lemma 2. One can also see from Panel (a) of Figure 6 that the optimal marginal taxes are positively jointed for low earners since the marginal tax $\frac{\partial}{\partial y_i} T^{m,*}(y_i, y_{-i})$ is increasing in y_{-i} for low values of y_i. This result is consistent with Lemma 5, which established optimality of positive jointness in the left tail. The same lemma also established that the optimal taxes must be negatively-jointed for high-earners.[19] This occurs at much higher earnings levels (>$8.5 mln) than the scale of the

[19] Spiritus et al. (2022) solved numerically a related optimal joint taxation problem and also found that optimal jointness is positive at the top and negative at the bottom.

x-axis we use. That being said, optimal jointness is very modest for all earnings levels, with marginal taxes for one spouse changing by, at most, several percentage points as a function earnings of the other spouse. Finally, Panel (b) of Figure 6 shows that the optimal marginal taxes on earnings y_i are decreasing in the value of the ratio y_i/y_{-i}. This is consistent with Lemma 6 in which we showed that secondary earners should face higher distortions if their productivity relative to that of primary earners is sufficiently low. Quantiatively, the same conclusion appears to hold more generally.

Comparison of the optimal and U.S. taxes reported in 6 reveals two striking differences between them. As can be seen from Panels (c) and (d), the U.S. tax schedule exhibits a substantial marriage penalty as the tax U.S. rates on singles are generally lower than those rates on married individuals with the same level of earnings. In contrast, our quantitative analysis suggests the planner should provide a marriage bonus setting lower tax rates on married persons.[20] Furthermore, this is consistent with our theoretical analysis in Section 3.2.1, which establishes that the optimality of lower taxes on married in a general non-parametric setting. Figure 6 also reveals that the optimal marginal taxes are steeper than the ones in the data. Qualitatively, this pattern is driven by two features of our calibration: the PLN marginal distribution of productivities, and the assumption that the Pareto weight of the richest individuals goes to zero. The PLN distribution has a thin left tail, $\lim_{t\to 0}\theta(t)=\infty$, which implies from (12) and (18) that optimal taxes should go to zero for the low earners. In contrast, the marginal taxes implied by the estimated HSV tax functions are positive for the low earners. The right tail of the PLN distribution is thick, with $\lim_{t\to\infty}\theta(t)=a<\infty$. This implies that the limiting average distortions in our calibrated economy, using formula (18), are

$$\lim_{t\to\infty}\mathbb{E}\left[\lambda_i^*\middle|w_i=t\right]=\frac{1}{\gamma a}\approx 1.35,$$

corresponding to the average tax rate on high-earner to be around 0.55. However, this limit is reached very slowly as the convergence rate of $\mathbb{E}\left[\alpha(w_{-i})|w_i\geq 0\right]$ to 0 is low. The slow speed of convergence explains why the optimal marginal tax rates in Figure 6 are substantially lower than this limit, even for individuals with earnings of \$300K.

An alternative way to represent the optimal tax schedule for married individuals is by the functional form $T^{m,*}(\mathbf{y})=T^{fam,*}\left(Y\left(\mathbf{y}\right),\iota\left(\mathbf{y}\right)\right)$, where $Y\left(\mathbf{y}\right)=y_1+y_2$ are the total family earnings and $\iota\left(\mathbf{y}\right)=\frac{\min\boldsymbol{y}}{Y(\mathbf{y})}$ is the share of the secondary earner. Family-earnings-based taxes are optimal if $T^{fam,*}$ does not depend on the second argument. In Figure 7, we plot $\frac{\partial}{\partial Y}T^{fam,*}\left(\cdot,\iota\right)$ for different values of ι. The purple dotted line shows the U.S. tax code

[20] This is aligned with the recent empirical findings in the family economics literature, e.g., Borella et al. (2023) argued that the U.S. may be suboptimal due to a positive marriage tax.

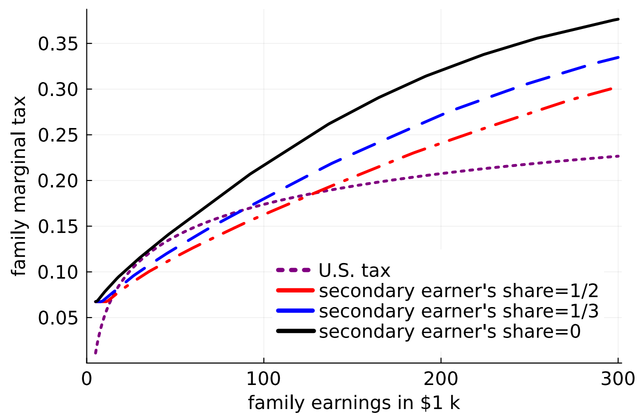


Figure 7: Marginal taxes on family earnings.

implied by the estimated HSV function. Since the U.S. tax schedule is family-earnings-based, $\frac{\partial}{\partial Y}T^{fam,US}(\cdot,\iota)$ is the same for all ι. Three other lines plot the optimal tax on family earnings for the same specifications that we used in Figures 6. The marginal tax rates vary substantially with the share of earnings of the secondary earner, with a higher share corresponding to a lower marginal family tax. In all cases, pure family earnings-based taxation is a poor approximation of the optimal tax code.

7 Conclusion

Multidimensional screening problems are ubiquitous in public finance applications. In this paper, we consider one of the simplest versions of such problems - optimal taxation of joint earnings of couples. We show that despite superficial similarity to multidimensional screening problems in industrial organization, our problem is much easier to analyze and can often be studied using the first-order approach.

We characterize the optimal taxes in these settings. Such taxes are a solution to a second-order partial differential equation, which is very complex and does not generally have an analytical solution. We show that this problem can be overcome by focusing on various conditional average moments of the optimal tax rates. These conditional moments provide significant insights into the economic mechanisms that drive the shape of the optimal tax schedule, both qualitatively and quantitatively.

In the calibrated economy, we find that the optimal taxes are negatively jointed at the bottom and positively at the top. However, this jointness is small, and the optimal taxes can be well approximated by separable, individual earnings-based taxation. In contrast, family earnings-based taxes provide a poor approximation to the optimal tax code, even when the planner's Pareto weights explicitly favor this type of taxation.

In the quantitative section, we only took first steps to connect our theoretical analysis with policy questions that arise in regards to taxation of families. We abstracted from many aspects of couple formation, labor supply decisions, and surplus division that we discussed in Section 5. Taxation of couples also has important implications for career choices of spouses and gender equality, fertility choices and child-rearing. It is impossible to do justice to these major issues within confines of one paper, and we leave a comprehensive quantitative investigation of these questions for future work.

References

Albanesi, Stefania, and Christopher Sleet. 2006. "Dynamic optimal taxation with private information." *Review of Economic Studies*, 73(1): 1–30.

Ales, Laurence, and Christopher Sleet. 2022. "Optimal Taxation of Income-Generating Choice." *Econometrica*, 90(5): 2397–2436.

Almas, Ingvild, Orazio Attanasio, and Pedro Carneiro. 2023. "Chapter 3 - Household decisions and intra-household distributions." In *Handbook of the Economics of the Family, Volume 1*. eds. by Shelly Lundberg, and Alessandra Voena, 1 of Handbook of the Economics of the Family: North-Holland, 111–149.

Alves, Cassiano Breno Machado, Carlos Eugenio da Costa, and Humberto Ataide Moreira. 2021. "Intrahousehold Inequality and the Joint Taxation of Household Earnings." FGV EPGE Economics Working Papers 825.

Armstrong, Mark. 1996. "Multiproduct Nonlinear Pricing." *Econometrica*, 64(1): 51–75.

Badel, Alejandro, Mark Huggett, and Wenlan Luo. 2020. "Taxing Top Earners: a Human Capital Perspective." *Economic Journal*, 130(629): 1200–1225.

Becker, Gary S. 1973. "A Theory of Marriage: Part I." *Journal of Political Economy*, 81(4): 813–846.

Becker, G.S. 1981. *A Treatise on the Family.*: Harvard University Press.

Boerma, Job, Aleh Tsyvinski, and Alexander P Zimin. 2022. "Bunching and Taxing Multidimensional Skills." Working Paper 30015, National Bureau of Economic Research.

Borella, Margherita, Mariacristina De Nardi, and Fang Yang. 2023. "Are marriage-related taxes and social security benefits holding back female labour supply?" *The Review of Economic Studies*, 90(1): 102–131.

Browning, M., P.A. Chiappori, and Y. Weiss. 2014. *Economics of the Family*. Cambridge Surveys of Economic Literature: Cambridge University Press.

Browning, Martin, Pierre-Andre Chiappori, and Arthur Lewbel. 2013. "Estimating Consumption Economies of Scale, Adult Equivalence Scales, and Household Bargaining Power." *Review of Economic Studies*, 80(4 (285)): 1267–1303.

Colombi, Roberto. 1990. "A New Model of Income Distribution: The pareto-lognormal Distribution." In *Income and wealth distribution, inequality and poverty.*: Springer, 18–32.

Diamond, Peter. 1998. "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates." *American Economic Review*, 88(1): 83–95.

Evans, Lawrence C. 2010. *Partial differential equations.* 19: American Mathematical Soc..

Evans, Lawrence C, and Ronald F Garzepy. 2018. *Measure theory and fine properties of functions.*: Routledge.

Farhi, Emmanuel, and Ivan Werning. 2013. "Insurance and Taxation over the Life Cycle." *Review of Economic Studies*, 80(2): 596–635.

Ferey, Antoine, Benjamin B. Lockwood, and Dmitry Taubinsky. 2022. "Sufficient Statistics for Nonlinear Tax Systems with General Across-Income Heterogeneity." CESifo Working Paper Series 9958, CESifo.

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

Gayle, George-Levi, and Andrew Shephard. 2019. "Optimal Taxation, Marriage, Home Production, and Family Labor Supply." *Econometrica*, 87(1): 291–326.

Golosov, Mikhail, Narayana Kocherlakota, and Aleh Tsyvinski. 2003. "Optimal Indirect and Capital Taxation." *Review of Economic Studies*, 70(3): 569–587.

Golosov, Mikhail, Maxim Troshkin, and Aleh Tsyvinski. 2016. "Redistribution and Social Insurance." *American Economic Review*, 106(2): 359–86.

Golosov, Mikhail, Maxim Troshkin, Aleh Tsyvinski, and Matthew Weinzierl. 2013. "Preference heterogeneity and optimal capital income taxation." *Journal of Public Economics*, 97 160–175.

Golosov, Mikhail, Aleh Tsyvinski, and Nicolas Werquin. 2014. "A Variational Approach to the Analysis of Tax Systems." Working Paper 20780, National Bureau of Economic Research.

Grisvard, P. 2011. *Elliptic Problems in Nonsmooth Domains.* Classics in Applied Mathematics: Society for Industrial and Applied Mathematics.

Guner, Nezih, Remzi Kaygusuz, and Gustavo Ventura. 2014. "Income Taxation of U.S. Households: Facts and Parametric Estimates." *Review of Economic Dynamics*, 17(4): 559–581.

Hammond, Peter. 1979. "Straightforward Individual Incentive Compatibility in Large Economies." *Review of Economic Studies*, 46(2): 263–282.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. 2017. "Optimal Tax Progressivity: An Analytical Framework." *Quarterly Journal of Economics*, 132(4): 1693–1754.

Hellwig, Christian, and Nicolas Werquin. 2022. "A Fair Day's Pay for a Fair Day's Work: Optimal Tax Design as Redistributional Arbitrage." TSE Working Papers 22-1284, Toulouse School of Economics (TSE).

Hua, Lei, and Harry Joe. 2014. "Strength of tail dependence based on conditional tail expectation." *Journal of Multivariate Analysis*, 123 143–159.

Jacquet, Laurence, Etienne Lehmann, and Bruno Van der Linden. 2013. "Optimal Redistributive Taxation with Both Extensive and Intensive Responses." *Journal of Economic Theory*, 148(5): 1770–1805.

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

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

Kushnir, Alexey, and Ali Shourideh. 2022. "Optimal Multi-Dimensional Mechanisms: A Heuristic Approach." working paper, Carnegie Mellon University.

Lazear, Edward P., and Robert T. Michael. 1980. "Family Size and the Distribution of Real Per Capita Income." *American Economic Review*, 70(1): 91–107.

Liebman, Jeffrey, and Daniel Ramsey. 2019. "Independent Taxation, Horizontal Equity, and Return-Free Filing." *Tax Policy and the Economy*, 33 109–130.

Lockwood, Benjamin B., and Matthew Weinzierl. 2015. "De Gustibus non est Taxandum: Heterogeneity in preferences and optimal redistribution." *Journal of Public Economics*, 124 74–80.

Luenberger, David G. 1997. *Optimization by vector space methods.*: John Wiley & Sons.

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

Mirrlees, James. 1976. "Optimal tax theory: A synthesis." *Journal of Public Economics*, 6(4): 327–358.

Mirrlees, James. 1986. "The Theory of Optimal Taxation." *Handbook of Mathematical Economics*, Volume 3 1197–1249.

Moser, Christian, and Pedro Olea de Souza e Silva. 2019. "Optimal Paternalistic Savings Policies." Opportunity and Inclusive Growth Institute Working Papers 17, Federal Reserve Bank of Minneapolis.

Ndiaye, Abdoulaye. 2018. "Flexible Retirement and Optimal Taxation." working paper, NYU.

Nelsen, Roger B. 2006. *An Introduction to Copulas.* New York, NY, USA: Springer, , 2nd edition.

Reid, M.G. 1934. *Economics of Household Production.* Books in home economics: J. Wiley & Sons, Incorporated.

Renes, Sander, and Floris Zoutman. 2017. "As Easy as ABC? Multidimensional Screening in Public Finance." CESifo Working Paper Series 6301, CESifo.

Rindler, Filip. 2018. *Calculus of Variations.*: Springer Cham.

Rochet, Jean. 1987. "A Necessary and Sufficient Condition for Rationalizability in a Quasilinear Context." *Journal of Mathematical Economics*, 16(2): 191–200.

Rochet, Jean-Charles, and Philippe Chone. 1998. "Ironing, Sweeping, and Multidimensional Screening." *Econometrica*, 66(4): 783–826.

Rockafellar, Ralph Tyrell. 2015. *Convex Analysis.*: Princeton university press.

Rothschild, Casey, and Florian Scheuer. 2013. "Redistributive Taxation in the Roy Model." *The Quarterly Journal of Economics*, 128(2): 623–668.

Rothschild, Casey, and Florian Scheuer. 2014. "A Theory of Income Taxation under Multidimensional Skill Heterogeneity." Working Paper 19822, National Bureau of Economic Research.

Rothschild, Casey, and Florian Scheuer. 2016. "Optimal Taxation with Rent-Seeking." *Review of Economic Studies*, 83(3): 1225–1262.

Shaked, M., and J.G. Shanthikumar. 2007. *Stochastic Orders.* Springer Series in Statistics: Springer New York.

Spiritus, Kevin, Etienne Lehmann, Sander Renes, and Floris T. Zoutman. 2022. "Optimal Taxation with Multiple Incomes and Types." IZA Discussion Papers 15036.

Stantcheva, Stefanie. 2017. "Optimal Taxation and Human Capital Policies over the Life Cycle." *Journal of Political Economy*, 125(6): 1931–1990.

Voena, Alessandra. 2015. "Yours, Mine, and Ours: Do Divorce Laws Affect the Intertemporal Behavior of Married Couples?" *American Economic Review*, 105(8): 2295–2332.

Appendix

8 Mathematical preliminaries

This section lists some basic mathematical concepts necessary to characterize a solution to the mechanism design problem. We refer the reader to Evans (2010) and Rindler (2018) for additional background reading.

Consider a measurable function $\varphi : \mathbb{R}^2_{++} \to \mathbb{R}$. It is said to be (i) *compactly supported* if $\phi|_D = 0$ for some compact $D\mathbb{R}^2_{++}$, (ii) *locally integrable* if $\int_D |\varphi|\, d\mathbf{w} < \infty$ for every compact $D \subset \mathbb{R}^2_{++}$ and *integrable* if $\int |\varphi|\, d\mathbf{w} < \infty$, (iii) *essentially bounded* if there exists a constant m such that $|\varphi(\mathbf{w})| \leq m$ a.e.. We will write $\mathscr{L}^1$ and $\mathscr{L}^\infty$ for the spaces of such functions that are integrable and essentially bounded, respectively.

A locally integrable function φ is *weakly differentiable* on U if there exists a locally integrable vector field $\nabla\varphi := \left(\frac{\partial\varphi}{\partial w_1}, \frac{\partial\varphi}{\partial w_2}\right) : \mathbb{R}^2_{++} \to \mathbb{R}^2$ such that for all infinitely differentiable ϕ with a compact support, we have

$$\int \frac{\partial\phi}{\partial w_i}\varphi d\mathbf{w} = -\int \phi\frac{\partial\varphi}{\partial w_i} d\mathbf{w}.$$

The vector field of partial derivatives of φ is called a *weak gradient*. It is unique up to a set of zero measure. If v is differentiable, it is weakly differentiable, and its weak gradient coincides with the classical one. It is well known that for a weakly differentiable $\phi \in \mathscr{L}^\infty$ with $\nabla\phi \in \mathscr{L}^\infty$ and a weakly differentiable $\varphi \in \mathscr{L}^1$ with $\nabla\varphi \in \mathscr{L}^1$, their product, $\phi\varphi$, is integrable, weakly differentiable with $\nabla(\phi\varphi) \in \mathscr{L}^1$ and satisfies $\nabla(\phi\varphi) = \varphi\nabla\phi + \phi\nabla\varphi$.

The *divergence theorem* is Theorem 1.5.3.1 in Grisvard (2011). Take a vector field $\mathbf{\Lambda} = (\Lambda_1, \Lambda_2) : \mathbb{R}^2_{++} \to \mathbb{R}^2$ of integrable, weakly differentiable functions with each $\nabla\Lambda_i \in \mathscr{L}^1$. In short, this theorem represents the integral of its divergence defined by $\nabla \cdot \mathbf{\Lambda} := \sum_{i=1}^n \frac{\partial\Lambda_i}{\partial w_i}$ over some set $U \subset \mathbb{R}^2_+$ as a certain integral of $\mathbf{\Lambda}$ along the boundary of U. An open, connected set U is said to have a Lipshitz boundary if it can be locally represented as a boundary of an epigraph of some Lipshitz continuous function after a rotation of coordinates, see Definition 1.2.1.1 in Grisvard (2011) and the figure below for an illustration.

Then, the divergence theorem asserts that for every bounded set U with a Lipshitz boundary, the following identity holds:

$$\int_U \nabla \cdot \mathbf{\Lambda} d\mathbf{w} = \int_{\partial U} \mathbf{\Lambda} \cdot \boldsymbol{n} d\sigma,$$

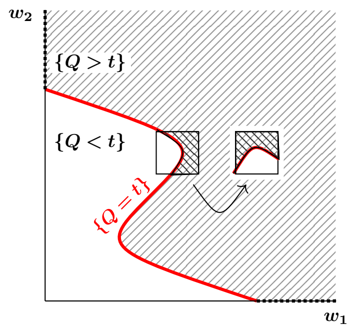


Figure 8: The solid red line is the level set $\{Q = t\}$, the hashed area is the superlevel set $\{Q > t\}$, and white region is the sublevel set $\{Q < t\}$. Two rectangles illustrate a rotation that can be used to represent a portion of the boundary as an epigraph of a Lipshitz continuous function.

where $n_i(\mathbf{w})$ is the i-th component of the outward unit vector to ∂U at $\mathbf{w}$ and σ is the Lebesgue measure on ∂U. The divergence theorem is a multi-dimensional generalization of the second fundamental theorem of calculus,

$$\int_a^b \Lambda'(w)\, dw = \Lambda(b) - \Lambda(a).$$

The *coarea formula* is Theorem 3.11 in Evans and Garzepy (2018). Take a measurable function $\varphi : \mathbb{R}^2_{++} \to \mathbb{R}$ and let $Q : \mathbb{R}^n_{++} \to \mathbb{R}_+$ be a locally Lipshitz onto function that satisfies (ii)-(iv) in Theorem 1 and such that $\varphi \|\nabla Q\|$ integrable. Then, the coarea formula asserts

$$\int_{\{Q>t\}} \varphi \|\nabla Q\| d\mathbf{w} = \int_t^\infty \left(\int_{\{Q=s\}} \phi d\sigma \right) ds,$$

where σ is the Lebesgue measure on ∂D_t. Hence, for a.e. t,

$$\frac{d}{dt} \int_{\{Q>t\}} \varphi \|\nabla Q\| d\mathbf{w} = - \int_{\{Q=t\}} \phi d\sigma.$$

The coarea formula generalizes the first fundamental theorem of calculus for strictly increasing Q,

$$\frac{d}{dt} \int_0^{Q^{-1}(t)} \varphi(w)\, Q'(w) dw = \varphi\left(Q^{-1}(t)\right).$$

It is worthwhile to remark that Theorem 3.11 in Evans (2010) doesn't requires (ii)-(iv) but makes a somewhat weaker statement. It is easy to see that their statement coincides with ours under the additional structure on Q that we impose here.

9 Optimal taxation as a mechanism design problem

We start with an auxiliary lemma that establishes properties of μ and tuples (v^s, c^s, y^s), (v^m, c^m, y_1^m, y_2^m) that satisfy the constraints of the mechanism design problem.

Lemma 9. *Consider μ and (v^s, c^s, y^s), (v^m, c^m, y_1^m, y_2^m) such that (2), (3), (4) and (5) hold. Then, both v^s and v^m are nondecreasing, bounded from below, locally Lipshitz, a.e. and weakly differentiable with weak derivatives given by (8). Moreover, $\mu \in (0,1)$, and*

(a) $v^s g, v^m f \in \mathscr{L}^1$,

(b) $w \frac{\partial v^s}{\partial w} g, w_1 \frac{\partial v^m}{\partial w_1} f, w_2 \frac{\partial v^m}{\partial w_2} f \in \mathscr{L}^1$.

Proof. Use (2) to substitute for c^s in (3) to obtain

$$v^s(w) \geqslant v^s(\widehat{w}) + \gamma \left(\frac{y^s(\widehat{w})}{\widehat{w}} \right)^{1/\gamma} \left(\left(\frac{\widehat{w}}{w} \right)^{1/\gamma} - 1 \right). \tag{51}$$

Monotonicity of v^s follows directly from (51) and nonnegativity of earnings. As a result, v^s is bounded from below by $v^s(0)$.

A further examination of (51) reveals that v^s is defined as a maximum of functions that are affine in $w^{-1/\gamma}$, thus v^s must be a convex function of $w^{-1/\gamma}$. Since the transformation $w \mapsto w^{-1/\gamma}$ is continuously differentiable on $\mathbb{R}_{++}$, Theorem 10.4 in Rockafellar (2015) implies that v^s is locally Lipshitz. By Theorem 6 on p. 296 (Rademacher Theorem) in Evans (2010), v^s is differentiable a.e. due to local Lipshitz continuity. Then, Theorem 5 and Remark on p. 295 in Evans (2010) imply that v^s is weakly differentiable. Finally, since v^s is differentiable a.e., the standard envelope argument applied to (51) together with the fact that the maximum on the right-hand side is attained at $\widehat{w} = w$ establishes that (8) holds at every point of differentiability.

Remark that the exactly same argument applies to v^m, hence it is also nondecreasing, bounded from below, locally Lipshitz, a.e. and weakly differentiable with weak derivatives given by (8).

We now show that $\mu \in (0,1)$ and two properties, (a) and (b). To begin, use (2) to rewrite (4) in terms of v^s and v^m as follows:

$$\frac{\mu}{2} \int \left(\sum_{i=1}^{2} \left(y_i^m - \gamma \left(\frac{y_i^m}{w_i} \right)^{1/\gamma} \right) - v^m \right) dF + (1-\mu) \int \left(y^s - \gamma \left(\frac{y^s}{w} \right)^{1/\gamma} - v^s \right) dG \geq 0. \tag{52}$$

Note that the value of $\max_{y \geq 0} \left(y - \gamma \left(\frac{y}{w} \right)^{1/\gamma} \right)$ is proportional to $w^{1/(1-\gamma)}$. Since v^s, v^m are bounded from below and $\int w^{1/(1-\gamma)} dG < \infty$, the left-hand side of (52) is finite.

It is immediate that, if $\mu \in (0,1)$, then Property (a) holds. If $\mu = 0$, then, since $\Phi(\mu) = -\infty$ and $\int v^m dF \geq v^m(0)$, we must have $\int v^s dG = \infty$ due to (5). This contradicts (52). The similar argument rules out $\mu = 1$, and hence Property (a) is established.

We now show that Property (b) holds. Consider the following auxiliary problem parameterized by $k \geq 0$:

$$\max_{y(\cdot) \geq 0} \int y dG - k \text{ s.t. } \int \left(\frac{y}{w} \right)^{1/\gamma} dG = k.$$

It is easy to see that the value of this problem diverges to $-\infty$ as $k \to \infty$. Substitute $\frac{\partial v^s}{\partial w}$ for y^s using (8) to obtain that $\int w \frac{\partial v^s}{\partial w} dG$ must be finite. The same argument applies to v^m, thus $\int w_i \frac{\partial v_i^m}{\partial w_i} dF < \infty$ for each coordinate i. $\square$

9.1 Relaxed problem

We now formally define and further simplify the relaxed problem introduced in the main text. Let $\mathscr{V}^s$ and $\mathscr{V}^m$ be the spaces of functions v^s and v^m satisfying the conditions listed in Lemma 9. Then, the relaxed problem is to select $\mu \in [0,1]$ and $(v^s, v^m) \in \mathscr{V}^s \times \mathscr{V}^m$ to maximize $\mathcal{W}$ defined in (6) subject to (5) and (9).

To make our analysis applicable to study extensions, we consider the setting developed in Section 5.1 in which the planner assigns social weights $\alpha^s(w)$ and $\beta(w_i|w_{-i})$ to single and married persons. This generalizes the benchmark economy by allowing $\int \alpha^s dG \neq 1$ and $\alpha^m(w_1, w_2) := \frac{\beta(w_1|w_2)+\beta(w_2|w_1)}{2} \neq \frac{\alpha^s(w_1)+\alpha^s(w_2)}{2}$ but still requires couple's weights α^m to be symmetric. The asymmetric case relevant for Sections 5.8 and 5.9 will be discussed separately.

Set $\eta := (\mu \mathbb{E}\alpha^m + (1-\mu)\mathbb{E}\alpha^s)^{-1}$. Remark that $\mathcal{W}$ can be rewritten as follows:

$$\mathcal{W} = \frac{\mu}{2} \int (\alpha^m - \mathbb{E}\alpha^m) v^m dF + (1-\mu) \int (\alpha^s - \mathbb{E}\alpha^s) v^s dG + \int_\mu^1 \Phi(\varepsilon) d\varepsilon +$$
$$+ \frac{1}{\eta} \left[\frac{\mu}{2} \int v^m dF + (1-\mu) \int v^s dG \right] + (\mathbb{E}\alpha^m - \mathbb{E}\alpha^s) \mu(1-\mu) \left[\frac{1}{2} \int v^m dF - \int v^s dG \right]. \quad (53)$$

Two terms in square brackets can be solved for from (5) and (9). It is immediate that the budget constraints must bind, thus the fomer term in the second line equals to $\mathcal{S}$, which is given by

$$\frac{\mu}{2} \int \sum_{i=1}^{2} \left(w_i^{1+\gamma} \left(\frac{\partial v^m}{\partial w_i} \right)^\gamma - \gamma w_i \frac{\partial v^m}{\partial w_i} \right) dF + (1-\mu) \int \left(w^{1+\gamma} \left(\frac{\partial v^s}{\partial w} \right)^\gamma - \gamma w \frac{\partial v^s}{\partial w} \right) dG. \quad (54)$$

The latter term equals to $\Phi(\mu)$ due to (5).

Putting all pieces together, the relaxed problem is

$$\max_{\substack{\mu \in [0,1] \\ (v^s, v^m) \in \mathscr{V}^s \times \mathscr{V}^m}} \frac{\mu}{2} \int (\alpha^m - \mathbb{E}\alpha^m) v^m dF + (1-\mu) \int (\alpha^s - \mathbb{E}\alpha^s) v^s dG + \int_\mu^1 \Phi(\varepsilon) d\varepsilon + \frac{1}{\eta} \mathcal{S} +$$
$$+ (\mathbb{E}\alpha^m - \mathbb{E}\alpha^s) \mu(1-\mu) \Phi(\mu). \quad (55)$$

It is worth to mention that the solution to (55) is defined up to two constants, $v^s(0)$ and $v^m(\mathbf{0})$. These constants are pinned by two binding constraints, (5), (9), so that $\int v^{s,*} dG = \mathcal{S}^* - \mu^* \Phi(\mu^*)$ and $\frac{1}{2} \int v^{m,*} dF = \mathcal{S}^* + (1-\mu^*)\Phi(\mu^*)$. Here, $\mathcal{S}^*$ is (54) evaluated at the optimum.

9.2 Optimality conditions

In this section, we formally derive conditions that are necessary and sufficient for optimality in the relaxed problem (55). Recall that θ is the tail statistics of G defined by $\theta(t) = \frac{tg(t)}{1-G(t)}$. In the proposition below, we use the shorthand notation θ_i to denote this statistics evaluated at $t = w_i$.

Proposition 2. *Consider $\mu \in (0,1)$ and (v^s, v^m) that satisfy (A1) $\lambda^s, \lambda_1^m, \lambda_2^m$ are continuous, (A2) $\underline{\lambda} \leq \lambda^s, \lambda_1^m, \lambda_2^m \leq \overline{\lambda}$ for some $-1 < \underline{\lambda} \leq \overline{\lambda} < \infty$, (A3) $\lambda^s, \lambda_1^m, \lambda_2^m$ are weakly differentiable, (A4) $\frac{\partial(w\lambda^s g)/\partial w}{g}, \frac{\sum_{i=1}^2 \partial\left(w_i \lambda_i^m f\right)/\partial w_i}{f} \in \mathscr{L}^\infty$ and (A5) $\lambda^s\theta, \lambda_1^m\theta_1, \lambda_2^m\theta_2 \in \mathscr{L}^\infty$.*

Set $\eta := (\mu\mathbb{E}\alpha^m + (1-\mu)\mathbb{E}\alpha^s)^{-1}$. Then, (v^s, v^m) is in $\mathscr{V}^s \times \mathscr{V}^m$ and maximizes the objective in (55) for fixed μ if and only if

$$\frac{\partial\left(\gamma w \lambda^s g\right)}{\partial w} = \eta\left(\alpha^s - \mathbb{E}\alpha^s\right) g, \tag{56}$$

$$\sum_{i=1}^{2} \frac{\partial\left(\gamma w_i \lambda_i^m f\right)}{\partial w_i} = \eta\left(\alpha^m - \mathbb{E}\alpha^m\right) f. \tag{57}$$

If $(v^{s,*}, v^{m,*})$ verifies (A1)-(A5), then (56), (57) hold and the following first-order condition w.r.t. μ is satisfied:

$$\frac{1-\gamma}{2} \int \sum_{i=1}^{2} w_i \left(\frac{w_i}{1+\lambda_i^{m,*}}\right)^{\gamma/(1-\gamma)} dF - (1-\gamma) \int w \left(\frac{w}{1+\lambda^{s,*}}\right)^{\gamma/(1-\gamma)} dG =$$
$$= \eta^*\Phi(\mu^*) + \eta^*\left(\mathbb{E}\alpha^s - \mathbb{E}\alpha^m\right)\left(\mathcal{S}^* + \frac{\partial\left[\mu^*(1-\mu^*)\Phi(\mu^*)\right]}{\partial\mu}\right). \tag{58}$$

Proposition 2 contains two parts. The first parts looks at the optimal choice of functions (v^s, v^m) in the relaxed problem with a fixed value of μ. This a concave problem; as a result, the differential equations in (56), (57) are necessary and sufficient for optimality of (v^s, v^m) satisfying the set of regularity conditions (A1)-(A5).

Recall that distortions are defined by (10). Condition (A2) means that marginal taxes are uniformly bounded with the upper bound strictly less than 1, and (A1) means that v^s, v^m are continuously differentiable. By (8), this is equivalent to the fact that earnings change continuously ruling out kinks in taxes. Condition (A3) ensures that (weak) derivatives in (56), (57) are well-defined. Then, (A4) and (A5) require that distortions and their derivatives are well-behaved on the boundary and at "infinity". In particular, (A5) means that $\lambda^s(w)\, wg(w)$ converges to 0 fast enough so that $\lim_{w\to\infty} \lambda^s(w)\, wg(w)\, \hat{v}^s(w) = 0$ for all $\hat{v}^s \in \mathscr{V}^s$. Finally, condition (A4) means that the sum of $\frac{\partial(wg)/\partial w}{g}\lambda^s(w)$ and $w\frac{\partial\lambda^s(w)}{\partial w}$ is bounded. Since $\frac{\partial(wg)/\partial w}{g} \sim -\theta(w)$ as $w \to \infty$, boundedness of the first-term is implied by (A2) and (A4). Hence, condition

(A4) reduces to the requirement that the derivative of λ^s doesn't explode as $w \to 0$ and converges to 0 fast enough as $w \to \infty$, which holds when this derivative is bounded and λ^s converges as $w \to \infty$. The interpretation of these condition for v^m is identical.

The second part of Proposition 2 gives the first-order necessary condition for μ^*. In general, there may be multiple solutions to (58) when $\mathbb{E}\alpha^m \neq \mathbb{E}\alpha^s$, because the relaxed problem is not jointly concave in μ and (v^s, v^m). However, in the benchmark economy or more generally when $\mathbb{E}\alpha^m = \mathbb{E}\alpha^s$, (58) pins down a unique value of μ^*, because (56), (57) do not depend on μ due to $\eta = 1$ for all μ.

Proof. We first show that (v^s, v^m) is in $\mathscr{V}^s \times \mathscr{V}^m$ provided that (A2) holds. Indeed, by the definition of λ^s in (10), $w\frac{\partial v^s}{\partial w} \leq \left(\frac{w}{1+\underline{\lambda}}\right)^{1/(1-\gamma)}$, which gives

$$v^s(w) - v^s(0) = \int_0^1 w\frac{\partial v^s(wt)}{\partial w}dt \leq (1-\gamma)\left(\frac{w}{1+\underline{\lambda}}\right)^{1/(1-\gamma)}.$$

Since the value of $\int w^{1/(1-\gamma)}dG$ is finite, both $v^s g$ and $w\frac{\partial v^s}{\partial w}g$ are integrable, thus $v^s \in \mathscr{V}^s$. The argument for v^m is identical as $w_i\frac{\partial v_i^m}{\partial w_i} \leq \left(\frac{w_i}{1+\underline{\lambda}}\right)^{1/(1-\gamma)}$ for $i = 1, 2$ implies

$$v^m(\boldsymbol{w}) - v^m(\boldsymbol{0}) = \int_0^1 \sum_{i=1}^2 w_i\frac{\partial v^m(\boldsymbol{w}t)}{\partial w_i}dt \leq (1-\gamma)\sum_{i=1}^2\left(\frac{w_i}{1+\underline{\lambda}}\right)^{1/(1-\gamma)},$$

and the result follows.

We now study optimality of (v^s, v^m) for fixed $\mu \in (0,1)$. Remark that v^s enters social welfare $\mathcal{W}$ only through the functional Υ^s defined by

$$\Upsilon^s(v^s) := \int \left(\eta\left(\alpha^s - \mathbb{E}\alpha^s\right)v^s + w^{1+\gamma}\left(\frac{\partial v^s}{\partial w}\right)^\gamma - \gamma w\frac{\partial v^s}{\partial w}\right)dG.$$

Since Υ^s is concave, $v^s \in \mathscr{V}^s$ satisfies $\Upsilon^s(v^s) \geq \Upsilon^s(\hat{v}^s)$ for all functions $\hat{v}^s$ in $\mathscr{V}^s$ if and only if the following "first-order condition" holds:

$$\lim_{t\to 0}\frac{\Upsilon^s((1-t)v^s + t\hat{v}^s) - \Upsilon^s(\hat{v}^s)}{t} \leq 0 \;\; \forall \hat{v}^s \in \mathscr{V}^s. \tag{59}$$

It is routine to verify using the monotone convergence theorem that the limit in (59) can be taken under the integral sign and that this condition is equivalent to

$$\int \left(\eta\left(\alpha^s - \mathbb{E}\alpha^s\right)\left(\hat{v}^s - v^s\right) + \gamma w\lambda^s\left(\frac{\partial \hat{v}^s}{\partial w} - \frac{\partial v^s}{\partial w}\right)\right)dG \leq 0 \;\; \forall \hat{v}^s \in \mathscr{V}^s. \tag{60}$$

Remark that $\mathscr{V}^s$ is a cone, thus $kv^s \in \mathscr{V}^s$ for every $k > 0$. It follows that

$$\int \left(\eta\left(\alpha^s - \mathbb{E}\alpha^s\right)v^s + \gamma w\lambda^s\frac{\partial v^s}{\partial w}\right)dG = 0, \tag{61}$$

which allows to eliminate v^s from (60).

Let $\hat{v}^s$ be a function in $\mathscr{V}^s$. Apply the divergence theorem (integration by parts) to obtain

$$\int_{\underline{t}}^{t} \gamma w \lambda^s \frac{\partial \hat{v}^s}{\partial w} dG = -\int_{\underline{t}}^{t} \frac{\partial\left(\gamma w \lambda^s g\right)/\partial w}{g} \hat{v}^s dG + \gamma \lambda^s(t)\theta(t)(1-G(t))\hat{v}^s(t) - \gamma\lambda^s(\underline{t})\underline{t}g(\underline{t})\hat{v}^s(\underline{t}).$$

(A2) ensures that the third term becomes 0 as $\underline{t} \to 0$. Since the expected value of $\hat{v}^s$ is finite, $(1-G(t))\hat{v}^s(t)$ converges to 0 as $t \to \infty$. It follows that the second term goes to 0 as $t \to \infty$ due to (A5). By (A2) and (A4), the dominated convergence theorem implies that (60) when combined with (62) can be rewritten as

$$\int \left[\eta\left(\alpha^s - \mathbb{E}\alpha^s\right) - \frac{\partial\left(\gamma w \lambda^s g\right)/\partial w}{g}\right] \hat{v}^s dG \leq 0. \tag{62}$$

By (A2) and the definition of λ^s in (10), $w\frac{\partial v^s}{\partial w} \geq \left(\frac{w}{1+\overline{\lambda}}\right)^{1/(1-\gamma)}$. Hence, the nonnegativity constraint on earnings is slack, i.e., in a neighborhood of every $w \in \mathbb{R}_{++}$, (62) can be freely varied by setting $\hat{v}^s = v^s \pm \phi$ for some smooth function ϕ that vanish outside this neighborhood. As a result, by the fundamental lemma of calculus of variations, (62) holds if and only if the integrand in the square brackets equals 0 for a.e. w.

The argument for married individuals is similar, and we will sketch it skipping intermediate steps. Recollect that v^m enters the relaxed problem only through the functional Υ^m defined by

$$\Upsilon^m(v^m) := \int \left(\eta\left(\alpha^m - \mathbb{E}\alpha^m\right) v^m + \sum_{i=1}^{2}\left(w_i^{1+\gamma}\left(\frac{\partial v^m}{\partial w_i}\right)^{\gamma} - \gamma w_i \frac{\partial v^m}{\partial w_i}\right)\right) dF.$$

Then, v^m solves the relaxed problem with fixed $\mu \in (0,1)$ if and only if

$$\int \left(\eta\left(\alpha^m - \mathbb{E}\alpha^m\right)\hat{v}^m + \sum_{i=1}^{2} \gamma w_i \lambda_i^m \frac{\partial \hat{v}^m}{\partial w_i}\right) dF \quad \leq \quad 0 \;\; \forall \hat{v}^m \in \mathscr{V}^m, \tag{63}$$

$$\int \left(\eta\left(\alpha^m - \mathbb{E}\alpha^m\right) v^m + \sum_{i=1}^{2} \gamma w_i \lambda_i^m \frac{\partial v^m}{\partial w_i}\right) dF \quad = \quad 0. \tag{64}$$

Let $\hat{v}^m$ be a function in $\mathscr{V}^m$. By the divergence theorem,

$$\int_{[\underline{t},t]^2} \sum_{i=1}^{2} \gamma w_i \lambda_i^m \frac{\partial \hat{v}^m}{\partial w_i} dF = -\int_{[\underline{t},t]^2} \frac{\sum \partial\left(\gamma w_i \lambda_i^m f\right)/\partial w_i}{f} \hat{v}^m dF +$$

$$+\sum_{i=1}^{2} \int_{\underline{t}}^{t} \gamma t \lambda_i^m(t, w_{-i})\hat{v}^m(t, w_{-i}) f(t, w_{-i}) dw_{-i} - \sum_{i=1}^{2} \int_{\underline{t}}^{t} \gamma \underline{t} \lambda_i^m(\underline{t}, w_{-i})\hat{v}^m(\underline{t}, w_{-i}) f(\underline{t}, w_{-i}) dw_{-i}.$$

Clearly, the second term in the second line converges to 0 as $\underline{t} \to 0$. We claim that the first term in the second line converges to 0 as $t \to \infty$. Indeed, by Hölder's inequality,

$$\sum_{i=1}^{2} \int_{0}^{t} \gamma t \left|\lambda_i^m(t, w_{-i})\right| \left|\hat{v}^m(t, w_{-i})\right| f(t, w_{-i}) dw_{-i} \leq \gamma \theta(t) \max_{i=1,2} \max_{w_{-i} \leq t} \left|\lambda_i^m(t, w_{-i})\right| \times$$
$$\times (1 - G(t)) \sum_{i=1}^{2} \int_{0}^{t} \left|\hat{v}^m(t, w_{-i})\right| f(w_{-i}|t) dw_{-i}.$$

The first term on the right-hand side is bounded due to (A5), the second one goes to 0 as $t \to \infty$. To see it, note that $1 - G(t) \leq \Pr(\max \boldsymbol{w} \geq t)$ and

$$\Pr(\max \boldsymbol{w} \geq t) \mathbb{E}\left[\hat{v}^m \middle| \max \boldsymbol{w} = t\right] \to 0 \text{ as t} \to \infty$$

due to integrability of $\hat{v}^m f$. The rest of the argument is exactly the same as for singles. To sum up, two differential equations, (56) and (57), are necessary and sufficient optimality conditions for fixed $\mu \in (0, 1)$.

It remains to show the necessary first-order condition for μ^*. Equation (58) directly follows from differentiating $\mathcal{W}$ w.r.t. μ and noting that

$$\Upsilon^s(v^{s,*}) = (1 - \gamma) \int w \left(\frac{w}{1 + \lambda^{s,*}}\right)^{\gamma/(1-\gamma)} dG, \tag{65}$$

$$\Upsilon^m(v^{m,*}) = \frac{1 - \gamma}{2} \int \sum_{i=1}^{2} w_i \left(\frac{w_i}{1 + \lambda_i^{m,*}}\right)^{\gamma/(1-\gamma)} dF. \tag{66}$$

□

We end this section by pointing out a certain well-known equivalence between $v^{m,*}$ and $\boldsymbol{\lambda}^{m,*}$. One can think equivalently of equations (57) and (15) either as a second-order partial differential equation describing the solution to the relaxed problem $v^{m,*}$, or as a system of joint first-order partial differential equations describing the optimal $\boldsymbol{\lambda}^{m,*}$ implied by that $v^{m,*}$. Formally: if $\left((1 + \lambda_i^m)^{1/(\gamma-1)} w_i^{\gamma/(1-\gamma)}\right)_{i=1,2}$ is continuously differentiable with derivatives that are uniformly continuous on bounded subsets and (15) holds, then there exists a unique (up to a constant) function v^m such that Equation (10) holds for these v^m and $\boldsymbol{\lambda}^m$.

9.3 Proof of Theorem 1

We assume here and throughout the rest of the paper that the first-order approach is valid, $\lambda^{s,*}$ and $\boldsymbol{\lambda}^{m,*}$ satisfy conditions (A1)-(A5) of Proposition 2.

Before proving the theorem, it is worthwhile to discuss the conditions that we impose on Q. First of all, the requirement that Q is onto ensures that every super- sublevel set is nonempty. It is made purely for simplicity of exposition. Recall that Q is locally Lipshitz if it is Lipshitz in a neighborhood of each point. It is well known that such functions are differentiable almost everywhere, e.g., see Evans (2010), hence $\frac{\partial \ln Q}{\partial \ln w_i}$ that appears in Equation (16) is well-defined. In addition, we require the following set of regularity conditions on Q: (Q1) $\mathbb{E}\left[\sum_{i=1}^{2} w_i \left|\partial Q/\partial w_i\right|\right]$ is finite, (Q2) $\{Q > t\}$, $\{Q < t\}$ are connected, (Q3) $\partial\{Q > t\}$, $\partial\{Q < t\}$ are Lipshitz, and (Q4) $\{Q = t\} = \partial\{Q > t\} \cap \partial\{Q < t\} \cap \mathbb{R}^2_{++}$. Condition (Q1) is needed to apply the divergence theorem to potentially unbounded super- and sublevel sets of Q. Conditions (Q2), (Q4) rule out situations in which a level set is "thick" or contains several connected components. Condition (Q3) means that closures of $\{Q > t\}$, $\{Q < t\}$ can be locally thought of as epigraphs of Lipshitz functions. These conditions are needed to apply the divergence theorem and the coarea formula as stated in the mathematical appendix. In principle, (Q2) can be dropped by applying the divergence to each of the connected components separately.

Condition (Q4) is satisfied automatically if Q is continuously differentiable and has no critical points, i.e., its derivatives don't vanish at the same time. Then, (Q3) is satisfied when $\{Q = t\}$ approach the boundary of $\mathbb{R}^2_{++}$ with no "casps". Finally, (Q2) holds if Q is either quasiconvex or quasiconcave.

Proof. Let's write (14) succinctly as $-\nabla \cdot (\mathbf{\Lambda} f) = \varphi f$, where $\varphi := 1 - \alpha^m$ is a scalar field and $\mathbf{\Lambda} := \left(\lambda_1^{m,*}\gamma w_1, \lambda_2^{m,*}\gamma w_2\right)$ is a vector field. First, define $D_t := \{Q > t\}$ for every $t > 0$. The properties of Q ensure that the boundary of D_t is the union of $\{Q = t\}$ and the portion of axes given by $\overline{D_t \cap \partial\mathbb{R}^2_{++}}$, which is illustrated with dots in Figure 1. Clearly, ∂D_t is a null set, hence $\int_{\{Q \geq t\}} \varphi f d\mathbf{w} = \int_{D_t} \varphi f d\mathbf{w}$ and $\Pr(Q \geq t) = \Pr(D_t)$.

Let B_r be an open ball of radius $r > 0$ centered at the origin such that $D_t \cap B_r \neq \emptyset$. Clearly, such an r exists because Q is continuous and onto. By the divergence theorem, Equation (14) can be transformed as

$$\int_{D_t \cap B_r} \varphi f d\boldsymbol{w} \quad = \quad -\int_{D_t \cap \partial B_r} \mathbf{\Lambda} f \cdot \boldsymbol{n} d\sigma - \int_{\partial D_t \cap B_r} \mathbf{\Lambda} f \cdot \boldsymbol{n} d\sigma. \tag{67}$$

Note the term on the left-hand side of (67) goes to $\int_{D_t} \varphi f d\boldsymbol{w}$ as $r \to \infty$ due to the dominated convergence theorem.

We now study two terms on the right-hand side of (67) as $r \to \infty$. On the boundary of B_r, the unit normal vector is given by $\boldsymbol{n} = \frac{w}{\|\boldsymbol{w}\|}$, hence $\int_{\partial B_r} |\mathbf{\Lambda} f \cdot \boldsymbol{n}| \, d\boldsymbol{w} \leq \overline{\lambda} r \int_{\partial B_r} f d\sigma$ due to

(A2). We claim that the upper bound goes to 0. Indeed, using the coarea formula, we obtain

$$\frac{d}{dr}\int_{B_r}\|\boldsymbol{w}\|fd\boldsymbol{w}=r\int_{\partial B_r}fd\sigma. \tag{68}$$

Since $\|\boldsymbol{w}\|\leq \max\boldsymbol{w}$ and the latter has a finite expected value, the limit of $\int_{B_r}\|\boldsymbol{w}\|fd\boldsymbol{w}$ is finite. As a result, the left-hand of (68) goes to 0 as claimed.

Recollect that the boundary of D_t can be partitioned into the union of $\{Q=t\}$ and $\overline{D_t\cap\partial\mathbb{R}^2_{++}}$. On the latter subset, $n_i=-1$ and $n_{-i}=0$ when $w_i=0$ and $w_{-i}>0$, thus $\boldsymbol{\Lambda}f\cdot\boldsymbol{n}=0$ due to (A2). On the former subset, the unit normal vector is given by $\boldsymbol{n}=-\frac{\nabla Q}{\|\nabla Q\|}$, hence $|\boldsymbol{\Lambda}f\cdot\boldsymbol{n}|\,\|\nabla Q\|\leq\overline{\lambda}\sum_{i=1}^2 w_i\frac{\partial Q}{\partial w_i}$. By assumption, the upper bound is integrable, thus $\boldsymbol{\Lambda}f\cdot\boldsymbol{n}\|\nabla Q\|$ is integrable as well. Using the coarea formula and passing r to ∞, which is valid due to the dominated convergence theorem, we obtain

$$\int_t^\infty\left(\int_{\{Q=s\}}\boldsymbol{\Lambda}f\cdot\frac{\nabla Q}{\|\nabla Q\|}d\sigma\right)ds=\int_{D_t}\boldsymbol{\Lambda}f\cdot\nabla Qd\boldsymbol{w}.$$

Differentiate this expression and combine it with (67) evaluated at $r\to\infty$ to get

$$\int_{D_t}\varphi fd\boldsymbol{w}=\int_{\{Q=t\}}\boldsymbol{\Lambda}f\cdot\frac{\nabla Q}{\|\nabla Q\|}d\sigma=-\frac{d}{dt}\int_{D_t}\boldsymbol{\Lambda}f\cdot\nabla Qd\boldsymbol{w}. \tag{69}$$

Finally, using the definition of conditional expectation, observe that the last term is simply

$$-\frac{d}{dt}\int_{D_t}\boldsymbol{\Lambda}f\cdot\nabla Qd\mathbf{w}=\mathbb{E}\left[\boldsymbol{\Lambda}\cdot\nabla Q|Q=t\right]\times\frac{-\partial\Pr\left(Q\geq t\right)}{\partial t}.$$

Combine these formulas and divide both sides by $\Pr\left(Q\geq t\right)$ to obtain (16). □

We end the proof by pointing an equivalent approach to establishing Theorem 1 that doesn't invoke the coarea formula. Instead of integrating (57) using the coarea formula, one may obtain the same conclusion directly from the variational optimality conditions (63), (64) by considering $\hat{v}$ of the form $\hat{v}(\boldsymbol{w})=v(\boldsymbol{w})+\psi(Q(\boldsymbol{w}))$ for some uni-dimensional function ψ. Under (Q1)-(Q4), we can combine these optimality conditions to obtain

$$\int\left(\left(1-\mathbb{E}\left[\alpha^m|Q=t\right]\right)\psi(t)+\mathbb{E}\left[\sum_{i=1}^2 t\gamma\lambda_i^m\frac{\partial\ln Q}{\partial\ln w_i}|Q=t\right]\psi'(t)\right)d\Pr(Q\leq t)=0. \tag{70}$$

The regularity conditions on Q ensure that all variations ψ are feasible and that $\frac{d\Pr(Q\leq t)}{dt}$ is well-defined and strictly positive. Integrating by parts (70), one can derive the formula in Theorem 1.

9.4 Properties of the optimal joint taxes

9.4.1 Proof of Lemma 2

Proof. By $F^a \leq_{PQD} F^b \leq_{PQD} \overline{F}$, the following first-order stochastic dominance relations hold:

$$\Pr^a\left(w_{-i} \geq t_{-i}|w_i \geq t_i\right) \leq \Pr^b\left(w_{-i} \geq t_{-i}|w_i \geq t_i\right) \leq \Pr\left(w_{-i} \geq t_{-i}|w_{-i} \geq t_i\right) \;\; \forall \mathbf{t}. \tag{71}$$

Remark that the second inequality in (71) holds on a set of positive measure because F^b is absolutely continuous and supported on $\mathbb{R}^2_+$. Since α is strictly decreasing, we have

$$\begin{aligned}
\mathbb{E}^a\left[\alpha^m|w_i \geq t\right] &= \frac{1}{2}\mathbb{E}\left[\alpha(w_i)|w_i \geq t\right] + \frac{1}{2}\mathbb{E}^a\left[\alpha(w_{-i})|w_i \geq t\right] \\
&\leq \frac{1}{2}\mathbb{E}\left[\alpha(w_i)|w_i \geq t\right] + \frac{1}{2}\mathbb{E}^b\left[\alpha(w_{-i})|w_i \geq t\right] = \mathbb{E}^b\left[\alpha^m|w_i \geq t\right] \\
&< \frac{1}{2}\mathbb{E}\left[\alpha(w_i)|w_i \geq t\right] + \frac{1}{2}\mathbb{E}\left[\alpha(w_i)|w_i \geq t\right] = \mathbb{E}\left[\alpha|w \geq t\right] \;\; \forall t
\end{aligned}$$

due to (71). Conclude that $\mathbb{E}^a\left[\lambda_i^{m,a,*}|w_i = t\right] \leq \mathbb{E}^b\left[\lambda_i^{m,b,*}|w_i = t\right]$ and $\mathbb{E}^b\left[\lambda_i^{m,b,*}|w_i = t\right] < \lambda^{s,b,*} = \lambda^{s,a,*}$.

If F^b is perfectly assortative, then $\mathbb{E}^b\left[\alpha^m|w_i \geq t\right]$ coincides with $\mathbb{E}\left[\alpha|w \geq t\right]$. As a result, we have $\mathbb{E}^b\left[\lambda_i^{m,b,*}|w_i = t\right] = \lambda^{s,b,*} = \lambda^{s,a,*}$.

If F^a is independent with marginals G, then $\mathbb{E}^a\left[\alpha^m|w_i \geq t\right] = \frac{1}{2}\mathbb{E}\left[\alpha|w \geq t\right] + \frac{1}{2}$. By monotonicity of α, the value of $\mathbb{E}\left[\alpha|w \geq t\right]$ is strictly less than than $\mathbb{E}\alpha = 1$ for all $t > 0$. It follows that $\mathbb{E}^a\left[\lambda_i^{m,a,*}|w_i = t\right] \geq \frac{1}{2}\frac{1-\mathbb{E}[\alpha|w \geq t]}{\gamma\theta(t)} > 0$ whenever F^a is positively dependent. □

9.4.2 Proof of Lemma 3

Proof. Recollect that the distribution with density $\alpha^a g$ first-order stochastically dominates the distribution with density $\alpha^b g$. It follows that $\mathbb{E}\left[\alpha^a|w \geq t\right] \geq \mathbb{E}\left[\alpha^b|w \geq t\right]$ for all t, hence $\lambda^{s,a,*} \leq \lambda^{s,b,*}$

As discussed in Chapter 6.E of Shaked and Shanthikumar (2007), under log-supermodularity of f, the distribution with density $\alpha^a(w_j)f$ first-order stochastically dominates the distribution with density $\alpha^b(w_j)f$ for $j = 1, 2$. As a result, $\mathbb{E}\left[\alpha^a(w_j)|w_i \geq t\right] \geq \mathbb{E}\left[\alpha^b(w_j)|w_i \geq t\right]$ for $j = 1, 2$ and for all t, thus $\mathbb{E}\left[\lambda_i^{a,*}|w_i = t\right] \leq \mathbb{E}\left[\lambda_i^{b,*}|w_i = t\right]$. □

9.4.3 Proof of Lemma 4

Proof. The case of random matching is discussed in the proof of Proposition 1 below, and so, here, we focus on the case of perfect assortative matching. Since the perfectly assortative distribution, i.e., $F(\boldsymbol{w}) = \min\{G(w_1), G(w_2)\}$, does not admit a density function, the regularity conditions of Proposition 14 are violated. To make progress, remark welfare $\mathcal{W}$ in (53) and surplus $\mathcal{S}$ in (54) depend on v, $\frac{\partial v}{\partial w_i}$ only through their values along the diagonal $w_1 = w_2 = w$. Letting $\widetilde{v}^m(w) := \frac{1}{2}v^m(w,w)$, we may rewrite these objects as

$$\mathcal{W} = \frac{\mu}{2}\int (1-\alpha)\,\widetilde{v}^m dG + (1-\mu)\int (1-\alpha)\,v^s dG + \int_\mu^1 \Phi(\varepsilon)d\varepsilon + \mathcal{S},$$

where

$$\mathcal{S} = \mu \int \sum_{i=1}^{2}\left(w^{1+\gamma}\left(\frac{\partial \widetilde{v}^m}{\partial w}\right)^\gamma - \gamma w \frac{\partial \widetilde{v}^m}{\partial w}\right) dF + (1-\mu)\int \left(w^{1+\gamma}\left(\frac{\partial v^s}{\partial w}\right)^\gamma - \gamma w\frac{\partial v^s}{\partial w}\right) dG.$$

Consider a further relaxation of the problem in which we directly select functions $\widetilde{v}^m$, v^s and a number μ to maximize $\mathcal{W}$. Repeating the argument that is used to establish Proposition 14, we obtain that the following optimal distortions:

$$\widetilde{\lambda}^{m,*}(t) = \lambda^{s,*}(t) = \frac{1-\mathbb{E}[\alpha|w\geq t]}{\gamma\theta(t)}.$$

By construction, $v^m(w_1,w_2) = \widetilde{v}^{m,*}(w_1) + \widetilde{v}^{m,*}(w_2)$ solves the planner's original relaxed problem. As a result, the optimal allocations with distortions $\lambda_i^{m,*}(\boldsymbol{w}) = \widetilde{\lambda}^{m,*}(w_i)$ can be implemented using separable taxation. □

9.4.4 Proof of Lemma 5

Proof. We start with a preliminary observation that will be useful to sign optimal jointnessss at the extremes. Set $F(w|t) := \Pr(w_{-i} \leq w | w_i = t)$ and unpack $\mathbb{E}\left[\lambda_i^{m,*}|w_i = t\right]$ conditioning on spouse $-i$ being more and less productive than spouse i to get

$$1 = \frac{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t \leq w_{-i}\right]}{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t\right]}(1-F(t|t)) + \frac{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t \leq w_{-i}\right]}{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t\right]}F(t|t). \tag{72}$$

Observe that $F(t|t) = \frac{1}{2}\frac{d(1-\Pr(\max \boldsymbol{w}\leq t))}{d(1-\Pr(w_i\leq t))}$. By L'Hôpital's rule and symmetry,

$$\lim_{t\to\infty} F(t|t) = \lim_{t\to\infty}\frac{1}{2}\frac{1-\Pr(\max\boldsymbol{w}\leq t)}{1-\Pr(w_i\leq t)} = 1 - \frac{1}{2}\lim_{t\to\infty}\Pr(w_{-i}\geq t|w_i\geq t).$$

For the Gaussian and FGM copulas, $\lim_{t\to\infty}\Pr(w_{-i}\geq t|w_i\geq t) = 0$, hence, by (72),

$$\lim_{t\to\infty}\frac{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t\leq w_{-i}\right]}{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t\right]} = 1.$$

As a result, Equation (20) implies

$$\lim_{t\to\infty} J(t) = \lim_{t\to\infty}\frac{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t\leq w_{-i}\right]}{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t\right]} - 1 = \frac{\lim_{t\to\infty}A(t)}{\lim_{t\to\infty}B(t)} - 1,$$

where $A(t) := \frac{1-\mathbb{E}[\alpha^m|\min \boldsymbol{w} \geq t]}{1-\mathbb{E}[\alpha^m|w_i \geq t]}$ and $B(t) := \frac{2\Pr(w_{-i} \geq t|w_i = t)}{\Pr(w_{-i} \geq t|w_i \geq t)}$.

For both copulas, the limit of $\mathbb{E}[\alpha^m|\boldsymbol{w} \geq (t,t)]$ equals to $\alpha(\infty) := \lim_{t\to\infty} \alpha(t)$. In the case of the Gaussian copula, the limit of $\mathbb{E}[\alpha^m|w_i \geq t]$ also equals to $\alpha(\infty)$ because α is continuous, bounded and $F(w|t)$ converges to the distribution that is degenerate at $w = \infty$, i.e., $\lim_{t\to\infty} F(w|t) = 0$ for all w. As a result, Term A converges to 1. It is well-known that the limit of Term B equals to $\frac{2}{1+\rho}$, see, e.g., Hua and Joe (2014). Since $\lim_{t\to\infty} J(t) < 0$, optimal jointness is negative for all sufficiently large values of t.

As for the FGM copula, since $F(w_1, w_2) = G(w_1)G(w_2)[1+\rho(1-G(w_1))(1-G(w_2))]$, direct computations give

$$\mathbb{E}[\alpha^m|w_i \geq t] = \frac{1}{2}\mathbb{E}\left[\alpha|w \geq t\right] + \frac{1}{2} - \frac{\rho}{2}G(t)\int \alpha d[G(1-G)]. \tag{73}$$

Since α is strictly decreasing, we have

$$\int \alpha d[G(1-G)] = -\int G(1-G)d\alpha > 0.$$

It follows from (73) that the limit of $\mathbb{E}[\alpha^m|w_i \geq t]$ is strictly smaller than $\frac{1}{2}\alpha(\infty) + \frac{1}{2}$, thus the limit of Term A is strictly less than 2. On the other hand, the reader can verify that Term B converges 2. We again obtain that $\lim_{t\to\infty} J(t) < 0$, and hence optimal jointness is negative for all sufficiently large values of t.

The case of the bottom left corner is similar to the previous one. First of all, remark that $\mathbb{E}\alpha = 1$, thus (18) can be equivalently expressed as

$$\mathbb{E}\left[\lambda_i^{m,*}|w_i = t\right] = \frac{\mathbb{E}[\alpha|w \leq t] - 1}{\gamma\underline{\theta}(t)},$$

where $\underline{\theta}(t) := \frac{\partial G(t)/\partial \ln t}{G(t)}$. Using the same argument as for the top corner, it can be shown that

$$\lim_{t\to 0} J(t) = \lim_{t\to 0} \frac{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t\right]}{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t \geq w_{-i}\right]} - 1 = \frac{\lim_{t\to 0} \underline{B}(t)}{\lim_{t\to 0} \underline{A}(t)} - 1,$$

where $\underline{A}(t) = \frac{\mathbb{E}[\alpha^m|\max \boldsymbol{w} \leq t]-1}{\mathbb{E}[\alpha^m|w_i \leq t]-1}$ and $\underline{B}(t) = \frac{2\Pr(w_{-i} \leq t|w_i = t)}{\Pr(w_{-i} \leq t|w_i \leq t)}$.

As before, these terms satisfy $\underline{A}(t) \to 1$ and $\underline{B}(t) \to \frac{2}{1+\rho}$ as $t \to 0$ for the Gaussian copula. In the case of the FGM copula, the limit of Term $\underline{A}$ is strictly less than 2, whereas Term $\underline{B}$ converges to 2. So, for both copulas, $\lim_{t\to\infty} J(t) > 0$, hence optimal jointness is positive for all sufficiently small values of t.

We now show the second part of the claim. Since f is symmetric and $\alpha^m(w_1, w_2) = \frac{1}{2}\alpha(w_1) + \frac{1}{2}\alpha(w_2)$, the equations characterizing $\mathbb{E}\left[\lambda_i^{m,*}|w_i = t \leq w_{-i}\right]$ and $\mathbb{E}\left[\lambda_i^{m,*}|w_i = t \geq w_{-i}\right]$,

(20) and (21), can be rewritten as

$$2\gamma t g\left(t\right)\mathbb{E}\left[\lambda_i^{m,*}|w_i = t \leq w_{-i}\right] = \frac{\int_t^\infty (1-\alpha(w))(1-F(t|w))dG(w)}{1-F(t|t)}, \quad (74)$$

$$2\gamma t g\left(t\right)\mathbb{E}\left[\lambda_i^{m,*}|w_i = t \geq w_{-i}\right] = \frac{\int_0^t (\alpha(w)-1)F(t|w)dG(w)}{F(t|t)}. \quad (75)$$

Set $\xi(t) = \int_0^t (\alpha(w) - 1)dG$. Since α is strictly decreasing and integrates to one, we have $\xi(t) > 0$ for all $t > 0$, $\xi(0) = \lim_{t\to\infty} \xi(t) = 0$, and $\xi(t) = \int_t^\infty (1-\alpha)dG$. Integrate by parts (74), (75) and use the properties of β to obtain

$$2\gamma t g\left(t\right)\mathbb{E}\left[\lambda_i^{m,*}|w_i = t \leq w_{-i}\right] = \xi(t) + \frac{\int_t^\infty \xi(w)[-F_2(t|w)]dw}{1-F(t|t)}, \quad (76)$$

$$2\gamma t g\left(t\right)\mathbb{E}\left[\lambda_i^{m,*}|w_i = t \geq w_{-i}\right] = \xi(t) + \frac{\int_0^t \xi(w)[-F_2(t|w)]dw}{F(t|t)}, \quad (77)$$

where F_2 stays for the partial derivative of the conditional cdf with respect to its second argument, i.e., $F_2(t|w) = \frac{\partial}{\partial w}\Pr\left(w_{-i} \leq t|w_i = w\right)$.

For the FGM copula with $\rho > 0$, we have

$$\begin{aligned} F(t|w) &= G(t)[1+\rho(1-G(t))(1-2G(w))], \\ -F_2(t|w) &= 2\rho G(t)(1-G(t))g(w). \end{aligned}$$

It follows from (76), (77) that the sign of $J(t)$ coincides with the sign of

$$1+\rho(1-G(t))(1-2G(t)) - \frac{\int_0^t \xi dG\,/G(t)}{\int_0^\infty \xi dG}. \quad (78)$$

Clearly, the first term in (78) equals to $1+\rho > 1$ at $t = 0$ and converges to 1 as $t \to \infty$. Note that

$$\frac{\partial}{\partial t}[(1-G(t))(1-2G(t))] = \rho\left(4G\left(t\right)-3\right)g\left(t\right) \gtrless 0 \iff G\left(t\right) \gtrless \tfrac{3}{4}.$$

Since $(1-G(t))(1-2G(t)) \gtrless 0$ if and only if $G(t) \lessgtr \frac{1}{2}$, the first term in (78) first monotonically decreases to some number below 1 and then monotonically increases to 1.

We now look at the second term in (78). It equals to $\xi(0) = 0$ at $t = 0$ and converges to 1 as $t \to \infty$. Using the definition of ξ, we can show

$$\frac{\int_0^t \xi dG}{G(t)} = \int_0^t \left(1 - \frac{G(w)}{G(t)}\right)(\alpha(w)-1)dG(w).$$

Since α is strictly decreases and integrates to one, the second term in (78) first increases to some number above 1 and then monotonically decreases to 1. Taking both pieces together, we conclude that there is a unique threshold $\widehat{t}$ such that

$$J(t) \gtreqless 0 \quad \Longleftrightarrow \quad 1+\rho(1-G(t))(1-2G(t)) - \frac{\int_0^t \xi dG \,/G(t)}{\int_0^\infty \xi dG} \gtreqless 0 \quad \Longleftrightarrow \quad t \lesseqgtr \widehat{t},$$

which concludes the proof. □

9.4.5 Proof of Lemma 6

Proof. Since $\mathbb{E}\alpha^m = 1$, Equation (26) can be rewritten using Bayes' rule as

$$\mathbb{E}\left[\lambda_{sec}^{m,*} - \lambda_{pr}^{m,*} | I = \iota\right] = \frac{\mathbb{E}\left[\alpha^m | I \leq \iota\right] - 1}{\gamma \frac{\partial \Pr(I\leq\iota)/\partial \ln\iota}{\Pr(I\leq\iota)}}. \tag{79}$$

For every value of ι, the weights α^m satisfy

$$\begin{aligned}\mathbb{E}\left[\alpha^m | I \leq \iota\right] &= \frac{1}{2}\mathbb{E}\left[\alpha\left(\max \boldsymbol{w}\right)| I \leq \iota\right] + \frac{1}{2}\mathbb{E}\left[\alpha\left(\max \boldsymbol{w}\cdot I\right)| I \leq \iota\right] \\ &\geq \frac{1}{2}\mathbb{E}\left[\alpha\left(\max \boldsymbol{w}\cdot I\right)| I \leq \iota\right] = \frac{1}{2}\mathbb{E}\left[\alpha\left(w_{-i}\right)| w_{-i} \leq \iota w_i\right],\end{aligned}$$

where the last equality is due to symmetry of F. The assumption ensures that w_{-i} conditional on $w_{-i} \leq \iota w_i$ converges almost surely to 0 as $\iota \to 0$. To see it note that $\mathbb{E}\left[w_{-i} | w_{-i} \leq \iota w_i\right] \leq \iota \mathbb{E}\left[w_i | w_{-i} \leq \iota w_i\right] \to_{\iota\to\infty} 0$. Since $\alpha(0) > 2$, $\mathbb{E}\left[\alpha^m | I \leq \iota\right] > 1$ for all small values of ι, thus $\mathbb{E}\left[\lambda_{sec}^{m,*} - \lambda_{pr}^{m,*} | I = \iota\right] > 0$ for all sufficiently small ι due to (79). □

9.5 Optimal taxation with random matching and the validity of the FOA

9.5.1 Proof of Proposition 1

Proof. The reader can verify that that $\left(\frac{1}{2}\lambda^{s,*}(w_1), \frac{1}{2}\lambda^{s,*}(w_2)\right)$, where $\lambda^{s,*}$ is given by Equation (12), satisfies conditions (A1)-(A5) of Proposition 2 provided that $\frac{1-G}{tg}$ and $\frac{G}{tg}$ converge to finite limits as $t \to \infty$ and $t \to 0$, respectively, and both are implied by finitness of $\lim_{t\to 0,\infty} \lambda^{s,*}(t)$. Moreover, these distortions also verify the necessary and sufficient optimality conditions listed in this proposition, i.e., (56) and (57). It follows that they characterize the solution to the relaxed problem.

By Proposition 2 in Rochet (1987), the first-order approach is valid if and only if $v^{s,*}$ and $v^{m,*}$ are convex functions of $x = w^{-1/\gamma}$. It is easy to see that $\frac{\partial v^{s,*}(x^{-\gamma})}{\partial x}$ and $\frac{\partial v^{m,*}\left(x_1^{-\gamma}, x_2^{-\gamma}\right)}{\partial x_i}$ are monotone transformations of $x \cdot (1 + \lambda^{s,*}(x^{-\gamma}))$ and $x_i \cdot \left(1 + \frac{1}{2}\lambda^{s,*}\left(x_i^{-\gamma}\right)\right)$, respectively. The fact that the first-order approach is more likely to hold in the bi-dimensional model than in the uni-dimensional setting can be seen from

$$x \cdot \left(1 + \frac{1}{2}\lambda^{s,*}\left(x^{-\gamma}\right)\right) - \widehat{x} \cdot \left(1 + \frac{1}{2}\lambda^{s,*}\left(\widehat{x}^{-\gamma}\right)\right) \geq \frac{x}{2} \cdot \left(1 + \lambda^{s,*}\left(x^{-\gamma}\right)\right) - \frac{\widehat{x}}{2} \cdot \left(1 + \lambda^{s,*}\left(\widehat{x}^{-\gamma}\right)\right) \quad \forall x \geq \widehat{x}.$$

□

10 Proofs for Section 5

Throughout this section, it is assumed that the first-order approach is valid. To streamline exposition, we also assume that the optimal distortions, $\lambda^{s,*}$ and $\boldsymbol{\lambda}^{m,*}$, are well-behaved in the sense the sense that satisfy analogs of (A1)-(A5) listed in Proposition 2.

10.1 The role of Pareto weights

Remark that (32) and (33) in the main text follow from the optimality conditions in 2. To obtain the former equation integrate (56) from $w = t$ to ∞; and, to obtain the latter one, apply the coarea formula with some Q to (57).

10.1.1 Proof of Corollary 2.

Setting $Q = \min \boldsymbol{w}$ and $Q = \max \boldsymbol{w}$ in (33), we obtain the following generalizations of (20) and (21):

$$\mathbb{E}\left[\lambda_i^{m,*}|w_i = t \geq w_{-i}\right] = \frac{\Pr\left(w_{-i} \geq t|w_i \geq t\right)}{2\Pr\left(w_{-i} \geq t|w_i = t\right)} \frac{\mathbb{E}\left[\frac{\alpha^m}{\mathbb{E}\alpha^m}|\max \boldsymbol{w} \leq t\right] - 1}{\gamma\theta(t)} \frac{\mathbb{E}\alpha^m}{\left(1-\mu^*\right)\mathbb{E}\alpha^s + \mu^*\mathbb{E}\alpha^m} \tag{80}$$

and

$$\mathbb{E}\left[\lambda_i^{m,*}|w_i = t \leq w_{-i}\right] = \frac{\Pr\left(w_{-i} \leq t|w_i \leq t\right)}{2\Pr\left(w_{-i} \leq t|w_i = t\right)} \frac{1 - \mathbb{E}\left[\frac{\alpha^m}{\mathbb{E}\alpha^m}|\min \boldsymbol{w} \geq t\right]}{\gamma\underline{\theta}(t)} \frac{\mathbb{E}\alpha^m}{\left(1-\mu^*\right)\mathbb{E}\alpha^s + \mu^*\mathbb{E}\alpha^m} \tag{81}$$

These equations are similar to the benchmark but allow for general Pareto weights.

Suppose first that the matching is random, i.e., $F = G^2$. Then, the ratios of probabilities in (80), (81) equal to $\frac{1}{2}$, hence

$$J(t) = \frac{\mathbb{E}\alpha^m - \mathbb{E}[\alpha^m|\min \boldsymbol{w} \geq t]}{\mathbb{E}[\alpha^m|\max \boldsymbol{w} \leq t] - \mathbb{E}\alpha^m} \frac{1-G(t)}{G(t)} - 1.$$

Using Bayes' rule, it is easy to see that this measure of jointness satisfies

$$J(t) \geq 0 \quad \Longleftrightarrow \quad \frac{2\left(\mathbb{E}[\alpha^m|w_i \leq t] - \mathbb{E}\alpha^m\right)}{\mathbb{E}[\alpha^m|\max \boldsymbol{w} \leq t] - \mathbb{E}\alpha^m} \geq 1.$$

The reader can verify that supermodularity of α^m is preserved in the sense that $\mathbb{E}[\alpha^m|\boldsymbol{w} \leq \boldsymbol{t}]$ is supermodular in $\boldsymbol{t}$ when α^m is supermodular. As a result, if α^m is supermodular, then

$$\mathbb{E}[\alpha^m|\max \boldsymbol{w} \leq t] - \mathbb{E}\alpha^m \geq 2\left(\mathbb{E}[\alpha^m|w_i \leq t] - \mathbb{E}\alpha^m\right).$$

Since $\mathbb{E}[\alpha^m|w_i = t]$ is strictly decreasing, $\mathbb{E}[\alpha^m|w_i \leq t] - \mathbb{E}\alpha^m > 0$. Combining all pieces together, we conclude that optimal jointness is negative for supermodular weights. The same argument can be used to establish that optimal jointnessss is positive for submodular weigths.

Suppose now F is characterized by the Gaussian copula with the correlation parameter $\rho > 0$. Recollect that the optimal distortions satisfy (80), (81). Following the same steps as in the proof of Lemma 5, we obtain $\lim_{t\to\infty} J(t) = \frac{1+\rho}{2} - 1 < 0$ and $\lim_{t\to 0} J(t) = \frac{2}{1+\rho} - 1 > 0$.

10.1.2 Relationship to Kleven et al. (2007)

In the working version of their paper, Kleven et al. (2007) (KKS for short) outlined how jointness can be signed at each productivity vector when all households are married, their productivities are independent and the social welfare criterion is given by $\int W(v^m)dF$, where W is a strictly increasing and strictly concave function. KKS used a stronger pointwise notion of jointness: a tax function T^m is *positively jointed at* $\boldsymbol{y}$ if $\frac{\partial^2}{\partial y_1 \partial y_2} T^m(\boldsymbol{y}) \geq 0$ and *negatively jointed at* $\boldsymbol{y}$ if $\frac{\partial^2}{\partial y_1 \partial y_2} T^m(\boldsymbol{y}) \leq 0$. This notion of jointness is closely related to the sign of cross-partials of spousal distortions. In particular, using the envelope condition and the definition of distortions (Equation (10)), one can show that, for each couple $\boldsymbol{w} \in \mathbb{R}^2_{++}$, taxes T^m are positively (or negatively) jointed at $\boldsymbol{y} = \boldsymbol{y}^{m,*}(\boldsymbol{w})$ when $\frac{\partial}{\partial w_{-i}} \lambda_i^{m,*}(\boldsymbol{w}) \geq 0$ (or $\frac{\partial}{\partial w_{-i}} \lambda_i^{m,*}(\boldsymbol{w}) \leq 0$). Here, since the signs $\frac{\partial}{\partial w_2} \lambda_1^{m,*}(\boldsymbol{w})$ and $\frac{\partial}{\partial w_1} \lambda_2^{m,*}(\boldsymbol{w})$ necessarily agree, we may take any spouse i's distortion. KKS argued that jointness in this pointwise sense is connected to the third derivative of the welfare function. Assuming that the optimal allocations are sufficiently smooth and regular, KKS showed that $\frac{\partial}{\partial w_{-i}} \lambda_i^{m,*}(\boldsymbol{w}) \geq 0$ (or $\frac{\partial}{\partial w_{-i}} \lambda_i^{m,*}(\boldsymbol{w}) \leq 0$) for all $\boldsymbol{w} \in \mathbb{R}^{\mathbf{2}}_{++}$ when the third derivative of W is strictly positive (or negative).

Building on the argument of KKS, we can strengthen Part (a) of Corollary 2 to the pointwise notion of jointness, separately for each type $\boldsymbol{w}$. In order to state the result formally, we require that g, α^m are twice continuously differentiable. We also assume that $v^{m,*}$ is three time continuously differentiable, the corresponding distortions are twice continuously differentiable with third derivatives that exist at least in the weak sense, and $\left\| \nabla \frac{\partial^2 v^{m,*}}{\partial w_1 \partial w_2} \right\| \neq 0$ when $\frac{\partial}{\partial w_{-i}} \lambda_i^{m,*} \neq 0$. Finally, we assume that the integral $\int_U \nabla \cdot \boldsymbol{\Lambda} d\boldsymbol{w}$, where $\boldsymbol{\Lambda} = (\Lambda_1, \Lambda_2)$ is defined by $\Lambda_i = \frac{\partial^2 \left(\lambda_i^{m,*} \gamma w_i g(w_i)\right) / \partial w_1 \partial w_2}{g(w_i)}$, can be transformed using the divergence theorem for any potentially unbounded open set U that has a Lipshitz boundary. The last assumption is satisfied when the optimal distortions are smooth, their third derivatives don't explode when productivities go to 0 and vanish sufficiently fast when productivities go to ∞.

Proposition 3. *Consider a general economy as described in Section 5.1 in which matching is random. For each* $\boldsymbol{w} \in \mathbb{R}^2_{++}$*, the optimal distortions for married satisfy*

$$\frac{\partial}{\partial w_{-i}} \lambda_i^{m,*}(\boldsymbol{w}) \frac{\partial^2}{\partial w_1 \partial w_2} \alpha^m(\boldsymbol{w}) \leq 0.$$

Therefore, the optimal tax schedule on married $T^{m,*}$ is positively (negatively) jointed at every $\boldsymbol{y}$ when the social weights on couples are strictly submodular (supermodular).

Proof. As explained in Proposition 2, under random matching, the optimal marginal taxes satisfy

$$\sum_{i=1}^{2} \frac{\partial \left(\lambda_i^{m,*} \gamma w_i g\left(w_i\right)\right) / \partial w_i}{g\left(w_i\right)} = \eta \left(\alpha^m - \mathbb{E}\alpha^m\right),$$

where $\eta = (\mu \mathbb{E}\alpha^m + (1-\mu)\mathbb{E}\alpha^s)^{-1}$. Differentiate this expression twice to obtain

$$\sum_{i=1}^{2} \nabla \cdot \boldsymbol{\Lambda} = \eta \frac{\partial^2 \alpha^m}{\partial w_1 \partial w_2}. \tag{82}$$

Recollect that the optimal distortions are related to the marital surplus via (10) that gives the following relationship between the cross partial of $v^{m,*}$ and $\frac{\partial \lambda_1^{m,*}}{\partial w_2}$, $\frac{\partial \lambda_2^{m,*}}{\partial w_1}$:

$$\frac{\partial^2 v^{m,*}}{\partial w_1 \partial w_2} = -\frac{1}{1-\gamma} \frac{w_1^{\frac{\gamma}{1-\gamma}}}{(1+\lambda_1^{m,*})^{\frac{2-\gamma}{1-\gamma}}} \frac{\partial \lambda_1^{m,*}}{\partial w_2} = -\frac{1}{1-\gamma} \frac{w_2^{\frac{\gamma}{1-\gamma}}}{(1+\lambda_2^{m,*})^{\frac{2-\gamma}{1-\gamma}}} \frac{\partial \lambda_2^{m,*}}{\partial w_1} \tag{83}$$

As shown by (83), the sign of $\frac{\partial^2 v^{m,*}}{\partial w_1 \partial w_2}$ is opposite to the sign of jointness when $w_1, w_2 > 0$. Furthermore, if $w_{-i} = 0 < w_i$, then this equation implies $\frac{\partial^2 v^{m,*}}{\partial w_1 \partial w_2} = \frac{\partial \lambda_i^{m,*}}{\partial w_{-i}} = 0$; however, $\frac{\partial \lambda_i^{m,*}}{\partial w_{-i}}$ may be nonzero on $\{w_i = 0\}$.

Let $U^- = \left\{\frac{\partial \lambda_1^{m,*}}{\partial w_2} \text{ and } \frac{\partial \lambda_2^{m,*}}{\partial w_1} < 0\right\}$ and $U^+ = \left\{\frac{\partial \lambda_1^{m,*}}{\partial w_2} \text{ and } \frac{\partial \lambda_2^{m,*}}{\partial w_1} > 0\right\}$ be productivities $\boldsymbol{w} \in \mathbb{R}_{++}^2$ at which both spouses' taxes are strictly negatively and positively jointed, respectively. Since $\left\|\nabla \frac{\partial^2 v^{m,*}}{\partial w_1 \partial w_2}\right\| \neq 0$ at each $\boldsymbol{w} \in \mathbb{R}_{++}^2$ in $\partial U^- \cup \partial U^+$, using (83), the reader can verify that under this assumption, ∂U^- and ∂U^+ are Lipshitz with $\boldsymbol{n}$ proportional to $\nabla \frac{\partial^2 v^{m,*}}{\partial w_1 \partial w_2}$ and $-\nabla \frac{\partial^2 v^{m,*}}{\partial w_1 \partial w_2}$, respectively.

Consider first the set U^-. It may have several connected components, so let U to be one of them. Using the definition of Λ_i, we can be unpack it as

$$\Lambda_i = \frac{\partial^2 \left(\lambda_i^{m,*} \gamma w_i g\left(w_i\right)\right) / \partial w_1 \partial w_2}{g\left(w_i\right)} = \gamma w_i \frac{\partial^2 \lambda_i^{m,*}}{\partial w_1 \partial w_2} + \gamma \frac{\partial (g(w_i) w_i) / \partial w_i}{g(w_i)} \frac{\partial \lambda_i^{m,*}}{\partial w_{-i}}. \tag{84}$$

Due to our assumption on $\boldsymbol{\Lambda}$, Equation (82) can be transformed using the divergence theorem, which gives

$$\eta \int_U \frac{\partial^2 \alpha^m}{\partial w_1 \partial w_2} d\boldsymbol{w} = \int_{U \cap \partial \mathbb{R}_{++}^2} \boldsymbol{\Lambda} \cdot \boldsymbol{n} d\sigma + \int_{\partial U \cap \mathbb{R}_{++}^2} \boldsymbol{\Lambda} \cdot \boldsymbol{n} d\sigma, \tag{85}$$

where, as usual, $\boldsymbol{n}$ stays for the outward unit normal vector.

Let $\boldsymbol{w} \in U \cap \partial\mathbb{R}^2_{++}$ be such that $w_i = 0 < w_{-i}$. Then, $n_i(\boldsymbol{w}) = 0$ and $n_{-i}(\boldsymbol{w}) = -1$, hence (84) gives

$$\boldsymbol{\Lambda} \cdot \boldsymbol{n} = -\Lambda_i = -\gamma \left(\lim_{w_i \to 0} \frac{\partial(g(w_i)w_i)/\partial w_i}{g(w_i)} \right) \frac{\partial \lambda_i^{m,*}}{\partial w_{-i}} \tag{86}$$

By L'Hopital's rule, $\lim_{w\to 0} \frac{g(w)w}{G(w)} = \lim_{w\to 0} \frac{\partial(g(w)w)/\partial w}{g(w)} \geq 0$, and, by definition of U, we have $\frac{\partial \lambda_i^{m,*}}{\partial w_{-i}} < 0$. It follows that the right-hand side in (86) is nonnegative, and so is $\int_{U\cap\partial\mathbb{R}^2_{++}} \boldsymbol{\Lambda}\cdot\boldsymbol{n}d\sigma$ in (85).

Now, let $\boldsymbol{w} \in \partial U \cap \mathbb{R}^2_{++}$. Remark that we necessarily have $\frac{\partial^2 v^{m,*}}{\partial w_1 \partial w_2} = \frac{\partial \lambda_1^{m,*}}{\partial w_2} = \frac{\partial \lambda_2^{m,*}}{\partial w_1} = 0$; furthermore, by (83), the i-th component of the outward unit normal is proportional to

$$-\frac{\partial}{\partial w_i}\frac{\partial^2 v^{m,*}}{\partial w_1 \partial w_2} = \frac{1}{1-\gamma}\frac{w_i^{\frac{\gamma}{1-\gamma}}}{(1+\lambda_i^{m,*})^{\frac{2-\gamma}{1-\gamma}}}\frac{\partial \lambda_i^{m,*}}{\partial w_1 \partial w_2}.$$

Hence, $\int_{\partial U\cap\mathbb{R}^2_{++}} \boldsymbol{\Lambda}\cdot\boldsymbol{n}d\sigma \geq 0$ in (85) due to (84).

Taking both cases together, we conclude that the left-hand side in (86) is nonnegative for every connected component U of U^-. The same argument can be used to establish that the left-hand side in (86) is nonpositive for every connected component U of U^+. These two facts and (83) imply

$$\frac{\partial^2}{\partial w_1 \partial w_2} v^{m,*}(\boldsymbol{w}) \frac{\partial^2}{\partial w_1 \partial w_2} \alpha^m(\boldsymbol{w}) \geq 0 \;\; \forall \boldsymbol{w}. \tag{87}$$

To sum up, Equations (87), (83) show that the sign of jointness is completely pinned down by modularity of α^m for couples with $w_1, w_2 > 0$. In particular, if $\frac{\partial^2 \alpha^m}{\partial w_1 \partial w_2} > 0$ ($\frac{\partial^2 \alpha^m}{\partial w_1 \partial w_2} < 0$), then the optimal taxes are negatively (positively) pointwise jointed at every $\boldsymbol{y} = \boldsymbol{y}^{m,*}(\boldsymbol{w})$ with $\boldsymbol{w} \in \mathbb{R}^2_{++}$. And, this conclusion extends to all $\boldsymbol{w} \in \mathbb{R}^2_+$ provided that $\frac{\partial}{\partial y_1 \partial y_2} T^{m,*}$ and $\boldsymbol{y}^{m,*}$ are continuous. $\square$

This proposition establishes the tight pointwise relationship between jointness of the optimal taxes on married and the modularity of couples' Pareto weights α^m in our framework in which α^m is exogenously specified. It turns out that the same point can be made in the setting of KKS because the third derivative of the social welfare function W is closely related to the modularity of implied weights. In particular, set $\alpha^m := W'(v^{m,*})$, where $v^{m,*}$ is the optimum when social welfare is given by $\int W(v^m)dF$. This function α^m should be though of as Pareto weights that the planner assigns to utilities of various couples; and, it is easy to see that, $v^{m,*}$ is also optimal when social welfare is given by $\int \alpha^m v^m dF$. Differentiating twice, we can rewrite (87) for these specific weights $\alpha^m = W'(v^{m,*})$ as

$$\frac{\partial^2 v^{m,*}}{\partial w_1 \partial w_2}\frac{\partial^2 \alpha^m}{\partial w_1 \partial w_2} = \frac{\partial v^{m,*}}{\partial w_1}\frac{\partial v^{m,*}}{\partial w_2}\frac{\partial^2 v^{m,*}}{\partial w_1 \partial w_2} W'''(v^{m,*}) + \left(\frac{\partial^2 v^{m,*}}{\partial w_1 \partial w_2}\right)^2 W''(v^{m,*}).$$

Due to concavity of W, the second term is nonpositive, and we arrive at

$$\frac{\partial^2}{\partial w_1 \partial w_2} v^{m,*}(\boldsymbol{w}) W^{'''}(v^{m,*}(\boldsymbol{w})) \geq 0 \ \ \forall \boldsymbol{w}. \tag{88}$$

Equations (87) and (88) show the one-to-one relationship between the sign of $W^{'''}$ and the modularity of the implied weights.

10.2 Public goods and economies of scale

This model is different from the benchmark in two ways. First, the relationship between the optimal distortions, which are still defined in Equation (10), and derivatives of $v^{s,*}$, $v^{m,*}$ now becomes

$$\lambda^{s,*}(w) := k^s \left(\frac{\partial v^{s,*}(w)}{\partial w} \right)^{\gamma - 1} w^{\gamma} - 1, \quad \lambda_i^{m,*}(\mathbf{w}) := k^m \left(\frac{\partial v^{m,*}(\mathbf{w})}{\partial w_i} \right)^{\gamma - 1} w_i^{\gamma} - 1.$$

Second, the resource constraint (9) now reads as

$$\mathcal{S} \geq \frac{\mu}{2} \int \frac{v^m}{k^m} dF + (1-\mu) \int \frac{v^s}{k^s} dG,$$

where $\mathcal{S}$ is our shorthand notation for the total economic output, that is now given by

$$\frac{\mu}{2} \int \sum_{i=1}^{2} \left(w_i^{1+\gamma} \left(\frac{\partial v^m}{\partial w_i} \right)^{\gamma} - \gamma w_i \frac{\partial v^m / \partial w_i}{k^m} \right) dF + (1-\mu) \int \left(w^{1+\gamma} \left(\frac{\partial v^s}{\partial w} \right)^{\gamma} - \gamma w \frac{\partial v^s / \partial w}{k^s} \right) dG.$$

Then, (5) and the modified resource constraint give

$$\int v^s dG = \frac{\mathcal{S} - \mu \Phi(\mu)/k^m}{\mu/k^m + (1-\mu)/k^s}, \ \ \frac{1}{2} \int v^m dF = \frac{\mathcal{S} + (1-\mu)\Phi(\mu)/k^s}{\mu/k^m + (1-\mu)/k^s}.$$

Following the same steps as in Section 9.1, welfare with additively separable weights, i.e., $\alpha^m(w_1, w_2) = \frac{\alpha(w_1)+\alpha(w_2)}{2}$, can be expressed as

$$\mathcal{W} = \frac{\mu}{2} \int (\alpha^m - 1) v^m dF + (1-\mu) \int (\alpha - 1) v^s dG + \int_{\mu}^{1} \Phi(\varepsilon) d\varepsilon + \frac{1}{\eta} \mathcal{S},$$

where $\eta = \mu/k^m + (1-\mu)/k^s$.

The argument in the proof of Proposition 2 gives that the following modified differential equations hold at the optimum:

$$\begin{aligned} \frac{\partial \left(\gamma w \lambda^{s,*} g \right)}{\partial w} &= \frac{\mu^*/k^m + (1-\mu^*)/k^s}{1/k^s} (\alpha - 1) g, \\ \sum_{i=1}^{2} \frac{\partial \left(\gamma w_i \lambda_i^{m,*} f \right)}{\partial w_i} &= \frac{\mu^*/k^m + (1-\mu^*)/k^s}{1/k^m} (\alpha^m - 1) f. \end{aligned}$$

Two expressions (34), (35) follow from integrating the first equation and applying the coarea formula with $Q = w_i$ to the second equation.

10.3 Home production and division of labor within families

We first list several useful properties of functions N^s and N^m. These properties will be used later on to show that the mechanism design problem is well-behaved and its optimal distortions satisfy (36) and (37). Recollect that the elasticity of labor supply for single households $\widetilde{e}^s$ and its transformation $\widetilde{\gamma}^s$ satisfy

$$\widetilde{\gamma}^s = \left(1 + \frac{\partial \ln\left(\partial N^s/\partial l\right)}{\partial \ln l}\right)^{-1} = \left(1 + (\widetilde{e}^s)^{-1}\right)^{-1}. \quad (89)$$

Similarly, the 2×2 elasticity matrix of labor supplies for married households and its transformation $\widetilde{\gamma}^m$ satisfy

$$\widetilde{\gamma}^m = \begin{bmatrix} 1 + \frac{\partial ln\left(\partial N_1^m/\partial l_1\right)}{\partial l_1} & \frac{\partial ln\left(\partial N_1^m/\partial l_1\right)}{\partial l_2} \\ \frac{\partial ln\left(\partial N_2^m/\partial l_1\right)}{\partial l_1} & 1 + \frac{\partial ln\left(\partial N_2^m/\partial l_2\right)}{\partial l_2} \end{bmatrix}^{-1} = \left(I + (\widetilde{e}^m)^{-1}\right)^{-1}. \quad (90)$$

Lemma 10. *(a) N^s and N^m are increasing and convex, (b) there exists some $\overline{x} > 0$ such that $\frac{\partial N^s(l)}{\partial l}, \frac{\partial N^m(\boldsymbol{l})}{\partial l_i}\Big|_{l_i=l} \leq l^{p-1}\left(l^p + \overline{x}\right)^{(1-p\gamma)/p\gamma}$ for all l, (c) $l\frac{\partial N^s}{\partial l}$ and $\left(l_1\frac{\partial N^m}{\partial l_1}, l_2\frac{\partial N^m}{\partial l_2}\right)$ are one-to-one on $\mathbb{R}_{++}$ and $\mathbb{R}^2_{++}$, (d) $\widetilde{\gamma}^s$, $\widetilde{\gamma}^m$ are uniformly bounded, (e) $\lim_{l\to\infty} \widetilde{\gamma}^s(l), \widetilde{\gamma}^m_{i,i}(\boldsymbol{l})\Big|_{l_i=l} = \gamma$ and $\lim_{l\to\infty} \widetilde{\gamma}^m_{i,-i}(\boldsymbol{l})\Big|_{l_i=l} = 0$, where convergence is uniform in l_{-i}.*

Proof. Recollect that N^s is defined as a minimum of a function that is jointly convex in (l, x), thus N^s is convex as well. The first order condition w.r.t. x can be expressed as

$$x^{\sigma+(1-\gamma)/\gamma} = (1-m)^{(1-p\gamma)/p\gamma}, \quad (91)$$

where $m := \frac{l^p}{x^p+l^p}$. This condition is also sufficient due to convexity. By the envelope Theorem, the derivative of N^s is given by

$$\frac{\partial N^s(l)}{\partial l} = l^{p-1}\left(x^p + l^p\right)^{(1-p\gamma)/p\gamma}. \quad (92)$$

It is immediate from (92) that N^s is strictly increasing which proves (a). (91) implies that $x \leq 1$ which proves (b). Part (c) follows from the previous observation and (92).

Totally differentiate (91) to obtain

$$\left[\frac{1-\gamma}{\gamma} - \frac{1-p\gamma}{\gamma}m + \sigma\right] d\ln x = -\frac{1-p\gamma}{\gamma} m d\ln l. \quad (93)$$

The term in the square brackets in (93) is bounded from below by $p - 1 + \sigma > 0$, thus the derivative of x is uniformly bounded. Then, totally differentiate (92) to obtain

$$d\ln\left(\frac{\partial N^s(l)}{\partial l}\right) = \left(p - 1 + \frac{1-p\gamma}{\gamma}m\right) d\ln l + \frac{1-p\gamma}{\gamma}(1-m) d\ln x.$$

By convexity and monotonicity of N^s, $\frac{\partial ln(\partial N^s(l)/\partial l)}{\partial l}$ is nonnegative. It follows that $\widetilde{\gamma}^s$ defined in (89) is uniformly bounded which proves (d). Finally, since the derivative of x is uniformly bounded and m goes to 1 as $l \to \infty$, $\widetilde{\gamma}^s$ converges to γ as $l \to \infty$.

Most of the arguments for married are analogous, i.e., N^m is convex due to joint convexity in $(\boldsymbol{l}, \boldsymbol{x})$. The first order condition w.r.t. x_i can be expressed as

$$x_i^{\sigma+(1-\gamma)/\gamma} = 2(1-r_i)^{1-q(1-\sigma)}(1-m_i)^{(1-p\gamma)/p\gamma}, \tag{94}$$

where $m_i := \frac{l_i^p}{x_i^p + l_i^p}$ and $r_i := \frac{x_{-i}^{1/q}}{x_1^{1/q}+x_2^{1/q}}$ for $i = 1, 2$. Again, (94) is sufficient due to convexity. By the envelope theorem, the derivative of N^m w.r.t. l_i is given by

$$\frac{\partial N^m(\boldsymbol{l})}{\partial l_i} = l_i^{p-1}\left(x_i^p + l_i^p\right)^{(1-p\gamma)/p\gamma}, \tag{95}$$

which proves (a). Equation (91) implies that $x \leq 2^{-\sigma-(1-\gamma)/\gamma}$ which proves (b). Part (c) follows from the previous observation and (96).

Totally differentiate (94) to obtain

$$\left[\frac{1-\gamma}{\gamma} - \frac{1-p\gamma}{\gamma}m_i - \frac{1-q(1-\sigma)}{q}r_i + \sigma\right] d\ln x_i = -\frac{1-q(1-\sigma)}{q}r_i d\ln x_j - \frac{1-p\gamma}{\gamma}m_i d\ln l_i. \tag{96}$$

The term in the square brackets in (96) is bounded from below by $p - 1 + \frac{1-q}{q} > 0$, thus the derivative of $\boldsymbol{x}$ is uniformly bounded. Then, totally differentiate (95) to obtain

$$d\ln\left(\frac{\partial N^m(l)}{\partial l_i}\right) = \left(p - 1 + \frac{1-p\gamma}{\gamma}m_i\right) d\ln l_i + \frac{1-p\gamma}{\gamma}(1-m_i)d\ln x_i.$$

By monotonicity and convexity of N^m, the determinant of the matrix in the square brackets in (90) is at least 1. Then, uniform boundedness of derivatives of $\boldsymbol{x}$ implies that $\widetilde{\gamma}^m$ is uniformly bounded as well which proves (d). Finally, observe that, by (94), m_i goes to 1 as $l_i \to \infty$ uniformly in l_j, which implies that $\widetilde{\gamma}^m_{i,i}(\boldsymbol{l})$ converges to γ and $\widetilde{\gamma}^m_{i,i}(\boldsymbol{l})$ converges to 0 as $l_i \to \infty$, where convergence is uniform in the other spouse labor supply. This shows (e) and concludes proof. □

Equipped with Lemma 10, we can formally establish the results in Section 5.3. Here, the notion of welfare is exactly the same as in the benchmark. To simplify exposition, it is convenient to define two auxiliary functions $\psi^s(l) := \frac{\partial N^s}{\partial \ln l}$ and $\psi_i^m := \frac{\partial N^m}{\partial \ln l_i}$. Then, the fist part of Lemma 9 extends, and the local incentive constraints can be succinctly expressed as

$$w\frac{\partial v^s}{\partial w} = \psi^s\left(\frac{y^s}{w}\right), \quad w_i\frac{\partial v^m}{\partial w_i} = \psi_i^m\left(\frac{y_1^m}{w_1}, \frac{y_2^m}{w_2}\right), \tag{97}$$

which is the exact analog of (8). To ensure properties (a), (b) of this lemma, we need to make certain assumptions. First, the expected first-best economic surplus from singles is finite, i.e.,

$$\int \max_{y\geq 0}\left(y - N^s\left(\frac{y}{w}\right)\right) dG < \infty;$$

second, the maximal expected economic surplus from singles diverges to $-\infty$ when the expected value of $\int \psi^s\left(\frac{y^s(w)}{w}\right) dG$ goes to $+\infty$. We impose the same assumptions on married with w_1, w_2 replaced by the maximum of spousal productivities. As a result, (a), (b) of Lemma 10 are implied by incentive compatibility, therefore we can formulate our relaxed problem in the same functional spaces, $\mathscr{V}^s$ and $\mathscr{V}^m$.

Finally, by Lemma 10, the local incentive constraints (97) can be inverted to solve for earnings as a function of derivatives v^s and v^m. Let $\phi^s := (\psi^s)^{-1}$ and $\phi^m := (\psi^m)^{-1}$. The relaxed problem is the same as in Section 9.1 but $\mathcal{S}$ is given by

$$\frac{\mu}{2}\int\left(\sum_{i=1}^{2} w_i\phi_i^m\left(w_1\frac{\partial v^m}{\partial w_1}, w_2\frac{\partial v^m}{\partial w_2}\right) - N^m\left[\boldsymbol{\phi}^m\left(w_1\frac{\partial v^m}{\partial w_1}, w_2\frac{\partial v^m}{\partial w_2}\right)\right]\right) dF + \\ + (1-\mu)\int\left(w\phi^s\left(w\frac{\partial v^s}{\partial w}\right) - N^s\left[\phi^s\left(w\frac{\partial v^s}{\partial w}\right)\right]\right) dG. \quad (98)$$

We now study the relaxed problem along the lines of the proof of Proposition 2. First of all, let λ^s and $\boldsymbol{\lambda^m}$ be defined as a function of marginal taxes, (10). It is straightforward to show that they are related to derivatives v^s and v^m by

$$\lambda^s = \frac{\phi^s\left(w\partial v^s/\partial w\right)}{\partial v^s/\partial w} - 1,\ \ \lambda_i^m = \frac{\phi_i^m\left(w_1\partial v^m/\partial w_1, w_2\partial v^m/\partial w_2\right)}{w_i\partial v^m/\partial w_i} - 1.$$

As before, we shall assume that the optimal distortions, $\lambda^{s,*}$ and $\lambda^{m,*}$, satisfy conditions (A1)-(A5) of Proposition 2. One important implication of (A2) is that labor supply is strictly positive and goes to ∞ as productivity goes to ∞. Indeed, since $\lambda^{s,*} \leq \overline{\lambda}$, by Lemma 10, we have

$$\frac{w}{1+\overline{\lambda}} \leq \psi^s\left(\frac{y^{s,*}(w)}{w}\right) \leq \left(\frac{y^{s,*}(w)}{w}\right)^{p-1}\left(\left(\frac{y^{s,*}(w)}{w}\right)^p + \overline{x}\right)^{(1-p\gamma)/p\gamma},$$

which shows that $y^{s,*}(w)$ is bounded away from zero on every compact subset of $\mathbb{R}_{++}$ and that $\frac{y^{s,*}(w)}{w} \to \infty$ as $w \to \infty$.

Clearly, singles and married individuals can be studied separately. We start with singles. To apply the variational argument from the proof of Proposition 2, we first need to investigate derivatives of ϕ^s. Consider the equation $l = \phi^s(x)$, where $x = wu$, that defines l as a function of u for fixed w. By definition, $l\frac{\partial N^s(l)}{\partial l} = wu$, thus $\left(\widetilde{\gamma}^s(l)\right)^{-1} d\ln l = d\ln u$ and $d\ln l = \frac{wu}{l}\frac{\partial\phi^s(wu)}{\partial x}d\ln u$, which gives

$$\frac{\partial\left(w\phi^s\left(wu\right) - N^s\left[\phi^s\left(wu\right)\right]\right)}{\partial u} = w\widetilde{\gamma}^s(l)\left(\frac{l}{u} - 1\right).$$

Observe that the term in the brackets is exactly $\lambda^{s,*}(w)$ when evaluated at $u = \frac{\partial v^{s,*}(w)}{\partial w}$.

Since Γ^s is uniformly bounded, (A1)-(A5) hold and nonnegativity of earnigns is slack, the argument in the proof of Proposition 2 is applicable. Specifically, the following differential equation is necessary for optimality:

$$\frac{\partial\left(\widetilde{\gamma}^{s,*}w\lambda^{s,*}g\right)}{\partial w} = (\alpha - 1)g, \tag{99}$$

where $\widetilde{\gamma}^{s,*}(w)$ is the value of $\widetilde{\gamma}^s(l)$ evaluated at $l = \frac{y^{s,*}(w)}{w}$. This equation is analogous to (56), and the only difference is that here $\widetilde{\gamma}^{s,*}$ is non-constant. Integrating this equation, we obtain the Diamond's ABC formula from Section 5.3, that is,

$$\widetilde{\gamma}^{s,*}(t)\lambda^{s,*}(t) = \frac{1 - \mathbb{E}\left[\alpha | w_i \geq t\right]}{\theta(t)}.$$

Since $\widetilde{\gamma}^s(l) \to \gamma$ as $l \to \infty$ and $\frac{y^{s,*}(w)}{w} \to \infty$ as $w \to \infty$, we get

$$\lim_{t\to\infty}\lambda^{s,*}(t) = \lim_{t\to\infty}\frac{1 - \mathbb{E}[\alpha | w \geq t]}{\gamma\theta(t)}.$$

We now look at married individuals. Again, the first step is to determine derivatives of $\boldsymbol{\phi}^m$. Consider the equation $\boldsymbol{l} = \boldsymbol{\phi}^m(\boldsymbol{x})$, where $\boldsymbol{x} = (w_1u_1, w_2u_2)$, that defines $\boldsymbol{l}$ as a function of $\boldsymbol{u}$ for fixed $\boldsymbol{w}$. By definition, $l_i\frac{\partial N^{\boldsymbol{m}}(\boldsymbol{l})}{\partial l_i} = w_iu_i$ for $i = 1, 2$, which gives $\left(\widetilde{\gamma}^m(\boldsymbol{l})\right)^{-1} d\ln\boldsymbol{l} = d\ln\boldsymbol{u}$ and

$$d\ln l_i = \frac{w_iu_i}{l_i}\frac{\partial\phi^m(w_1u_1, w_2u_2)}{\partial x_i}d\ln u_i + \frac{w_{-i}u_{-i}}{l_i}\frac{\partial\phi^m(w_1u_1, w_2u_2)}{\partial x_{-i}}d\ln u_{-i}.$$

Combining these expressions, we obtain

$$\frac{\partial\left(w_1\phi_1^m\left(w_1u_1, w_2u_2\right) + w_2\phi_2^m\left(w_1u_1, w_2u_2\right) - N^m\left[\boldsymbol{\phi}^m\left(w_1u_1, w_2u_2\right)\right]\right)}{\partial u_i} =$$

$$= w_i\widetilde{\gamma}^m_{i,i}(\boldsymbol{l})\left(\frac{l_i}{u_i} - 1\right) + w_{-i}\frac{u_{-i}}{u_i}\widetilde{\gamma}^m_{-i,i}(\boldsymbol{l})\left(\frac{l_i}{u_{-i}} - 1\right) = w_{-i}\frac{u_{-i}}{u_i}\widetilde{\gamma}^m_{i,-i}(\boldsymbol{l}) = w_i\widetilde{\gamma}^m_{i,-i}(\boldsymbol{l}), \tag{100}$$

where the second equality is due to the definition of $\widetilde{\gamma}^m$ and $\frac{l_{-i}\partial N^{\boldsymbol{m}}(\boldsymbol{l})/\partial l_{-i}}{l_i\partial N^{\boldsymbol{m}}(\boldsymbol{l})/\partial l_i} = \frac{u_{-i}w_{-i}}{u_iw_i}$.

Conditions (A1)-(A5) of Proposition 2, uniform boundedness of Γ^m and the fact that earnings are strictly positive, permits us to apply the same argument as in the proof of Proposition 2.

Observe that $\left(\frac{l_1}{u_1}-1, \frac{l_2}{u_2}-1\right)$ equals to $\boldsymbol{\lambda}^{m,*}(\boldsymbol{w})$ when evaluated at $\boldsymbol{w} = \left(\frac{\partial v^{m,*}(\boldsymbol{w})}{\partial w_1}, \frac{\partial v^{m,*}(\boldsymbol{w})}{\partial w_2}\right)$, thus, by (100), the following differential equation is necessary for optimality:

$$\sum_{i=1}^{2} \frac{\partial\left(\widetilde{\gamma}_{i,i}^{m,*} w_i \lambda_i^{m,*} f + \widetilde{\gamma}_{i,-i}^{m,*} w_i \lambda_{-i}^{m,*} f\right)}{\partial w_i} = (\alpha^m - 1) f, \tag{101}$$

where $\widetilde{\gamma}^{m,*}(\boldsymbol{w})$ stays for $\widetilde{\gamma}^m(\boldsymbol{l})$ evaluated at $\boldsymbol{l} = \left(\frac{y_1^{m,*}(\boldsymbol{w})}{w_1}, \frac{y_2^{m,*}(\boldsymbol{w})}{w_2}\right)$. Integrate (101) using the coarea formula with $Q = w_i$ and $Q = \min \boldsymbol{w}$ to obtain the following conditional moments of optimal distortions:

$$\mathbb{E}\left[\widetilde{\gamma}_{i,i}^{m,*}\lambda_i^{m,*} + \widetilde{\gamma}_{i,-i}^{m,*}\lambda_{-i}^{m,*} | w_i = t\right] = \frac{1-\mathbb{E}\left[\alpha^m | w_i \geq t\right]}{\theta(t)},$$

$$\mathbb{E}\left[\widetilde{\gamma}_{i,i}^{m,*}\lambda_i^{m,*} + \widetilde{\gamma}_{i,-i}^{m,*}\lambda_{-i}^{m,*} | w_i = t\right] = \frac{\Pr(w_j \geq t | w_i \geq t)}{2\Pr\left(w_j \geq t | w_i = t\right)} \frac{1-\mathbb{E}\left[\alpha^m | \boldsymbol{w} \geq (t,t)\right]}{\theta(t)}.$$

By (A2), the optimal distortions are uniformly bounded. Since $\widetilde{\gamma}_{i,i}^m(\boldsymbol{l}) \to \gamma$ and $\widetilde{\gamma}_{i,-i}^m(\boldsymbol{l}) \to 0$ as $l_i \to \infty$ uniformly in l_{-i} and $\frac{y_i^{m,*}(\boldsymbol{w})}{w_i} \to \infty$ as $w_i \to \infty$, we conclude that

$$\lim_{t\to\infty} \mathbb{E}\left[\lambda_i^{m,*} | w_i = t\right] = \lim_{t\to\infty} \frac{1-\mathbb{E}\left[\alpha^m | w_i \geq t\right]}{\gamma\theta(t)}.$$

10.4 Bargaining and the allocation of resources within couples

We shall first study optimal taxation with generalized Nash bargaining as defined in the main text and then look at the special case of equal bargaining powers. Recollect than each spouse i's utility in a couple with productivities $\boldsymbol{w}$ is given by

$$U^m(w_i|w_{-i}) = \eta(w_i|w_{-i}) v^m(\boldsymbol{w}) + \left[\eta(w_{-i}|w_i)\left(v^s(w_i) - \varrho\right) - \eta(w_i|w_{-i})\left(v^s(w_{-i}) - \varrho\right)\right].$$

In this expression, $\eta(w_i|w_{-i})$ is a number in $[0,1]$ and $\mathbb{E}\eta = \frac{1}{2}$. Using symmetry of v^m and f, it is easy to see that $\mathbb{E}U^m = \frac{1}{2}\mathbb{E}v^m$. So, the resource constraint and the participation constraint, Equation (5), are exactly the same as in Section 9.1, and only the welfare criterion, $\mathcal{W}$, is a bit more complex. Using symmetry and U^m defined above, we can express $\mathcal{W}$ as

$$\begin{aligned}
\mathcal{W} &= (1-\mu)\mathbb{E}\left[\alpha v^s\right] + \frac{\mu}{2}\mathbb{E}\left[\alpha(w_1)U^m(w_1|w_2) + \alpha(w_2)U^m(w_2|w_1)\right] + \int_\mu^1 \Phi(\varepsilon)d\varepsilon + const \\
&= (1-\mu)\mathbb{E}\left[\left(\alpha - \frac{\mu}{1-\mu}\xi\right)v^s\right] + \frac{\mu}{2}\mathbb{E}\left[\overline{\alpha}^m v^m\right] + \int_\mu^1 \Phi(\varepsilon)d\varepsilon + const, \qquad (102)
\end{aligned}$$

where

$$\overline{\alpha}^m(w_1,w_2) = \eta(w_1|w_2)\alpha(w_1) + \eta(w_2|w_1)\alpha(w_2),$$

$\xi(w) = \mathbb{E}\left[\left(\alpha(w_{-i}) - \alpha(w_i)\right)\eta(w_{-i}|w_i) | w_i = w\right]$, and *const* is the term that only depends on ϱ.

Following the argument used in Section 9.1, we can solve for the pair of expected pecuniary utilities, $\frac{1}{2}\mathbb{E}v^m = \mathcal{S} + (1-\mu)\Phi(\mu)$ and $\mathbb{E}v^s = \mathcal{S} - \mu\Phi(\mu)$, where $\mathcal{S}$ is given by (54), and substitute them into $\mathcal{W}$. The reader can verify that $(1-\mu)\mathbb{E}\alpha - \mu\mathbb{E}\xi + \frac{\mu}{2}\mathbb{E}\overline{\alpha}^m = 1$, hence (102) can be simplified as

$$\mathcal{W} = (1-\mu)\mathbb{E}\left[\left(\alpha - 1 + \frac{\mu}{1-\mu}\left(\mathbb{E}\xi - \xi\right)\right)v^s\right] + \frac{\mu}{2}\mathbb{E}\left[\left(\overline{\alpha}^m - \mathbb{E}\overline{\alpha}^m\right)v^m\right] + \int_\mu^1 \Phi(\varepsilon)d\varepsilon + \mathcal{S} + \\ + \int_\mu^1 \Phi(\varepsilon)d\varepsilon + \left(\mathbb{E}\overline{\alpha}^m + 1 - \frac{\mu}{1-\mu}\mathbb{E}\xi\right)\mu(1-\mu)\Phi(\mu) + const. \quad (103)$$

So, the relaxed problem is to maximize the objective in (103).

Using the same variational argument as was invoked in the context of the benchmark model, see Proposition 2, we obtain

$$\frac{\partial\left(\gamma w \lambda^{s,*} g\right)}{\partial w} = \alpha - 1 + \frac{\mu^*}{1-\mu^*}\left(\mathbb{E}\xi - \xi\right)g,$$
$$\sum_{i=1}^{2} \frac{\partial\left(\gamma w_i \lambda_i^{m,*} f\right)}{\partial w_i} = \left(\overline{\alpha}^m - \mathbb{E}\overline{\alpha}^m\right)f.$$

The optimal taxes on married stated in the main text follows from the coarea formula with $Q = w_i$ applied to the second differential equation. As for single person households, remark that the first differential equation implies

$$\lambda^{s,*}(t) = \frac{1 - \mathbb{E}[\alpha|w \geq t]}{\gamma\theta(t)} + \frac{\mu^*}{1-\mu^*}\frac{\mathbb{E}[\xi|w \geq t] - \mathbb{E}\xi}{\gamma\theta(t)}.$$

In the case of equal bargaining powers, i.e., $\eta = \frac{1}{2}$, the expected value of ξ equals to 0, and we recover Equation (40). Since α is decreasing and $F \leq_{PQD} \overline{F}$, we have

$$\mathbb{E}[\alpha(w_{-i})|w_i \geq t] \geq \mathbb{E}[\alpha(w_i)|w_i \geq t],$$

which shows that $\mathbb{E}[\xi|w \geq t] > 0$ for all $t > 0$. Conclude that the second term in (40) is positive as claimed in the main text.

10.5 Extensive margin of labor supply

The model with extensive margin is quite complex, and we shall only outline key steps of how it can be addressed skipping some technical details. First of all, we need to introduce some auxiliary notations. Consider a single person who can obtain v^s by participating in the labor force. Let $\pi^s(v^s)$ and $L^s(v^s)$ be their expected optimal pecuniary utility net lump-sum payments and probability of working, that is

$$\pi^s(v^s) := \mathbb{E}\max\{v^s-\kappa,0\} = \int_0^{v^s}(v^s-\kappa)dH, \tag{104}$$

$$L^s(v^s) := \Pr(v^s-\kappa\geq 0) = H(v^s). \tag{105}$$

Similarly, consider a married couple who can obtain v^m when both spouses participates in the labor force and $\widetilde{v}_i^m$ when only spouse i works. Denote their expected optimal marital surplus net lump-sum payments and choice probabilities by $\pi^m\left(v^m,\widetilde{v}_1^m,\widetilde{v}_2^m\right)$ and $L^m\left(v^m,\widetilde{v}_1^m,\widetilde{v}_2^m\right)$, $\widetilde{L}_1^m\left(v^m,\widetilde{v}_1^m,\widetilde{v}_2^m\right)$, $\widetilde{L}_2^m\left(v^m,\widetilde{v}_1^m,\widetilde{v}_2^m\right)$. Formally, these objects are given by

$$\pi^m\left(v^m,\widetilde{v}_1^m,\widetilde{v}_2^m\right) := \mathbb{E}\max\left\{v^m-\kappa_1-\kappa_2,\widetilde{v}_1^m-\kappa_1,\widetilde{v}_1^m-\kappa_2,0\right\}, \tag{106}$$

$$L^m\left(v^m,\widetilde{v}_1^m,\widetilde{v}_2^m\right) := \Pr\left(v^m-\kappa_1-\kappa_2\geq\max\left\{\widetilde{v}_1^m-\kappa_1,\widetilde{v}_2^m-\kappa_2,0\right\}\right), \tag{107}$$

$$\widetilde{L}_i^m\left(v^m,\widetilde{v}_1^m,\widetilde{v}_2^m\right) := \Pr\left(\widetilde{v}_i^m-\kappa_i\geq\max\left\{v^m-\kappa_1-\kappa_2,\widetilde{v}_{-i}^m-\kappa_{-i},0\right\}\right). \tag{108}$$

In general, couples' choice probabilities are rather complex. However, if $v^m=\widetilde{v}_1^m+\widetilde{v}_2^m$, then spouse i's participation decision is independent of κ_{-i}, and we obtain $L^m=H(\widetilde{v}_1^m)H(\widetilde{v}_2^m)$, $\widetilde{L}_i^m=H(\widetilde{v}_i^m)\left(1-H(\widetilde{v}_{-i}^m)\right)$.

Given these definitions, we now in position to define the relaxed problem that we use to study optimal taxation with extensive margin responses. Recollect that a single person with a productivity w obtains the expected pecuniary benefit of $U^s(w)=\pi^s(v^s(w))+b^s$. The planner collects $\mathcal{R}^s:=\int T^s(y^s)L^s(v^s)dG-b^s$ from single persons in expectations. As explained in the main text, their taxes are related to v^s via the envelope and first-order conditions,

$$T^s(y^s(w)) = w^{1+\gamma}\left(\frac{\partial v^s(w)}{\partial w}\right)^{\gamma}-\gamma w\frac{\partial v^s(w)}{\partial w}-v^s(w), \tag{109}$$

$$\frac{\frac{\partial}{\partial y}T^s(y^s(w))}{1-\frac{\partial}{\partial y}T^s(y^s(w))} = \left(\frac{\partial v^s(w)}{\partial w}\right)^{\gamma-1}w^{\gamma}-1. \tag{110}$$

Since married couples share their marital surplus equally, a married person in a couple with productivities $\boldsymbol{w}$ receives $U^m(w_i|w_{-i})=\frac{1}{2}\pi^m\left(v^m(\boldsymbol{w}),\widetilde{v}^m(w_1),\widetilde{v}^m(w_2)\right)+\frac{1}{2}b^m$. The planner collects

$$\mathcal{R}^m:=\int\left[T^m\left(\boldsymbol{y}^m\right)L^m\left(v^m,\widetilde{v}_1^m,\widetilde{v}_2^m\right)+\sum_{i=1}^{2}T^m\left(\widetilde{y}_i^m\right)\widetilde{L}_i^m\left(v^m,\widetilde{v}_1^m,\widetilde{v}_2^m\right)\right]dF-b^m,$$

from married couples in expectations, where we used shorthand notations $\widetilde{v}_i=\widetilde{v}^m(w_i)$ and $\widetilde{y}_i^m=\widetilde{y}^m(w_i)$. The relations between taxes T^m on married and their pecuniary benefits from working implied by the envelope theorem and first-order conditions are

$$T^m\left(\boldsymbol{y}^m(\boldsymbol{w})\right) = \sum_{i=1}^{2}\left(w_i^{1+\gamma}\left(\frac{\partial v^m(\boldsymbol{w})}{\partial w_i}\right)^{\gamma} - \gamma w_i \frac{\partial v^m(\boldsymbol{w})}{\partial w_i}\right) - v^m(\boldsymbol{w}), \quad (111)$$

$$\frac{\frac{\partial}{\partial y_i}T^m\left(\boldsymbol{y}^m(\boldsymbol{w})\right)}{1-\frac{\partial}{\partial y_i}T^m\left(\boldsymbol{y}^m(\boldsymbol{w})\right)} = \left(\frac{\partial v^m(\boldsymbol{w})}{\partial w_i}\right)^{\gamma-1} w_i^{\gamma} - 1. \quad (112)$$

Furthermore, $\widetilde{T}^m$ and $\widetilde{v}^m$ satisfy analogs of (109) and (110).

We use the same notion of social welfare as in the benchmark, that is

$$\mathcal{W} = \mu\mathbb{E}\left[\alpha^m U^m\right] + (1-\mu)\mathbb{E}\left[\alpha U^s\right] + \int_{\mu}^{1}\Phi(\varepsilon)d\varepsilon. \quad (113)$$

So, the relaxed problem of the planner is to maximize social welfare $\mathcal{W}$ subject to the resource constraint, $\frac{\mu}{2}\mathcal{R}^m + (1-\mu)\mathcal{R}^s \geq 0$, and to the marriage market participation constraint, $\Phi(\mu) = \mathbb{E}U^m - \mathbb{E}U^s$. It turns out that the latter constraint is always slack in this economy due to $\mathbb{E}\alpha = \mathbb{E}\alpha^m$. The planner can use lump-sum payments b^m, b^s to simultaneously balance the budget and implement any desired marriage rate μ without burning social surplus. Clearly, the resource constraint must hold as an equality at the optimum. Putting all pieces together, the relaxed program can be rewritten as the unconstrained maximization problem with the following objective:

$$\frac{\mu}{2}\int\left[\alpha^m\pi^m\left(v^m,\widetilde{v}_1^m,\widetilde{v}_2^m\right) + T^m\left(\boldsymbol{y}^m\right)L^m\left(v^m,\widetilde{v}_1^m,\widetilde{v}_2^m\right) + \sum_{i=1}^{2}T^m\left(\widetilde{y}_i^m\right)\widetilde{L}_i^m\left(v^m,\widetilde{v}_1^m,\widetilde{v}_2^m\right)\right]dF+$$
$$+(1-\mu)\int\left[\alpha\pi^s(v^s) + T^s(y^s)L^s(v^s)\right]dG + \int_{\mu}^{1}\Phi(\varepsilon)d\varepsilon. \quad (114)$$

We now derive the set of optimality conditions for the relaxed problem. Recollect that (114) depends on v^s only through the expression in the second line. The necessary condition for optimality can be found using usual the same variational technique as in the benchmark economy, though this condition may be insufficient due to lack of concavity. Be definition of π^s in (104), we have $\frac{\partial\pi^s(v^s)}{\partial v^s} = H(v^s) = L^s(v^s)$. Then, under conditions analogous to (A1)-(A5), we can repeat the argument used in the proof of Proposition 2 to obtain

$$\frac{\partial\left(\gamma w\lambda^{s,*}L^s(v^{s,*})g\right)}{\partial w} = (\alpha-1)\,L^s(v^{s,*})g + T^{s,*}(y^{s,*})\frac{\partial L^s(v^{s,*})}{\partial v^s}g. \quad (115)$$

where all objects are evaluated at the productivity w. It is worthwhile to mention that this condition completely determines $v^{s,*}$. First, remark that (A2) is equivalent to uniform boundedness of $\lambda^{s,*} = \frac{\frac{\partial}{\partial y}T^{s,*}(y^{s,*})}{1-\frac{\partial}{\partial y}T^{s,*}(y^{s,*})}$, hence the term in the brackets on the left-hand of (115) goes to

0 as $w \to 0, \infty$. As a result, the value of $v^{s,*}(0)$ is pinned down by equating the integral of the right-hand side to 0, that is

$$\int \left((\alpha - 1) L^s(v^{s,*}) + T^{s,*}(y^{s,*}) \frac{\partial L^s(v^{s,*})}{\partial v^s} \right) dG = 0.$$

Second, remark that $L^s(v^{s,*}) = H(v^{s,*})$ and $\frac{\partial L^s(v^{s,*})}{\partial v^s} = h(v^{s,*})$. Therefore, setting $H^s(w) := H(v^{s,*})$ and $g^s(w) := \frac{H^s(w)g(w)}{\int H^s dG}$, we can integrate (115) from $w = t$ to $w = \infty$ to obtain (44), which pins down $\frac{\partial v^{s,*}}{\partial w}$ and the corresponding marginal tax rates.

Optimal taxation of married persons has many features in common with the problem of single households but it is considerably more involved. To begin, note that the objective in (114) depends on v^m, $\widetilde{v}^m$ only through the expression in the first line. By definition of π^m in (106), we have

$$\frac{\partial \pi^m(v^m, \widetilde{v}_1^m, \widetilde{v}_2^m)}{\partial v^m} = L^m(v^m, \widetilde{v}_1^m, \widetilde{v}_2^m), \quad \frac{\partial \pi^m(v^m, \widetilde{v}_1^m, \widetilde{v}_2^m)}{\partial \widetilde{v}_i^m} = \widetilde{L}_i^m(v^m, \widetilde{v}_1^m, \widetilde{v}_2^m).$$

Again, under appropriate versions of (A1)-(A5), we can apply the argument used in the proof of Proposition 2 separately to v^m and $\widetilde{v}^m$ and obtain

$$\sum_{i=1}^{2} \frac{\partial \left(\gamma w_i \lambda_i^{m,*} L^m(v^{m,*}, \widetilde{v}_1^{m,*}, \widetilde{v}_2^{m,*}) f\right)}{\partial w_i} = (\alpha^m - 1) L^m(v^{m,*}, \widetilde{v}_1^{m,*}, \widetilde{v}_2^{m,*}) f +$$
$$+ T^{m,*}(\boldsymbol{y}^{m,*}) \frac{\partial L^m(v^{m,*}, \widetilde{v}_1^{m,*}, \widetilde{v}_2^{m,*})}{\partial v^m} f + \sum_{i=1}^{2} \widetilde{T}^m(\widetilde{y}_i^{m,*}) \frac{\partial \widetilde{L}_i^m(v^{m,*}, \widetilde{v}_1^{m,*}, \widetilde{v}_2^{m,*})}{\partial v^m} f \quad (116)$$

and

$$\sum_{i=1}^{2} \frac{\partial \left(\gamma w_i \widetilde{\lambda}_i^{m,*} \int \widetilde{L}_i^m(v^{m,*}, \widetilde{v}_1^{m,*}, \widetilde{v}_2^{m,*}) f dw_{-i}\right)}{\partial w_i} = \int \Bigg[(\alpha^m - 1) \widetilde{L}_i^m(v^{m,*}, \widetilde{v}_1^{m,*}, \widetilde{v}_2^{m,*}) f +$$
$$+ T^{m,*}(\boldsymbol{y}^{m,*}) \frac{\partial L^m(v^{m,*}, \widetilde{v}_1^{m,*}, \widetilde{v}_2^{m,*})}{\partial \widetilde{v}_i^{m,*}} f + \sum_{j=1}^{2} \widetilde{T}^m(\widetilde{y}_j^{m,*}) \frac{\partial \widetilde{L}_j^m(v^{m,*}, \widetilde{v}_j^{m,*}, \widetilde{v}_j^{m,*})}{\partial \widetilde{v}_i^{m,*}} f \Bigg] dw_{-i}. \quad (117)$$

In Equations (116), (117) all objects are evaluated at $\boldsymbol{w}$ and w_i, respectively. As can been seen from these equations, extensive responses for married couples are much more complex. In particular, we need to take into an account changes in all participation probabilities as revealed by the second lines in (116), (117). These equations jointly determine derivatives of $v^{m,*}$, $\widetilde{v}^{m,*}$ and their intercepts.

We now study the case of random matching, i.e., $F = G^2$. We shall show that the optimality conditions for married hold for separable taxes that takes the form $T^m(y_1, y_2) =$

$\widetilde{T}^m(y_1) + \widetilde{T}^m(y_2)$. For such taxation, we have $v^m(w_1, w_2) = \widetilde{v}^m(w_1) + \widetilde{v}^m(w_2)$ for all spousal productivities, and hence their participation decisions are independent from each other. In order to verify that separable taxation satisfies (116) and (117), we first need to compute derivatives of their choice probabilities. To this end, let v^m, $\widetilde{v}_1^m$, $\widetilde{v}_2^m$ be a tuple of numbers such that $v_m = \widetilde{v}_1^m + \widetilde{v}_2^m$. By (107) and (108), we have

$$L^m(v^m, \widetilde{v}_1^m, \widetilde{v}_2^m) = H(\widetilde{v}_1^m)H(\widetilde{v}_2^m), \quad \widetilde{L}_i^m(v^m, \widetilde{v}_1^m, \widetilde{v}_2^m) = H(\widetilde{v}_i^m)\left(1 - H(\widetilde{v}_{-i}^m)\right) \tag{118}$$

and

$$\frac{\partial L^m(v^m, \widetilde{v}_1^m, \widetilde{v}_2^m)}{\partial v^m} = H(\widetilde{v}_1^m)h(\widetilde{v}_2^m) + h(\widetilde{v}_1^m)H(\widetilde{v}_2^m), \tag{119}$$

$$\frac{\partial L^m(v^m, \widetilde{v}_1^m, \widetilde{v}_2^m)}{\partial \widetilde{v}_i^m} = -H(\widetilde{v}_i^m)h(\widetilde{v}_{-i}^m), \tag{120}$$

$$\frac{\partial \widetilde{L}_i^m(v^m, \widetilde{v}_1^m, \widetilde{v}_2^m)}{\partial v^m} = -H(\widetilde{v}_i^m)h(\widetilde{v}_{-i}^m), \tag{121}$$

$$\frac{\partial \widetilde{L}_i^m(v^m, \widetilde{v}_1^m, \widetilde{v}_2^m)}{\partial \widetilde{v}_i^m} = H(\widetilde{v}_i^m)h(\widetilde{v}_{-i}^m) + (1 - H(\widetilde{v}_{-i}^m))h(\widetilde{v}_i^m), \tag{122}$$

$$\frac{\partial \widetilde{L}_{-i}^m(v^m, \widetilde{v}_1^m, \widetilde{v}_2^m)}{\partial \widetilde{v}_i^m} = 0. \tag{123}$$

As explained above, under separable taxation the choice probabilities that appear in (116), (117) satisfy (118) and (119)-(123) for every productivity vector.

Consider the tax function $T^{m,*}(y_1, y_2) = \widetilde{T}^{m,*}(y_1) + \widetilde{T}^{m,*}(y_2)$ described in the main text and characterized by (45). If $\widetilde{\lambda}_i^{m,*} = \frac{\frac{\partial}{\partial y}\widetilde{T}^m(\widetilde{y}_i^{m,*})}{1 - \frac{\partial}{\partial y}\widetilde{T}^m(\widetilde{y}_i^{m,*})}$ is bounded, then, since $\int \alpha dG = 1$, Equation (44) implies

$$\int \left[\left(\frac{\alpha - 1}{2}\right)(1 - H(\widetilde{v}^{m,*})) - \widetilde{T}^m(\widetilde{y}^{m,*})h(\widetilde{v}^{m,*})\right] dG = 0. \tag{124}$$

Using $\alpha(w_1, w_2) = \frac{1}{2}\alpha(w_1) + \frac{1}{2}\alpha(w_2)$ and $f(w_1, w_2) = g(w_1)g(w_2)$, Equations (118), (119)-(123), and the boundary condition (124), it is routine to verify that both (116), (117) are satisfied for such taxes.

10.6 Proof of Lemma 7

The optimality conditions that we devised in the previous section are only necessary because the relaxed problem defined in (114) is not concave in v^s, v^m, $\widetilde{v}^m$. So, even though (44) and (45) verify these optimality conditions, there can be other tax schedules that do better. In addition, since the optimal taxes appear on both sides of (115) and (116) when persons can respond at the extensive margin, it is difficult to compare the marginal tax rates on single and married. In this section, we fill these gaps in the special case of our setting with i.i.d. types and non-random $\kappa = \underline{\kappa} = \overline{\kappa}$.

Proof. To overcome the challenge of lack of concavity, we consider an extended problem in which the planner can observe labor force participation probabilities and can condition taxes on them. Now, the planner selects tax schedules $T^s(y|e)$ and $T^m(\boldsymbol{y}|\boldsymbol{e})$ for single and married households, where $e \in [0,1]$ and $\boldsymbol{e} = (e_1, e_2) \in [0,1]^2$. The interpretation is straightforward, e.g., if a single person joins the labor force with probability e and supplies y units of labor, then that person pays $T(y|e)$ dollars to the planner. The original model is subsumed by imposing additional restrictions on taxes: $T^s(y|e) = T^s(y)e + b^s$ and

$$T^m(\boldsymbol{y}|e) = T^m(\boldsymbol{y})e_1e_2 + \widetilde{T}^m(y_1)e_1(1-e_2) + \widetilde{T}^m(y_2)e_2(1-e_1) + b^m. \tag{125}$$

As we shall show, the extended taxation problem is concave and it admits a solution that satisfies these constraints.

Let v^s, v^m be the solution to the optimization problems of single and married households when faced with such generalized taxes, that is

$$\begin{aligned} v^s(w) &:= \max_{e,y} \; T^s(y|e) - \left(\gamma\left(\frac{y}{w}\right)^{1/\gamma} + \kappa\right) e \text{ s.t. } e \in [0,1], y \geq 0, \\ v^m(\boldsymbol{w}) &:= \max_{\{e_i,y_i\}_{i=1}^2} \; T^m(\boldsymbol{y}|\boldsymbol{e}) - \sum_{i=1}^{2}\left(\gamma\left(\frac{y_i}{w_i}\right)^{1/\gamma} + \kappa\right) e_i \text{ s.t. } e_1, e_2 \in [0,1], y_1, y_2 \geq 0. \end{aligned}$$

The notion of welfare and the marriage market participation constraints are identical to the benchmark. Then, the characterization of incentive constraints is similar to the benchmark, e.g., the envelope theorem applied to v^s yields $w\frac{\partial v^s}{\partial w} = \left(\frac{y^s}{w}\right)^{1/\gamma} e^s$. Hence, the planner can collect

$$\mathcal{R}^s = \int \left[w^{1+\gamma}\left(\frac{\partial v^s}{\partial w}\right)^{\gamma}(e^s)^{1-\gamma} - \kappa e^s - \gamma w\frac{\partial v^s}{\partial w}\right] dG \tag{126}$$

from singles in expectations. Clearly, for fixed v^s, it is optimal to select e^s that pointwise maximizes (126) as it improves total revenues available for redistribution. The reader can verify that

$$\begin{aligned} \psi\left[w^{1+\gamma}\left(\frac{\partial v^s}{\partial w}\right)^{\gamma}\right] &:= \max_{e^s\in[0,1]} w^{1+\gamma}\left(\frac{\partial v^s}{\partial w}\right)^{\gamma}(e^s)^{1-\gamma} - \kappa e^s = \\ &= \begin{cases} \gamma\left(\frac{1-\gamma}{\varrho}\right)^{(1-\gamma)/\gamma} w^{(1+\gamma)/\gamma}\frac{\partial v^s}{\partial w}, & w^{1+\gamma}\left(\frac{\partial v^s}{\partial w}\right)^{\gamma} \leq \frac{\kappa}{1-\gamma}, \\ w^{1+\gamma}\left(\frac{\partial v^s}{\partial w}\right)^{\gamma} - \kappa, & w^{1+\gamma}\left(\frac{\partial v^s}{\partial w}\right)^{\gamma} \geq \frac{\kappa}{1-\gamma}. \end{cases} \end{aligned} \tag{127}$$

The exactly same construction applies to married, thus the economic output $\mathcal{S}$ (Equation (54) in the benchmark) can be succinctly expressed as

$$\frac{\mu}{2}\int\sum_{i=1}^{2}\left(\psi\left[w_i^{1+\gamma}\left(\frac{\partial v^m}{\partial w_i}\right)^{\gamma}\right] - \gamma w_i\frac{\partial v^m}{\partial w_i}\right) dF + (1-\mu)\int\left(\psi\left[w^{1+\gamma}\left(\frac{\partial v^s}{\partial w}\right)^{\gamma}\right] - \gamma w\frac{\partial v^s}{\partial w}\right) dG. \tag{128}$$

Following Section 9.1, the relaxed problem is exactly as that section but now $\mathcal{S}$ is given by (128). The advantage of pre-solving for e^s and $\boldsymbol{e}^m$ is that the relaxed problem becomes a concave program in v^s , v^m and can be studied along the lines of the proof of Proposition 2.

We now derive and further analyze the set of necessary and sufficient conditions for optimality. As in the proof of Proposition 2, marginal taxes and distortions of singles and married can be studied in isolation. Observe that ψ is continuously differentiable. Set λ^s to be

$$\lambda^s := \begin{cases} \left(\frac{1-\gamma}{\kappa}\right)^{(1-\gamma)/\gamma} w^{1/\gamma} - 1, & w^{1+\gamma}\left(\frac{\partial v^s}{\partial w}\right)^{\gamma} \leq \frac{\kappa}{1-\gamma}, \\ w^{\gamma}\left(\frac{\partial v^s}{\partial w}\right)^{\gamma-1} - \kappa, & w^{1+\gamma}\left(\frac{\partial v^s}{\partial w}\right)^{\gamma} \geq \frac{\kappa}{1-\gamma}, \end{cases}$$

and define $\boldsymbol{\lambda}^m$ analogously as a function of v^m. With these notations, the variational conditions listed in the proof of Proposition 2, that is (60), (63), (60), (64), are necessary and sufficient provided that conditions (A1)-(A5) hold.

We claim that the optimum coincides with the benchmark solution above certain thresholds. Specifically, there are numbers $\underline{w}^s$ for singles and $\underline{w}^m$ for married so that a single (married) person works if and only if $w_i \geq \underline{w}^s$ ($w_i \geq \underline{w}^m$, resp.); moreover, above these cut-offs distortions are exactly as in the benchmark. Recollect that, by Proposition 4, the optimal benchmark distortions are $\lambda^{s,*}(w)$ and $\left(\frac{1}{2}\lambda^{s,*}(w_1), \frac{1}{2}\lambda^{s,*}(w_2)\right)$, where $\lambda^{s,*}$ is defined in (12). Since the first-order approach is valid, that is both conditions of Proposition 1 hold, there are unique thresholds such that welfare gains in (65), (66) at the margin are exactly κ, i.e.,

$$\kappa = (1-\gamma)\underline{w}^s\left(\frac{\underline{w}^s}{1+\lambda^{s,*}(\underline{w}^s)}\right)^{\gamma/(1-\gamma)}, \quad \kappa = (1-\gamma)\underline{w}^m\left(\frac{\underline{w}^m}{1+\frac{1}{2}\lambda^{s,*}(\underline{w}^m)}\right)^{\gamma/(1-\gamma)}. \tag{129}$$

Consider v^s so that $\frac{\partial v^s}{\partial w}(w) = 0$ for $w < \underline{w}^s$, which gives $\lambda^s(w) = \left(\frac{1-\gamma}{\kappa}\right)^{(1-\gamma)/\gamma} w^{1/\gamma} - 1$, and $\lambda^s(w) = \lambda^{s,*}(w)$ otherwise. Then, λ^s constructed in this way satisfies (A1)-(A5); moreover, since $\lambda^s(w) \leq \lambda^{s,*}(w)$ for all w, (60) and (61) hold. Indeed, for any alternative function $\hat{v}^s \in \mathscr{V}^s$,

$$\int \gamma w \lambda^s \frac{\partial \hat{v}^s}{\partial w} dG + \int (\alpha - 1)\hat{v}^s dG = \int \gamma w \left(\underbrace{\lambda^s - \lambda^{s,*}}_{\leq 0}\right) \underbrace{\frac{\partial \hat{v}^s}{\partial w}}_{\geq 0} dG \leq 0,$$

which shows (61). By construction, $\left(\lambda^s(w) - \lambda^{s,*}(w)\right)\frac{\partial v^s(w)}{\partial w} = 0$ for all w, thus (60) is satisfied as well.

The argument for married individuals is identical. Consider $v^m(\boldsymbol{w}) = \widetilde{v}^m(w_1) + \widetilde{v}^m(w_2)$ for $\widetilde{v}^m$ that satisfies $\frac{\partial \widetilde{v}^m(w_i)}{\partial w_i} = 0$ for $w_i < \underline{w}^m$, which gives $\lambda_i^m(\boldsymbol{w}) = \left(\frac{1-\gamma}{\kappa}\right)^{(1-\gamma)/\gamma} w_i^{1/\gamma} - 1$, and $\lambda_i^m(\boldsymbol{w}) = \frac{1}{2}\lambda^{s,*}(w_i)$ otherwise. These distortions satisfy (A1)-(A5). Furthermore, the condition for validity of the first-order approach in Proposition 1 implies that $\lambda_i^m(w_i) \le \frac{1}{2}\lambda^{s,*}(w_i)$ for all w_i. As a result, for any function $\hat{v}^m \in \mathscr{V}^m$, potentially non-separable,

$$\int \sum_{i=1}^{2} \gamma w_i \lambda_i^m \frac{\partial \hat{v}_i^m}{\partial w_i} dF + \int (\alpha^m - 1)\hat{v}^m dF = \int \sum_{i=1}^{2} \gamma w_i \left(\underbrace{\lambda_i^m - \frac{1}{2}\lambda^{s,*}}_{\le 0} \right) \underbrace{\frac{\partial \hat{v}^m}{\partial w_i}}_{\ge 0} dF \le 0,$$

which shows (64). By construction, $\left(\lambda_i^m(\boldsymbol{w}) - \frac{1}{2}\lambda^{s,*}(w_i)\right) \frac{\partial v^m(\boldsymbol{w})}{\partial w_i} = 0$ for all $\boldsymbol{w}$, thus (63) is satisfied as well.

To sum up, we identified the solution to the extended problem in which the planner can condition taxes on labor force participation probabilities. In this problem, there are two thresholds, $\underline{w}^s$ and $\underline{w}^m$, such that a person works if and only if his/her productivity is above that person's threshold. Furthermore, since $\lambda^{s,*}$ is nonnegative due to monotonicity of α, examination of (129) makes it clear that the threshold for married is lower than one for singles due to lower distortions. The optimal marginal taxes on those who work are exactly as in the benchmark with random matching.

Since the optimal labor participation decisions are integral and a person who doesn't participate in the labor force selects zero earnings, the mechanism identified above also solves the original model with extensive margin, i.e., (125) holds. To see it more formally, let $T^{s,*}(y)$ and $T^{m,*}(y_1, y_2) = \widetilde{T}^{m,*}(y_1) + \widetilde{T}^{m,*}(y_2)$ be taxes that solve the benchmark problem in Lemma 4. Under such taxation, the marital surplus of couples is also separable , i.e., $v^{m,*}(w_1, w_2) = \widetilde{v}^{m,*}(w_1) + \widetilde{v}^{m,*}(w_2)$. Denote the earnings obtained from $v^{s,*}$, $\widetilde{v}^{m,*}$ via the envelope condition by $y^{s,*}$, $\widetilde{y}^{m,*}$ and set $\Delta^s := v^{s,*}(\underline{w}^s) - \kappa$, $\Delta^m := v^{m,*}(\underline{w}^m) - \kappa$. Consider $T^s(y|e) = (T^{s,*}(y) + \Delta^s)e + b^s$ and $T^m(\boldsymbol{y}|e) = (\widetilde{T}^{m,*}(y_1) + \Delta^m)e_1 + (\widetilde{T}^{m,*}(y_1) + \Delta^m)e_2 + b^m$. By construction, each single (married) persons works if and only if $w_i \ge \underline{w}^s$ ($w_i \ge \underline{w}^m$, resp.) in which case that persons chooses earnings $y^{s,*}$ ($\widetilde{y}^{m,*}$, resp.). We can choose lump-sum payments b^s, b^m to balance the budget and ensure that the desired fraction of individuals gets married by solving

$$\mu^* \int_{\underline{w}^m}^{\infty} (\widetilde{T}^{m,*}(y) + \Delta^m) dG + (1 - \mu^*) \int_{\underline{w}^s}^{\infty} (T^{s,*}(y) + \Delta^s) dG = \mu^* b^m + (1 - \mu^*) b^s,$$

$$\int_{\underline{w}^m}^{\infty} \left(\widetilde{v}^{m,*} - \kappa - \Delta^m\right) dG - \int_{\underline{w}^s}^{\infty} \left(v^{s,*} - \kappa - \Delta^s\right) dG = b^m - b^s,$$

which are exactly the resource and marriage market participation constraints in the original model of the previous section. This concludes Parts (a) and (b) of the lemma.

We now show Part (c). The reader can verify that the optimal marriage rate μ^* is pinned down by the following analog of Equation (17):

$$\int \max\left\{w\left(\frac{w}{1+\frac{1}{2}\lambda^{s,*}}\right)^{\gamma/(1-\gamma)} - \kappa, 0\right\} dG - \int \max\left\{w\left(\frac{w}{1+\lambda^{s,*}}\right)^{\gamma/(1-\gamma)} - \kappa, 0\right\} dG = \frac{\Phi(\mu^*)}{1-\gamma}.$$

□

Since $\lambda^{s,*}(w) > \frac{1}{2}\lambda^{s,*}(w) > 0$ for all $t > 0$ due to strict monotonicity of α, the left-hand side of this expression is decreasing in κ. At $\kappa = 0$, the marriage rate is the same as in the bencmark model. As $\kappa \to \infty$, the left-hand side goes to 0, and hence the marriage rate monotonically decreases to μ^{LF} that solves $\Phi(\mu^{\text{LF}}) = 0$. Recollect that the marriage market participation constraint requires reads as $\Phi(\mu^*) = \mathbb{E}U^{m,*} - \mathbb{E}U^{s,*}$, thus $\mathbb{E}U^{m,*} - \mathbb{E}U^{s,*}$ is strictly decreasing in κ as well, and goes to 0 as $\kappa \to \infty$.

10.7 Selection into marriage

The main conceptual difference here is that there are two marriage cutoffs, μ_l and μ_h. Equation (5) has to hold for each cutoff individually, that is

$$\Phi(\mu_q) = \frac{1}{2}\int\int v^m(\boldsymbol{w})dH_q(w_1)dH_q(w_2) - \int v^s(w)dH_q \quad \text{for} \quad q = l, h. \tag{130}$$

One important implication of differential cutoffs is that the distributions of productivities, G^s and F, are endogenous. According to Bayes' rule, they satisfy

$$\begin{aligned} (1-\mu)G^s(w) &= \frac{1-\mu_l}{2}H_l(w) + \frac{1-\mu_h}{2}H_h(w), \\ (1-\mu)F(\boldsymbol{w}) &= \frac{\mu}{2}H_l(w_1)H_l(w_2) + \frac{\mu_h}{2}H_h(w_1)H_h(w_2), \end{aligned}$$

where $\mu = \frac{\mu_l+\mu_h}{2}$ is the economy-wide marriage rate. To ensure that the relaxed problem is well-defined, we require

$$\int\int \max\left\{w_1^{1/(1-\gamma)}, w_2^{1/(1-\gamma)}\right\} dH_q(w_1)dH_q(w_2) < \infty$$

for every signal q. Then, it is easy to see that, each marriage rate must be interior and conditions (a), (b) of Lemma 9 hold for each $q = l, h$.

We now study the relaxed problem for fixed $\mu_l, \mu_h \in (0,1)$. In contrast to Section 9.1, the resource constraint and (5) cannot be eliminated, we therefore use the Lagrange multiplier approach. For fixed $\mu_l, \mu_h \in (0,1)$, the problem is concave. So, let δ_l, δ_h be Lagrange multipliers on (5) and η be a multiplier on (52). Existence of δ_l, δ_h and η is standard, e.g., see Chapter

8 in Luenberger (1997); moreover, it is immediate that $\eta = 1$. To sum up, ignoring the terms that don't depend on v^s, v^m, the Lagrangian can be written as follows:

$$\frac{\mu}{2}\int(\alpha^m(\boldsymbol{w})-1)v^m(\boldsymbol{w})dF+(1-\mu)\int(\alpha(w)-1)v^s(w)dG^s+\mathcal{S}+$$
$$+\sum_{q=l,h}\delta_q\left(\int v^s(w)dH_q-\frac{1}{2}\int\int v^m(\boldsymbol{w})dH_q(w_1)dH_q(w_2)\right),$$

where $\mathcal{S}$ is defined in (54).

Our analysis of the necessary conditions for v^s, v^m in the proof of Proposition 2 goes without changes, and it gives the following analogs of (56), (57):

$$\frac{\partial\left(\gamma w\lambda^s(w)g^s(w)\right)}{\partial w} = \left(\alpha(w)-1\right)g^s(w)+\frac{\sum_{q=l,h}\delta_q h_q(w)}{1-\mu^*}, \tag{131}$$
$$\sum_{i=1}^{2}\frac{\partial\left(\gamma w_i\lambda_i^m(\boldsymbol{w})f(\boldsymbol{w})\right)}{\partial w_i} = \left(\alpha^m(\boldsymbol{w})-1\right)f(\boldsymbol{w})-\frac{\sum_{q=l,h}\delta_q h_q(w_1)h_q(w_2)}{\mu^*}. \tag{132}$$

Equations (131), (132), which are necessary for optimality for fixed marriage rates, and the coarea formula for $Q = w_i$ imply that the optimal distortions satisfy two equations in the text.

10.8 Optimality of taxation of family earnings

10.8.1 Proof of Lemma 8

Proof. By Proposition (2), the optimal tax is family-earnings-based if and only if $\widetilde{\lambda}(r) := \frac{1-\mathbb{E}\left[\alpha^m|R\geq r\right]}{\gamma\theta_r(r)}$ verifies (57). Solve for (w_1, w_2) as a function of (r,ι) to obtain

$$\max\boldsymbol{w}=\frac{r}{\left(1+\iota^{1/(1-\gamma)}\right)^{(1-\gamma)}},\quad \min\boldsymbol{w}=\frac{r}{\left(1+\iota^{1/(\gamma-1)}\right)^{(1-\gamma)}}.$$

It is routine to verify that $d\boldsymbol{w} = \frac{w_1w_2}{r\iota}drd\iota$. Thus, f and $\widetilde{f}$, which is the density of (r,ι), are related by $f = \frac{1}{2}\frac{r\iota}{w_1w_2}\widetilde{f}$. Since $R(t\boldsymbol{w}) = tR(\boldsymbol{w})$ and $I(t\boldsymbol{w}) = I(\boldsymbol{w})$ for all $t > 0$, we have

$$\sum_{i=1}^{2}\frac{\partial\left(w_i\widetilde{\lambda}(R(\boldsymbol{w}))f(\boldsymbol{w})\right)}{\partial w_i}=\left.\frac{\partial\left(t^2\widetilde{\lambda}(R(t\boldsymbol{w}))f(t\boldsymbol{w})\right)}{\partial t}\right|_{t=1}=\frac{1}{2}\frac{R(\boldsymbol{w})I(\boldsymbol{w})}{w_1w_2}\left.\frac{\partial\left(r\widetilde{\lambda}(r)\widetilde{f}(r,\iota)\right)}{\partial r}\right|_{r=R(\boldsymbol{w})}. \tag{133}$$

It follows from (133) that $\widetilde{\lambda}$ satisfies (57) if and only if

$$\frac{\partial\left(\gamma r\widetilde{\lambda}\widetilde{f}\right)}{\partial r}=\left(\alpha-1\right)\widetilde{f}. \tag{134}$$

Divide this equation by the marginal density of ι and integrate to see the claim in the first part of the lemma.

We now show the second part of the lemma. Here, we assume that α is measurable only w.r.t. R, that is $\alpha^m(\boldsymbol{w}) = \widetilde{\alpha}(R(\boldsymbol{w}))$ for some function $\widetilde{\alpha}$. In this case, Condition (134), which is necessary and sufficient, can be unpacked as follows:

$$\frac{\partial \ln \widetilde{f}(r,\iota)}{\partial r} = \widetilde{\alpha}(r) - 1 - \frac{\partial(\gamma r \widetilde{\lambda}(r))}{\partial r} = \frac{\partial \ln g_r(r)}{\partial r}, \tag{135}$$

where g_r is the density of $r = R(\boldsymbol{w})$. The last equality is due to the definition of $\widetilde{\lambda}$. Clearly, $\widetilde{f}/g_r$ must be independent of r for (135) to hold, which is equivalent to independence of (r,ι). □

We end this section with an example of comparative statics that can be performed using the methodology that we developed in the main text. As shown in Section 3.2, the optimal distortions satisfy

$$\mathbb{E}\big[\lambda_{sec}^{m,*} - \lambda_{pr}^{m,*}|I = \iota\big] = \frac{1 - \mathbb{E}[\alpha^m|I \geq \iota]}{\gamma\theta_\iota(\iota)}.$$

Suppose that $\alpha^m(\boldsymbol{w}) = \widetilde{\alpha}(R(\boldsymbol{w}))$ for some decreasing function $\widetilde{\alpha}$. Let $\widetilde{F}^a, \widetilde{F}^b$ be two distributions of transformed variables (r,ι). If $\widetilde{F}^a \leq_{PQD} \widetilde{F}^b$, then we have

$$\Pr{}^a\left(R \geq r|I \geq \iota\right) \leq \Pr{}^b\left(R \geq r|I \geq \iota\right) \quad \forall\left(r,\iota\right). \tag{136}$$

Since α^m is measurable only with respect to r and decreasing in this variable, the first-order stochastic dominance relationship in (136) gives $\mathbb{E}^a\left[\alpha^m|I \geq \iota\right] \geq \mathbb{E}^b\left[\alpha^m|I \geq \iota\right]$. Thus, $\mathbb{E}^a\big[\lambda_{sec}^{m,a,*} - \lambda_{pr}^{m,a,*}|I = \iota\big] \leq \mathbb{E}^b\big[\lambda_{sec}^{m,b,*} - \lambda_{pr}^{m,b,*}|I = \iota\big]$.

10.9 Gender differences

Since there may be different numbers of males and females on the marriage market, we allow for rationing to clear it. Specifically, we return agents with the highest values of preference shocks of the "surplus" gender back to the singlehood. As before, let μ be the marriage rate and suppose that o^* is the "deficit" gender, which simply means that $\Delta := \int v_{o^*}^s d\hat{G}_{o^*} - \int v_{-o^*}^s d\hat{G}_{-o^*} \geq 0$. Then, the marriage rate is given by

$$\Phi(\mu) = \frac{1}{2}\int v^m d\hat{F} - \int v_{o^*}^s d\hat{G}_{o^*}. \tag{137}$$

The resource constraints reads as

$$\mathcal{S} \geq \frac{\mu}{2}\int v^m d\hat{F} + (1-\mu)\int v_{o^*}^s d\hat{G}_{o^*} - (1-\mu)\frac{\Delta}{2},$$

and $\mathcal{S}$ is defined as in the benchmark (Equation (54)) but allowing for differential treatment of single households from different genders.

It is easy to see that the modified resource constraint must bind, thus when combined with (137), it can be solved uniquely for expected utilities of married and singles o as a function of $\Phi(\mu)$, $\mathcal{S}$ and Δ. Substituting these expected utilities into the welfare criterion, we obtain

$$\mathcal{W} = \frac{\mu}{2} \int (\hat{\alpha}^m - 1)\, \hat{v}^m d\hat{F} + \frac{1-\mu}{2} \sum_{o=1,2} \int (\hat{\alpha}_o - 1)\, v_o^s dG_o + \mathcal{S} + \int_\mu^1 \Phi(\varepsilon) d\varepsilon, \tag{138}$$

which is independent of Δ. We conclude that it is immaterial which gender is in "deficit", and there is always a solution in which the market clears exactly, i.e., $\Delta = 0$, when taxes are allowed to be gender-specific.

The rest of the argument is exactly the same as in the proof of Proposition 2, and it immediately implies that that the optimal gender-specific distortions $\{\hat{\lambda}_o^{s,*}, \hat{\lambda}_o^{m,*}\}_{o\in\{1,2\}}$ satisfy (47).

We now look at the case of gender-neutral taxation, i.e., $v_o^s = v^s$ for both genders $o = 1, 2$ and v^m is symmetric. In contrast to gender-specific taxation discussed above, rationing will play a role to clear the marriage market, i.e., it is not longer the case that there are multiple values of Δ that are consistent with the optimum.

In order to derive the optimal taxes, we first "symmetrize" the economy as discussed in the main text. It is routine to verify that if v^m is a symmetric function and $v_o^s = v^s$ for both $o = 1, 2$, then $\mathcal{W}$ defined in (138) can be expressed as

$$\frac{\mu}{2} \int (\alpha^m - 1)\, v^m dF + (1-\mu) \int (\alpha - 1)\, v^s dG + \int_\mu^1 \Phi(\varepsilon) d\varepsilon + \mathcal{S}.$$

Recollect that under gender-neutrality, we are effectively back to the symmetric setting of Section 9.1. So, Proposition 2 can be directly applied, and it gives that the optimal gender-neutral distortions $\{\lambda^{s,*}, \lambda_o^{m,*}\}_{o\in\{1,2\}}$ (48). Equation (49) follows from direct calculations using (47) and (48), e.g.,

$$\sum_{o=1,2} \hat{\lambda_o}^{s,*}(t) t\gamma\omega_o(t) = \sum_{o=1,2} \frac{\int_t^\infty (1-\hat{\alpha}_o(w)) g_o(w) dw}{g_o(t)} \frac{1}{2} \frac{g_o(t)}{g(t)} = \hat{\lambda_o}^{s,*}(t) t\gamma.$$

Calculations for married are identical.

10.10 Optimal restricted taxation

In this section, we explore optimal taxation under additional restrictions. First, we study the case of separable gender-neutral taxation, i.e., $T^m(w_1, w_2) = \widetilde{T}^m(y_1) + \widetilde{T}^m(y_2)$ for some tax function $\widetilde{T}$. Under such taxation, spouse i's labor supply decision is necessarily independent of

w_{-i}. Hence, the marital surplus takes the form $v^m(\boldsymbol{w}) = \widetilde{v}^m(w_1) + \widetilde{v}^m(w_2)$ for some function $\widetilde{v}^m$. The reader can verify that for such additively separable and symmetric v^m, we have

$$\int (\alpha^m - 1)\, v^m dF = 2 \int \left(\mathbb{E}\left[\alpha^m | w_i = w\right] - 1\right) \widetilde{v}^m dG.$$

Clearly, v^m enters the total economic output $\mathcal{S}$ only through $\widetilde{v}^m$ as $\frac{\partial v^m(w_1,w_2)}{\partial w_i} = \frac{\partial \widetilde{v}^m(w_i)}{\partial w_i}$. So, we reduced the analysis of individual earnings-based taxation of married individuals to the analysis of singles. Following the argument in Proposition 2, we obtain

$$\widetilde{\lambda}^{m,*}(t) = \frac{1 - \mathbb{E}\left[\alpha^m | w_i \geq t\right]}{\gamma \theta(t)},$$

which equals to the the optimal unrestricted distortions $\mathbb{E}\left[\lambda_i^{m,*} | w_i = t\right]$ as shown in (18).

Suppose that taxes are family-earnings-based. As explained in Section 5.7, under such taxation, v^m is measurable only w.r.t. to r, i.e., $v^m(\boldsymbol{w}) = \widetilde{v}^m(R(\boldsymbol{w}))$ for some $\widetilde{v}^m$, hence

$$\int (\alpha^m - 1)\, v^m dF = \int \left(\mathbb{E}\left[\alpha^m | R = r\right] - 1\right) \widetilde{v}^m dG_r,$$

where $\widetilde{v}$ is the distribution of $r = R(\boldsymbol{w})$. In order to make our previous argument applicable, we need to show that v^m enters the total economic output $\mathcal{S}$ only through $\widetilde{v}^m$. Indeed, since $w_1 \frac{\partial R}{\partial w_1} + w_2 \frac{\partial R}{\partial w_2} = r$,

$$\sum_{i=1}^{2} \left(w_i^{1+\gamma} \left(\frac{\partial v^m}{\partial w_i} \right)^{\gamma} - \gamma w_i \frac{\partial v^m}{\partial w_i} \right) = r^{1+\gamma} \left(\frac{\partial \widetilde{v}^m(r)}{\partial r} \right)^{\gamma} - \gamma r \frac{\partial \widetilde{v}^m(r)}{\partial r}.$$

Similarly to the previous part, we reduced the analysis of family-earnings-based taxation of married individuals to the analysis of singles. Following the argument what was used to prove Proposition 2, we obtain

$$\lambda^{m,fam,*}(r) = \frac{1 - \mathbb{E}\left[\alpha^m | R \geq r\right]}{\gamma \theta_r(r)},$$

which is exactly the expression in (46).

11 Quantitative analysis

11.1 Calibration

We use data from the 2020 CPS survey. In our dataset, we have pre-tax earnings of 11087 couples, each consisting of two individuals who (a) have a spouse in the same household, (b) worked for at least 20 weeks in 2020, (c) are 25-65 years old. Our measure of earnings includes only wage earnings. The sample is representative of approximately 42 million people.

We suppose that the data comes from a symmetric environment with $\gamma = 1/4$; thus, we symmetrize the dataset by creating one more copy of every household in which the identities of two spouses are interchanged. This gives us 2×11087 couples with identical distributions of earnings for each spouse and the same dependence patterns as before. We normalize earnings by 100 thousand so that the average value of individual earnings in the dataset equals 0.75.

Following Guner et al. (2014) and Heathcote et al. (2017) we assume that the data is generated with the following tax function: $T(y_1, y_2) = (y_1 + y_2) - \nu(y_1 + y_2)^{1-\tau}$. Guner et al. (2014) estimated (τ, ν) for married couples using the IRS data in which earnings are normalized by 53 thousand. Since we normalize earnings by 100 thousand, we adjust their estimate, which is $\tau = 0.06$ and $\nu = 0.91$, so that total tax bills in dollar terms are identical. The parameter τ doesn't need any adjustment but $\nu = 0.91 \times (\frac{53}{100})^{\tau}$.

Given the assumed log-linear tax schedule, each couple solves

$$\max_{(y_1,y_2)\geq \mathbf{0}} \nu \left(y_1 + y_2\right)^{1-\tau} - \sum_{i=1}^{2} \gamma \left(\frac{y_i}{w_i}\right)^{1/\gamma},$$

which allows us to express unobserved productivites as a function of observed earnings (Equation (50)) and construct their empirical distribution.

We calibrate the marginal distribution of productivities and their copula separately. Recall that the marginal G is assumed to follow a PLN distribution with parameters $(a, \eta, \sigma) \in \mathbb{R}_{++} \times \mathbb{R} \times \mathbb{R}_{++}$, that is

$$G(t) = \Phi\left(\frac{\ln t - \eta}{\sigma}\right) - t^{-a} \exp\left(a\eta + a^2\sigma^2/2\right) \Phi\left(\frac{\ln t - \eta - a\sigma^2}{\sigma}\right).$$

Our first target moment is the Pareto statistic (computed with 183 observations at t that corresponds to 99% percentile of the empirical cdf). In our sample this moment equals to 2.95, and since

$$\lim_{t\to\infty} \frac{\mathbb{E}\left[w_i|w_i \geq t\right]}{\mathbb{E}\left[w_i|w_i \geq t\right] - t} = a,$$

we set a to 2.95. The second target moment is the Gini coefficient. It equals to 0.31 in the dataset. It can be shown that (e.g., see Colombi (1990)) for a PLN distribution, it is given by

$$2\Phi\left(\frac{\sigma}{\sqrt{2}}\right) - 1 + 2\frac{e^{a(a-1)\sigma^2}}{2a-1}\Phi\left(\frac{(1-2a)\sigma}{\sqrt{2}}\right),$$

where Φ is the standard normal distribution. This gives us $\sigma = 0.4$. Our final target moment is the mean value of individual productivities that equals 0.81 in the sample. Using the closed form expression

$$\mathbb{E}w_i = \frac{a}{a-1}e^{\mu+\sigma^2/2},$$

we get $\mu = -0.71$.

As for the copula of (w_1, w_2), we calibrate it using the Kendell's tau dependence coefficient (see Chapter 5 in Nelsen (2006)), which is a rank measure of concordance, theoretically:

$$\Pr\left(\left(w_1 - \widetilde{w}_1\right)\left(w_2 - \widetilde{w}_2\right) > 0\right) - \Pr\left(\left(w_1 - \widetilde{w}_1\right)\left(w_2 - \widetilde{w}_2\right) < 0\right),$$

where (w_1, w_2) and $(\widetilde{w}_1, \widetilde{w}_2)$ are independent copies of productivities. Clearly, this statistic only depends on the underlying copula, not on G, and closed form expressions are available for many copulas. In our dataset, it equals to 0.21. We tried several copulas and found that the Gaussian one fits the data very well. For the Gaussian copula, Kendell's tau is given by $\frac{2\arcsin\rho}{\pi}$, where ρ is its correlation parameter. This gives us $\rho = 0.33$.

11.2 Numerical approach

In this section, we overview the numerical approach that we used to find the optimal taxes. We first discretize the problem using a finite logarithmic grid of 399 equally spaced productivities. The grid is logarithmic in the sense that a ratio of two consecutive points is constant. This allows to improve accuracy at the left tail and capture the thick right tail. Let $\left\{w^1, \ldots, w^{400}\right\}$, where $w^1 = 0.12$ and $w^{400} = 10$, be this grid. The 400th point is added to ensure that our discretized relaxed problem can approximate the original relaxed problem in which the domain is unbounded. It will be convenient to also define $w^0 := 0$.

We numerically solve a relaxed problem that only contains downward incentive constraints, one for each spouse, that is

$$\max_{v,y_1,y_2 \geq \mathbf{0}} \sum_{n_1,n_2=1}^{400} v\left(w^{n_1}, w^{n_2}\right)\left(\alpha^m\left(w^{n_1}, w^{n_2}\right) - 1\right) f\left(w^{n_1}, w^{n_2}\right) +$$
$$+ \sum_{i=1}^{2} \sum_{n_1,n_2=1}^{400} \left(y_i\left(w^{n_1}, w^{n_2}\right) - \gamma\left(\frac{y_i\left(w^{n_1}, w^{n_2}\right)}{w^{n_i}}\right)^{1/\gamma}\right) f\left(w^{n_1}, w^{n_2}\right)$$

subject to the following set of incentive constraints: for all $n_i = 2, \ldots, 400$, $n_{-i} = 1, \ldots, 400$ and $i = 1, 2$,

$$v\left(w^{n_i}, w^{n_{-i}}\right) \geq v\left(w^{n_i - 1}, w^{n_{-i}}\right) + \gamma y_i^{1/\gamma}\left(w^{n_i - 1}, w^{n_{-i}}\right)\left(\left(w^{n_i - 1}\right)^{-1/\gamma} - \left(w^{n_i}\right)^{-1/\gamma}\right).$$

In this problem, f is set to be

$$f\left(w^{n_1}, w^{n_2}\right) = \begin{cases} \Pr\left(w^{n_i - 1} < w_i \leq w^{n_i} \quad \forall i\right), & n_i, n_{-i} < 400; \\ \Pr\left(w^{n_i - 1} < w_i, w^{n_{-i} - 1} < w_{-i} \leq w^{n_{-i}}\right), & n_i = 400 > n_{-i}; \\ \Pr\left(w^{n_i - 1} < w_i \quad \forall i\right), & n_i = n_{-i} = 400. \end{cases}$$

And, α^m is normalized so that $\sum_{n_1,n_2=1}^{400} \alpha^m (w^{n_1}, w^{n_2}) f (w^{n_1}, w^{n_2}) = 1$.

The solution to the relaxed problem is easy to find, and it is always the case that all incentive constraints are binding. Given this solution, we then numerically verify all remaining (global) incentive constraints. In all cases, we found that the first-order approach holds.