The Optimal Taxation of Couples[*]

Mikhail Golosov
University of Chicago

Ilia Krasikov
HSE University

September 2023

Abstract

We consider optimal joint nonlinear earnings taxation of couples. We use multi-dimensional mechanism design techniques to study this problem and show that the first-order approach – that restricts attention to only local incentive constraints – is valid for a broad set of primitives. Optimal taxes are characterized by the solution to a certain second-order partial differential equation. Using the Coarea Formula, we solve this equation for various conditional averages of optimal tax rates and identify key forces that determine the optimal tax rates; show how these rates depend on earnings of each spouse, correlation in spousal earnings, and redistributive objectives of the planner; compare optimal rates for primary and secondary earners; identify both the conditions under which simple tax systems are optimal and the sources of welfare gains from more sophisticated taxes when those conditions are not satisfied. Optimal tax rates for married individuals are increasing in correlation of spousal earnings but are lower than the tax rates for single individuals, and the marginal rates for one spouse increase (decrease) in the earnings of the other when both spouses have low (high) earnings.

1   Introduction

A significant part of income taxes in developed countries is paid by households that consist of several adult members. For example, in the U.S., over 70 percent of federal individual income taxes are collected from married couples. Yet, the theory of optimal taxation of family income is poorly understood. What economic forces determine the shape and magnitude of the optimal joint earnings tax schedule? How should one member's taxes depend on the earnings of the other member? Is it ever optimal to tax each individual in a couple separately or to use total family income as a base for earnings taxation?

In this paper, we take a step to answer these questions. We use the mechanism design approach. The optimal tax problem can be equivalently cast as a problem of a fictitious mechanism designer who chooses allocations for all households based on reports about their productivities subject to the incentive compatibility constraints. Optimal taxes for married households are characterized by the solution to a certain second-order non-linear partial differential equation. While there are no known techniques to solve this equation explicitly, we show that this roadblock can be partially side-stepped using a mathematical result known as the Coarea formula. It allows us to derive closed-form expressions for various conditional moments that these taxes must satisfy. The Coarea formula shows that, in a very broad sense, optimal taxes for married coples are determined by a trade-off between benefits from redistribution, captured by Pareto weights, and deadweight costs of taxation, captured by elasticities of labor supplies of the spouses and statistics summarizing the joint distribution of earnings of the spouses.

We first develop our approach in the simplest or *benchmark* economy. All individuals in the benchmark economy are ex-ante identical and have the same preferences, that for tractability we assume to be quasi-linear with constant elasticity of labor supply. All individuals draw their productivities from the same distribution. Before their productivity draws are realized, they decide whether to stay single or go to the marriage market. Individuals on the marriage market receive signals about their future productivity, form couples with other individuals on the marriage market based on these signals, and agree to share their marital surplus equally. After productivity draws are realized, married and single individuals work, pay taxes, and consume. The social planner chooses taxes for single and married households to maximize Pareto-weighted sum of expected utilities of individuals.

While simple, the benchmark economy allows us to focus on some key characteristics of taxation. In our economy, single and married co-exist allowing us to compare optimal taxes on one- and two-person households in the same setting. Our marriage formation process

admits arbitrary correlation in earnings within couples, which is important since in the data spousal earnings are positively but imperfectly correlated. The social planner only cares about redistribution across individuals with different expected utilities and does not have any inherent preference for or against marriage.

The benchmark economy allows us also to investigate an important technical question in multi-dimensional taxation: whether the first-order approach (FOA) is valid when agents characteristics are multi-dimensional. The FOA simplifies the mechanism design problem by restricting attention to only local incentive constraints. The FOA has been the standard technique to analyze uni-dimensional tax models since the seminal work of Mirrlees (1971) but there has been a lingering suspicion that it might generically fail in higher dimensions.[1] We derive necessary and sufficient conditions for validity of FOA for both single and married households in our benchmark economy when matching on thew marriage market is random. Comparison of these conditions reveals that conditions for validity of FOA for married couples are *less stringent* for for single households, so that the FOA is *more likely* to hold in these bi-dimensional settings.[2]

Using this insight, we analyze properties of optimal taxes for single and married households under FOA. We derive closed-form expressions for the optimal tax distortions[3] for single and married households. The optimal tax distortions for single persons are exactly the same as in the classical Mirrleesian economy and can be represented by the famous ABC formula due to Diamond (1998). The Coarea formula gives closed-form expressions for a rich set of conditional moments that characterize optimal taxes for married couples.

One of the key insight of the analysis is that the planner, even though she does not have an inherent stance for or against marriage, sets taxes that encourage couples formation. Optimal

---

[1] For example, in their classic study of the optimal taxation of couples, Kleven et al. (2009) write (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."

[2] This result also allows us to explain that the failure of the FOA in multi-dimensional settings that was observed by Armstrong (1996) and Rochet and Chone (1998) is driven not by incentive constraints per se but by their interaction with participation constraints, which are typically absent in public finance applications.

[3] As it is standard in the optimal tax literature since the work of Mirrlees (1971), we characterize properties of the monotone transformation of the marginal tax rates, $\frac{\partial}{\partial y_i} T(y_1, y_2) \left/ \left(1 - \frac{\partial}{\partial y_i} T(y_1, y_2)\right)\right.$, where T is the tax function and y_i is earnings of spouse i. We refer to this object as the optimal tax distortion.

tax rates for married persons are lower than for single, and the marriage rate is higher than in the laissez-faire. The social planner cares about redistribution between more and less productive individuals, and uses the tax system to provide this redistribution. Taxation, however, has deadweight costs. When individuals form couples, they pool their resources, implicitly redistributing resources within families. It is costly to crowd out this intra-family redistribution through distortionary taxation. As a result, the optimal tax rates are lower for married households, which also incentivizes marriages.

Using our analytical formulas we can study how optimal distortions of a married household depends on the degree of correlation in productivities of married couples. When matching into couples is random, so that those productivities are uncorrelated, the optimal tax distortions for married persons are exactly one half of those for single persons. The cost of tax distortion is exactly the same for single and married persons but random matching implies that the benefits of redistribution are cut in half, leading to this result. The more correlated spousal productivities are the higher are optimal distortions are on average, but they always remain below optimal distortions for single persons unless matching is perfectly assortative, in which case the two coincide.

Within couples, secondary earners – spouses with much lower productivity then their partners – generally face higher distortions than primary earners under the optimal tax code. The planner who sufficiently cares about redistribution wants to target transfers to the poorest individuals. Those transfers need to be phased out as secondary earner's income increases, leading to high labor distortions.

We also study how marginal taxes on one spouse dependent on earnings of the other, to which we refer to as jointness. When matching is random, the optimal taxes for married households are disjointed, so that taxes on one spouse do not depend on earnings of the other. More generally, optimal jointness depends on two terms: how much more redistributed the jointed tax system provides relative to a disjointed tax system, and how much extra distortions it imposes. Positive jointness allows the planner to target taxes to the richest couples, negative jointness allows the planner to target transfers to the poorest couples. Both types of jointness exacerbate distortions from the best dis-jointed tax system. In the tails, for very productive and very unproductive couples, the effect from exacerbation of distortions dominates redistributory benefits from jointness for a large class of commonly used joined distributions. As the result, the optimal jointness is negative at the top and positive at the bottom.

Once we complete our investigation of the optimal taxation in benchmark economy, we consider a number of extensions. We show implications for optimal taxation of home production

and economies of scale in consumption that marriage allows, bargaining over allocation of resources within couples, extensive marginal labor supply adjustment, heterogeneous selection into marriage, gender differences. We also explore implications for optimal taxation when the social planner has alternative social objective functions or when restricted to use simpler taxes. The Coarea formula approach that we developed in the benchmark economy allows us to derive a number of sharp implications of these extensions.

In the last part of the paper, we characterize optimal taxes numerically. 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 a Gaussian copula with Pareto-lognormal marginal distributions can well approximate this distribution. We find that our analytical formulas provide excellent guidance about numerical properties of the optimal tax schedule. In the U.S. data, spousal productivities are positively but not perfectly dependent, so 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). The Gaussian copula is tail independent, so consistent with our analytical results, optimal jointness is positive for low earners and negative for high earners. The quantitative magnitude of this optimal jointness is small, so individual earnings-based taxes provide a good approximation to the optimal unrestricted tax schedule. In contrast, taxation based only on total family earnings is generally quite far from the unrestricted optimum, even when Pareto weights are chosen to explicitly favor family earnings-based taxation.

Our paper is related to several strands of literature. Small literature in public finance uses the multi-dimensional mechanism design approach to study optimal taxation. Mirrlees (1976) was perhaps the first to derive the partial differential equation that characterizes such taxes under the assumption that the FOA is valid but noted that it is much more difficult to solve than its uni-dimensional analog. Several authors imposed additional assumptions to simplify the multi-dimensional tax environment. 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 the generalization of their ideas of redistributional arbitrage to multi-dimensional type spaces. In a series of papers, Roth-

schild 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 properties of optimal taxes in a fairly unrestricted multi-dimensional environment and shed light on economic forces that are hard to see in more specialized settings.[4]

The most closely related to our study is the unpublished Section 3 of the working paper by Kleven et al. (2007), henceforth KKS. In that section, KKS considered an economy populated by married couples and made several very insightful observations. They noted that one should expect the FOA to hold in optimal tax settings when the planner is close to utilitarian, derived a formula that is analogous to our formula for the optimal average distortion, and characterized the sign of optimal jointness under the assumption that spousal productivity draws are independent. Our work systematically builds on these insights and allows us to obtain many novel results, such as a comparison of conditions for the validity of FOA in uni- and bi-dimensional settings, comparative statics results, implications of productivity dependence on optimal taxation, comparison of optimal tax rates for single and married individuals or spouses in the same couple, and implications of incorporation of various additional mechanisms in benchmark model, such as home production or economies of scale, that have been extensively studies in the family economics literature.

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.[5] 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 for our purposes 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's structural parameters affect optimal taxes. In contrast, our formulas are in terms of exogenous primitives, which allows us to prove sharp theoretical results. That being said, we show in the paper how our formulas can be obtained using variational techniques by constructing perturbations that allows one to express optimal

---

[4]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.

[5]More recently, some ideas of this approach have been successfully applied by Bierbrauer et al. (2023) to detect Pareto inefficiencies in the joint taxation of couples income.

taxes in terms of model's primitives. Those perturbations differ from ones typically considered in the literature and should be of independent interest.

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. Analytical results we derive provide insights about forces that determine optimal taxes that are often hard to see with numerical work.

The rest of the paper is organized as follows. In Section 2, we present our environment. In Section 3, we describe the mechanism design approach and conditions for the validity of the FOA. Section 4 characterizes optimal taxes in the benchmark economy. Section 5 considers its various extensions. Section 6 provides calibration and quantitative analysis. Section 7 concludes.

2   The benchmark environment

In this section, we describe the benchmark economy. This economy allows us to present our main ideas and results in the most transparent way. Later in the paper, in Section 5, we consider various extensions and modifications of this economy to highlight implications of additional economic forces from which the benchmark economy abstracts.

The benchmark economy is populated by measure one of ex-ante identical individuals or persons. Each individual has utility $c - \gamma l^{1/\gamma}$, where c and l denote consumption and labor, and $\gamma \in (0, 1)$ is the parameter capturing the elasticity of labor supply.[6] Each person decides whether to stay single or get married, and how much to work and consume. All decisions occur in three stages.

Stage 1. Each person draws a preference shock ε that captures non-pecuniary benefits of marriage. Let $\mathbb{E}U^s$ and $\mathbb{E}U^m$ be expected utilities from consumption and labor that person i obtains if that person decides to stay single and get married. The person stays single if $\mathbb{E}U^s + \varepsilon \geq \mathbb{E}U^m$, and goes to the marriage market otherwise.

Preference shocks ε are i.i.d. across agents and independent of all other shocks. They are drawn from an absolutely continuous probability distribution that has support on $\mathbb{R}$ and strictly positive density. We use Φ to denote the inverse of the cdf of this distribution.

Stage 2. Each person who went on the marriage market receives a signal q. This signal is observable to all individuals on the marriage market. Signal q is uncorrelated across individuals but may be correlated with productivity w, which is realized in Stage 3. Individuals on the

---

[6]In particular, the relationship between the elasticity of labor supply e and the parameter γ is $1/\gamma = 1 + 1/e$.

marriage market use these signals to form married couples. In particular, each person on the marriage market marries another person on that market who has the same signal, and they agree to share their marital surplus equally.

We use μ to denote the marriage rate, i.e., the fraction of individuals who get married.

Stage 3. Each person, married or single, draws productivity w from a cumulative probability distribution G. A person with productivity w who works l hours earns income $y = wl$. After productivity draws are realized, single and married households decide how much to work and consume taking into account taxes T^s and T^m for single and married households. The decision problem of a single person household is

$$v^s(w) := \max_{(c,y)} c - \gamma \left(\frac{y}{w}\right)^{1/\gamma} \quad s.t. c \le y - T^s(y), y \ge 0.$$

The decision problem of a married household is

$$v^m(w_1, w_2) := \max_{(c_i, y_i)_{i=1}^2} \sum_{i=1}^2 \left(c_i - \gamma \left(\frac{y_i}{w_i}\right)^{1/\gamma}\right) \quad s.t. \sum_{i=1}^2 c_i \le \sum_{i=1}^2 y_i - T^m(y_1, y_2), (y_1, y_2) \ge 0.$$

Let $U^s(w) = v^s(w)$ be utility from consumption and labor of a single person with productivity w, and $U^m(w_i|w_{-i})$ be utility of a married person with productivity w_i whose partner has productivity w_{-i}. Equal surplus division agreed in Stage 2 implies that $U^m(w_1|w_2) = U^m(w_2|w_1) = \frac{1}{2} v^m(w_1, w_2)$.[7]

We use $F(w_1, w_2)$ to denote the joint distribution of productivities of couples, which we can assume, without loss of generality, to be symmetric. Let $F(w_j|w_i)$ to denote the conditional distribution of productivities of a married individual with productivity w_i. Given these definitions, expected utilities $\mathbb{E} U^s$ and $\mathbb{E} U^m$ used in Stage 1 are given by

$$\mathbb{E} U^s = \int U^s(w) G(dw), \qquad \mathbb{E} U^m = \int \left(\int U^m(w_i|w_{-i}) F(dw_{-i}|w_i)\right) G(dw_i).$$

To streamline our exposition and simplify technical details, we assume that G has domain $\mathbb{R}_+$, density g, and satisfies $\int w^{1/(1-\gamma)} dG < \infty$. In our benchmark economy, signals q do not play an important role and we do not need to specify how they are correlated with w; it suffices to take F as the primitive of this economy. We assume that F has density denoted by f. Our assumptions imply that F is symmetric with marginals G, and the distribution of productivities of single persons is also G.

The benchmark economy is purposefully set up to isolate key implications of optimal taxation of single and married households in the simplest settings. It is also very amendable to

---

[7]Maximization problem that defines v^m pins down only the sum $C = c_1 + c_2$. Surplus division rule allocates that C between the two spouses.

extensions since, by slightly changing details of the events in each stage or the order at which different stages occur, we can incorporate a variety of economic mechanisms related to marriage that has been emphasized by the empirical labor literature. For now, we review some of the main features of our benchmark settings. All individuals are ex-ante identical. They marry before they know their productivities and share their realized marital surplus equally. Marriages are potentially assortative, so that spousal productivities might be correlated. Signals q can be interpreted as education or other observable socio-economic indicators that individuals use when selecting marriage partners. As mentioned above, these signals in the benchmark economy merely provides a micro-foundation for F and equal surplus division, and play no role otherwise; they will be helpful once we allow for marriage rates to be correlated with productivities in Section 5.7. Marriage decisions are influenced both by taxes, that affect pecuniary utilities $\mathbb{E}U^s$ and $\mathbb{E}U^m$, and by non-pecuniary benefits ε ("love"). Distribution Φ captures the relative importance of the two forces and the responsiveness of marriage rates to taxes. Our assumptions imply that $\mu \in (0,1)$ and Φ is differentiable with strictly positive derivative, so that the elasticity of marriage to changes in pecuniary benefits is not zero. This streamlines our discussion as we do not need to worry about corner solutions. Exogenous marriages, that do not respond to taxes, can be modeled as a distribution of ε that assigns positive probabilities only to two values $\bar{\varepsilon}$ and $-\bar{\varepsilon}$ for some large $\bar{\varepsilon}$. While this case does not fit into out framework, it can be approximated arbitrarily closely by choosing appropriate Φ.

Earnings of married spouses are positively correlated in the data. The joint distribution F allows us to capture correlation. It will be useful to review general statistical notions of dependence.[8] We say that productivities are independent, or *matching* is *random,* if $F(w_1, w_2) = G(w_1)G(w_2)$ for all w_1, w_2, or $F = G^2$ for short. Productivities are *positively dependent* if $F \geq G^2$. Positive dependence (also known in statistics as positive quadrant dependence or PQD) is equivalent to the condition that $\mathbb{C}ov\left(\phi_1\left(w_1\right), \phi_2\left(w_2\right)\right) \geq 0$ for any two increasing functions ϕ_1 and ϕ_2. Distribution F^b is *more dependent* than F^a if F^a and F^b have the same marginals and $F^b \geq F^a$. We denote this by $F^b \geq_{PQD} F^a$. In statistics, this is known as the positive quadrant dependence order, and it allows us to conduct comparative statics analysis with respect to a degree of assortativeness in productivities in general, non-parametric settings. Any F satisfies bounds $F^* \geq_{PQD} F \geq_{PQD} F_*$, where F^* and F_* are distributions under perfect positive and negative assortative matching. We allow our economy to include the case of perfect assortative matching as it provides a useful bound even though strictly speaking it does not satisfy our stated assumptions as F^* does not have density.

---

[8]See Nelsen (2006); Shaked and Shanthikumar (2007) for textbook discussion of dependence concepts for bi-variant random variables.

The main focus of our paper is on characterizing properties of optimal taxation. Before we study it, it will be useful to describe the equilibrium of our model in the absence of taxes. We refer to this economy as *laissez-faire* and use superscripts LF to denote allocations there.

Lemma 1. *In the laissez-faire,* $\mathbb{E}U^{m,LF} = \mathbb{E}U^{s,LF}$*, and both* $\mathbb{E}U^{m,LF}$ *and* μ^{LF} *are independent of* F *(holding* G *fixed).*

In the laissez-faire, the expected utility of consumption and labor is the same for all persons, married and single, and this utility is independent of assortativeness of matching in the marriage market. In laissez-faire economy, the total amount of resources available for single and married persons is the same. Marriage and assortativeness of matches introduces uncertainty into how those resources get eventually allocated but individuals are risk neutral and this uncertainty does not affect their marriage decisions and ex-ante utility.

3 Optimal taxation as a mechanism design problem

We now turn to characterizing optimal taxes. We first need to specify the social objective that the planner, who sets these taxes, maximizes. In the benchmark economy, we assume that this social objective is given by $\mathcal{W} := \int \alpha(w)\mathbb{E}\left[U|w\right]G(dw)$, where $\mathbb{E}\left[U|w\right]$ is the expected utility of a person with productivity w and function α captures Pareto weights. We assume that α is non-negative, strictly decreasing, bounded, continuous function normalized so that $\int \alpha(w)G(dw) = 1$.

It will be useful to re-write $\mathcal{W}$ in terms of v^s and v^m that we introduced at Stage 3. Let function $\widetilde{\alpha}$ be defined by

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

and observe that

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

Substitute this expression into the definition of $\mathcal{W}$ and re-arrange terms to see that welfare can be written as

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

The planner chooses taxes T^s and T^m to maximize this welfare. Taxes must be budget-feasible, so that total tax revenues are non-negative, but can be arbitrary functions otherwise. We refer to them as optimal (unrestricted) taxes. In Section 5.9 we consider optimal taxes that have additional restrictions imposed on them and show that there is a close relationship

between optimal unrestricted and restricted taxation. Note that the domain of T^m is $\mathbb{R}^2_+$, so that it is a bi-dimensional function.

3.1 The mechanism design problem

We study our optimal tax problem using the mechanism design approach. Our steps in setting up the mechanism design problem, simplifying it, and deriving its optimality condition are standard, and we present them in the body of the paper heuristically. At the same time, some care is needed if one wants to use these conditions to verify sufficiency, i.e., whether tax functions that satisfy these conditions are indeed optimal. For this reason, in the appendix we state our mechanism design problem formally, being explicit about functional spaces in which various functions live, bounds that they satisfy, and notions of differentiability that we apply.

We use bold letters, such as $\mathbf{w}$, to denote pairs (w_1, w_2) for a married couple. Using standard Taxation principle arguments (see, e.g., Hammond (1979)) one can show that T^s, T^m are budget feasible if and only if there exists $\mu \in [0, 1]$ and tuples (v^s, c^s, y^s) and $(v^m, \mathbf{c}^m, \mathbf{y}^m)$ that satisfy

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

$$v^s(w) \geq c^s(\widehat{w}) - \gamma \left(\frac{y^s(\widehat{w})}{w} \right)^{1/\gamma}, v^m(\mathbf{w}) \geq \sum_{i=1}^{2} \left(c_i^m(\widehat{\mathbf{w}}) - \gamma \left(\frac{y_i^m(\widehat{\mathbf{w}})}{w_i} \right)^{1/\gamma} \right) \forall w, \widehat{w}, \mathbf{w}, \widehat{\mathbf{w}} \quad (4)$$

$$\frac{\mu}{2} \int \sum_{i=1}^{2} (y_i^m(\mathbf{w}) - c_i^m(\mathbf{w})) F(d\mathbf{w}) + (1 - \mu) \int (y^s(w) - c^s(w)) G(dw) \geq 0, \quad (5)$$

$$\Phi(\mu) = \frac{1}{2} \int v^m(\mathbf{w}) F(d\mathbf{w}) - \int v^s(w) G(dw). \quad (6)$$

Equation (3) gives definitions of utilities of married and single households, and equation (4) are the incentive constraints. Equation (5) says that total consumption cannot exceed total earnings. Finally, equation (6) shows that the fraction of married households is determined by the value of ε at which a person is indifferent between being married and staying single.[9] From any allocation that satisfies these constraints one can construct budget-feasible tax functions T^s and T^m that decentralize this allocation as households optimal choices given these taxes. Thus, the optimal tax problem can be stated as a choice of μ, (v^s, c^s, y^s), $(v^m, \mathbf{c}^m, \mathbf{y}^m)$ that maximize welfare (2) subject to constraints (3) − (6).

---

[9]The marriage rate μ is given by equation $\mu = \Pr(\varepsilon \leq \mathbb{E}U^m - \mathbb{E}U^s)$. Inverting $\Pr(\cdot)$, we obtain $\Phi(\mu) = \mathbb{E}U^m - \mathbb{E}U^s$, which is equation (6) written in terms of v^m and v^s.

This problem can be simplified. Note that the incentive constraints (4) imply, due the Envelope Theorem, the following relationships:

$$\frac{\partial v^s(w)}{\partial w} = \frac{(y^s(w))^{1/\gamma}}{w^{1+1/\gamma}}, \quad \frac{\partial v^m(\mathbf{w})}{\partial w_i} = \frac{(y_i^m(\mathbf{w}))^{1/\gamma}}{w_i^{1+1/\gamma}} \text{ for } i = 1, 2. \tag{7}$$

We can use (7) to substitute out for y^s, $\mathbf{y}^m$ and (3) to substitute for c^s, $c_1^m + c_2^m$ in the constraint set of the mechanism design problem. This allows to write the mechanism design problem as a choice of μ, v^s, v^m that maximize welfare (2) subject to (6),

$$\frac{\mu}{2} \int \left(\sum_{i=1}^{2} \left(w_i^{1+\gamma} \left(\frac{\partial v^m(\mathbf{w})}{\partial w_i} \right)^{\gamma} - \gamma w_i \frac{\partial v^m(\mathbf{w})}{\partial w_i} \right) - v^m(\mathbf{w}) \right) F(d\mathbf{w}) +$$

$$+ (1-\mu) \int \left(w^{1+\gamma} \left(\frac{\partial v^s(w)}{\partial w} \right)^{\gamma} - \gamma w \frac{\partial v^s(w)}{\partial w} - v^s(w) \right) G(dw) \geq 0, \tag{8}$$

and, for all $w, \widehat{w}, \mathbf{w}, \widehat{\mathbf{w}}$,

$$v^s(w) \;\; \geqslant \;\; v^s(\widehat{w}) + \gamma \widehat{w} \frac{\partial v^s(\widehat{w})}{\partial w} \left(\left(\frac{\widehat{w}}{w} \right)^{1/\gamma} - 1 \right), \tag{9}$$

$$v^m(\mathbf{w}) \;\; \geqslant \;\; v^m(\widehat{\mathbf{w}}) + \sum_{i=1}^{2} \gamma \widehat{w}_i \frac{\partial v^m(\widehat{\mathbf{w}})}{\partial w_i} \left(\left(\frac{\widehat{w}_i}{w_i} \right)^{1/\gamma} - 1 \right). \tag{10}$$

Observe that all the local incentive constraints are already inside of the objective function (8), so constraints (9), (10) can affect the solution to this maximization problem only if some of the non-local constraints bind. The *relaxed problem* is the mechanism design problem in which these constraints are dropped. Let $(\mu^*, v^{s,*}, v^{m,*})$ be the solution to this relaxed problem. We say that *the first-order approach (FOA) is valid* if $(\mu^*, v^{s,*}, v^{m,*})$ is also the solution to the original problem. We say that the *FOA is valid for single households* if the solution to the relaxed problem satisfies (9) for single but not necessarily (10) for married households, and the *FOA is valid for married households* in the opposite case.

3.2   Validity of the FOA

The analysis of the mechanism design problem substantially simplifies when the FOA is valid. In uni-dimensional settings, essentially all papers that characterize optimal taxes analytically (e.g., Mirrlees (1971), Diamond (1998), Saez (2001)) assume that the FOA holds. While it is known that the FOA may fail in uni-dimensional settings for some parameter values, those cases appear to be rare in realistic applications that verify FOA validity numerically (see, e.g., Farhi and Werning (2013), Golosov et al. (2016), or Heathcote and Tsujiyama (2021)). One

common concern, exemplified by the quote from Kleven et al. (2009) given in the introduction, is whether the FOA is ever valid in multi-dimensional settings.

In this section, we examine conditions for validity of the FOA in our benchmark economy when matching is random. In this case, we can derive explicitly necessary and sufficient conditions for the validity of FOA for both single and married households, and compare them. It turns out that conditions for validity of FOA for married households are less stringent than for single households, and so the FOA is more likely to hold in the bi-dimensional case.

Proposition 1. *Consider the benchmark economy with random matching. Define* $\lambda^{\#}(t) := \frac{\int_t^{\infty}(1-\alpha(w))g(w)dw}{\gamma t g(t)}$ *and assume that it is bounded, continuously differentiable with bounded derivatives.*

The FOA for single households is valid if and only if

$$x \cdot \left(1 + \lambda^{\#}\left(x^{-\gamma}\right)\right) \quad \text{is increasing in } x. \tag{11}$$

The FOA for married households is valid if and only if

$$x \cdot \left(1 + \frac{1}{2}\lambda^{\#}\left(x^{-\gamma}\right)\right) \quad \text{is increasing in } x. \tag{12}$$

In particular, (12) holds whenever (11) holds.

The proof of this proposition is given in the appendix. It consists of two steps. First, we reformulate our mechanism design problem in transformed type variables $x = w^{-1/\gamma}$. Using this transformation has an advantage because the global incentive constraints (9) are linear in x. The results of Rochet (1987) for such problems imply that validity of FOA is equivalent to the convexity of solution to the relaxed problem. The second step is to solve the relaxed problem explicitly, which it possible to do when matching is random, and verify conditions for convexity directly. Solutions to the relaxed problems for single and married are convex if and only if equations (11) and (12) hold. Examination of these equations reveals that the FOA is more likely to hold for married than for single households.

It is insightful to consider the economic interpretation of equations (11) and (12). Condition (11) can equivalently be written as

$$\left[1 + \lambda^{\#}(w)\left(1 + \gamma\frac{\partial \ln(wg(w))}{\partial \ln w}\right)\right] + [1 - \alpha(w)] \geq 0 \text{ for all } w. \tag{13}$$

The term in the first square brackets is typically positive; the term in the second square brackets is negative for low w and positive for high w. Thus, inequality (13) 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 (13). The analogous conditions for married households is similar but allows for a larger set of weights α. We discuss the intuition for this result once we characterize optimal taxes in Section 4.

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. The surplus extraction problem is isomorphic to a very particular, highly redistributive set of social weights for which indeed the FOA fails in more than one dimension. But the FOA still holds for a wide class of Pareto weights used in applied work.[10]

Motivated by this finding, we assume for the rest of the paper that the FOA is valid and describe properties of the optimal taxes under this assumption. In our quantitative analysis in Section 6 we numerically check validity of the FOA and finds that it holds in all cases that we consider in the calibrated economy.

3.3 Optimality conditions to the relaxed problem

We now characterize optimality conditions to the relaxed mechanism design problem. It will be convenient to state them not in terms of $v^{s,*}$, $v^{m,*}$ but transformations of these functions $\lambda^{s,*}$ and $\boldsymbol{\lambda}^{m,*} = \left(\lambda_1^{m,*}, \lambda_2^{m,*}\right)$ defined by

$$\lambda^{s,*}(w) := \left(\frac{\partial v^{s,*}(w)}{\partial w}\right)^{\gamma-1} w^{\gamma} - 1, \quad \lambda_i^{m,*}(\mathbf{w}) := \left(\frac{\partial v^{m,*}(\mathbf{w})}{\partial w_i}\right)^{\gamma-1} w_i^{\gamma} - 1. \qquad (14)$$

These transformations are closely related to the optimal taxes $T^{s,*}$, $T^{m,*}$ via

$$\frac{\frac{\partial}{\partial y}T^{s,*}(y^*(w))}{1 - \frac{\partial}{\partial y}T^{s,*}(y^*(w))} = \lambda^{s,*}(w), \qquad \frac{\frac{\partial}{\partial y_i}T^{m,*}(\mathbf{y}^*(\mathbf{w}))}{1 - \frac{\partial}{\partial y_i}T^{m,*}(\mathbf{y}^*(\mathbf{w}))} = \lambda_i^{m,*}(\mathbf{w}), \qquad (15)$$

where $y^*(w)$, $\mathbf{y}^*(\mathbf{w})$ correspond to optimal choices of earnings by single and married households under the optimal tax system. Thus, $\lambda^{s,*}$ and $\boldsymbol{\lambda}^{m,*}$ are monotone transformations of optimal marginal tax rates, and we refer to $\lambda^{s,*}$ and $\boldsymbol{\lambda}^{m,*}$ as *optimal distortions*. $\boldsymbol{\lambda}^{m,*}$ is also closely related to the cross-partial derivative of $T^{m,*}$ and satisfies $sign\left(\frac{\partial^2}{\partial y_1 \partial y_2}T^{m,*}(\mathbf{y}^*(\mathbf{w}))\right) = sign\left(\frac{\partial}{\partial w_{-i}}\lambda_i^*(\mathbf{w})\right)$ for all i. We say that $T^{m,*}$ is *positively (negatively) jointed* at $\mathbf{w}$ if this

---

[10]We discuss these conditions and present formal results for a broad class of multi-dimensional tax problems in Golosov and Krasikov (2023). See also Kleven et al. (2007) for a related discussion.

sign is positive (negative), $T^{m,*}$ is *separable* or *disjointed* if this cross partial is always zero. Separable taxes can be written as $T^{m,*}(y_1, y_2) = \widetilde{T}^{m,*}(y_1) + \widetilde{T}^{m,*}(y_2)$.

We now derive the optimality conditions for the relaxed mechanism design problem. It is easy to show, using the fact that $\int \alpha(w)G(dw) = \int \alpha^m(\boldsymbol{w})F(d\boldsymbol{w}) = 1$, that the Lagrange multiplier on (8) is equal to one. Standard variational arguments can be used to show that $v^{s,*}$ satisfies

$$\frac{\partial}{\partial w}\left(\lambda^{s,*}(w)\gamma w g(w)\right) = (\alpha(w) - 1)g(w) \tag{16}$$

with a boundary condition

$$\lim_{w \to 0,\infty} \lambda^{s,*}(w)wg(w) = 0. \tag{17}$$

Similarly, the optimality conditions for $v^{m,*}$ are given by

$$\sum_{i=1}^{2}\frac{\partial}{\partial w_i}\left(\lambda_i^{m,*}(\mathbf{w})\gamma w_i f(\mathbf{w})\right) = (\alpha^m(\mathbf{w}) - 1)f(\mathbf{w}), \tag{18}$$

with a boundary condition

$$\lim_{w_i \to 0,\infty} \lambda_i^{m,*}(\mathbf{w})w_i f(\mathbf{w}) = 0 \text{ for all } w_{-i}. \tag{19}$$

Moreover, the cross-partials of $v^{m,*}$ must agree, in the sense that $\frac{\partial^2 v^{m,*}}{\partial w_1 \partial w_2} = \frac{\partial^2 v^{m,*}}{\partial w_2 \partial w_1}$, which can be written in terms of $\boldsymbol{\lambda}^{m,*}$ as

$$\frac{\partial}{\partial \ln w_2}\left(\frac{w_1}{1 + \lambda_1^{m,*}(\mathbf{w})}\right)^{1/(1-\gamma)} = \frac{\partial}{\partial \ln w_1}\left(\frac{w_1}{1 + \lambda_1^{m,*}(\mathbf{w})}\right)^{1/(1-\gamma)}. \tag{20}$$

The first order condition with respect to μ can be expressed as

$$\frac{1-\gamma}{2}\int \sum_{i=1}^{2} w_i \left(\frac{w_i}{1 + \lambda_i^{m,*}(\mathbf{w})}\right)^{\gamma/(1-\gamma)} F(d\mathbf{w}) - (1-\gamma)\int w\left(\frac{w}{1 + \lambda^{s,*}(w)}\right)^{\gamma/(1-\gamma)} G(dw) =$$

$$= \Phi(\mu^*). \tag{21}$$

These conditions are both necessary and sufficient for characterization of optimal taxes under mild boundedness assumptions stated in the appendix.[11] sufficient

4 Optimal distortions for single and married households

In this section, we use equations (16) – (21) to characterize properties of optimal taxes. As in Section 3, we focus in the body of our paper on showing the economic insights. Some results,

---

[11]Those conditions are similar to the conditions on $\widetilde{\lambda}$ that we used in Proposition 1. In fact, it can be shown that $\lambda^{s,*} = \widetilde{\lambda}$ in that economy.

mainly those requiring us to take limits $w \to 0, \infty$, require additional regularity conditions F, G must satisfy. Stating and discussing those conditions interrupts the flow of the presentation and we relegate them to the appendix.

Observe that there are no linkages between the ordinary differential equation (16) – (17) that characterizes $\lambda^{s,*}$, and a system of partial differential equations (18) – (20) that characterize $\boldsymbol{\lambda}^{m,*}$. This is due to the fact that the Lagrange multiplier on (6) is equal to zero. This result has a natural economic interpretation. Constraint (6) captures how much the planner wants to redistribute between single and married. In our benchmark economy, the planner uses the same Pareto weight on any individual irrespective of their marital status, and thus constraint (6) is slack.

Since there are no linkages in the optimality conditions for single and married, the corresponding optimal distortions can be solved independently of each other, and that they do not depend on the marriage rate μ^*. Distortions $\lambda^{s,*}$ and $\boldsymbol{\lambda}^{m,*}$ fully characterize optimal marginal taxes $\frac{\partial}{\partial y} T^{s,*}$ and $\left(\frac{\partial}{\partial y_i} T^{m,*} \right)_{i=1}^2$. The optimal average taxes also depend on intercepts $T^{s,*}(0)$ and $T^{m,*}(0,0)$ that are determined, together with μ^*, by equation (21).

Because optimal distortions for single households are independent of the marriage rate, they coincide with optimal distortions in the economy that has only single households, which is isomorphic to an environment that has been studied extensively in public finance. Integrate equation (16) from t to ∞, using the boundary conditions (19), to find the closed-form expression for $\lambda^{s,*}$:

$$\lambda^{s,*}(t) = \frac{1 - \mathbb{E}\left[\alpha | w \geq t \right]}{\gamma \theta(t)}, \tag{22}$$

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

$$\theta(t) := \frac{t g(t)}{1 - G(t)}. \tag{23}$$

Equation (22) is the optimal tax formula that is well-known from the work of Diamond (1998) and the textbook treatment in Salanie (2003). The optimal tax on singles features a familiar trade-off between benefits of redistribution and costs of taxation. A higher marginal tax rate on a single person with productivity t, returned back uniformly to all single households, increases average taxes for all $w > t$. The social value of a dollar in the hands of a single person with productivity w is $\alpha(w)$, the social value of a dollar in the hands of an average single person is $\int \alpha(w) G(dw) = 1$. Therefore, the numerator of (22) captures the benefits of redistribution. The denominator of (22) captures the cost of tax distortions. Those distortions arise because individuals with productivity $w = t$ reduce their earnings and hence tax revenues collected

from those households. The reduction in tax revenues is determined by elasticity parameter of labor supply γ, productivity t, and the mass of agents affected by this perturbation $g(t)$.

The optimality conditions for married persons is represented by a system of non-linear partial differential equations (18) – (20). Unfortunately, solving such equations in general is hard.[12] 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 (18) is a relatively tractable linear differential equation in $\boldsymbol{\lambda}^{m,*}$. As such, conditions (18) can be integrated over various subsets of $\mathbb{R}^2$ to find a variety of conditional averages of distortions $\boldsymbol{\lambda}^{m,*}$. The key mathematical tool that we will use is the Coarea Formula.[13] This formula states that for any function $Q : \mathbb{R}^2_{++} \to \mathbb{R}_{++}$ that satisfies mild technical restriction equations (18) and (19) imply

$$\mathbb{E}\left[\sum_{i=1}^{2} \lambda_i^{m,*} \frac{\partial \ln Q}{\partial \ln w_i} \Big| Q = t\right] = \frac{1 - \mathbb{E}\left[\alpha^m | Q \geq t\right]}{\gamma t \frac{-\partial \Pr(Q \geq t)}{\partial t} / \Pr\left(Q \geq t\right)}. \tag{24}$$

Formula (24) shows that optimal distortions for married households can be represented in the very general sense by the same trade-off between the benefits from redistribution and the costs of distortions the determines optimal taxes for single households. Equation (24) holds for any function Q, and by considering various such functions, one can obtain rich characterization of optimal distortions. The intuition for (24) can be obtained by studying perturbations of the joint tax function T whereby we increase tax levels by \$1 for all couples $\mathbf{w}$ that satisfy $Q(\mathbf{w}) > q$ and adjust the lump sum tax component to satisfy the government budget constraint.

As we pointed out above, the benchmark social planner values single and married persons equally and does not have inherent desire to redistribute between them. Nonetheless, optimal taxes do affect utilities of single and married persons differently, and hence they also affect the marriage rate. The next result compares those in the optimum and the laissez-faire.

Lemma 2. *In the optimum, $\frac{1}{2}v^{m,*} = \mathbb{E}U^{m,*} \geq \mathbb{E}U^{s,*} = v^{s,*}$ and $\mu^* \geq \mu^{LF}$. $\mathbb{E}U^{m,*}$, $\mathbb{E}U^{s,*}$ and μ^* depend on F.*

---

[12]Renes and Zoutman (2017) show that if $\boldsymbol{\lambda}^{m,*}$ is a conservative vector field (i.e., it is a gradient of some function so that $\frac{\partial}{\partial w_2}\lambda_1^{m,*} = \frac{\partial}{\partial w_1}\lambda_2^{m,*}$) then it can be characterized in some cases using so-called Green functions. Unfortunately, there is no reason in general for $\boldsymbol{\lambda}^{m,*}$ to be such a field; instead, the relevant auxiliary condition is (20) which equivalent to requiring that the gradient of $v^{m,*}$ is a conservative vector field. This equation is highly non-linear which makes it difficult to find $\boldsymbol{\lambda}^{m,*}$ or to use Green functions.

[13]The Coarea Formula is a generalization of Fubini's theorem that expresses an integral of a function in terms of integrals over the level sets of another function. We use it in conjunction with the Divergence Theorem, which is a multidimensional analogue of integration by parts, to integrate (18) over upper contour sets of Q and express the value of that integral in terms of conditional averages of the optimal distortions on level sets of Q. Different functions Q allows us to study different subsets, i.e., upper contour sets of Q. This becomes particularly convenient when subsets that we are interested in have relatively complicated structure. Those subsets emerge when we compare optimal distortions within couples or study conditions under which taxation of total family income $y_1 + y_2$ is optimal.

The key take-away from this lemma is that the social planner implicitly encourages marriages in the optimum. The marriage rate μ^* depends on how assortative the matching into marriage is, and hence on the implicit incentives for marriage that the tax system provides. Our analysis below will shed light on why this is the case. Before we study our economy for arbitrary F, it will be useful to focus on the special case when F is independent. This is the same economy we considered in Section 3.2.

Lemma 3. *In the benchmark economy with random matching, $\lambda_i^{m,*}$ is independent of w_{-i} and satisfies*

$$\lambda_i^{m,*}(t, w_{-i}) = \frac{1}{2}\lambda^{s,*}(t) \text{ for all } t, w_{-i}.$$

The optimal taxes $T^{m,*}$ are separable.

Lemma 3 shows that in the economy with random matching optimal taxes for married persons take a very simple form. A married person faces a tax on their earnings y of the form $\widetilde{T}^{m,*}(y)$, which is independent of earnings of their spouse. The marginal taxes for married persons are lower than for single, with optimal distortions for married persons being exactly one half of distortions for single persons with the same productivity.

To understand the intuition for this result, it is helpful to repeat the same perturbational thought experiment that we gave after equation (22) but now adapting it to $\widetilde{T}^{m,*}$. Recall that in our benchmark economy the distribution of productivities of single and married persons is the same, G. Thus, costs of tax distortions are the same for single and married, and are given by $\gamma\theta(t)$. In contrast, the benefits from redistribution are different for single and married persons. Married persons share resources with their spouses. Thus, the burden of extra \$1 of taxes on a married person is split within the couple. Therefore, the benefits of redistribution depend not only on the social weight of the statutory tax-payer but also on the social weight of their spouse. In the economy with random matching, a person of any productivity marries, in expectation, a spouse with average productivity. This implies that the benefits from redistribution for married persons are cut in half, leading to the formula in Lemma 3.

The key insight of this discussion is that some redistribution of resources occurs within families. The planner values this redistribution and it is costly to crowd it out through taxation. Thus, the planner keeps the burden of taxation for married persons lower, implicitly incentivizing selection in the marriages as shown in Lemma 2.

When matching is non-random, optimal taxes for married households are, generally, non-separable and their characterization is more involved. However, one can show that the insights from our random matching economy continue to apply "on average" and further sharpen our

characterization of this result. Towards this end, set $Q(\mathbf{w}) = w_i$ in formula (24) to obtain an expression for the optimal average distortion for a married person with productivity t:

$$\mathbb{E}\left[\lambda_i^{m,*}|w_i = t\right] = \frac{1 - \mathbb{E}\left[\alpha^m|w_i \geq t\right]}{\gamma\theta(t)}. \tag{25}$$

Equation (25) is remarkably similar to equation (22). The intuition for it can be understood by considering a perturbation that raises \$1 of tax revenues from couples in which a spouse has productivity $w > t$. Since the distribution of productivities for married persons is G, the cost of tax distortion from the perturbation is the same as for single persons and given by $\gamma\theta(t)$. The benefit of redistribution depend on $\mathbb{E}\left[\alpha^m|w_i \geq t\right]$, which captures both the social weight of the person with productivity w and the weight of their spouse. One can show that this term is decreasing in the degree of assortativeness of marriage matching, leading to the following comparative statics result.

Lemma 4. *Consider two economies, a and b, which are identical in all respects except the joint distribution of productivities, and assume that $F^a \leq_{PQD} F^b$. Then the relationship between optimal distortions in the two economies is*

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

The second inequality becomes equality if F^b is perfectly assortative.
$\mathbb{E}^a\left[\lambda_i^{m,a,*}|w_i = t\right] \geq 0$ *for all t if F^a is positively dependent.*

Lemma 4 shows that optimal average distortions are ranked by dependence, so stronger dependence implies higher average distortions. The average distortions are highest in the economy with perfect assortative matching,[14] and in which case distortions for single and married persons coincide. In general, optimal distortions for a married person are lower on average than for a single person of the same productivity. Some redistribution occurs within couples even in the absence of taxation, and it is costly to crowd it out via distortionary taxation.

The last part of Lemma 4 extends the insight of Mirrlees (1971) that in the optimum labor distortions are non-negative. In bi-dimensional settings, this result requires, in addition to the assumption that Pareto weights are decreasing, that productivities are positively dependent. It is easy to construct examples with optimal negative distortions for some types when productivities are negatively correlated.[15]

---

[14]The same arguments also imply that distortions are lowest economies with perfect negative assortative matching.

[15]Just like in the uni-dimensional settings, the optimal labor distortions may be negative if Pareto weights are locally increasing for some w.

Our benchmark economy is summarized by three objects: the joint distribution F, the elasticity parameter γ, and Pareto weights α. Lemma 4 provides comparative statics for optimal average distortions with respect to the dependence embedded in F. It is immediate to use equation (25) to observe that $\mathbb{E}\left[\lambda_i^{m,*}|w_i = t\right]$ is decreasing in γ, so that higher labor supply elasticity leads to lower taxes. The next lemma provides a comparative statics with respect to α.

Lemma 5. *Consider two economies, a and b, which are identical in all respects except the Pareto weights, and assume that $\alpha^b(w)/\alpha^a(w)$ 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.*

Pareto weights α^b are more redistributory than α^a in the sense that they assign higher weights on lower types.[16] In a uni-dimensional model, more redistributive Pareto weights imply higher optimal distortions. In bi-dimensional settings, the same result holds of, in addition, f is log-supermodular. Supermodularity is satisfied by many commonly used positively-dependent distributions, e.g., by the bivariate log-normal distribution with a non-negative correlation parameter (see Karlin and Rinott (1980) for other examples).

Before we proceed with further analysis, it will be useful to point out that equations (22) and (25) can equivalently be written as

$$\lambda^{s,*}(t) = \frac{\mathbb{E}\left[\alpha|w \leq t\right] - 1}{\gamma t g(t)/G(t)}, \quad \mathbb{E}\left[\lambda_i^{m,*}\Big|w_i = t\right] = \frac{\mathbb{E}\left[\alpha^m|w_i \leq t\right] - 1}{\gamma t g(t)/G(t)}. \tag{26}$$

These equations highlight an alternative way to think about the trade-off that determines optimal distortions. Consider a perturbation in which the planner increases transfers by \$1 to the poorest persons (either single or married) that are then phased-out at some productivity level t, and the whole tax schedule (for single or married) is adjusted accordingly to be budget-neutral. The benefit of this redistribution is represented by the numerators of the expressions in (26). Phasing out of transfers distorts labor supply, and the cost of this distortion is captured by the denominators in (26). Equations (26) are mathematically equivalent to (22) and (25) as both raising taxes from the richest and giving transfers to the poorest creates the same changes in the tax schedule. This equivalence will no longer hold once we consider more sophisticated perturbations to discuss optimal jointness.

Equation (25) sheds light on how optimal distortions compare between married individuals with different productivities. Since it takes the same form as (22), the analysis of Diamond

---

[16]In mathematics, our definition of "more redistributory" weights is known as the uni-variant likelihood ratio order. See Chapters 1.C and 6.E of Shaked and Shanthikumar (2007).

(1998) applies directly. Another insightful way to describe optimal distortions for married persons, that does not have an analogue in uni-dimensional analysis, is to compare distortions between two spouses in the same couple. Let $\iota = w_2/w_1$ be the relative productivity of the two spouses in a couple, and G_ι be the cumulative distribution of ι implied by F. We use θ_ι to denote the tail statistics of G_ι defined by analogy with (23). By setting $Q(\mathbf{w}) = w_2/w_1$ in formula (24) we obtain

$$\mathbb{E}\left[\lambda_2^{m,*} - \lambda_1^{m,*} | w_2/w_1 = \iota\right] = \frac{1 - \mathbb{E}\left[\alpha^m | w_2/w_1 \geq \iota\right]}{\gamma \theta_\iota(\iota)}. \tag{27}$$

In this formula, $\lambda_2^{m,*}(\mathbf{w}) - \lambda_1^{m,*}(\mathbf{w})$ is the difference of distortions between the two spouses in the same couple, and $\mathbb{E}\left[\lambda_2^{m,*} - \lambda_1^{m,*} | w_2/w_1 = \iota\right]$ is the average of this difference across all couples whose relative productivity w_2/w_1 is equal to ι. Equation (27) shows that this average relative distortion is determined by a formula very similar to (25), except that the cost of distortions is summarized by θ_ι rather than θ. The next lemma derives implications of this equation for relative distortions between two spouses who are sufficiently different in their productivity.

Lemma 6. *If $\alpha(0) > 2$ and $\lim_{\iota \to \infty} \mathbb{E}[w_2 | w_2/w_1 \geq \iota] < \infty$, then $\mathbb{E}\left[\lambda_2^{m,*} - \lambda_1^{m,*} | w_2/w_1 = \iota\right] < 0$ for all sufficiently large ι.*

This lemma focuses on the case when the planner is sufficiently redistributive to low earners, in the sense that $\alpha(0) > 2$. The condition rules out extreme negative dependence when the most unproductive type, $w_1 = 0$, is matched to very productive types. This condition is satisfied if productivities are independent or log-normally distributed with correlation parameter $\rho \geq 0$.

The lemma shows that a spouses who have much lower productivity than their partner faces, on average, higher distortions. This result has a natural interpretation that secondary earners (i.e., spouses with lower earnings in a couple) face higher marginal tax rates than primary earners. The intuition for this result is that the redistributive planner targets transfers to couples with low earners. These transfers are being phased out as earnings of the secondary earner increase, leading to the high implicit marginal tax rates for such spouses. Note that this result also implies that family-earnings based taxation, i.e., a tax schedule of the form $T^m(y_1, y_2) = \widetilde{T}^m(y_1 + y_2)$, is generally suboptimal in the benchmark economy.

4.1 Optimal jointness

In this section, we study optimal jointness of couples taxation. To make our exposition most transparent, we focus on a particular measure of jointness, defined as

$$J(t) = \frac{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t, w_{-i} \geq t\right]}{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t, w_{-i} \leq t\right]} - 1, \tag{28}$$

which we refer to as *average jointness*. This measure is positive if a person's distortion are higher when married to a more productive spouse than when married to a less productive spouse.

As the first step to characterize properties of J, set $Q(w_1, w_2) = \min\{w_1, w_2\}$ in formula (24) and re-arrange to obtain

$$\mathbb{E}\left[\lambda_i^{m,*}|w_i = t, w_{-i} \geq t\right] = \frac{\Pr(w_{-i} \geq t|w_i \geq t)}{2\Pr(w_{-i} \geq t|w_i = t)} \frac{1 - \mathbb{E}\left[\alpha^m|\mathbf{w} \geq (t,t)\right]}{\gamma\theta(t)}. \tag{29}$$

This formula provides a closed-form expression for the average distortions conditional on being married to a more productive partner. By comparing this equation to the unconditional average, equation (25), we can derive the closed form expressions for J. In particular, it is easy to see that

$$J(t) \geq 0 \qquad \Longleftrightarrow \qquad \overline{A}(t) \times \overline{B}(t) \geq 1 \tag{30}$$

where

$$\overline{A}(t) := \frac{1 - \mathbb{E}\left[\alpha^m|\mathbf{w} \geq (t,t)\right]}{1 - \mathbb{E}\left[\alpha^m|w_i \geq t\right]}, \quad \overline{B}(t) := \frac{\Pr(w_{-i} \geq t|w_i \geq t)}{2\Pr(w_{-i} \geq t|w_i = t)}.$$

This expression shows that the sign of the jointness is determined by two forces, $\overline{A}$ and $\overline{B}$. As we explain below, $\overline{A}$ captures additional redistributive benefits of taxation that positive jointness gives to the planner, and this expression is greater than one; $\overline{B}$ captures additional distortions from positive jointness, and it is less than one. Thus, jointness is positive if the redistributive benefits outweight additional costs, and negative otherwise.

To develop the intuition for this expression, consider the following thought experiment. Take a disjointed tax schedule $T^m(y_1, y_2) = \widetilde{T}^m(y_1) + \widetilde{T}^m(y_2)$ and slightly perturb it by introducing jointness. Specifically, increase the level of taxes by some amount ϵ for all couples whose productivities satisfy $\min\{w_1, w_2\} > t$ and adjust the lump sum component of the tax schedule to satisfy the government budget constraint. Such perturbation increases marginal taxes for all individuals on the boundary of orthant $\{\mathbf{w} : \mathbf{w} > (t,t)\}$ by $d\tau$. Jointness is positive if $d\tau > 0$ and negative if $d\tau < 0$. Equivalently, this perturbation can be described as increasing tax rates on earnings in the interval $[y(t), y(t) + y'(t)dt]$ by $d\tau$.

This perturbation has redistributory benefits and distortionary costs. Positive jointness increases taxes for the richest couples, those located in the $\{\mathbf{w} : \mathbf{w} > (t,t)\}$ orthant, and redistributes them to an average couple. Thus, the redistributory benefits are given by

$$\mathfrak{B} = \nabla y(t) dt d\tau \int_t^\infty \int_t^\infty (1 - \alpha^m(\mathbf{w})) f(\mathbf{w}) d\mathbf{w}.$$

The distortionary costs arise because couples on the interior boundary of $\{\mathbf{w} : \mathbf{w} > (t,t)\}$ reduce their earnings due to higher tax rates. To calculate these costs, observe that, since we started with a disjointed tax schedule, the reduction in earnings of every distorted individual equals to $\Delta d\tau$, where $\Delta = \gamma t y'(t)/(1 - \frac{\partial}{\partial y}\widetilde{T}^m(y(t)))$.[17] There are $2g(t)\Pr(w_{-i} \geq t | w_i = t)dt$ of distorted individuals. Therefore, the total cost of distortions is

$$\mathfrak{C} = y'(t)\, dt d\tau \gamma t 2g(t) \Pr(w_{-i} \geq t | w_i = t) \frac{\frac{\partial}{\partial y}\widetilde{T}^m(y(t))}{1 - \frac{\partial}{\partial y}\widetilde{T}^m(y(t))}.$$

The net welfare effect of introducing jointness is the difference between benefits and costs:

$$\mathfrak{B} - \mathfrak{C} = y'(t) dt d\tau \Pr(\mathbf{w} \geq (t,t)) \times \tag{31}$$

$$\times \left[(1 - \mathbb{E}[\alpha^m | \mathbf{w} \geq (t,t)]) - \underbrace{\frac{2\Pr(w_{-i} \geq t)\Pr(w_j \geq t | w_i = t)}{\Pr(w_{-i} \geq t, w_i \geq t)}}_{=1/\overline{B}(t)} \underbrace{\frac{t g(t)}{1 - G(t)}}_{=\theta(t)} \gamma \frac{\frac{\partial}{\partial y}\widetilde{T}^m(y(t))}{1 - \frac{\partial}{\partial y}\widetilde{T}^m(y(t))} \right].$$

Formula (31) applies to an arbitrary separable tax schedule. To see the connection to equation (30), evaluate (31) at the *optimal* separable tax, $\widetilde{T}^{m,*}$. While we study such taxes formally in Section 5.9, it is easy to use perturbational arguments to show that $\widetilde{T}^{m,*}$ must satisfy

$$\frac{\frac{\partial}{\partial y}\widetilde{T}^{m,*}(y(t))}{1 - \frac{\partial}{\partial y}\widetilde{T}^{m,*}(y(t))} = \frac{1 - \mathbb{E}[\alpha^m | w_i \geq t]}{\gamma\theta(t)}. \tag{32}$$

Substitute equation (32) into (31) and re-arrange to obtain that $\mathfrak{B} - \mathfrak{C}$ is proportional to $\overline{A} \times \overline{B} - 1$. Thus, jointness is positive if and only if $\overline{A} \times \overline{B} \geq 1$, as shown in equation (30).

We now turn to understanding properties of $\overline{A}$ and $\overline{B}$. First, observe that if productivities are independent then $\overline{A}(t) = 2$ and $\overline{B}(t) = 1/2$ for all t, which implies $J(t) = 0$. This is the average version of the result shown in Lemma 3 that optimal jointness is zero when productivities are independent. For positively dependent distributions, $\overline{A} \leq 2$ and $\overline{B} \geq 1/2$, so that dependence both decreases benefits of better targeting taxes and costs of distortions

---

[17]To show this result, differentiate the optimality condition $1 - \frac{\partial}{\partial y}\widetilde{T}^m(y(w)) = (y(w))^{1/\gamma - 1} w^{-1/\gamma}$.

from jointness. Therefore, whether optimal average jointness is positive or negative depends on which of these effects dominates.

To get some intuition for the more general result consider, first the case when F is a bi-variant log-normal distribution with correlation of log productivities $\rho > 0$. It is easy to see that for log-normal distribution both $\mathbb{E}\left[\alpha^m | \mathbf{w} \geq (t, t)\right]$ and $\mathbb{E}\left[\alpha^m | w_i \geq t\right]$ converge to $\alpha(\infty) := \lim_{t \to \infty} \alpha(t)$. This implies that benefits from better targeting taxes disappear, $\lim_{t \to \infty} \overline{A}(t) = 1$. At the same time, direct calculations show that $\lim_{t \to \infty} \overline{B}(t) = \frac{1+\rho}{2} < 1$. Therefore, $\overline{A}(t) \times \overline{B}(t)$ is less than one for large t and optimal average jointness is negative for high-earning couples.

This result carries out to many other joint distributions used in applied work. Consider limits $\overline{A}(\infty) := \lim_{t \to \infty} \overline{A}(t)$ and $\overline{B}(\infty) := \lim_{t \to \infty} \overline{B}(t)$. Let $F(\cdot | \infty)$ denote the limit of $F(\cdot | t)$ as $t \to \infty$ in the weak convergence sense. Taking the limit of $\overline{A}(t)$ we obtain

$$\overline{A}(\infty) = \frac{1 - \frac{1}{2}\alpha(\infty) - \frac{1}{2}\alpha(\infty)}{1 - \frac{1}{2}\alpha(\infty) - \frac{1}{2}\int \alpha(w)F(dw|\infty)}. \tag{33}$$

It is immediate to see that if $F(\cdot | \infty)$ strictly dominates (in first order stochastic dominance sense) the unconditional distribution of productivities $G(\cdot)$ then $\overline{A}(\infty) < 2$; moreover, if $F(\cdot | \infty)$ is degenerate, in the sense that $F(w | \infty) = 0$ for all w, then $\overline{A}(\infty) = 1$.

To find the limit of $B(t)$, use the L'Hopital's rule to show that

$$\overline{B}(\infty) = \overline{\kappa} := \lim_{t \to \infty} \frac{\ln \Pr(w_1 \geq t)}{\ln \Pr(w_1, w_2 \geq t)}.$$

Statistics $\overline{\kappa}$ is extensively studied in the theory of extreme values. It measures the speed of convergence to asymptotic right-tail independence and its values are known for many commonly used joint distributions.[18] A common way to generate joint distributions in applied work is by using parametric families of copulas. Gaussian copula (which generalizes correlation properties of a bi-variant normal distribution to distributions with arbitrary marginals, and includes bi-variant log-normal distributions as a special case) has $\overline{\kappa} = (1 + \rho)/2$. Many other commonly

---

[18]Recall (see, e.g., Nelsen (2006)) that a symmetric distribution is *right-tail independent* if $\lim_{t \to \infty} \Pr(w_{-i} \geq t | w_i \geq t) = 0$. Intuitively, tail-independent distributions rule out asymptotic emergence of mass points on the 45 degree line in the (w_1, w_2) space. Ledford and Tawn (1996, 1998) showed that under weak conditions, the right-tail behavior of a symmetric distribution could be written as

$$\Pr(w_1 \geq t, w_2 \geq t) \sim L(t) \cdot (\Pr(w_i \geq t))^{1/\overline{\kappa}} \quad \text{as } t \to \infty,$$

where L is some function such that $L\left(G^{-1}\left(e^{-1/t}\right)\right)$ is "slowly varying", which is a certain generalization of functions that converge as $t \to \infty$. As can be seen from the above expression, larger values of $\overline{\kappa} \in [1/2, 1)$ mean slower convergence to independence. Coles et al. (1999); Heffernan (2000); Hua and Joe (2014) discuss the theoretical properties and empirical estimation of these statistics. Hua and Joe (2011, 2014) report values of $\overline{\kappa}$ for a large number of copulas.

used families of copulas, such as FGM, Pareto, Frank, Clayton, Ali-Mikhail-Haq, have $\overline{\kappa} = 1/2$. We can summarize this discussion in the following lemma.

Lemma 7. *Optimal average jointness is negative at the top, in the sense that $\lim_{t \to \infty} J(t) < 0$, if either (i) $\overline{\kappa} = \frac{1}{2}$ and $F(\cdot|\infty) >_{FOSD} G(\cdot)$, or (ii) $\overline{\kappa} < 1$ and $F(\cdot|\infty)$ is degenerate.*

All copulas mention in the previous paragraph satisfy conditions of this lemma as long as their correlation parameters are strictly positive (see appendix for details).

To analyze jointness at the bottom, it is more insight to use not equation (30) but rather

$$J(t) \leq 0 \qquad \Longleftrightarrow \qquad \underline{A}(t) \times \underline{B}(t) \geq 1 \tag{34}$$

where

$$\underline{A}(t) := \frac{\mathbb{E}\left[\alpha^m | \mathbf{w} \leq (t,t)\right] - 1}{\mathbb{E}\left[\alpha^m | w_i \leq t\right] - 1}, \quad \underline{B}(t) := \frac{\Pr(w_j \leq t, w_i \leq t)}{2\Pr(w_{-i} \leq t | w_i = t)}.$$

This equation can be obtained by deriving expression for $\mathbb{E}\left[\lambda_i^{m,*} | w_i = t, w_{-i} \leq t\right]$ by analogy with (29). Term $\underline{A}$ in this expression captures the benefits of better targeting transfers to the poorest couples using jointness; term $\underline{B}$ captures additional distortions from this targeting. Equation (34) appears similar to (30), but it has one important difference. While positive jointness helps redistribution by targeting taxes to the richest couples, negative jointness helps redistribution by targeting transfers to the poorest couples. Given this observation, it follows that conclusions of Lemma 7 flip sign once it is extended to the poorest couples. In particular, let $\underline{\kappa} := \lim_{t \to 0} \frac{\ln \Pr(w_1 \leq t)}{\ln \Pr(w_1, w_2 \leq t)}$ be the parameter capturing the speed of convergence to the left-tail independence. We have

Lemma 8. *Optimal average jointness is positive at the bottom, in the sense that $\lim_{t \to 0} J(t) > 0$, if either (i) $\underline{\kappa} = \frac{1}{2}$ and $F(\cdot|0) <_{FOSD} G(\cdot)$, or (ii) $\underline{\kappa} < 1$ and $F(0|0) = 1$.*

5 Extensions

In Section 2 we presented the simplest model of taxation of single and married households, and developed techniques to characterize properties of optimal taxes in this model. While a natural starting point, that model is stylized. It abstracts from home production, economies of scale in marriage, extensive margin in labor supply decisions, bargaining over allocation of resources within couples, selection into marriage, gender differences. Moreover, a policy maker may have different objective from the benchmark case. For example, the planner may have inherent preference for single or married households, or may want to use simpler taxes that those implied by the unrestricted optimal tax schedule.

The goal of this section is to show how the approach that we developed in Section 3.3 can be adapted to study such extensions. Since the number of possible extensions is large, we cannot discuss them in-depth within confines of one paper. For this reason we organize this section as follows. Each subsection contains an extension of our benchmark model in one of the directions mentions above. To save space, we skip derivations and focus in each subsection on only one or two new insights that emerge from that extension. Appendix contains derivations and additional details.

5.1 Social weights on single and married

In this section, we allow the planner to have richer preferences than those we considered in the benchmark economy. The purpose of this extension is two-fold. First, this is an important question in its own right as a policymaker may evaluate social welfare differently from the way we considered in Section 2. Second, several other extensions resemble, in a reduced form, the social planner with such weights. Throughout this section, we assume that $\alpha(w)$ is a Pareto weight on a single household with productivity w, with normalization $\int \alpha(w)G(dw) = 1$.

Consider first the case when the planner uses Pareto weights $k\alpha^m(\mathbf{w})$ for couple $\mathbf{w}$, where $k > 0$ is a scalar and α^m as given in (1). The planner has preference for married individuals when $k > 1$ and single individuals when $k < 1$. Proceeding as in Section 3.3, it is easy to derive formulas for the optimal average distortions in this economy:

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

$$\mathbb{E}\left[\lambda_i^{m,*}|w_i = t\right] = \frac{1 - \mathbb{E}\left[\alpha^m|w_i \geq t\right]}{\gamma\theta(t)}\frac{k}{1 - \mu^* + \mu^*k}. \tag{36}$$

These formulas are almost the same as (22) and (25) except for terms $\frac{1}{1-\mu^*+\mu^*k}$ and $\frac{k}{1-\mu^*+\mu^*k}$. It is insightful to understand the economic intuition behind these terms and their implications.

For concreteness, suppose that $k > 1$. It is immediate to verify that in this case married persons face higher distortions than in the benchmark economy while single persons face lower distortions. This result seems surprising. Why is it optimal for the planner to increase distortions for those households that she values more? The answer to this is that by increasing redistribution (and, hence, distortions) for more preferred households and decreasing it for the less preferred ones the social planner can redistribute resources with smaller behavioral responses. To understand why this is the case, suppose that the planer merely increases taxes lump-sum by $1 on single households and returns proceeds lump-sum to married households. This policy reform would incentivize some hitherto single individuals to get married. This is

costly both because these individuals have non-pecuniary preference to stay single, and because it reduces the amount of transfers available for married couples. The planner can reduce this behavioral response by simultaneously increasing redistribution within married households and reducing redistribution within single households, which makes it more attractive for the marginal individual to stay single.

Consider now implications of non-separable Pareto weights for couples. Suppose that the planner assigns a weight $\alpha^m(\mathbf{w}) \neq \frac{\alpha(w_1)+\alpha(w_2)}{2}$ to a married couple $\mathbf{w}$. To separate the economics of non-separable weights on couples from desire to redistribute between married and single, we focus on the case $\int \alpha^m(\boldsymbol{w})F(d\boldsymbol{w}) = 1$.

It is easy to see that none of the derivations in formula (24) required separability of α^m and, therefore, this formula as well as its implications, such as equations (25), (27) or (29), remain unchanged. We now generalize the comparative statics results that we presented in Lemmas 4, 6, 7, and 8. To state those results, we want to maintain the assumption that average Pareto weight on married persons is the same as for single even as we change, for example, the degree of dependence. We say that $\alpha^{m,b} \sim \alpha^{m,a}$ if $\alpha^{m,a}$, $\alpha^{m,b}$ are the same up to the normalization constant, that is if the ratio $\alpha^{m,b}(\mathbf{w})/\alpha^{m,a}(\mathbf{w})$ is independent of $\mathbf{w}$.

Corollary 1. Consider economies, a and b s.t. $\int \alpha^{m,b}(\boldsymbol{w})F^b(d\boldsymbol{w}) = \int \alpha^{m,a}(\boldsymbol{w})F^a(d\boldsymbol{w}) = 1$.

(a). Suppose $F^a = F^b$. Conclusions of Lemma 5 hold provided that $\alpha^{m,b}$ is more redistributory than $\alpha^{m,a}$ in the multivariate likelihood ratio order sense,[19] and conclusions of Lemma 6 hold for $(F^a, \alpha^{m,a})$ provided that $\alpha^{m,a}(0, w) > 1$ for all w.

(b) Suppose $F^a \leq_{PQD} F^b$, $\alpha^{m,b} \sim \alpha^{m,a}$, and $\alpha^{m,a}$ is supermodular. Then Lemma 7 holds for $(F^a, \alpha^{m,a})$, and Lemma 4 holds at the top, in the sense that $\lim_{t \to \infty} \frac{\mathbb{E}^b\left[\lambda_i^{m,b,*}|w_i=t\right]}{\mathbb{E}^a\left[\lambda_i^{m,a,*}|w_i=t\right]} \geq 1$.

(c). Suppose $F^a \leq_{PQD} F^b$, $\alpha^{m,b} \sim \alpha^{m,a}$, and $\alpha^{m,a}$ is submodular. Then Lemma 8 holds for $(F^a, \alpha^{m,a})$, and Lemma 4 holds at the bottom, in the sense that $\lim_{t \to 0} \frac{\mathbb{E}^b\left[\lambda_i^{m,b,*}|w_i=t\right]}{\mathbb{E}^a\left[\lambda_i^{m,a,*}|w_i=t\right]} \geq 1$.

Part (a) of this proposition provides a way to compare redistributiveness of Pareto weights in general, bi-variate settings. Parts (b) and (c) extend results about jointness and average optimal distortions for non-separable weights α^m. The cross-partial derivatives of α^m play an important role. When these derivatives are positive (so that α^m are supermodular), the planer has a stronger desire to extract resources from couples in orthants $\{\mathbf{w} : \mathbf{w} > \mathbf{t}\}$ with large $\mathbf{t}$,

---

[19]$\alpha^{m,b}$ dominates $\alpha^{m,a}$ in the multivariate likelihood ratio order if $\alpha^{m,a}(\boldsymbol{w}')\alpha^{m,b}(\boldsymbol{w}'') \geq \alpha^{m,b}(\boldsymbol{w}' \vee \boldsymbol{w}'')\alpha^{m,a}(\boldsymbol{w}' \wedge \boldsymbol{w}'')$ for all $\boldsymbol{w}', \boldsymbol{w}''$, and it is a generalization of the uni-variant likelihood ratio order that we used in Lemma 5. Under auxiliary regularity assumptions on either $\alpha^{m,a}$ or $\alpha^{m,b}$ it is equivalent to the statement that $\alpha^{m,b}(\mathbf{w})/\alpha^{m,a}(\mathbf{w})$ is decreasing in $\mathbf{w}$. See Chapter 6.E of Shaked and Shanthikumar (2007) for discussion and proofs.

relative to the benchmark economy case. This amplifies the mechanisms behind results in Lemma 4 and 7 for high-earners, but has ambiguous effect for low-earners. When the cross-partial derivatives of α^m are negative (so that α^m are submodular), the planner has a stronger desire to transfer resources to couples in orthants $\{\mathbf{w} : \mathbf{w} < \mathbf{t}\}$ with small $\mathbf{t}$. This amplifies the mechanisms emphasized in the previous section for lower earners.

Parts (b) and (c) of this corollary are related to findings of Kleven et al. (2007). Those authors consider optimal jointness in the economy in which productivities are independent, all households are married, and social welfare is given by $\int W(v^m)\,dF$, where W is the social welfare function. They prove that in that economy the optimal jointess is negative (positive) if the third derivative of W is positive (negative). The same result can shown in our settings in terms of modularity of α^m: the optimal jointness is negative (positive) if α^m is supermodular (submodular) and matching is random (see appendix for the proof). Corollary 1 implies that even small degree of dependence can break this result. Suppose for concreteness that α^m is supermodular and F is log-normal with correlation ρ. The optimal jointness is negative for all $\mathbf{w}$ if productivities are independent, $\rho = 0$. However, conditions of Lemma 8 are satisfied for any positive correlation $\rho > 0$, so that optimal jointness at the bottom is strictly positive for any degree dependence.

5.2 Bargaining and allocation of resources within couples

Married couples in our benchmark economy act as a single unit whereby they maximize their joint surplus and split it equally between spouses. A sizable literature in labor economics (see, e.g., a handbook chapter by Almas et al. (2023) for an overview) has shown that resource division within families often departs from this idealized view. For example, a spouse with higher income often have greater control over surplus division. In this section, we study how optimal taxation is affected.

We consider a simple model of bargaining within households in the spirit of Manser and Brown (1980) and McElroy and Horney (1981). In particular, we modify the description of Stages 2 and 3 of our benchmark model and assume that married couples bargain over surplus division after their productivities $\mathbf{w}$ are realized but before they supply labor on the market. When bargaining, each spouse uses the threat of a divorce, at a personal cost $\varrho > 0$, in which case both spouses become single and cannot remarry. Thus, the outside option of spouse i with productivity w_i is $v^s(w_i) - \varrho$. Cost ϱ is assumed to be sufficiently high so that it is not socially efficient for couples to get divorced. Using symmetric Nash bargaining as a solution

concept, the utility of a married person can be shown to be equal to

$$U^m\left(w_i|w_{-i}\right) = \frac{1}{2}v^m(\mathbf{w}) + \frac{v^s(w_i) - v^s(w_{-i})}{2}.$$

Since $v^s(\cdot)$ is increasing, this equation shows that a more productive spouse gets a bigger share of family surplus, consistent with the motivation given in the beginning of this section.

Our approach from Section 3.3 can be adapted to this economy with minimal changes. It is easy to show that optimal average distortions for married persons are given by (25), while distortions for single persons satisfy

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

The first term on the right hand side is the same as (22), the second term is new and it captures the effect of bargaining. This term is positive when F is positively dependent so that bargaining over marital surplus increases distortions for single persons and leaves distortions for married persons unchanged.

To understand the intuition for this result, observe that bargaining reduces the amount of redistribution that occurs within couples, relative to the benchmark economy. While the planner can compensate for this reduction by providing more redistribution for married households via T^m, a more efficient response is to increase redistribution for single households via T^s. This decreases inequality both among single households, directly through tax system, and among married households, indirectly by making spouses' outside options and, therefore, surplus division more equal.

5.3 Optimality of taxation of family earnings

In many countries, notably the U.S., married couples pay taxes based on their total family earnings, i.e., the tax function T^m takes form $T^m\left(y_1, y_2\right) = \widetilde{T}^m\left(y_1 + y_2\right)$. We call this *family-earnings based taxation*. In this section, we explore conditions under which such form of taxation is optimal.

Let $Y = y_1 + y_2$ be family earnings. To motivate our approach, observe that if couple $\mathbf{w}$ faces a family earnings-based tax then one can write their optimality conditions as

$$\frac{Y\left(\mathbf{w}\right)^{1-\gamma}}{\left(1 - \frac{\partial}{\partial Y}\widetilde{T}^m\left(Y\left(\mathbf{w}\right)\right)\right)^\gamma} = \left(w_1^{1/(1-\gamma)} + w_2^{1/(1-\gamma)}\right)^{1-\gamma} := R\left(\mathbf{w}\right).$$

This equation shows that both spouses in all couples $\mathbf{w}$ for whom $R\left(\mathbf{w}\right) = r$ face the same marginal tax. Thus, the question whether the unrestricted tax system is family-earnings based can be reformulated as whether optimal distortions only depend on $R(\mathbf{w})$.

It will be convenient to change coordinates to study this question. Recall function $I(\mathbf{w}) = w_2/w_1$ that we defined in Section 3.3. R and I allow us to change coordinates of the space describing joint productivities of couples from $\mathbf{w}$ to $(r, \iota) = (R(\mathbf{w}), I(\mathbf{w}))$. These coordinates have a natural economic interpretation. Coordinate r is a measure of total family earnings, while ι captures how family earnings are allocated between the two spouses. Let $\widetilde{F}(r, \iota)$ be the distribution of productivities in these new coordinates, $\widetilde{F}(\cdot|\iota)$ the distribution of r conditional on a given value of ι, and $\widetilde{f}(\cdot|\iota)$ its density. Finally, let $G_r(r)$ be the marginal distribution of r and $\theta_r(r)$ be its tail statistics.

It is clear from Lemma 9 that family-earnings based taxation is unlikely to be optimal if weights α^m on couples are additively separable in $\mathbf{w}$. To make analysis interesting, we consider arbitrary weights α^m that satisfy $\int \alpha^m(\boldsymbol{w})F(d\boldsymbol{w}) = 1$. We say that α^m is *measurable only w.r.t.* r if it can be written in the form $\alpha^m(R(\mathbf{w}))$. Using our general formula (24), we can prove the following result.

Lemma 9. *Let* $\widetilde{\lambda}(r, \iota) := \frac{1 - \mathbb{E}[\alpha^m | R \geq r, I = \iota]}{\gamma r \widetilde{f}(r|\iota)}$.

(a). The optimal tax is family-earnings based if and only if $\widetilde{\lambda}(r, \iota)$ *is independent of* ι.

(b). Suppose that α *is measurable only w.r.t.* r, *and* r *and* ι *are independent. Then, the optimal tax system is family earnings-based and it satisfies*

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

(c). Consider any two economies, a *and* b, *that are identical except distribution* $\widetilde{F}$, *and weights* α^m *and suppose that social weights for couples are only measurable w.r.t.* r. *If* $\widetilde{F}^a \leq_{PQD} \widetilde{F}^b$, $\alpha^{m,b} \sim \alpha^{m,a}$, *then* $\mathbb{E}^b\left[\lambda_2^{m,b,*} - \lambda_1^{m,b,*}|I = \iota\right] \geq \mathbb{E}^a\left[\lambda_2^{m,a,*} - \lambda_1^{m,a,*}|I = \iota\right]$ *for all* ι.

Part (a) of this lemma provides necessary and sufficient conditions, summarized by function $\widetilde{\lambda}$, under which family-earnings based taxation is optimal. As can be expected, these conditions are knife-edge and not very informative in general. Part (b) provides one specific case that has a natural economic interpretation. Family-earnings based taxation is optimal if the planner values equally all couples with the same r and, in addition, if distribution $\widetilde{F}$ is such that r is independent of ι. The interpretation of this result is that family-earnings based taxation is optimal if the planner has inherent preference for such taxation and there is no correlation between family earnings and the share earned by the primary earner. Note that the optimal distortions in this case closely resemble equation (22), except that distortions are captured by statistics θ_r rather than θ. Pure family-earnings based taxation is suboptimal for such a

planner when r and ι are correlated because the planner can exploit information contained in ι to provide better redistribution across couples with different r. Part (c) of Lemma 9 sheds light on how this can be done by providing comparative statistics with respect to the degree of dependence in (r, ι).

Note that parts (b) and (c) of Lemma 9 have close parallels to our discussion of conditions under which separable taxation is optimal in the benchmark economy. In that economy, each individual with productivity w is weighted with an individual-specific weight $\alpha(w)$ that does not depend on whether or whom this individual marries. Thus, the planner has inherent preference for individual-earnings based taxation, so that taxes of one person do not depend on earnings of their spouse. Despite such preference, individual-earnings based taxation was generally suboptimal unless productivities $\mathbf{w}$ are independent.

5.4 Public goods and economies of scale in marriage

Married persons share costs of many goods that simultaneously provide benefits to both spouses, from housing and child-rearing to Netflix subscriptions. In this section we explore implications of these economies of scale for optimal taxation.

We model them as a simple public good. In particular, we assume that each person's utility is $\phi(c^{pr}, c^{pub}) - \frac{1}{\gamma} l^{\gamma}$, where ϕ is a constant returns to scale function, and c^{pr} and c^{pub} are private and within-household public good. Without loss of generality, we set prices of both goods to be equal to one, so that single and married households solve, respectively,

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

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

Since the same c^{pub} enters utility of both spouses, marriage leads to efficiency gains in this settings. The rest of the model is as in Section 2.

Consumption expenditures can be conveniently aggregated. Households' maximization problems can be written as

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

$$v^m(\mathbf{w}) := \max_{\left(C, (y_i)_{i=1}^2\right)} b^m C - \frac{1}{\gamma}\sum_{i=1}^2 \left(\frac{y_i}{w_i}\right)^{\gamma} \text{ s.t. } C \leq \sum_{i=1}^2 y_i - T^m(y_1, y_2), \mathbf{y} \geq \mathbf{0.}$$

where

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

It is easy to see that $b^m > b^s$ when ϕ is strictly increasing in c^{pub} so that the marginal utility of total consumption expenditures C is greater for married persons due to economies of scale. The analysis of this economy is very similar to the economy in considered in Section 5.1, in which the planner had exogenously higher weight on married households. In particular, it is easy to show that optimal distortions satisfy

$$\lambda^{s,*}(t) = \frac{1 - \mathbb{E}\left[\alpha | w \geq t\right]}{\gamma \theta(t)} \left(1 - \mu^* + \mu^* \frac{b^s}{b^m}\right), \tag{37}$$

$$\mathbb{E}\left[\lambda_i^{m,*} \middle| w_i = t\right] = \frac{1 - \mathbb{E}\left[\alpha^m | w_i \geq t\right]}{\gamma \theta(t)} \left(1 - \mu^* + \mu^* \frac{b^s}{b^m}\right). \tag{38}$$

The intuition for this result is the same that we discussed in Section 5.1.

5.5 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. This requires households to decide how to allocate effort between home and market work and, in two-person households, how to divide these tasks between spouses. In this section, we incorporate home production and division of labor in our economy and consider their implications for optimal taxation.

Let d be consumption of the home good and x be the effort 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 home good is $D^s(x) = x$ and their maximization problem is

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

For married households, home production technology is $D^m(\mathbf{x}) = \left(x_1^{1/q} + x_2^{1/q}\right)^q$ and we assume that home good as a public good. Married households solve

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

where $\mathbf{y}, \mathbf{x} \geq \mathbf{0}$. Parameter $p \in (1, 1/\gamma)$ captures the elasticity of substitution between hours at home and at work, with $p \to 1$ represents the limit of perfect substitution. Parameter $\sigma \in [0,1)$ captures curvature in the utility of consumption of the home produced good. Parameter $q \in [1, 1/(1-\sigma)]$ allows home produced goods by the two married spouses to be imperfect substitutes. Restrictions on parameters ensures that all choices are interior and we can abstract from corner solutions.

31

As in the model in Section 5.4, aggregation simplifies analysis. In particular, define functions $N^s(l)$ and $N^m(\mathbf{l})$ by

$$N^s(l) := \min_{x \geq 0} -\frac{x^{1-\sigma}}{1-\sigma} + \gamma (l^p + x^p)^{1/(\gamma p)}, \quad N^m(\mathbf{l}) := \min_{\mathbf{x} \geq 0} -2\frac{\left(\sum_{i=1}^2 x_i^{1/q}\right)^{q(1-\sigma)}}{1-\sigma} + \gamma \sum_{i=1}^2 (l_i^p + x_i^p)^{1/(\gamma p)}.$$

Using these definitions, single and married household problems can be written as

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

$$v^m(\mathbf{w}) := \max_{(C,\mathbf{y})} C - N^m\left(\frac{y_1}{w_1}, \frac{y_2}{w_2}\right) \text{ s.t. } C \leq \sum_{i=1}^2 y_i - T^m(y_1, y_2), \mathbf{y} \geq \mathbf{0}.$$

This model is now identical to the benchmark economy except that the disutility of labor is given by general functions N^s and N^m. The key economic difference from the benchmark environment is that elasticities of labor supply at work are no longer constant.

We now describe properties of the solution of this model. Recall that in our benchmark economy, the elasticity of labor supply e is constant, and the relationship between the elasticity parameter γ and e is given by $\gamma = (1 + 1/e)^{-1}$. Consider now preferences for single households. The elasticity of their labor supply at labor level l is $e(l) = \frac{\partial N^s(l)/\partial l}{l \cdot \partial^2 N^s(l)/\partial l^2}$, and define $\Gamma^s(l)$ by

$$\Gamma^s(l) := (1 + 1/e(l))^{-1},$$

which generalizes γ in the case when elasticity is not constant. Similarly, for couples define a 2×2 matrix $\Gamma^m(\mathbf{l})$ as

$$\Gamma^m(\mathbf{l}) := \begin{bmatrix} 1 + \frac{l_1 \cdot \partial^2 N^m(\mathbf{l})/\partial l_1^2}{\partial N^m(\mathbf{l})/\partial l_1} & \frac{l_2 \cdot \partial^2 N^m(\mathbf{l})/\partial l_1 \partial l_2}{\partial N^m(\mathbf{l})/\partial l_1} \\ \frac{l_1 \cdot \partial^2 N^m(\mathbf{l})/\partial l_1 \partial l_2}{\partial N^m(\mathbf{l})/\partial l_2} & 1 + \frac{l_2 \cdot \partial^2 N^m(\mathbf{l})/\partial l_2^2}{\partial N^m(\mathbf{l})/\partial l_2} \end{bmatrix}^{-1}.$$

With a bit of algebra, it can be shown that $\Gamma^m = (1 + E^{-1})^{-1}$, where E is the 2×2 matrix of all (cross-)elasticities of $\mathbf{l}$ with respect to after-tax wage rates for both spouses. We use Γ^m_{ij} to denote the ij^{th} element of matrix Γ^m.

Using these definitions, we can study optimal distortions in this economy. We start with single households. Direct adaptation of our approach yields the formula

$$\lambda^{s,*}(t) = \frac{1 - \mathbb{E}[\alpha|w \geq t]}{\Gamma^s(y^*(t)/t)\,\theta(t)}. \tag{39}$$

This is a version of the well-known Diamond's ABC formula, who also considered the case of non-constant elasticity of labor supply. The intuition and interpretation for it exactly the same that we gave after equation (22).

Equation (39) holds for any general disutility of labor and does not depend on home production per se. The home production structure contains additional insights about how single persons allocate effort at work and at home. Consider a person with productivity t who faces marginal taxes that are bounded away from 1 from above. As t increases, the opportunity cost of working at home increases too, and the person starts to allocate more hours at work. We have $\lim_{l \to \infty} \Gamma^s(l) = \gamma$ and, therefore, we obtain

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

This result shows that optimal distortions for high productive individuals are determined by the same parameters with and without home production.

We now turn to optimal distortions for married persons. The Coarea formula (24) adapted to our settings implies

$$\mathbb{E}\left[\Gamma_{ii}^m \lambda_i^{m,*} + \Gamma_{ij}^m \lambda_j^{m,*}|w_i = t\right] = \frac{1 - \mathbb{E}\left[\alpha^m|w_i \geq t\right]}{\theta(t)}. \tag{41}$$

This equation generalizes equation (25) for the case when labor supply elasticities are interdependent between spouses. The planner now needs to take into account that a higher marginal tax on person i not only affects labor supply of that person (captured by Γ_{ii}^m in this formula) but also the labor supply of their spouse (captured by Γ_{ij}^m).

Equation (41), just like equation (39), holds for arbitrary joint utility of labor supply $N^m(\mathbf{l})$. Home production has additional implications. If a spouse i is more productive, they supply more effort at work and less at home. This diminishes the importance of home production and wages of the other spouse. In particular, we have $\lim_{l_i \to \infty} \Gamma_{ii}^m(l_i, l_j) = \gamma$ and $\lim_{l_i \to \infty} \Gamma_{ij}^m(l_i, l_j) = 0$ for all l_j, which implies that

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

This formula shows that optimal distortions for both single and married high-earners remain the same as in the benchmark economy and many of the comparative statics results derived in Section 3.3. Furthermore, the analysis of jointness in the right tail proceeds with no changes, extending the results of Lemma 7 to this economy.

5.6 Extensive margin

In our benchmark economy, all labor supply adjustments are done along the intensive margin. In this section, we briefly discuss implications of also adding the extensive margin to our analysis. We consider a simple extension of our benchmark environment by assuming that any

person who supplies $l_i > 0$ incurs an additional disutility cost $\varrho > 0$. We focus on the case of random matching, since it is easy to fully characterize the optimum in this case and highlight the role that extensive margin plays in the analysis.

Lemma 10. *In the model with extensive margin and random matching, optimal distortions for single persons are given by (22), and for married are as given in Lemma 3. There are productivity cut-offs $\underline{w}^s, \underline{w}^m$ for single and married, with $\underline{w}^s > \underline{w}^m$, so that a person works positive hours if and only if their productivity is above those cut-offs.*

The optimal tax formulas remain the same as in the economy with only intensive margin. The only difference is that now individuals with low productivity choose not to supply any labor and exit employment. Single individuals face higher distortions under the optimal system than married persons and, therefore, are more likely to choose not to work.

5.7   Selection into marriage

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

The simplest way to introduce correlation of marriage rates and productivities in our economy is to switch the order of Stages 1 and 2. That is, assume that individuals first observes singles q about their productivity and then decide whether to enter the marriage market. The rest of the set up is as in Section 2.

To highlight the key differences that this set up introduces we assume that there are two signals, q_h and q_l that occur with equal probability. Let H_h and H_l be distribution of productivities of individuals who receive signals q_h and q_l, and μ_h and μ_l be their marriage rates. The average marriage rate μ satisfies $\mu = \frac{1}{2}(\mu_h + \mu_l)$.

If signals q are correlated with w, marriage rates will differ across w. Let G_s and G_m be distributions of productivities among single and married persons, and $\theta_s(t)$ and $\theta_m(t)$ denote their tail statistics defined as in (23). Following the same steps as in baseline economy, one can show that the optimal distortions for single and married persons satisfy

$$\lambda^{s,*}(t) = \frac{1 - \mathbb{E}^s[\alpha | w_i \geq t]}{\gamma \theta_s(t)} + \frac{1}{1 - \mu^*} \frac{\delta_h(1 - H_h(t)) + \delta_l(1 - H_l(t))}{\gamma \theta_s(t)},$$

$$\mathbb{E}[\lambda_i^{m,*} | w_i = t] = \frac{1 - \mathbb{E}^m[\alpha^m | w_i \geq t]}{\gamma \theta_m(t)} - \frac{1}{\mu^*} \frac{\delta_h(1 - H_h(t)) + \delta_l(1 - H_l(t))}{\gamma \theta_m(t)},$$

where δ_h and δ_l are the Lagrange multipliers on the two analogues of equation (6), for q_h and q_l. These multipliers must satisfy $\mu^* \delta_h + (1 - \mu^*) \delta_l = 0$. Examining these equations reveals

34

that the optimal tax formula consists of two terms. The first terms on the right hand sides of of both equations are the same as (22) and (25), except now they take into account that tail statistics θ summarizing marginal distribution of productivities differs between single and married individuals. The second terms on the right hand side of these equations capture the additional effect from responses of marriages of persons with different productivities to taxes. The planner recognizes that taxes affect these marriage decisions and takes these behavioral responses into account when choosing tax rates. These terms take opposite signs for single and married persons and have economic intuition similar to the implications of exogenously different Pareto weights in Section 5.1 or public goods in Section 5.4.

5.8 Gender differences

Family economics literature has documented systematic differences in earnings among males and females. In this section, we incorporate this heterogeneity in our model of optimal taxation.

We assume that there are two fixed genders, M and F, each of measure one half, and that married couples must consist of one spouse of each gender. Persons M and F draw their productivities from distributions G_M and G_F, with tail statistics θ_M and θ_F. Spouses of each gender are the same in every other respect[20] and description of the matching process is as in Section 2 with one modification. Since M and F 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 return individuals with highest values of ε of the "surplus" gender on the marriage market back to the singlehood.

We assume that social welfare function is

$$\mathcal{W} := \frac{1}{2} \int \alpha_M(w) \mathbb{E}\left[U_M | w\right] G_M(dw) + \frac{1}{2} \int \alpha_F(w) \mathbb{E}\left[U_F | w\right] G_F(dw),$$

where weights α_M and α_F are not necessarily the same, and set $\int \alpha_j(w) G_j(dw) = 1$ for $j = M, F$. Let $\alpha^m(w_M, w_F) := \frac{1}{2}\alpha_M(w_M) + \frac{1}{2}\alpha_F(w_F)$. For now we consider optimal tax system that can use gender as a marker to tax liability. This implies that single persons face taxes that differ by gender, and taxes of married couples depends not only on vector of earnings but also on the genders of earners. In the next section, we reconsider this question by imposing gender-neutrality on the tax system.

Our analysis in Section 3.3 extends to this economy with minimal changes. In particular,

---

[20]Empirical estimates of labor supply elasticities often differ between genders. Those differences emerge endogenously if we incorporate home production along the lines of Section 5.5. It is also trivial to extend our analysis to the case when genders differ in Φ and γ.

that the optimal distortions satisfy

$$\lambda_j^{s,*}(t) = \frac{1 - \mathbb{E}[\alpha_j | w_j \geq t]}{\gamma \theta_j(t)}, \quad \mathbb{E}[\lambda_j^{m,*} | w_j = t] = \frac{1 - \mathbb{E}[\alpha^m | w_j \geq t]}{\gamma \theta_j(t)} \text{ for } j = M, F.$$

These formulas is a direct generalization of our results in Section 3.3 and Sections 5.1, with all the intuition as described in those sections.

5.9 Optimal restricted taxation

One observation about the optimal unrestricted taxes in Section 5.8 is that they depend on the gender of the tax payer. This result is not surprising. The usual tagging logic (see Akerlof (1978)) indicates that it is generally optimal to use any observable tag, such as gender, if it is correlated with unobservable productivities to reduce costs of taxation. It is natural to ask how optimal taxes should look like when a policy maker does not want to use such taxes. This motivates our analysis of optimal restricted taxes.

In this section, we use the model of Section 5.8 and consider optimal taxation when taxes have additional restrictions. We focus on three types of restricted tax forms that are usually used in practice: (i) *gender-neutral* taxes, i.e., taxes that satisfy $T_M^s(y) = T_F^s(y)$ and $T^m(y_1, y_2) = T^m(y_2, y_1)$ for all y_1, y_2, (ii) individual-earnings based taxation of couples, $T^m(y_1, y_2) = \widetilde{T}^m(y_1) + \widetilde{T}^m(y_2)$, and (iii) family-earnings based taxation of couples, $T^m(y_1, y_2) = \widetilde{T}^m(y_1 + y_2)$. Note that couples taxation is automatically gender-neutral in (ii) and (iii).

Incorporating these restrictions is fairly easy into our mechanism design problem. Restriction (i) is equivalent to requiring that $v_M^s = v_F^s$ and v^m is symmetric; restriction (ii) is equivalent to requiring that v^m is symmetric and additively separable in w_1, w_2; and restriction (iii) is equivalent to imposing that v^m is measurable only w.r.t. r. It turns out that the optimal restricted taxes is easy to characterize and there is a remarkably close connection to the optimal unrestricted taxes.

Lemma 11. *Let* $\omega_j(t) := \frac{g_j(t)}{g_M(t) + g_F(t)}$ *for* $j = M, F$.

(a) The optimal distortions in the gender-neutral tax system, $\lambda_i^{s,nrl,*}$, $\lambda_i^{m,nrl,*}$, *satisfy*

$$\lambda^{s,nrl,*}(t) = \sum_{j=M,F} \omega_j(t) \lambda_j^{s,*}(t), \quad \mathbb{E}[\lambda_i^{m,nrl,*} | w_i = t] = \sum_{j=M,F} \omega_j(t) \mathbb{E}[\lambda_j^{m,*} | w_j = t];$$

(b) The optimal distortions in the individual-earnings based taxation of couples, $\lambda^{m,ind,*}$, *satisfy*

$$\lambda^{m,ind,*}(t) = \sum_{j=M,F} \omega_j(t) \mathbb{E}[\lambda_j^{m,*} | w_j = t];$$

(c) The optimal distortions in the family-earnings based taxation of couples, $\lambda^{m,fam,*}$, satisfy

$$\lambda^{m,fam,*}(r) = \mathbb{E}[\lambda_j^{m,*}|R=r] = \frac{1 - \mathbb{E}[\alpha^m|R \geq r]}{\gamma\theta_r(r)} \; for \; j \in \{M,F\}.$$

This lemma provides closed-form expressions for optimal distortions of all three restricted tax systems. It also shows a remarkable close connection between optimal restricted and unrestricted taxation. In all three cases, the distortions under optimal restricted taxes are equal, on average, to the distortions under optimal unrestricted taxes, where averages are taken along dimensions that restricted taxes cannot use. The unrestricted and restricted optimal taxes are chosen to equalize benefits from redistribution to costs of distortions. The main difference is that the unrestricted planner can allocates those costs and benefits more efficiently for each couple while restricted planners can do that only on average.

6 Quantitative analysis

In this section, we illustrate theoretical implications of our analysis using a quantitative model.

6.1 Calibration

To calibrate our model, we use data on earnings of couples from 2020 CPS survey. We restrict attention to couples in which both individuals are between 25 and 65 years old and worked at least 20 weeks in 2019. We invert the joint distribution of earnings to obtain the distribution of productivities. In order to do so, we assume that the environment is symmetric and set $\gamma = 1/4$, so that the implied labor supply elasticity of $1/3$ is the mid-range of values considered by Diamond (1998). Following Guner et al. (2014) and Heathcote et al. (2017), who argue 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. With such taxes, the relationship between an observed vector of earnings $\mathbf{y}$ and an unobserved vector of productivities $\mathbf{w}$ is given by

$$w_i^{1/\gamma} = \frac{1}{(1-\tau)\nu} y_i^{1/\gamma - 1} (y_1 + y_2)^\tau. \tag{43}$$

To invert this mapping, we use the value of (τ, ν) that Guner et al. (2014) estimate for the U.S. married couples.

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

the marginal distribution of productivities G can be well approximated by a Pareto lognormal (PLN) distribution.[21] We choose its three parameters (η, σ, a) to match the mean level of productivity, its Gini coefficient, and the tail parameter. These three moments can be expressed analytically in terms of parameters (η, σ, a) so that these parameters can be obtained by a simple inversion of those equations (see appendix for the details). Panel (a) of Figure 1 shows the empirical and calibrated distribution G.
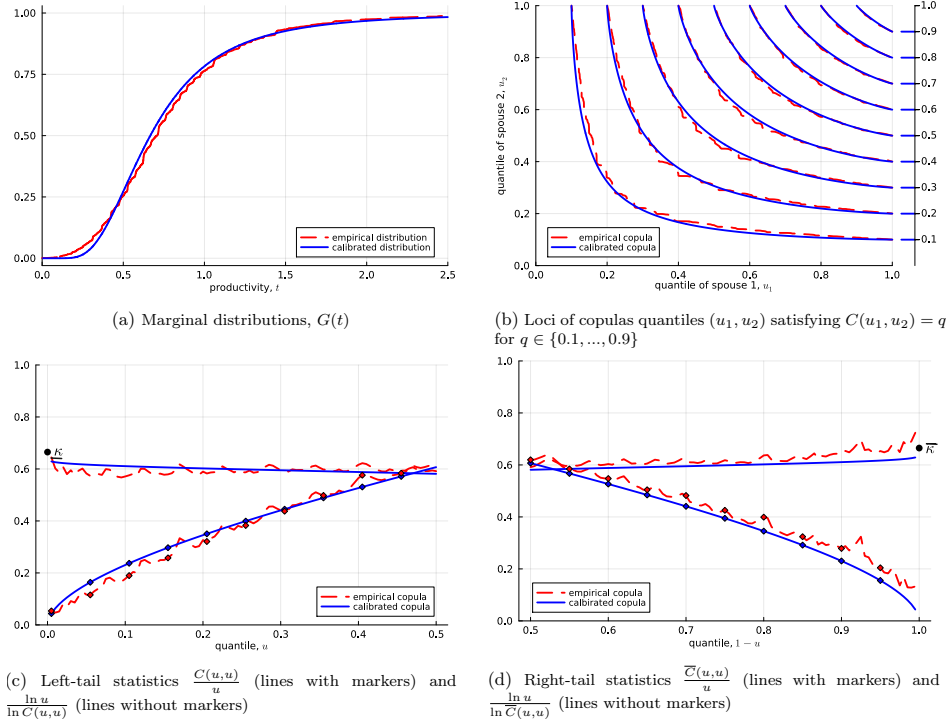

(a) Marginal distributions, $G(t)$

(b) Loci of copulas quantiles (u_1, u_2) satisfying $C(u_1, u_2) = q$ for $q \in \{0.1, ..., 0.9\}$

(c) Left-tail statistics $\frac{C(u,u)}{u}$ (lines with markers) and $\frac{\ln u}{\ln C(u,u)}$ (lines without markers)

(d) Right-tail statistics $\frac{\overline{C}(u,u)}{u}$ (lines with markers) and $\frac{\ln u}{\ln \overline{C}(u,u)}$ (lines without markers)

Figure 1: Empirical and calibrated joint distributions of productivities

Our theory emphasizes several key statistics of the joint distribution F that determine the optimal shape of taxes: the degree of dependence in productivities, right- and left-tail (in)dependence, and the speeds of convergence $\underline{\kappa}, \overline{\kappa}$. In the data, both observed earnings and backed-out productivities are positively dependent, with Kendell's tau measure of dependence for productivities equaling 0.21.[22] The joint distribution appears to be both left- and right-tail independent, but the rate of convergence to independence is fairly low. Panels (c) and (d)

---

[21]The PLN family was introduced in Colombi (1990) as 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.

[22]Kendell's tau is the standard measure of strength dependence of two variables (see Chapter 5 in Nelsen (2006)). Kendell's tau measure has an advantage over Pearson's correlation coefficient because it is independent of the marginal distributions. In our data, Pearson's correlation coefficients of earnings and productivities are 0.21 and 0.25, respectively.

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

Table 1: Calibrated parameters

illustrate this by plotting in dashed lines the value of the empirical copula[23] $C(u, u)/u$ for difference percentiles of productivities (red line), and the value of $\ln u/\ln C(u, u)$ (blue lines). Consistent with left tail-independence, $C(u, u)/u$ approaches zero for low u; consistent with the slow speed of convergence $\ln u/\ln C(u, u)$ remains much above $1/2$. Panel (d) plots similar statistics summary statistics for the right tails, using moments of the empirical survival copula.

We experimented with different families of copulas to capture these patterns and found that the Gaussian copula fits the data very well. Its parameter ρ, when chosen to match the Kendell's tau dependence coefficient,[24] also fits well the measures of left- and right-tail dependence and speeds of convergence. This can be seen from panels (c) and (d) of Figure 1, where in solid lines we plot the counterparts of calibrated Gaussian copula of the empirical objected plotted in dashed lines. Black dots show the speeds of convergence for the Gaussian copula, $\underline{\kappa} = \overline{\kappa} = \frac{1+\rho}{2}$. Panel (b) of Figure 1 shows the "isoquants" of both empirical and calibrated joint distribution, where each line plots, for a given q, all pairs (u_1, u_2) that satisfy $C(u_1, u_2) = q$. Table 1 summarizes all our parameters and their empirical counterparts.

We want to make several remarks. If one were to assume that the HSV tax schedule applies to individual rather than family earnings, then calibration of the unobservable distribution of productivities using the observed distribution of earnings is particularly simple and transparent. If the joint distribution of earnings is PLN-Gaussian and individuals face HSV taxes, then the joint distribution of productivities is also PLN-Gaussian. Moreover, the parameters

---

[23]Recall that any distribution F has corresponding functions C and $\overline{C}$, called *copula* and *survival copulas*, that satisfy $C(G_1(t_1), G_2(t_2)) = \Pr(\mathbf{w} \leq \mathbf{t})$ and $\overline{C}(1 - G_1(t_1), 1 - G_2(t_2)) = \Pr(\mathbf{w} \geq \mathbf{t})$ for all $\mathbf{t}$, respectively. Conceptually, $C(u_1, u_2)$ is the joint probability that the productivity of spouse 1 is in the u_1^{th} quantile of her marginal distribution and the productivity of spouse 2 is in the u_2^{th} quantile. Copulas allow one to isolate dependence properties of F from properties of marginal distributions G_1, G_2 in general settings.

[24]The relationship between Kendell's tau and parameter ρ of the Gaussian copula is given by Kendell's tau $= 2\frac{\arcsin \rho}{\pi}$.

of (η, σ, a, ρ) of this distribution of productivities can be expressed in closed form as functions of mean, Gini, Pareto tail, and Kendell's tau measures of the joint distribution of earnings, as well as parameters (τ, ν) of the HSV tax function. Thus, the entire calibration can be done using the moments of the raw data directly, sidestepping the need for the inversion described in equation (43). The calibration approach we present in the text is more general and uses a more realistic specification of the tax function. In any case, we tried both approaches and obtained very similar results.

Secondly, in the appendix we show how the optimal taxes would change if the joint distribution were given by the FGM rather than the Gaussian copula, calibrated to match the same Kendell's tau coefficient. The calibrated FGM copula, like the Gaussian copula we use, fits empirical isoquants reasonably well and is tail independent. However, it converges to tail independence much faster, with $\underline{\kappa}, \overline{\kappa}$ equal to $\frac{1}{2}$ that are clearly rejected by our data. Thus, comparing the results reported in the text for the Gaussian copula with those reported in the appendix for the FGM copula highlights the role of the speed of convergence of tail dependence for optimal taxation.

6.2 Optimal taxes in the calibrated economy

We focus on the benchmark economy. We assume that Pareto weights are given by $\alpha(w) = const \times e^{-mw^{1/(1-\gamma)}}$, where $const$ is chosen so that they integrate to one. Recall from our discussion in Section 3.3 that optimal marginal taxes are fully characterized by F, γ and α. We set the parameter m to 0.35, which in our calibration implies that the average optimal marginal tax rate coincides with the average marginal tax rate in 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, which we report here and in the appendix, we found that the FOA was valid. We provide additional computational details in the appendix. We summarize properties of the optimal marginal taxes in two sets of figures. The first set reports the marginal taxes $\frac{\partial}{\partial y_i} T^*(y_i, y_{-i})$ as a function of y_i, holding y_{-i} fixed at different levels. The second set reports $\frac{\partial}{\partial y_i} T^*(y_i, by_i)$ as a function of y_i, holding ratio of earnings y_{-i}/y_i fixed at different values of b. For ease of comparison, Figure 2 reports these statistics of the U.S. tax schedule implied by the estimated HSV functional form.

We now discuss the optimal taxes. We start with a benchmark economy, which restricts $k = 1$. We set the parameter m to 0.35, which in our calibration implies that the average optimal marginal tax rate coincides with the average marginal tax rate in the data. This way, the total amount of redistribution is similar in our model and the data. Figure 3 reports the
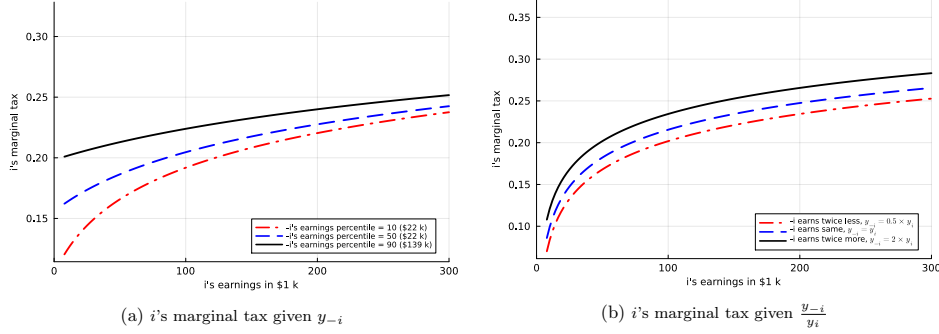
(a) i's marginal tax given y_{-i} (b) i's marginal tax given $\frac{y_{-i}}{y_i}$

Figure 2: U.S. taxes implied by the estimated HSV schedule

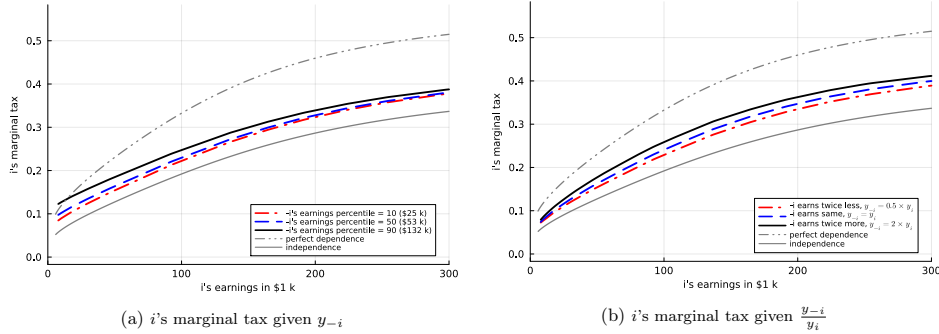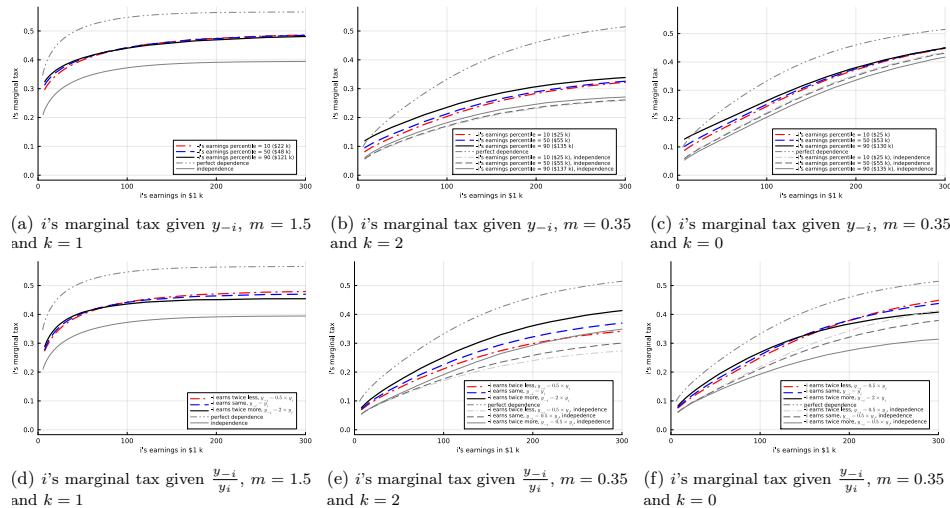(a) i's marginal tax given y_{-i} (b) i's marginal tax given $\frac{y_{-i}}{y_i}$

Figure 3: Optimal taxes, $m = 0.35$

optimal taxes in the same format as Figure 2. It also plots, in gray lines, optimal tax rates for two alternative assumptions about dependence of productivities: perfect dependence (dashed-dotted) and independence (solid). As discussed in Section 3.3, the optimal taxes under perfect dependence also coincide with taxes on single individuals. The optimal taxes with independent types are separable and thus do not depend on the other spouse's earnings.

Optimal taxes in the calibrated economy lie between the two gray lines, consistent with the comparative statics results we established in Proposition 4. One can also see from Figure 2 that the optimal marginal taxes are positively jointed for low earners since the marginal tax $\frac{\partial}{\partial y_i} T^* (y_i, y_{-i})$ is increasing in y_{-i} for low values of y_i. This result follows from Lemma 8. The same proposition also established that the optimal taxes must be negatively-jointed for high-earners.[25] This occurs at much higher earnings levels ($>$8.5 mln) than the scale of the 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. This feature is driven by the properties of the Gaussian copula. In the appendix, we plot the optimal taxes for the FGM copula and show that optimal jointness

---

[25]Spiritus et al. (2022) solved numerically a related optimal joint taxation problem and also found that optimal jointness may be positive or negative.

(a) i's marginal tax given y_{-i}, $m = 1.5$ and $k = 1$

(b) i's marginal tax given y_{-i}, $m = 0.35$ and $k = 2$

(c) i's marginal tax given y_{-i}, $m = 0.35$ and $k = 0$

(d) i's marginal tax given $\frac{y_{-i}}{y_i}$, $m = 1.5$ and $k = 1$

(e) i's marginal tax given $\frac{y_{-i}}{y_i}$, $m = 0.35$ and $k = 2$

(f) i's marginal tax given $\frac{y_{-i}}{y_i}$, $m = 0.35$ and $k = 0$

Figure 4: Optimal taxes, robustness to m and k

is much more pronounced in that case. This is consistent with our discussion in Section 4.1, where we showed that a slower pace of convergence to tail independence implies a smaller force for positive jointness at the bottom (negative jointness at the top), that eventually switched to negative (positive) jointness for tail-dependence distributions. The Gaussian copula, with its lower speed of convergence, implies smaller jointness than the FGM copula.

The optimal tax schedule shares properties implied by our analytical formulas, both qualitatively and quantitatively. In particular, in the appendix, we plot the optimal distortions $\lambda^*(\cdot, w_{-i})$ where wages of the spouse w_{-i} are held at the 50th percentile of the productivity distribution. These distortions align very closely with the analytical expression for the average distortions $\mathbb{E}\left[\lambda_i^* \middle| w_i = t\right]$ that we derived in equation (25).

In Figure 4, we explore sensitivity of these results to parameters governing Pareto weights. In the top row we plot the optimal taxes with a much more redistributive planner ($m = 1.5$). The optimal taxes are higher, reflecting a stronger desire for redistribution. Negative jointness now occurs at much lower earnings levels than in Figure 3 and can be easily seen on the graphs. Nonetheless, the level of optimal jointness remains low.

The second and third columns in Figure 4 we consider optimal taxes under non-separable Pareto weights $\alpha^m(\mathbf{w}) = const \times \left[\frac{1}{2}\alpha(w_1)^k + \frac{1}{2}\alpha(w_2)^k\right]^{1/k}$. Our benchmark weights correspond to the special case of $k = 1$. Weights α^m are submodular when $k \geq 1$ and supermodular when $k \leq 1$. The case $k = 0$ corresponds to weights α^m being measurable only w.r.t. to r, i.e., the case when the planner has an inherent preference for family earnings-based taxation. We report in the last two columns of Figure 4 optimal taxes for $(m, k) = (0.35, 2)$ and for

42

$(m, k) = (0.35, 0)$. The level of the optimal marginal tax rates in these two cases is broadly similar to our starting point, $(m, k) = (0.35, 1)$. Consistent with Corollary 1, submodularity amplifies the benefits of positive jointness at the bottom, so the positive jointness is now much more visible and substantial, especially in panel (b). On the other hand, supermodularity decreases these benefits and amplifies gains from the negative jointness at the top. As a result, both negative jointness at the bottom and positive jointness at the top are now clearly visible. Nonetheless, the optimal magnitude of this jointness is very low, and the optimal taxation is close to the individual earnings-based tax schedule. This remains the case even when $k = 0$, i.e., social weights explicitly favor family earnings-based taxation.

Figure 4 also plots the optimal tax rate under perfect dependence and independence. The optimal tax schedule is no longer separable with independent types when $k \neq 1$. Therefore, in the last two columns of Figure 4, we use several gray lines to plot optimal tax rates under independent types. One can easily see from these graphs that under independence optimal taxes are positively (negatively) jointed if α is submodular (supermodular), consistent with the results in Kleven et al. (2007) and our discussion in Section 4.1.

6.2.1 Is family-earning based taxation optimal?

In this section we consider how well family-earnings based taxes approximate the unconstrained optimal. To answer this question, we represent the optimal taxes as $T^{m,*}(\mathbf{y}) = T^{fam,*}(Y(\mathbf{y}), \varrho(\mathbf{y}))$, where $Y(\mathbf{y}) = y_1 + y_2$ are the total family earnings and $\varrho(\mathbf{y}) = \frac{\min\{y_1, y_2\}}{Y(\mathbf{y})}$ is the share of a secondary earner in total earnings. Family-earnings based taxes are optimal if $T^{fam,*}$ does not depend on the second argument.

In Figure 5 we plot $\frac{\partial}{\partial Y} T^{fam,*}(\cdot, \varrho)$ for different values of ϱ. Panel (a) uses this representation to show the U.S. tax code implied by the estimated HSV function. Since U.S. tax schedule is family earnings-based, $\frac{\partial}{\partial Y} T^{fam,US}(\cdot, \varrho)$ is the same for all ϱ. In panels (b)-(f) we plot the optimal tax on family earnings for the same specifications that we used in Figures 3 and 4. 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 - the optimal taxation of joint earnings of couples. We show that despite superficial similarity to multidimensional screening

(a) U.S. taxes implied by the estimated HSV model

(b) Optimum, $m = 0.35$ and $k = 1$

(c) Optimum, $m = 1.5$ and $k = 1$

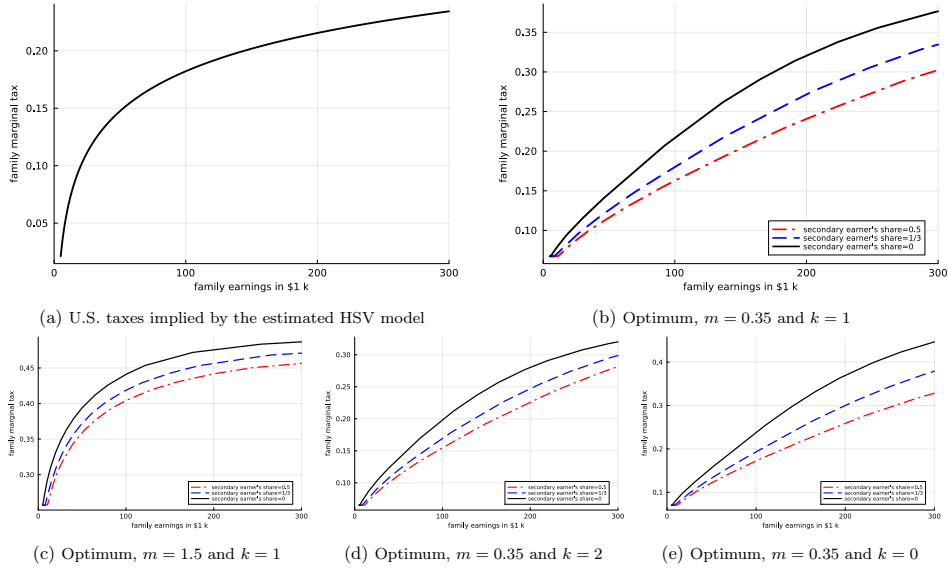(d) Optimum, $m = 0.35$ and $k = 2$

(e) Optimum, $m = 0.35$ and $k = 0$

Figure 5: Marginal taxes on family earnings

problems in industrial organization, our problem is much easier to analyze and can often be studied using the first-order approach. We identify the lack of participation constraints in our application as the key reason for this simplification.

We also 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 taxes. These conditional moments are very illuminating about 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 Pareto weights explicitly favor this type of taxation.

44

References

Akerlof, George A. 1978. "The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning." *American Economic Review*, 68(1): 8–19.

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.

Bierbrauer, Felix J, Pierre C Boyer, Andreas Peichl, and Daniel Weishaar. 2023. "The Taxation of Couples." working paper.

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

Coles, Stuart, Janet Heffernan, and Jonathan Tawn. 1999. "Dependence Measures for Extreme Value ABCnalyses." *Extremes*, 2(4): 339–365.

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, and Ilia Krasikov. 2023. "Multi-dimensional taxation." Technical report.

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.

Hajłasz, Piotr. 1999. *Sobolev mappings, co-area formula and related topics.*: Institut Mittag-Leffler.

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.

Heathcote, Jonathan, and Hitoshi Tsujiyama. 2021. "Optimal Income Taxation: Mirrlees Meets Ramsey." *Journal of Political Economy*, 129(11): 3141–3184.

Heffernan, Janet E. 2000. "A Directory of Coefficients of Tail Dependence." *Extremes*, 3(3): 279–290.

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. 2011. "Tail order and intermediate tail dependence of multivariate copulas." *Journal of Multivariate Analysis*, 102(10): 1454–1471.

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

Karlin, Samuel, and Yosef Rinott. 1980. "Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions." *Journal of Multivariate Analysis*, 10(4): 467–498.

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.

Ledford, Anthony W, and Jonathan A Tawn. 1996. "Statistics for Near Independence in Multivariate Extreme Values." *Biometrika*, 83(1): 169–187.

Ledford, Anthony W, and Jonathan A Tawn. 1998. "Concomitant Tail Behaviour for Extremes." *Advances in applied Probability*, 30(1): 197–215.

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.

Manser, Marilyn, and Murray Brown. 1980. "Marriage and Household Decision-Making: A Bargaining Analysis." *International Economic Review*, 21(1): 31–44.

McElroy, Marjorie B., and Mary Jean Horney. 1981. "Nash-Bargained Household Decisions: Toward a Generalization of the Theory of Demand." *International Economic Review*, 22(2): 333–349.

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.

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.

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.

Saez, Emmanuel. 2001. "Using Elasticities to Derive Optimal Income Tax Rates." *Review of Economic Studies*, 68(1): 205–229.

Salanie, Bernard. 2003. *The Economics of Taxation.* MIT Press: The MIT Press.

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.

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.

Let $U \subseteq \mathbb{R}^n$ be open, where $n \geq 1$. A function $v : U \to \mathbb{R}$ is compactly supported in U if it is zero outside some compact $C \subset U$. A measurable $v : U \to \mathbb{R}$ is *locally integrable* on U if $\int_C |v(\mathbf{w})| \, d\mathbf{w} < \infty$ for every compact $C \subset U$, and it is called *integrable* on U if $\int_U |v(\mathbf{w})| \, d\mathbf{w} < \infty$. A measurable $v : U \to \mathbb{R}$ is said to be *essentially bounded* on U if there exists m such that $|v(\mathbf{w})| \leq m$ a.e. on U. We will write $\mathscr{L}^1$ and $\mathscr{L}^\infty$ for the spaces of integrable and essentialle bounded functions, respectively.

A locally integrable function $v : U \to \mathbb{R}$ is *weakly differentiable* on U if there exists a locally integrable vector field $\frac{\partial v}{\partial \mathbf{w}} : U \to \mathbb{R}^n$ such that for all infinitely differentiable ϕ with a compact support in U,

$$\int_U \frac{\partial \phi(\mathbf{w})}{\partial w_i} v(\mathbf{w}) \, d\mathbf{w} = - \int_U \phi(\mathbf{w}) \frac{\partial v(\mathbf{w})}{\partial w_i} d\mathbf{w}.$$

The vector field of partial derivatives of v 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 weakly differentiable $\phi \in \mathcal{L}^\infty$ with $\frac{\partial \phi}{\partial \mathbf{w}} \in \mathcal{L}^\infty$ and weakly differentiable $v \in \mathcal{L}^1$ with $\frac{\partial v}{\partial \mathbf{w}} \in \mathcal{L}^1$, the product ϕv is weakly differentiable with $\frac{\partial (\phi v)}{\partial \mathbf{w}} = \frac{\partial \phi}{\partial \mathbf{w}} v + \frac{\partial v}{\partial \mathbf{w}} \phi$; moreover, ϕv and its weak gradient is integrable. Then, the following identity, known as the *Divergence Theorem* (Theorem 1.5.3.1 in Grisvard (2011)), is satisfied for any bounded U with a Lipshitz boundary:

$$\int_{\overline{U}} \frac{\partial \left[\phi(\mathbf{w}) v(\mathbf{w}) \right]}{\partial \mathbf{w}} d\mathbf{w} = \int_{\partial U} \phi(\mathbf{w}) v(\mathbf{w}) n_i(\mathbf{w}) \sigma(d\mathbf{w}),$$

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 *Coarea Formula* is Theorem 11 in Hajłasz (1999). Let $Q : \mathbb{R}^n_{++} \to \mathbb{R}^m$ be weakly differentiable, where $m \leq n$. Then, for every (Borel) measurable $v : \mathbb{R}^n_+ \to \mathbb{R}$ that is either non-negative or is such that $v \cdot \mathcal{J}Q$ is integrable, where $\mathcal{J}Q(\mathbf{w})$ is the Jacobian of Q at $\mathbf{w}$,

$$\int v(\mathbf{w}) \cdot \mathcal{J}Q(\mathbf{w}) d\mathbf{w} = \int \left(\int_{Q^{-1}(\boldsymbol{t})} v(\mathbf{w}) \mathcal{H}^{n-m}(d\mathbf{w}) \right) d\boldsymbol{q},$$

where $\mathcal{H}^{n-m}$ is the so called $(n-m)$-dimensional Hausdorff measure. As explained in Chapter 2 of Evans and Garzepy (2018), the Hausdorff measure coincides with the standard Lebesgue measure for "nice" sets, e.g., $n = 2$, $m = 1$ and $Q^{-1}(t)$ is a plane curve. On the other hand, if $n = m$ and Q is injective, then the Hausdorf measure is a counting measure; thus, the inner integral on the right-hand side is $v\left(Q^{-1}(t)\right)$, and we recover the standard change of variables formula.

9 Proofs for Section 3.1

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

Lemma 12. *Consider μ and (v^s, c^s, y^s), $(v^m, \mathbf{c}^m, \mathbf{y}^m)$ such that (3), (4), (5) and (6) 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 (7). Moreover, $\mu \in (0, 1)$ and (a) $v^s g, v^m f \in \mathscr{L}^1$ and (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 (3) to substitute for c^s in (4) to obtain

$$v^s(w) \geqslant v^s(\widehat{w}) + \gamma \left(\frac{y^s(\widehat{w})}{\widehat{w}}\right)^{1/\gamma} \left(\left(\frac{\widehat{w}}{w}\right)^{1/\gamma} - 1\right). \tag{44}$$

Monotonicity of v^s follows directly from (44) and nonnegativity of earnings. As a result, v^s is bounded from below by $v^s(0)$.

A further examination of (44) 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 (44) together with the fact that the maximum on the right-hand side is attained at $\widehat{w} = w$ establishes that (7) 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 (7).

We now show properties $\mu \in (0, 1)$ and (a), (b). To begin, use (3) to rewrite (5) in terms

of v^s and v^m as follows:

$$\frac{\mu}{2} \int \left(\sum_{i=1}^{2} \left(y_i^m(\boldsymbol{w}) - \gamma \left(\frac{y_i^m(\boldsymbol{w})}{w_i} \right)^{1/\gamma} \right) - v^m(\mathbf{w}) \right) F(d\mathbf{w})+$$

$$+ (1-\mu) \int \left(y^s(w) - \gamma \left(\frac{y^s(w)}{w} \right)^{1/\gamma} - v^s(w) \right) G(dw) \geq 0. \quad (45)$$

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 $\int w^{1/(1-\gamma)} G(dw) < \infty$, the left-hand side of (45) is finite.

It is immediate that, if $\mu \in (0,1)$, then property (a) holds. If $\mu = 0$, then, since $\Phi(\mu) = -\infty$ and $\int v^m(\boldsymbol{w}) F(d\boldsymbol{w}) \geq v^m(0)$, we must have $\int v^s(w) G(dw) = \infty$ due to (6). This contradicts (45). The similar argument rules out $\mu = 1$, and hence property (a) is established.

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

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

It is easy to see that the value of this problem diverges to $-\infty$ as $b \to \infty$. Substitute $\frac{\partial v^s}{\partial w}$ for y^s using (7) to obtain that $\int w \frac{\partial v^s(w)}{\partial w} G(dw)$ must be finite. The same argument applies to v^m, thus $\int w_i \frac{\partial v_i^m(\boldsymbol{w})}{\partial w_i} F(d\boldsymbol{w}) < \infty$ for $i = 1, 2$. $\qquad \square$

9.1 Relaxed problem

We now formally define and further simplify the relaxed problem introduced in the main text. Let $\mathcal{V}^s$ and $\mathcal{V}^m$ be the spaces of functions v^s and v^m satisfying the conditions listed in Lemma 12. Then, the relaxed problem is to select $\mu \in [0,1]$ and $(v^s, v^m) \in \mathcal{V}^s \times \mathcal{V}^m$ to maximize $\mathcal{W}$ defined in (2) subject to (6) and (8).

To make our analysis applicable to study extensions in Section 5, we allow $\alpha^m(w_1, w_2) \neq \frac{\alpha(w_1) + \alpha(w_2)}{2}$ but still require this function to be symmetric and $\int \alpha(w) G(dw) = 1$. The asymmetric case relevant for Sections 5.8 and 5.9 will be discussed separately.

Remark that $\mathcal{W}$ can be rewritten as follows:

$$\mathcal{W} = \frac{\mu}{2} \int \left(\alpha^m(\boldsymbol{w}) - \mathbb{E}\left[\alpha^m \right] \right) v^m(\boldsymbol{w}) F(d\boldsymbol{w}) + (1-\mu) \int \left(\alpha(w) - 1 \right) v^s(w) G(dw)+$$

$$+ \int_{\mu}^{1} \Phi(\varepsilon) d\varepsilon + \left(\mu \mathbb{E}\left[\alpha^m \right] + (1-\mu) \right) \left[\frac{\mu}{2} \int v^m(\boldsymbol{w}) F(\boldsymbol{w}) + (1-\mu) \int v^s(w) G(dw) \right] +$$

$$+ \left(\mathbb{E}\left[\alpha^m \right] - 1 \right) \mu (1-\mu) \left[\frac{1}{2} \int v^m(\boldsymbol{w}) F(\boldsymbol{w}) - \int v^s(w) G(dw) \right]. \quad (46)$$

Two terms in square brackets can be solved for from (6) and (8). It is immediate that the budget constraints must bind, thus the term in the second line equals to $\mathcal{S}$, which is given by

$$\mathcal{S} = \frac{\mu}{2} \int \sum_{i=1}^{2} \left(w_i^{1+\gamma} \left(\frac{\partial v^m(\mathbf{w})}{\partial w_i} \right)^\gamma - \gamma w_i \frac{\partial v^m(\mathbf{w})}{\partial w_i} \right) F(d\mathbf{w}) +$$

$$+ (1-\mu) \int \left(w^{1+\gamma} \left(\frac{\partial v^s(w)}{\partial w} \right)^\gamma - \gamma w \frac{\partial v^s(w)}{\partial w} \right) G(dw). \quad (47)$$

The term in the square brackets in the third line equals to $\Phi(\mu)$ due to (6).

Putting all pieces together, the relaxed problem is

$$\max_{\substack{\mu \in [0,1] \\ (v^s, v^m) \times \mathcal{V}^s \times \mathcal{V}^m}} \frac{\mu}{2} \int (\alpha^m(\mathbf{w}) - \mathbb{E}[\alpha^m]) \, v^m(\mathbf{w}) F(d\mathbf{w}) + (1-\mu) \int (\alpha(w) - 1) \, v^s(w) G(dw) +$$

$$+ \int_\mu^1 \Phi(\varepsilon) d\varepsilon + (\mu \mathbb{E}[\alpha^m] + (1-\mu)) \, \mathcal{S} + (\mathbb{E}[\alpha^m] - 1) \, \mu(1-\mu)\Phi(\mu). \quad (48)$$

It is worth to mention that the solution to (48) is defined up to constants $v^s(0)$ and $v^m(\mathbf{0})$. These constants are pinned by two binding constraints (6), (8) so that $\int v^{s,*}(w) G(dw) = \mathcal{S}^* - \mu^* \Phi(\mu^*)$ and $\frac{1}{2} \int v^{m,*}(\mathbf{w}) F(\mathbf{w}) = \mathcal{S}^* + (1-\mu^*)\Phi(\mu^*)$. Here, $\mathcal{S}^*$ is (47) 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 (48). 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) $\frac{\partial (w\lambda^s g)/\partial w}{g}, \frac{\sum_{i=1}^{2} \partial(w_i \lambda_i^m f)/\partial w_i}{f} \in \mathscr{L}^\infty$ and (A5) $\lambda^s \theta, \lambda_1^m \theta_1, \lambda_2^m \theta_2 \in \mathscr{L}^\infty$.*

Set $\eta := (\mu \mathbb{E}[\alpha^m] + (1-\mu))^{-1}$. Then, (v^s, v^m) is in $\mathcal{V}^s \times \mathcal{V}^m$ and maximizes the objective in (48) for fixed μ if and only if

$$\frac{\partial (\gamma w \lambda^s(w) g(w))}{\partial w} = \eta (\alpha(w) - 1) g(w), \quad (49)$$

$$\sum_{i=1}^{2} \frac{\partial (\gamma w_i \lambda_i^m(\mathbf{w}) f(\mathbf{w}))}{\partial w_i} = \eta (\alpha^m(\mathbf{w}) - \mathbb{E}[\alpha^m]) f(\mathbf{w}). \quad (50)$$

If $(v^{s,*}, v^{m,*})$ verifies (A1)-(A5), then (49), (50) hold and the following first-order conditions w.r.t. μ is satisfied:

$$\frac{1-\gamma}{2} \int \sum_{i=1}^{2} w_i \left(\frac{w_i}{1+\lambda_i^{m,*}(\boldsymbol{w})}\right)^{\gamma/(1-\gamma)} F(d\boldsymbol{w}) - (1-\gamma) \int w \left(\frac{w}{1+\lambda^{s,*}(w)}\right)^{\gamma/(1-\gamma)} G(dw) =$$

$$= \Phi(\mu^*) + \eta^*\left(1 - \mathbb{E}\left[\alpha^m\right]\right)\left(\mathcal{S}^* + \frac{\partial\left[\mu^*(1-\mu^*)\Phi(\mu^*)\right]}{\partial\mu}\right). \quad (51)$$

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 (49), (50) 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 (14) and satisfy (15). 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 (7), this is equivalent to the fact that earnings change continuously ruling out kinks in taxes. Condition (A3) ensures that (weak) derivatives in (49), (50) 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 $\hat{v}^s \in \mathcal{V}^s$. 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 (51) when $\mathbb{E}[\alpha^m] \neq 1$, because the relaxed problem is not jointly concave in μ and (v^s, v^m). However, in the benchmark economy or more generally when $\mathbb{E}[\alpha^m] = 1$, (51) pins down a unique value of μ^*, because (49), (50) do not depend on μ due to $\eta = 1$ for all μ.

Proof. We first show that (v^s, v^m) is in $\mathcal{V}^s \times \mathcal{V}^m$ provided that (A2) holds. Indeed, by (14), $w\frac{\partial v^s(w)}{\partial w} \leq \left(\frac{w}{1+\overline{\lambda}}\right)^{1/(\gamma-1)}$, which gives

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

Since the value of $\int w^{1/(1-\gamma)}G(dw)$ is finite, both $v^s g$ and $w\frac{\partial v^s}{\partial w}g$ are integrable, thus $v^s \in \mathcal{V}^s$.

The argument for v^m is identical as $w_i \frac{\partial v_i^m(\boldsymbol{w})}{\partial w_i} \leq \left(\frac{w_i}{1+\overline{\lambda}}\right)^{1/(\gamma-1)}$ for $i = 1, 2$ implies

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

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

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

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:

$$\lim_{t \to 0} \frac{\Upsilon^s((1-t)v^s + t\hat{v}^s) - \Upsilon^s(\hat{v}^s)}{t} \leq 0 \ \ \forall \hat{v}^s \in \mathscr{V}^s. \tag{52}$$

It is routine to verify using the Monotone Convergence Theorem that the limit in (52) can be taken under the integral sign and that this condition is equivalent to

$$\int \left(\eta \left(\alpha(w) - 1 \right) (\hat{v}^s(w) - v^s(w)) + \gamma w \lambda^s(w) \left(\frac{\partial \hat{v}^s(w)}{\partial w} - \frac{\partial v^s(w)}{\partial w} \right) \right) G(dw) \leq 0 \ \ \forall \hat{v}^s \in \mathscr{V}^s. \tag{53}$$

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

$$\int \left(\eta \left(\alpha(w) - 1 \right) v^s(w) + \gamma w \lambda^s(w) \frac{\partial v^s(w)}{\partial w} \right) G(dw) = 0, \tag{54}$$

which allows to eliminate v^s from (53).

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

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

(A2) ensures that the second 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 on right-hand side goes to 0 as $t \to \infty$ due to (A5). By (A2) and (A4), the Dominated Convergence Theorem implies that (53) when combined with (55) can be rewritten as

$$\int \left[\eta \left(\alpha(w) - 1 \right) - \frac{\partial \left(\gamma w \lambda^s(w) g(w) \right) / \partial w}{g(w)} \right] \hat{v}^s(w) G(dw) \leq 0. \tag{55}$$

By (A2), $w\frac{\partial v^s(w)}{\partial w}] \geq \left(\frac{w}{1+\underline{\lambda}}\right)^{1/(\gamma-1)}$. Hence, nonnegativity constraints on earnings are slack, i.e., in a neighborhood of every $w \in \mathbb{R}_{++}$, (55) 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, (55) 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

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

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

$$\int \left(\eta \left(\alpha^m(w) - \mathbb{E}\left[\alpha^m\right] \right) \hat{v}^m(\boldsymbol{w}) + \sum_{i=1}^{2} \gamma w_i \lambda_i^m(\boldsymbol{w}) \frac{\partial \hat{v}^m(\boldsymbol{w})}{\partial w_i} \right) F(d\boldsymbol{w}) \leq 0 \ \forall \hat{v}^m \in \mathscr{V}^m, \ (56)$$

$$\int \left(\eta \left(\alpha^m(w) - \mathbb{E}\left[\alpha^m\right] \right) v^m(\boldsymbol{w}) + \sum_{i=1}^{2} \gamma w_i \lambda_i^m(\boldsymbol{w}) \frac{\partial v^m(\boldsymbol{w})}{\partial w_i} \right) F(d\boldsymbol{w}) = 0. \quad (57)$$

Let $\hat{v}^m$ be a function in $\mathscr{V}^m$.

By the Divergence Theorem,

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

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

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,

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

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

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 $1 - G(t) \leq \Pr(\boldsymbol{w} \geq (t, t))$ and $\Pr(\boldsymbol{w} \geq (t,t))\mathbb{E}\left[\hat{v}^m | \max\{w_1, w_2\} = t\right] \to 0$ as $t \to \infty$.

The rest of the argument is exactly the same as for singles. To sum up, two differential equations (49), (50) are necessary and sufficient optimality conditions for fixed $\mu \in (0,1)$.

It remains to show the necessary first-order condition for μ^*. Equation (51) directly follows from differentiating $\mathcal{W}$ w.r.t. μ and noting that

$$\Upsilon^s(v^{s,*}) = (1-\gamma) \int w \left(\frac{w}{1+\lambda^{s,*}(w)} \right)^{\gamma/(1-\gamma)} G(dw), \tag{58}$$

$$\Upsilon^m(v^{m,*}) = \frac{1-\gamma}{2} \int \sum_{i=1}^{2} w_i \left(\frac{w_i}{1+\lambda_i^{m,*}(\boldsymbol{w})} \right)^{\gamma/(1-\gamma)} F(d\boldsymbol{w}). \tag{59}$$

$\square$

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 (50) and (20) 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 $\left((1+\lambda_i^m(\mathbf{w}))^{1/(\gamma-1)} w_i^{\gamma/(\gamma-1)} \right)_{i=1,2}$ is continuously differentiable with derivatives that are uniformly continuous on bounded subsets and (20) holds, then there exists a unique (up to a constant) function v^m such that equation (14) holds for these v^m and $\boldsymbol{\lambda}^m$.

9.3 Proof of Proposition 1

Proof. It is easy to see that $\lambda^s := \lambda^\#$ and $\boldsymbol{\lambda}^m := \left(\frac{1}{2}\lambda^\#(w_1), \frac{1}{2}\lambda^\#(w_2) \right)$ satisfy conditions (A1)-(A5) of Proposition 2 provided that $\overline{\theta}^{-1}(t)$ and $\underline{\theta}^{-1}(t)$ converges to a finite limit as $t \to \infty$ and $t \to 0$, respectively, which is implied by finiteness of $\lim_{t\to 0,\infty} \lambda^\#(t)$. Moreover, these distortions also verify the necessary and sufficient optimality conditions listed in this proposition, i.e., (49) and (50). 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 $\frac{\partial v^{s,*}(x^{-\gamma})}{\partial x}$ is proportional to $x \cdot \left(1+\lambda^\#(x^{-\gamma}) \right)$ and $\frac{\partial v^{m,*}(x_1^{-\gamma}, x_2^{-\gamma})}{\partial x_i}$ is proportional to $x_i \cdot \left(1+\frac{1}{2}\lambda^\#\left(x_i^{-\gamma} \right) \right)$. Then, 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

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

$\square$

10 Proofs for Section 4

Throughout this section, it is assumed that the first-order approach is valid, $\lambda^{s,*}$ and $\boldsymbol{\lambda}^{m,*}$ satisfy conditions (A1)-(A5) of Proposition 2.

10.1 Coarea formula

We first formally state and prove equation (24) that computes averages of optimal marginal taxes conditional on level curves of some function Q. We will provide this result for the general symmetric case with α^m potentially different from $\frac{\alpha(w_1)+\alpha(w_2)}{2}$ so that the Coarea formula will be directly applicable in Section 5. In the special case of our benchmark economy, $\alpha^m(w_1, w_2) = \frac{\alpha(w_1)+\alpha(w_2)}{2}$, thus the value of η^* in Proposition 2 equals one.

Proposition 3. *Let $Q : \mathbb{R}_{++}^2 \to \mathbb{R}_{++}$ be a continuous function with piecewise continuously differentiable level curves such that*

$$\mathbb{E}\left[\sum_{i=1}^2 w_i \left|\frac{\partial Q(\boldsymbol{w})}{\partial w_i}\right| \bigg| Q \le t\right] < \infty. \tag{60}$$

Then,

$$\mathbb{E}\left[\sum_{i=1}^2 \gamma \lambda_i^{m,*} \frac{\partial \ln Q}{\partial \ln w_i} \bigg| Q = t\right] = \eta^* \frac{\mathbb{E}[\alpha^m] - \mathbb{E}[\alpha^m|Q \ge t]}{-\partial \ln Pr(Q \ge t)/\partial \ln t} = \eta^* \frac{\mathbb{E}[\alpha^m|Q \le t] - \mathbb{E}[\alpha^m]}{\partial \ln Pr(Q \le t)/\partial \ln t}. \tag{61}$$

Proof. Since Q has piecewise continuously differentiable level curves, it is weakly differentiable and the boundary of $\{\mathbf{w}|Q(\mathbf{w}) \le t\}$ is Lipshitz. Then, our assumptions on $\boldsymbol{\lambda}^{m,*}$ ensure that (50) can be transformed by the Divergence Theorem. Thus, we obtain

$$\int_{Q^{-1}(t)} \sum_{i=1}^2 \gamma w_i \lambda_i^{m,*}(\mathbf{w}) f(\mathbf{w}) n_i(\mathbf{w}) \sigma(d\mathbf{w}) = \int_{\{\mathbf{w}|Q(\mathbf{w}) \le t\}} \sum_{i=1}^2 \frac{\partial(\gamma w_i \lambda_i^{m,*}(\mathbf{w}) f(\mathbf{w}))}{\partial w_i} d\mathbf{w},$$

$$= \Pr(Q \le t) \eta^* (\mathbb{E}[\alpha^m|Q \le t] - \mathbb{E}[\alpha^m]), \tag{62}$$

where $\mathbf{n}(\mathbf{w})$ is the outward unit normal to $\{\widehat{\mathbf{w}}|Q(\widehat{\mathbf{w}}) \le t\}$ at $\mathbf{w}$.

By (60), $\sum_{i=1}^2 \gamma w_i \lambda_i^{*,m} \frac{\partial Q}{\partial w_i} f$ is integrable on $\{\mathbf{w}|Q(\mathbf{w}) < t\}$. The Coarea Formula discussed in the mathematical appendix implies that

$$\mathbb{E}\left[\sum_{i=1}^2 \gamma w_i \lambda_i^{*,m}(\boldsymbol{w}) \frac{\partial Q(\boldsymbol{w})}{\partial w_i} \bigg| Q \le t\right] = \frac{\int_0^t \left(\int_{(Q)^{-1}(s)} \sum_{i=1}^2 \gamma w_i \lambda_i^{m,*}(\boldsymbol{w}) \frac{\partial Q(\boldsymbol{w})}{\partial w_i} f(\mathbf{w}) \frac{\sigma(d\mathbf{w})}{\|\partial Q(\boldsymbol{w})/\partial \boldsymbol{w}\|}\right) ds.}{\Pr(Q \le t)}$$

Then, by the definition of conditional expectation, for a.e. q,

$$\mathbb{E}\left[\sum_{i=1}^2 \gamma w_i \lambda_i^{m,*}(\boldsymbol{w}) \frac{\partial Q(\boldsymbol{w})}{\partial w_i} \bigg| Q = t\right] = \frac{\int_{(Q)^{-1}(t)} \sum_{i=1}^2 \gamma w_i \lambda_i^{m,*}(\boldsymbol{w}) \frac{\partial Q(\boldsymbol{w})}{\partial w_i} f(\mathbf{w}) \frac{\sigma(d\mathbf{w})}{\|\partial Q(\boldsymbol{w})/\partial \boldsymbol{w}\|}}{d\Pr(Q \le t)\big/ dt}. \tag{63}$$

It is easy to see that the unit normal vector on the boundary of $\{\mathbf{w}|Q(\mathbf{w}) \leq t\}$ is given by $n_i(\mathbf{w}) = \frac{\partial Q(\boldsymbol{w})/\partial w_i}{\|\partial Q(\boldsymbol{w})/\partial \boldsymbol{w}\|}$. Divide each side of (63) by Q to establish the second part of equation (61).

To see the first part of equation (61), use Bayes rule to obtain

$$\Pr(Q \geq t)(\mathbb{E}[\alpha^m] - \mathbb{E}[\alpha|Q \geq t]) = \Pr(Q \leq t)(\mathbb{E}[\alpha|Q \leq t] - \mathbb{E}[\alpha^m]).$$

And, the result follows from the second part of equation (61). $\qquad\square$

10.2 Proof of Lemma 2

Proof. Since $x \mapsto (1+x)^{\gamma/(\gamma-1)}$ is decreasing and convex, for all t,

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

where the second inequality is due to Lemma 4. Then, (51) implies that $\Phi(\mu^*) \geq 0$; as a result, $\mu^* \geq \mu^{LF}$ and

$$\frac{1}{2}\int v^{m,*}(\boldsymbol{w})F(d\boldsymbol{w}) - \int v^{s,*}(w)G(dw) = \Phi(\mu^*) \geq 0.$$

$\qquad\square$

10.3 Average distortions

10.3.1 Proof of Lemma 3

Proof. The argument is the same as in the proof of Proposition 1. $\qquad\square$

10.3.2 Proof of Lemma 4

Proof. By $F^a \leq_{PQD} F^b$,

$$\Pr^a(w_{-i} \geq t_{-i}|w_i \geq t_i) \leq Pr^b(w_{-i} \geq t_{-i}|w_i \geq t_i) \quad \forall \mathbf{t}. \tag{64}$$

Since α is decreasing, the first-order stochastic dominance gives $\mathbb{E}^a[\alpha^m|w_i \geq t] \geq \mathbb{E}^b[\alpha^m|w_i \geq t]$, thus $\mathbb{E}^a\left[\lambda_i^{m,a,*}|w_i = t\right] \leq \mathbb{E}^b\left[\lambda_i^{m,b,*}|w_i = t\right]$.

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

If F^b is perfectly assortative, then $\mathbb{E}^b[\alpha^m|w_i \geq t]$ coincides with $\mathbb{E}[\alpha|w \geq t]$ which implies that $\mathbb{E}^b\left[\lambda_i^{m,b,*}|w_i = t\right] = \lambda^{s,b,*}(t) = \lambda^{s,a,*}(t)$ for all t. $\qquad\square$

10.3.3 Proof of Lemma 5

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

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, $\mathbb{E}[\alpha^a(w_j)|w_i \geq t] \geq \mathbb{E}[\alpha^b(w_j)|w_i \geq t]$ for $j = 1, 2$, thus $\mathbb{E}[\lambda_i^{a,*}|w_i = t] \leq \mathbb{E}[\lambda_i^{b,*}|w_i = t]$ for all t. $\square$

10.3.4 Proof of Lemma 6

Proof. Since α is nondecreasing, for every value of ι,

$$\mathbb{E}[\alpha^m|w_2/w_1 \geq \iota] \geq \frac{1}{2}\mathbb{E}[\alpha(w_1)|w_2/w_1 \geq \iota].$$

The assumption ensures that w_1 conditional on $w_2/w_1 \geq \iota$ converges almost surely to 0 as $\iota \to \infty$. To see it note that $\mathbb{E}[w_1|w_2/w_1 \geq \iota] \leq \frac{1}{\iota}\mathbb{E}[w_2|w_2/w_1 \geq \iota] \to_{\iota \to \infty} 0$. Since $\alpha(0) > 2$, $\mathbb{E}[\alpha^m|w_2/w_1 \geq \iota] > 1$ for all large values of ι, thus $\mathbb{E}[\lambda_2^{m,*} - \lambda_1^{m,*}|w_2/w_1 \geq \iota] < 0$ for all sufficiently large ι due to (27). $\square$

10.4 Average jointness

It is assumed throughout that the following coefficients measuring tail-dependence and speed of convergence to are well-defined: $\overline{\chi} := \lim_{t \to \infty} \Pr(w_{-i} \geq t|w_i \geq t)$, $\underline{\chi} := \lim_{\to 0} \Pr(w_{-i} \geq t|w_i \geq t)$ and $\overline{\kappa} := \lim_{t \to \infty} \frac{\ln \Pr(w_i \geq t)}{\ln \Pr(\boldsymbol{w} \geq (t,t))}$, $\underline{\kappa} := \lim_{u \to 0} \frac{\ln \Pr(w_i \leq t)}{\ln \Pr(\boldsymbol{w} \leq (t,t))}$. Moreover, both limiting distributions $F(\cdot|\infty)$ and $F(\cdot|0)$ exist.

10.4.1 Proof of Lemma 7

Proof. We start with a preliminary observation that will be useful to sign average jointness at the extremes. Unpack $\mathbb{E}[\lambda_i^{m,*}|w_i = t]$ conditioning on spouse j being more and less productive than spouse i to get

$$1 = \frac{\mathbb{E}[\lambda_i^{m,*}|w_i = t \leq w_{-i}]}{\mathbb{E}[\lambda_i^{m,*}|w_i = t]}\Pr(w_{-i} \geq t|w_i = t) + \frac{\mathbb{E}[\lambda_i^{m,*}|w_i = t \leq w_{-i}]}{\mathbb{E}[\lambda_i^{m,*}|w_i = t]}\Pr(w_{-i} \leq t|w_i = t).$$
(65)

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

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

which clearly equals to $1 - \frac{\overline{\chi}}{2}$. Then, dividing each side of (65) by $\mathbb{E}\left[\lambda_i^{m,*}|w_i = t\right]$ and taking t to ∞, we obtain

$$1 = \lim_{t \to \infty} \frac{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t \leq w_{-i}\right]}{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t\right]} \frac{\overline{\chi}}{2} + \lim_{t \to \infty} \frac{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t \geq w_{-i}\right]}{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t\right]} \left(1 - \frac{\overline{\chi}}{2}\right). \tag{66}$$

We are now in position to prove Lemma 7. Suppose $\overline{\kappa} < 1$, thus $\overline{\chi} = 0$. As discussed in the text, $\overline{B}(\infty) = \overline{\kappa}$, and $\overline{A}(\infty) = \frac{1 - \frac{1}{2}\alpha(\infty) - \frac{1}{2}\alpha(\infty)}{1 - \frac{1}{2}\alpha(\infty) - \frac{1}{2}\int \alpha(w)F(dw|\infty)}$.

Case (i). If $F(\cdot|\infty)$ strictly first-order stochastically dominates G, then $\overline{A}(\infty) < 2$ due to $\int \alpha(w)F(dw|\infty) < \int \alpha(w)G(dw) = 1$, thus the product $\overline{A}(\infty) \times \overline{B}(\infty) < 1$ when $\overline{\kappa} = \frac{1}{2}$.

Case (ii). If $F(\cdot|\infty)$ is degenerate, then $\overline{A}(\infty) = 1$, thus the product $\overline{A}(\infty) \times \overline{B}(\infty) < 1$ when $\overline{\kappa} < 1$. $\qquad\square$

10.4.2 Proof of Lemma 8

Proof. For the left corner, the analogue of (66) is

$$1 = \lim_{t \to 0} \frac{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t \leq w_{-i}\right]}{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t\right]} \left(1 - \frac{\underline{\chi}}{2}\right) + \lim_{t \to 0} \frac{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t \geq w_{-i}\right]}{\mathbb{E}\left[\lambda_i^{m,*}|w_i = t\right]} \frac{\underline{\chi}}{2}.$$

And, the result follows from the same argument as in the proof of Lemma 7. $\qquad\square$

10.4.3 Relationship to Kleven et al. (2007)

In their working version of the paper, Kleven et al. (2007) (KKS for short) outlined how jointness can be signed at each productivity vector under the assumption of independent types. We now briefly (and heuristically) review their argument in the notations of the present paper focusing for simplicity on a symmetric economy with supermodular Pareto weights. The argument for submodular weights is identical.

When types are independent, equation (50) can be rewritten as

$$\sum_{i=1}^{2} \frac{\partial\left(\lambda_i^{m,*}(\boldsymbol{w})\gamma w_i g(w_i)\right)/\partial w_i}{g(w_i)} = \alpha^m(\boldsymbol{w}) - 1. \tag{67}$$

Differentiate twice (67) to obtain

$$\sum_{i=1}^{2} \frac{\partial}{\partial w_i}\left(\frac{\partial^2\left(\lambda_i^{m,*}(\boldsymbol{w})\gamma w_i g(w_i)\right)/\partial w_1 \partial w_2}{g(w_i)}\right) = \frac{\partial^2 \alpha^m(\boldsymbol{w})}{\partial w_1 \partial w_2}. \tag{68}$$

Define the set of types U for which jointness is strictly positive as

$$U := \left\{\boldsymbol{w}|\frac{\partial\left(\lambda_i^{m,*}(\boldsymbol{w})\gamma w_i g(w_i)\right)}{\partial w_{-i}} > 0\right\}.$$

Due to the boundary conditions discussed in the text, the set U is contained in the interior of $\mathbb{R}^2_+$.

KKS suggested to integrate (68) over U, assuming that it is non-empty, and then use the Divergence Theorem to obtain

$$
\begin{aligned}
\int_U \frac{\partial^2 \alpha(\boldsymbol{w})}{\partial w_1 \partial w_2} d\boldsymbol{w} &= \int_U \sum_{i=1}^2 \frac{\partial}{\partial w_i} \left(\frac{\partial^2 \left(\lambda_i^{m,*}(\boldsymbol{w}) \gamma w_i g(w_i)\right)/\partial w_1 \partial w_2}{g(w_i)} \right) = \\
&= \int_{\partial U} \sum_{i=1}^2 \left(\frac{\partial^2 \left(\lambda_i^{m,*}(\boldsymbol{w}) \gamma w_i g(w_i)\right)/\partial w_1 \partial w_2}{g(w_i)} \right) n_i(\boldsymbol{w}) \sigma(d\boldsymbol{w}), \quad (69)
\end{aligned}
$$

where, as usual, $\boldsymbol{n}(\boldsymbol{w})$ is the outward unit normal to ∂U at $\boldsymbol{w}$. Clearly, the expression on the left-hand side is non-negative due to supermodularity of α. On the other hand, since on $\frac{\partial\left(\lambda_i^{m,*}(\boldsymbol{w})\gamma w_i g(w_i)\right)}{\partial w_{-i}} > 0$ on U but $\frac{\partial\left(\lambda_i^{m,*}(\boldsymbol{w})\gamma w_i g(w_i)\right)}{\partial w_{-i}} < 0$ on the interior of its complement, we must have $n_i(\boldsymbol{w}) \propto -\frac{\partial^2\left(\lambda_i^{m,*}(\boldsymbol{w})\gamma w_i g(w_i)\right)}{\partial w_1 \partial w_2}$. It follows that the second line in (69) is non-positive. Conclude $\boldsymbol{n}(\boldsymbol{w}) = \boldsymbol{0}$, which a contradiction.

The KKS's argument is powerful but assumes quite a bit of smoothness and regularity. Using our techniques, under much milder smoothness, jointness can be shown to be negative on average when α is strictly supermodular. The reader can verify that under independence of $\boldsymbol{w}$, supermodularity of α^m translates into supermodularity of its conditional expectation, that is

$$
\frac{\partial^2}{\partial t_1 \partial t_2} \mathbb{E}\left[\alpha^m | \boldsymbol{w} \le \boldsymbol{t}\right] \ge 0 \ \ \forall t.
$$

Then, since $\underline{B}(t) = 1$ and $\mathbb{E}[\alpha^m] = 1$,

$$
\frac{\frac{1}{2}\mathbb{E}\left[\alpha^m | \boldsymbol{w} \le (t,t)\right] - 1}{\mathbb{E}\left[\alpha^m | w_i \le t\right] - 1} \ge 1 \ \ \forall t,
$$

thus average jointness is non-positive.

11 Proofs for Section 5

Throughout this section, it is assumed that the first-order approach is valid, $\lambda^{s,*}$ and $\boldsymbol{\lambda}^{m,*}$ satisfy conditions (A1)-(A5) of Proposition 2.

11.1 Social weights on single and married

11.1.1 Proof of Corollary 1

Proof. Part (a). The argument for comparative statics with respect to Pareto weights is identical to the proof of Lemma 5. Specifically, as explained in Chapter 6.E of Shaked and

Shanthikumar (2007), under log-supermodularity of f, the distribution with density $\alpha^a f$ first-order stochastically dominates the distribution with density $\alpha^b f$. Hence,

$$\int_W \mathbf{1}_{[t,\infty)}(w_i)\alpha^a(\mathbf{w}) f(\mathbf{w}) d\mathbf{w} \geq \int_W \mathbf{1}_{[t,\infty)}(w_i)\alpha^b(\mathbf{w}) f(\mathbf{w}) d\mathbf{w},$$

which implies $\mathbb{E}\left[\lambda_i^{m,a,*}|w_i = t\right] \leq \mathbb{E}\left[\lambda_i^{m,a,*}|w_i = t\right]$ for all t.

As shown in the proof of Lemma 6, $\iota^{-1}w_2$ conditionally on $w_2/w_1 \geq \iota$ converges to 0 almost surely as $\iota \to 0$. Since α is continuous, monotone and bounded, it is uniformly continuous, thus

$$\liminf_{\iota \to \infty} \mathbb{E}^a\left[\alpha^{m,a}|w_2/w_1 \geq \iota\right] = \liminf_{\iota \to \infty} \mathbb{E}^a\left[\alpha^{m,a}(\iota^{-1}w_2, w_2)|w_2/w_1 \geq \iota\right] > 1.$$

Part (b). Since $\alpha^{m,a} \sim \alpha^{m,b}$, they satisfy $\alpha^{m,a}(\infty, w_j) = \mathbb{E}^b\left[\alpha^{m,a}\right]\alpha^{m,b}(\infty, w_j)$, where

$$\mathbb{E}^b\left[\alpha^{m,a}\right] = \underbrace{\mathbb{E}^a\left[\alpha^{m,a}\right]}_{=1} + \int \alpha^{m,a}(\mathbf{w})\left(F^a(d\mathbf{w}) - F^b(d\mathbf{w})\right) = \tag{70}$$

$$= 1 + \int \left(\underbrace{F^b(d\mathbf{w}) - F^a(d\mathbf{w})}_{\geq 0}\right)\alpha^{m,a}(d\mathbf{w}).$$

Note that for supermodular $\alpha^{m,a}$, the term $\mathbb{E}^b\left[\alpha^a\right]$ is more than one. Since $F^a \leq_{PQD} F^b$, (64) implies that the distribution $F^a(\cdot|\infty)$ first-order stochastically dominates $F^b(\cdot|\infty)$. As a result, $\mathbb{E}^b\left[\alpha^{m,b}(\infty, w_{-i})|w_i = \infty\right] \leq \mathbb{E}^a\left[\alpha^{m,b}(\infty, w_{-i})|w_i = \infty\right]$ due to monotonicity of $\alpha^{m,b}$. Taking all pieces together, we obtain

$$\lim_{t \to 0} \frac{\mathbb{E}^b\left[\lambda_i^{m,b,*}|w_i = t\right]}{\mathbb{E}^a\left[\lambda_i^{m,b,*}|w_i = t\right]} = \frac{1 - \mathbb{E}^b\left[\alpha^{m,b}(\infty, w_{-i})|w_i = \infty\right]}{1 - \mathbb{E}^b\left[\alpha^{m,a}\right]\mathbb{E}^a\left[\alpha^{m,b}(\infty, w_{-i})|w_i = \infty\right]} \geq 1.$$

We now show that Lemma 7 extends as well. Consider the right tail and note that separability of $\alpha^{m,a}$ was used in the proof of Lemma 7 only when $\overline{\kappa} = \frac{1}{2}$. So (ii) of this Lemma directly applies to non-separable α^m. As for (i), under supermodularity and symmetry of $\alpha^{m,a}$,

$$\alpha^{m,a}(\infty) \geq -\mathbb{E}^a\left[\alpha^{m,a}\right] + 2\mathbb{E}^a\left[\alpha^{m,a}(\infty, w_{-i})\right].$$

Since $\mathbb{E}^a\left[\alpha^{m,a}\right] = 1$ and $\mathbb{E}^a\left[\alpha^{m,a}(\infty, w_{-i})\right] < \mathbb{E}^a\left[\alpha^{m,a}|w_i = \infty\right]$ due to $F^a(\cdot|\infty)$ strictly first-order stochastically dominating G, we can conclude

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

Part (c). The arguments here are analogous to Part (b). $\qquad\square$

11.2 Bargaining and allocation of resources within couples

In this model, the relaxed problem is exactly the same as in benchmark but welfare $\mathcal{W}$ is now given by

$$\mathcal{W} = \frac{\mu}{2}\mathbb{E}\left[\alpha^m v^m\right] + (1-\mu)\mathbb{E}\left[\alpha v^s\right] + \int_\mu^1 \Phi\left(\varepsilon\right)d\varepsilon + \mu\mathbb{E}\left[\left(\alpha(w_i) - \alpha(w_{-i})\right)v^s(w_i)\right],$$

where the last term is the adjustment due to bargaining. It is immediate to verify using the arguments in the proof of Proposition 2 that $\boldsymbol{\lambda}^{m,*}$ are exactly the same as in the benchmark and

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

Under positive dependence, monotonicity of α and (64) jointly imply that the last term is nonnegative.

11.3 Optimality of taxation of family earnings

11.3.1 Proof of Lemma 9

Proof. Part(a). Proposition (3) with $Q = R$ implies that

$$\mathbb{E}\left[\sum_{i=1}^2 \frac{w_i^{1/(1-\gamma)}}{w_1^{1/(1-\gamma)} + w_2^{1/(1-\gamma)}}\lambda_i^{m,*}\,\middle|\,R = r\right] = \frac{1 - \mathbb{E}\left[\alpha^m|R \geq r\right]}{\gamma\theta_r\left(r\right)}, \tag{71}$$

where θ_r is the tail statistics of r. By Proposition (2), the optimal tax is family-earnings based if and only if $\widehat{\lambda}(r) := \frac{1 - \mathbb{E}\left[\alpha^m|R \geq r\right]}{\gamma\theta_r(r)}$ verifies (50).

Solve for (w_1, w_2) as a function of (r, ι) to obtain

$$w_1 = \frac{r}{\left(1 + \iota^{1/(1-\gamma)}\right)^{(1-\gamma)}}, \quad w_2 = \frac{r}{\left(1 + \iota^{1/(\gamma-1)}\right)^{(1-\gamma)}}.$$

It is routine to verify that $d\boldsymbol{w} = \frac{w_1 w_2}{r\iota}drd\iota$, thus f and $\widetilde{f}$, which is the density of (r, ι), are related by $f = \frac{r\iota}{w_1 w_2}\widetilde{f}$. Since $w_1\frac{\partial R}{\partial w_1} + w_2\frac{\partial R}{\partial w_2} = r$ and $w_1\frac{\partial I}{\partial w_1} + w_2\frac{\partial I}{\partial w_2} = 0$,

$$\sum_{i=1}^2 \frac{\partial\left(w_i\widehat{\lambda}(R(\boldsymbol{w}))f(\boldsymbol{w})\right)}{\partial w_i} = \frac{R(\boldsymbol{w})I(\boldsymbol{w})}{w_1 w_2}\frac{\partial\left(r\widehat{\lambda}(r)\widetilde{f}(r,\iota)\right)}{\partial r}. \tag{72}$$

It follows from (72) that $\widehat{\lambda}$ satisfies (50) if and only if

$$\frac{\partial\left(\gamma r\widehat{\lambda}(r)\widetilde{f}(r,\iota)\right)}{\partial r} = \left(\alpha\left(\boldsymbol{w}(r,\iota)\right) - 1\right)\widetilde{f}(r,\iota).$$

Divide this equation by the marginal density of ι and integrate to see the claim in Part (a), that is, family taxes is optimal if and only if $\widetilde{\lambda} = \widehat{\lambda}$.

Part (b). It immediately follows from the argument in Part (a).

Part (c). By $\widetilde{F}^a \leq_{PQD} \widetilde{F}^b$,

$$\mathrm{Pr}^a\left(R \geq r | I \geq \iota\right) \leq \mathrm{Pr}^b\left(R \geq r | I \geq \iota\right) \quad \forall\,(r, \iota). \tag{73}$$

Since α^m is measurable only with respect to r and decreasing in this variable, the first-order stochastic dominance relationship in (73) gives $\mathbb{E}^a\left[\alpha^m | I \geq \iota\right] \geq \mathbb{E}^b\left[\alpha^m | I \geq \iota\right]$. Thus, by (27), $\mathbb{E}^a\left[\lambda_2^{m,a,*} - \lambda_1^{m,a,*} | I = \iota\right] \leq \mathbb{E}^b\left[\lambda_2^{m,b,*} - \lambda_1^{m,b,*} | I = \iota\right]$. $\qquad\square$

11.4 Public goods and economies of scale in marriage

This model is different from the benchmark in two ways. First, the relationship between the optimal distortions, which are still defined in (15), and derivatives of $v^{s,*}$, $v^{m,*}$ now becomes

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

Second, the resource constraint (8) now reads as

$$\mathcal{S} \geq \frac{\mu}{2} \int v^m(\mathbf{w}) F(\boldsymbol{w})/b^m + (1-\mu) \int v^s(w) G(dw)/b^s,$$

where $\mathcal{S}$ is our shorthand notation for the total economic output, i.e.,

$$\mathcal{S} := \frac{\mu}{2} \int \sum_{i=1}^{2} \left(w_i^{1+\gamma} \left(\frac{\partial v^m(\mathbf{w})}{\partial w_i}\right)^\gamma - \gamma w_i \frac{\partial v^m(\mathbf{w})}{\partial w_i}/b^m\right) F(d\mathbf{w}) +$$

$$+ (1-\mu) \int \left(w^{1+\gamma} \left(\frac{\partial v^s(w)}{\partial w}\right)^\gamma - \gamma w \frac{\partial v^s(w)}{\partial w}/b^s\right) G(dw).$$

Then, (6) and the modified resource constraint give

$$\int v^s(w) G(dw) = \frac{\mathcal{S} - \mu\Phi(\mu)/b^m}{\mu/b^m + (1-\mu)/b^s}, \quad \frac{1}{2} \int v^m(\boldsymbol{w}) G(d\boldsymbol{w}) = \frac{\mathcal{S} + (1-\mu)\Phi(\mu)/b^s}{\mu/b^m + (1-\mu)/b^s}.$$

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

$$\mathcal{W} = \frac{\mu}{2} \int (\alpha^m(\boldsymbol{w}) - 1)\, v^m(\boldsymbol{w}) F(d\boldsymbol{w}) + (1-\mu) \int (\alpha(w) - 1)\, v^s(w) G(dw) +$$

$$+ \int_\mu^1 \Phi(\varepsilon) d\varepsilon + \frac{1}{\mu/b^m + (1-\mu)/b^s} \mathcal{S}.$$

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

$$\frac{\partial \left(\gamma w \lambda^{s,*}(w) g(w)\right)}{\partial w} = \frac{\mu^*/b^m + (1-\mu^*)/b^s}{1/b^s} \left(\alpha(w) - 1\right) g(w),$$

$$\sum_{i=1}^{2} \frac{\partial \left(\gamma w_i \lambda_i^{m,*}(\boldsymbol{w}) f(\boldsymbol{w})\right)}{\partial w_i} = \frac{\mu^*/b^m + (1-\mu^*)/b^s}{1/b^m} \left(\alpha^m(\boldsymbol{w}) - 1\right) f(\boldsymbol{w}).$$

Two expressions (37), (38) follow from integrating the first equation and applying the Coarea Formula with $Q = w_i$ as in Proposition 3 to the second equation.

11.5 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 (40) and (42).

Lemma 13. *(a) N^s and N^m are increasing and convex, (b) there exists some $\bar{x} > 0$ such that $\frac{\partial N^s(l)}{\partial l}, \frac{\partial N^m(\boldsymbol{l})}{\partial l_i}\Big|_{l_i=l} \leq l^{p-1} \left(l^p + \bar{x}\right)^{(1-p\gamma)/p\gamma}$ 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) Γ^s and Γ^m are uniformly bounded, (e) $\lim_{l\to\infty} \Gamma^s(l), \Gamma_{ii}^m(\boldsymbol{l})\Big|_{l_i=l} = \gamma$ and $\lim_{l\to\infty} \Gamma_{ij}^m(\boldsymbol{l})\Big|_{l_i=l} = \gamma$, where convergence is uniform in l_j.*

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 for w.r.t. x can be expressed as

$$x^{\sigma + (1-\gamma)/\gamma} = (1-m)^{(1-p\gamma)/p\gamma}, \tag{74}$$

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

$$\frac{\partial N^s(l)}{\partial l} = l^{p-1} \left(x^p + l^p\right)^{(1-p\gamma)/p\gamma}. \tag{75}$$

It is immediate from (75) that N^s is strictly increasing which proves (a). (74) implies that $x \leq 1$ which proves (b). Part (c) follows from the previous observation and (75).

Totally differentiate (74) to obtain

$$\left[\frac{1-\gamma}{\gamma} - \frac{1-p\gamma}{\gamma}m + \sigma\right] d\ln x = -\frac{1-p\gamma}{\gamma}m d\ln l. \tag{76}$$

The term in the square brackets in (76) is bounded from below by $p - 1 + \sigma > 0$, thus the derivative of x is uniformly bounded. Then, totally differentiate (75) 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 Γ^s defined by

$$\Gamma^s(l) = \left[1 + \frac{\partial ln\left(\partial N^s(l)/\partial l\right)}{\partial l}\right]^{-1}$$

is uniformly bounded which proves (d). Finally, since the derivative of x is uniformly bounded and m goes to 1 as $l \to \infty$, Γ^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 (l, x). The first order condition for w.r.t. x_i can be expressed as

$$x_i^{\sigma+(1-\gamma)/\gamma} = 2(1-r_i)^{1-q(1-\sigma)}(1-m_i)^{(1-p\gamma)/p\gamma}, \tag{77}$$

where $m_i := \frac{l_i^p}{x_i^p+l_i^p}$ and $r_i := \frac{x_j^{1/q}}{x_1^{1/q}+x_2^{1/q}}$ for $i = 1, 2$. Again, (77) is sufficient due to convexity. By the Envelope Theorem, the derivative of N^m w.r.t. l_i is given by

$$\frac{\partial N^m(l)}{\partial l_i} = l_i^{p-1}\left(x_i^p + l_i^p\right)^{(1-p\gamma)/p\gamma}, \tag{78}$$

which proves (a). (74) implies that $x \leq 2^{-\sigma-(1-\gamma)/\gamma}$ which proves (b). Part (c) follows from the previous observation and (79).

Totally differentiate (77) to obtain

$$\left[\frac{1-\gamma}{\gamma} - \frac{1-p\gamma}{\gamma}m_i - \frac{1-q(1-\sigma)}{q}r_i + \sigma\right]d\ln x_i = -\frac{1-q(1-\sigma)}{q}r_i d\ln x_j - \frac{1-p\gamma}{\gamma}m_i d\ln l_i. \tag{79}$$

The term in the square brackets in (79) is bounded from below by $p - 1 + \frac{1-q}{q} > 0$, thus the derivative of x is uniformly bounded. Then, totally differentiate (78) to obtain

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

Recall that Γ^m is defined to be

$$\Gamma^m(l) = \left[\begin{array}{cc} \frac{\partial ln\left(\partial N_1^m(l)/\partial l_1\right)}{\partial l_1} + 1 & \frac{\partial ln\left(\partial N_1^m(l)/\partial l_1\right)}{\partial l_2} \\ \frac{\partial ln\left(\partial N_2^m(l)/\partial l_1\right)}{\partial l_1} & \frac{\partial ln\left(\partial N_2^m(l)/\partial l_2\right)}{\partial l_2} + 1 \end{array}\right]^{-1}.$$

By monotonicity and convexity of N^m, the determinatnt of the matrix in the square brackets is at least 1. Then, uniform boundedness of derivatives of x implies that Γ^m is uniformly bounded as well which proves (d). Finally, observe that, by (77), m_i goes to 1 as $l_i \to \infty$ uniformly in l_j, which implies that $\Gamma_{ii}^m(l)$ converges to γ and $\Gamma_{ij}^m(l)$ converges to 0 as $l_i \to \infty$, where convergence is uniform in the other spouse labor supply. This shows (e) and concludes proof. $\square$

Equipped with Lemma 13, we can formally establish the results in Section 41. Here, the notion of welfare is exactly as in the benchmark. To simplify exposition, it is convenient to define 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 12 extends, and the local incentive constraints can be succinctly expressed as

$$w \frac{\partial v^s(w)}{\partial w} = \psi^s \left(\frac{y^s(w)}{w} \right), \quad w_i \frac{\partial v^m(\boldsymbol{w})}{\partial w_i} = \psi_i^m \left(\frac{y_1^m(\boldsymbol{w})}{w_1}, \frac{y_2^m(\boldsymbol{w})}{w_2} \right), \tag{80}$$

which is the exact analogue of (7). 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.,

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

second, the maximal expected economic surplus from singles diverges to $-\infty$ when the expected value of $\int \psi^s \left(\frac{y^s(w)}{w} \right) G(dw)$ goes to $+\infty$. We impose the same assumptions on married. As a result, (a), (b) of Lemma 13 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 13, the local incentive constraints (80) 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

$$\frac{\mu}{2} \int \left(\sum_{i=1}^2 w_i \phi_i^m \left(w_1 \frac{\partial v^m(\boldsymbol{w})}{\partial w_1}, w_2 \frac{\partial v^m(\boldsymbol{w})}{\partial w_2} \right) - N^m \left[\phi^m \left(w_1 \frac{\partial v^m(\boldsymbol{w})}{\partial w_1}, w_2 \frac{\partial v^m(\boldsymbol{w})}{\partial w_2} \right) \right] \right) F(d\mathbf{w}) +$$

$$+ (1 - \mu) \int \left(w \phi^s \left(w \frac{\partial v^s(w)}{\partial w} \right) - N^s \left[\phi^s \left(w \frac{\partial v^s(w)}{\partial w} \right) \right] \right) G(dw). \tag{81}$$

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, (15). It is straightforward to show that they are related to derivatives v^s and v^m by

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

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 13, 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 $(\Gamma^s(l))^{-1} d \ln l = d \ln u$ and $d \ln l = \frac{wu}{l} \frac{\partial \phi^s(wu)}{\partial x} d \ln u$, which gives

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

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:

$$\frac{\partial \left(\Gamma^s \left(\frac{y^{s,*}(w)}{w} \right) w \lambda^{s,*}(w) g(w) \right)}{\partial w} = (\alpha(w) - 1) g(w). \tag{82}$$

This equation is analogous to (49), and the only difference is that here Γ^s is non-constant. Integrating this equation, we obtain the Diamond's ABC formula from Section 5.5, that is,

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

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

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

We now look at married individuals. Again, the first step is to determine derivatives of ϕ^m. Consider the equation $\boldsymbol{l} = \boldsymbol{\phi}^m(\boldsymbol{x})$, where $\boldsymbol{x} = (w_1 u_1, w_2 u_2)$, that defines $\boldsymbol{l}$ as a function of $\boldsymbol{u}$ for fixed $\boldsymbol{w}$. By definition, $l_i \frac{\partial N^m(\boldsymbol{l})}{\partial l_i} = w_i u_i$ for $i = 1, 2$, which gives $(\Gamma^m(\boldsymbol{l}))^{-1} d \ln \boldsymbol{l} = d \ln \boldsymbol{u}$ and

$$d \ln l_i = \frac{w_i u_i}{l_i} \frac{\partial \phi^m(w_1 u_1, w_2 u_2)}{\partial x_i} d \ln u_i + \frac{w_j u_j}{l_i} \frac{\partial \phi^m(w_1 u_1, w_2 u_2)}{\partial x_j} d \ln u_j.$$

Combining these expressions we obtain

$$\frac{\partial \left(w_1 \phi_1^m(w_1 u_1, w_2 u_2) + w_2 \phi_2^m(w_1 u_1, w_2 u_2) - N^m[\boldsymbol{\phi}^m(w_1 u_1, w_2 u_2)] \right)}{\partial u_i} =$$

$$= w_i \Gamma_{ii}^m(\boldsymbol{l}) \left(\frac{l_i}{u_i} - 1 \right) + w_j \frac{u_j}{u_i} \Gamma_{ji}^m(\boldsymbol{l}) \left(\frac{l_i}{u_j} - 1 \right) = w_j \frac{u_j}{u_i} \Gamma_{ij}^m(\boldsymbol{l}) = w_i \Gamma_{ij}^m(\boldsymbol{l}), \quad (83)$$

69

where the second equality is due to the definition of Γ^m and $\frac{l_j \partial N^m(l)/\partial l_j}{l_i \partial N^m(l)/\partial l_i} = \frac{u_j w_j}{u_i w_i}$.

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 (83), the following differential equation is necessary for optimality:

$$\sum_{i=1}^{2} \frac{\partial \left(\Gamma_{ii}^m \left(\frac{y_1^{m,*}(\boldsymbol{w})}{w_1}, \frac{y_2^{m,*}(\boldsymbol{w})}{w_2} \right) w_i \lambda_i^{m,*}(\boldsymbol{w}) f(\boldsymbol{w}) + \Gamma_{ij}^m \left(\frac{y_1^{m,*}(\boldsymbol{w})}{w_1}, \frac{y_2^{m,*}(\boldsymbol{w})}{w_2} \right) w_i \lambda_j^{m,*}(\boldsymbol{w}) f(\boldsymbol{w}) \right)}{\partial w_i} =$$

$$= (\alpha^m(\boldsymbol{w}) - 1) f(\boldsymbol{w}). \quad (84)$$

Integrate (84) using the Coarea Formula with $Q = w_i$ and $Q = \min\{w_1, w_2\}$ to obtain the following conditional moments of optimal distortions:

$$\mathbb{E}\left[\Gamma_{ii}^m \lambda_i^{m,*} + \Gamma_{ij}^m \lambda_j^{m,*} | w_i = t \right] = \frac{1 - \mathbb{E}\left[\alpha^m | w_i \geq t\right]}{\theta(t)},$$

$$\mathbb{E}\left[\Gamma_{ii}^m \lambda_i^{m,*} + \Gamma_{ij}^m \lambda_j^{m,*} | w_i = t \right] = \frac{\Pr(w_j \geq t | w_i \geq t)}{2\Pr(w_j \geq t | w_i = t)} \frac{1 - \mathbb{E}\left[\alpha^m | w_i \geq t\right]}{\theta(t)}.$$

By condition (A2), the optimal distortions are uniformly bounded. Since $\Gamma_{ii}^m(\boldsymbol{l}) \to \gamma$ and $\Gamma_{ij}^m(\boldsymbol{l}) \to 0$ as $l_i \to \infty$ uniformly in l_j and $\frac{y_i^{m,*}(\boldsymbol{w})}{w_i} \to \infty$ as $w_i \to \infty$, we conclude that

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

and average jointness at the top is exactly as in the benchmark.

11.6 Extensive margin

In general, the model with extensive margin is quite complex and cannot be directly studied using variational arguments. To overcome this challenge, we consider a linear relaxation in which individual preferences and output take the following form: $v = c - \left(\gamma \left(\frac{y}{w} \right)^{1/\gamma} + \varrho \right) z$ and yz, respectively, where $z \in [0, 1]$ is an additional choice variable. The original model is subsumed by these functional forms with the additional constraint on z, i.e., it is an indicator function of $y > 0$.

The notion of welfare is exactly as in the benchmark, and the characterization of incentive constraints is similar to the benchmark: $w \frac{\partial v^s(w)}{\partial w} = \left(\frac{y^s(w)}{w} \right)^{1/\gamma} z^s(w)$. Solve for (c^s, y^s) as a function v^s, $\frac{\partial v^s}{\partial w}$ and z^s to obtain that the following expression for tax revenues from singles with type w:

$$y^s(w)z^s(w) - c^s(w) = w^{1+\gamma} \left(\frac{\partial v^s(w)}{\partial w} \right)^\gamma (z^s(w))^{1-\gamma} - \varrho z^s(w) - \gamma w \frac{\partial v^s(w)}{\partial w}. \quad (85)$$

Clearly, for fixed v^s, it is optimal to select z^s that pointwise maximizes (85) as it improves total revenues available for redistribution. The reader can verify that

$$\psi\left[w^{1+\gamma}\left(\frac{\partial v^s(w)}{\partial w}\right)^\gamma\right] = \max_{z^s \in [0,1]} w^{1+\gamma}\left(\frac{\partial v^s(w)}{\partial w}\right)^\gamma (z^s)^{1-\gamma} - \varrho z^s =$$

$$= \begin{cases} \gamma\left(\frac{1-\gamma}{\varrho}\right)^{(1-\gamma)/\gamma} w^{(1+\gamma)/\gamma}\frac{\partial v^s(w)}{\partial w}, & w^{1+\gamma}\left(\frac{\partial v^s(w)}{\partial w}\right)^\gamma \leq \frac{\varrho}{1-\gamma}, \\ w^{1+\gamma}\left(\frac{\partial v^s(w)}{\partial w}\right)^\gamma - \varrho, & w^{1+\gamma}\left(\frac{\partial v^s(w)}{\partial w}\right)^\gamma \geq \frac{\varrho}{1-\gamma}. \end{cases} \quad (86)$$

The exactly same construction applies to married, thus the economic output $\mathcal{S}$ (equation (47) in the benchmark) can be succinctly expressed as follows:

$$\mathcal{S} = \frac{\mu}{2}\int\sum_{i=1}^2\left(\psi\left[w_i^{1+\gamma}\left(\frac{\partial v^m(\boldsymbol{w})}{\partial w_i}\right)^\gamma\right] - \gamma w_i\frac{\partial v^m(\boldsymbol{w})}{\partial w_i}\right)F(d\mathbf{w})+$$

$$+ (1-\mu)\int\left(\psi\left[w^{1+\gamma}\left(\frac{\partial v^s(w)}{\partial w}\right)^\gamma\right] - \gamma w\frac{\partial v^s(w)}{\partial w}\right)G(dw). \quad (87)$$

Following Section 9.1, the relaxed problem is exactly as in 9.1 but now $\mathcal{S}$ is given by (87). The advantage of pre-solving for z^s and $\boldsymbol{z^m}$ 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 for singles and married can be studied in isolation. Observe that ψ is continuously differentiable. Set λ^s to be

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

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 (53), (56), (53), (57), are necessary and sufficient provided that conditions (A1)-(A5) hold.

We claim that the optimum coincides with the benchmark 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 3, the optimal benchmark distortions are $\lambda^\#(w)$ and $\left(\frac{1}{2}\lambda^\#(w_1), \frac{1}{2}\lambda^\#(w_2)\right)$, where $\lambda^\#$ is defined in Proposition 1. Since the first-order approach is valid, that is both conditions of Proposition 1 hold, there are unique thresholds such that welfare gains in (58), (59) at the margin are exactly ϱ, i.e.,

$$\varrho = (1-\gamma)\underline{w}^s\left(\frac{\underline{w}^s}{1+\lambda^\#(\underline{w}^s)}\right)^{\gamma/(1-\gamma)}, \quad \varrho = (1-\gamma)\underline{w}^m\left(\frac{\underline{w}^m}{1+\frac{1}{2}\lambda^\#(\underline{w}^m)}\right)^{\gamma/(1-\gamma)}. \quad (88)$$

Consider v^s so that $\frac{\partial v^s(w)}{\partial w} = 0$ for $w < \underline{w}^s$, which gives $\lambda^s(w) = \left(\frac{1-\gamma}{\varrho}\right)^{(1-\gamma)/\gamma} w^{1/\gamma} - 1$, and $\lambda^s(w) = \lambda^{\#}(w)$ otherwise. Then, λ^s constructed in this way satisfies (A1)-(A5); moreover, since $\lambda^s(w) \leq \lambda^{\#}(w)$ for all w, (53) and (54) hold. Indeed, for any function $\hat{v}^s \in \mathscr{V}^s$,

$$\int \gamma w \lambda^s(w) \frac{\partial \hat{v}^s(w)}{\partial w} G(dw) + \int (\alpha(w) - 1)\hat{v}^s(w) G(dw) =$$

$$= \int \gamma w \left(\underbrace{\lambda^s(w) - \lambda^{\#}(w)}_{\leq 0}\right) \underbrace{\frac{\partial \hat{v}^s(w)}{\partial w}}_{\geq 0} G(dw) \leq 0,$$

which shows (54). By construction, $\left(\lambda^s(w) - \lambda^{\#}(w)\right) \frac{\partial v^s(w)}{\partial w} = 0$ for all w, thus (53) 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 $\lambda_i^m(\boldsymbol{w}) = \left(\frac{1-\gamma}{\varrho}\right)^{(1-\gamma)/\gamma} w_i^{1/\gamma} - 1$, and $\lambda_i^m(\boldsymbol{w}) = \frac{1}{2}\lambda^{\#}(w_i)$ otherwise. These distortions satisfy (a)-(e). Furthermore, the condition for validity of the first-order approach in Proposition 1 implies that $\lambda_i^m(w_i) \leq \frac{1}{2}\lambda^{\#}(w_i)$ for all w_i. As a result, for any function $\hat{v}^m \in \mathscr{V}^m$, potentially non-separable,

$$\int \sum_{i=1}^{2} \gamma w_i \lambda_i^m(\boldsymbol{w}) \frac{\partial \hat{v}_i^m(\boldsymbol{w})}{\partial w_i} G(d\boldsymbol{w}) + \int (\alpha^m(\boldsymbol{w}) - 1)\hat{v}^m(\boldsymbol{w}) G(d\boldsymbol{w}) =$$

$$= \int \sum_{i=1}^{2} \gamma w \left(\underbrace{\lambda_i^m(\boldsymbol{w}) - \frac{1}{2}\lambda^{\#}(w_i)}_{\leq 0}\right) \underbrace{\frac{\partial \hat{v}^m(\boldsymbol{w})}{\partial w_i}}_{\geq 0} F(d\boldsymbol{w}) \leq 0,$$

which shows (57). By construction, $\left(\lambda_i^m(\boldsymbol{w}) - \frac{1}{2}\lambda^{\#}(w_i)\right) \frac{\partial v^m(\boldsymbol{w})}{\partial w_i} = 0$ for all $\boldsymbol{w}$, thus (53) is satisfied as well.

To sum up, we identified the solution of the linear relaxation in which 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 own threshold. Since the optimal labor participation decisions are integral, this mechanism also solves the original model with extensive margin. Furthermore, since $\lambda^{\#}$ is nonnegative due to monotonicity of α, examination of (88) 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.

11.7 Selection into marriage

The main conceptual difference here is that there are two marriage cut-offs, μ_l and μ_h. (6) has to hold for each cutoff individually, that is

$$\Phi(\mu_q) = \frac{1}{2} \int v^m(\boldsymbol{w}) H_q(dw_1) H_q(dw_2) - \int v^s(w) H_q(dw) \quad \text{for} \quad q = l, h. \tag{89}$$

On top of the distributions of productivities are endogenous. According to Bayes rule, they satisfy

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

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 $\int w^{1/(1-\gamma)} H_q(dw) < \infty$ for each signal. Then, it is easy to see that, each marriage rate must be interior and conditions (a), (b) of Lemma 12 hold for each signal, i.e., with $G = H_q$ and $F = H_q^2$.

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 (6) 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 to be Lagrange multipliers equations on (6) and η be the multiplier on (45). 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 do not depend on v^s, v^m, the Lagrangian can be written as follows:

$$\frac{\mu}{2} \int (\alpha^m(\boldsymbol{w}) - 1) v^m(\boldsymbol{w}) F(d\boldsymbol{w}) + (1 - \mu) \int (\alpha(w) - 1) v^s(w) G^s(w) + \mathcal{S} +$$
$$+ \sum_{q=l,h} \delta_q \left(\int v^s(w) H_q(dw) - \frac{1}{2} \int v^m(\boldsymbol{w}) H_q(dw_1) H_q(dw_2) \right),$$

where $\mathcal{S}$ is defined in (47).

Our analysis of the necessary conditions on v^s, v^m in the proof of Proposition 2 goes without changes, and it gives the following analogs of (49), (50):

$$\frac{\partial (\gamma w \lambda^s(w) g^s(w))}{\partial w} = (\alpha(w) - 1) g^s(w) + \frac{\sum_{q=l,h} \delta_q h_q(w)}{1 - \mu}, \tag{90}$$

$$\sum_{i=1}^{2} \frac{\partial (\gamma w_i \lambda_i^m(\boldsymbol{w}) f(\boldsymbol{w}))}{\partial w_i} = (\alpha^m(\boldsymbol{w}) - 1) f(\boldsymbol{w}) - \frac{\sum_{q=l,h} \delta_q h_q(w_1) h_q(w_2)}{\mu}. \tag{91}$$

Differential equations (90) and (91), which are necessary for optimality for fixed marriage rates, imply that the optimal distortions satisfy two equations in the text:

$$\lambda^{s,*}(t) = \frac{1 - \mathbb{E}^s\left[\alpha|w_i \geq t\right]}{\gamma\theta_s\left(t\right)} + \frac{1}{1 - \mu^*}\frac{\delta_h(1 - H_h(t)) + \delta_l(1 - H_l(t))}{\gamma\theta_s\left(t\right)},$$

$$\mathbb{E}[\lambda_i^{m,*}|w_i = t] = \frac{1 - \mathbb{E}^m\left[\alpha^m|w_i \geq t\right]}{\gamma\theta_m\left(t\right)} - \frac{1}{\mu^*}\frac{\delta_h(1 - H_h(t)) + \delta_l(1 - H_l(t))}{\gamma\theta_m\left(t\right)},$$

where we used the Coarea Formula with $Q = w_i$ to obtain the second expression.

11.8   Gender differences

To ensure that the relaxed problem is well-defined, we require $\int w^{1/(1-\gamma)}G_i(dw) < \infty$ for $i = 1, 2$. 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 j is the "deficit" gender, which simply means that $\Delta := \int v_j^s(w_j)G_j(dw_j) - \int v_{-j}^s(w_{-j})G_{-j}(dw_{-j}) \geq 0$. Then, the marriage rate is given by

$$\Phi(\mu) = \frac{1}{2}\int v^m(\boldsymbol{w})F(d\boldsymbol{w}) - \int v_j^s(w_j)G_j(dw_j). \tag{92}$$

The resource constraints reads as

$$\mathcal{S} \geq \frac{\mu}{2}\int v^m(\boldsymbol{w})F(dw) + (1 - \mu)\int v_j^s(w_j)G(dw_j) - (1 - \mu)\frac{\Delta}{2},$$

and $\mathcal{S}$ is defined as in the benchmark (equation (47)) but allowing for differential treatment of single males and single females.

It is easy to see that the modified resource constraint must bind, thus when combined with (92), it can be solved uniquely for expected utilities of married and singles j as a function of $\Phi(\mu)$, $\mathcal{S}$ and Δ. Substituting these expected utilities into the welfare criterion, we obtain

$$\mathcal{W} = \frac{\mu}{2}\int\left(\alpha^m(\boldsymbol{w}) - 1\right)v^m(\boldsymbol{w})F(d\boldsymbol{w}) + \int_\mu^1\Phi(\varepsilon)d\varepsilon + \frac{1 - \mu}{2}\int\left(\alpha_j(w_j) - 1\right)v_j^s(w_j)G_j(dw_j) +$$

$$+ \frac{1 - \mu}{2}\int\left(\alpha_{-j}(w_{-j}) - 1\right)v_{-j}^s(w_{-j})G_{-j}(dw_{-j}) + \mathcal{S}, \tag{93}$$

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.

11.9 Optimal restricted taxation

11.9.1 Proof of Lemma 11

Proof. Part (a). Since taxes are gender-neutral, $v_M^s = v_F^s$. In contrast to Section 11.8, 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. Define the symmetrized versions of distributions F^{nrl} and G^{nrl} by permuting genders at random, that is $F^{nrl}(w_M, w_F) = \frac{F(w_M, w_F) + F(w_F, w_M)}{2}$ and $G^{nrl}(w) := \frac{G_M(w) + G_F(w)}{2}$. Then, define the symmetrized versions of Pareto weights $\alpha^{m,nrl}$ and α^{nrl} by permuting genders at random, that is

$$\alpha^{nrl}(w) = \frac{\alpha_M(w)g_M(w) + \alpha_F(w)g_F(w)}{2g^{nrl}(w)},$$

$$\alpha^{m,nrl}(w_M, w_F) = \frac{\alpha^m(w_M, w_F)f(w_M, w_F) + \alpha^m(w_F, w_M)f(w_F, w_M)}{2f^{nrl}(w_M, w_F)}.$$

It is routine to verify that if v^m is a symmetric function and $v_M^s = v_F^s$, welfare $\mathcal{W}$ defined in (93) can be expressed as

$$\frac{\mu}{2} \int \left(\alpha^{m,nrl}(\boldsymbol{w}) - 1 \right) v^m(\boldsymbol{w}) F^{nrl}(d\boldsymbol{w}) + (1-\mu) \int \left(\alpha^{nrl}(w) - 1 \right) v^s(w) G^{nrl}(dw) + \int_\mu^1 \Phi(\varepsilon) d\varepsilon + \mathcal{S}.$$

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

$$\lambda^{s,nrl,*}(t) = \frac{1 - \mathbb{E}^{nrl}\left[\alpha^{nrl}|w \geq t\right]}{\gamma\theta^{nrl}(t)}, \quad \mathbb{E}\left[\lambda_j^{m,nrl,*}|w_i = t\right] = \frac{1 - \mathbb{E}^{nrl}\left[\frac{\alpha_M(w_M) + \alpha_F(w_F)}{2}|w_j \geq t\right]}{\gamma\theta^{nrl}(t)}.$$

Then, direct verification concludes Part (a) of the lemma.

Part (b). In addition, to gender-neutrality we require taxes to be separable. This simply means that $v^m(w_M, w_F) = \hat{v}^m(w_M) + \hat{v}^m(w_F)$ for some function $\hat{v}$. Recall that $\alpha^m(w_M, w_F) = \frac{\alpha_M(w_M) + \alpha_F(w_F)}{2}$; thus, for additively separable and symmetric v^m we obtain

$$\int \left(\alpha^{m,nrl}(\boldsymbol{w}) - 1 \right) v^m(\boldsymbol{w}) F^{nrl}(d\boldsymbol{w}) =$$

$$= 2 \int \left(\mathbb{E}^{nrl}\left[\frac{\alpha_M(w_M) + \alpha_F(w_F)}{2}|w_j = w \right] - 1 \right) \hat{v}^m(w) G^{nrl}(dw).$$

Clearly, v^m enters the total economic output $\mathcal{S}$ only through $\hat{v}^m$ as $\frac{\partial v^m(w_M, w_F)}{\partial w_j} = \frac{\partial \hat{v}^m(w_j)}{\partial w_j}$. 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

$$\lambda^{m,ind,*}(t) = \frac{1 - \mathbb{E}^{nrl}\left[\frac{\alpha_M(w_M)+\alpha_F(w_F)}{2}|w_j \geq t\right]}{\gamma\theta^{nrl}(t)}.$$

Then, direct verification concludes Part (b) of the lemma.

Part (c). If taxes are family-earnings based, then v^m is measurable only w.r.t. to r, i.e., $v^m(\boldsymbol{w}) = \hat{v}^m(R(\boldsymbol{w}))$ for some $\hat{v}$, hence

$$\int (\alpha^m(\boldsymbol{w}) - 1)\, v^m(\boldsymbol{w})F(d\boldsymbol{w}) = \int (\mathbb{E}\left[\alpha^m|R = r\right] - 1)\, \hat{v}^m(r)\widetilde{F}(dr).$$

By construction, $\mathbb{E}\left[\alpha^m|R = r\right]$ and $\widetilde{F}$ are symmetric, so there is no need to further symmetrize these objects. In order to make our previous argument aplicable, we need to show that v^m enters the total economic output $\mathcal{S}$ only through $\hat{v}^m$. Indeed, since $w_M\frac{\partial R}{\partial w_M} + w_F\frac{\partial R}{\partial w_F} = r$,

$$\sum_{j=1}^{2}\left(w_j^{1+\gamma}\left(\frac{\partial v^m(\mathbf{w})}{\partial w_j}\right)^\gamma - \gamma w_j\frac{\partial v^m(\mathbf{w})}{\partial w_j}\right) = r^{1+\gamma}\left(\frac{\partial \hat{v}^m(r)}{\partial r}\right)^\gamma - \gamma r\frac{\partial \hat{v}^m(r)}{\partial r}.$$

Similarly to Part (b), 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

$$\lambda^{m,fam,*}(r) = \frac{1 - \mathbb{E}\left[\alpha^m|R \geq r\right]}{\gamma\theta_r(r)}.$$

Finally, note that the first equality in this part of the lemma is (71) can be derived using the Coarea formula with $Q = R$. $\qquad\square$

12   Quantitative analysis

12.1   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

$$\max_{(y_1, y_2) \geq \mathbf{0}} \nu \, (y_1 + y_2)^{1-\tau} - \sum_{i=1}^{2} \gamma \left(\frac{y_i}{w_i}\right)^{1/\gamma},$$

which allows us to express unobserved productivites as a function of observed earnings (equation (43)) and construct the empirical distribution of productivities.

We calibrate a 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 (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

$$\mathbb{E}\left[w_i\right] = \frac{a}{a-1} e^{\mu + \sigma^2/2},$$

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:

$$\Pr\left((w_1 - \widetilde{w}_1)(w_2 - \widetilde{w}_2) > 0\right) - \mathbb{P}\left((w_1 - \widetilde{w}_1)(w_2 - \widetilde{w}_2) < 0\right),$$

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 $\rho = 0.33$. In Section 12.4 of the appendix we re-calibrate the model to the FGM copula, i.e., $C(u_1, u_2) = u_1 u_2 (1 + \rho(1 - u_1)(1 - u_2))$, for which Kendell's tau is given by $\frac{2\rho}{9}$, thus $\rho = 0.96$.

12.2   Numerical approach

In this section, we overview the numerical approach that we used to find the optimal taxes. First of all, we 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 $\{w^1, \ldots, w^{400}\}$, 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

$$
\max_{v, \boldsymbol{y} \geq \boldsymbol{0}} \sum_{n_1, n_2 = 1}^{400} v\left(w^{n_1}, w^{n_2}\right) \left(\alpha^m\left(w^{n_1}, w^{n_2}\right) - 1\right) f\left(w^{n_1}, w^{n_2}\right) +
$$

$$
+ \sum_{i=1}^{2} \sum_{n_1, n_2 = 1}^{400} \left(y_i\left(w^{n_1}, w^{n_2}\right) - \gamma \left(\frac{y_i\left(w^{n_1}, w^{n_2}\right)}{w^{n_i}} \right)^{1/\gamma} \right) f\left(w^{n_1}, w^{n_2}\right)
$$

subject to the following set of incentive constraints: for all $n_i = 2, \ldots, 400$, $n_{-i} = 1, \ldots, 400$ and $i = 1, 2$,

$$
v\left(w^{n_i}, w^{n-i}\right) \geq v\left(w^{n_i-1}, w^{n-i}\right) + \gamma y_i\left(w^{n_i-1}, w^{n-i}\right) \left(\left(w^{n_i-1}\right)^{-1/\gamma} - \left(w^{n_i}\right)^{-1/\gamma} \right).
$$

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 $\sum_{n_1, n_2 = 1}^{400} \alpha^m\left(w^{n_1}, w^{n_2}\right) f\left(w^{n_1}, w^{n_2}\right) = 1$.

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.

12.3 Comparison of distortions

Figure 6 plots the optimal distortions $\lambda_i^* (\cdot, w_{-i})$ when w_{-i} is fixed at its median value, i.e., 0.66, against the average distortions $\mathbb{E}[\lambda_i^* | w_i = t]$. For both the Gaussian and FGM copulas, the average distortions (red dashed lines) are very close to the optimal distortion with $w_{-i} = 0.66$ (solid blue lines).
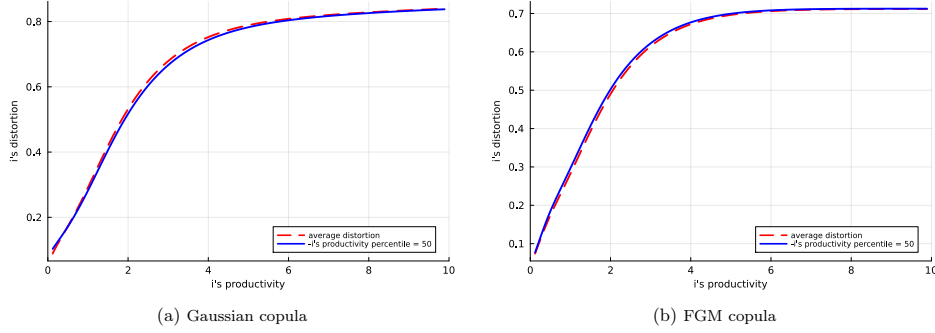

(a) Gaussian copula   (b) FGM copula

Figure 6: Comparison of average and pointwise distortions at 50th percentile

12.4 FGM copula

In this section, we report the optimal taxes when the empirical distribution of productivities is calibrated to the FGM copula. Recall that the FGM copula is defined to be $C(u_1, u_2) = u_1 u_2 (1 + \rho(1 - u_1)(1 - u_2))$, where $\rho \in [-1, 1]$. We calibrate parameter ρ to be 0.96, matching the Kendell's tau statistic. Figure 7 shows the FGM copula fits isoquants of the empirical distribution fairly well (Panel (b)), but it does not match the speed of convergence to tail-independence (Panels (c) and (d)).
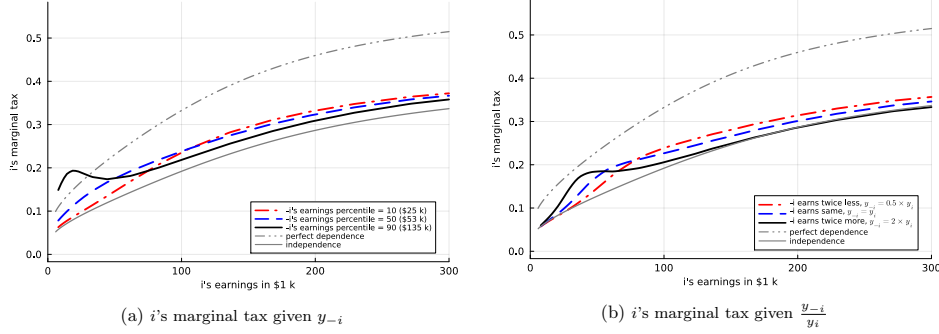
(a) i's marginal tax given y_{-i}

(b) i's marginal tax given $\frac{y_{-i}}{y_i}$

Figure 8: Optimal taxes, $m = 0.35$ and $k = 1$

(a) Marginal distributions, $G(t)$

(b) Loci of copulas quantiles (u_1, u_2) satisfying $C(u_1, u_2) = q$ for $q \in \{0.1, ..., 0.9\}$

(c) Left-tail statistics $\frac{C(u,u)}{u}$ (lines with markers) and $\frac{\ln u}{\ln C(u,u)}$ (lines without markers)

(d) Right-tail statistics $\frac{\overline{C}(u,u)}{u}$ (lines with markers) and $\frac{\ln u}{\ln \overline{C}(u,u)}$ (lines without markers)
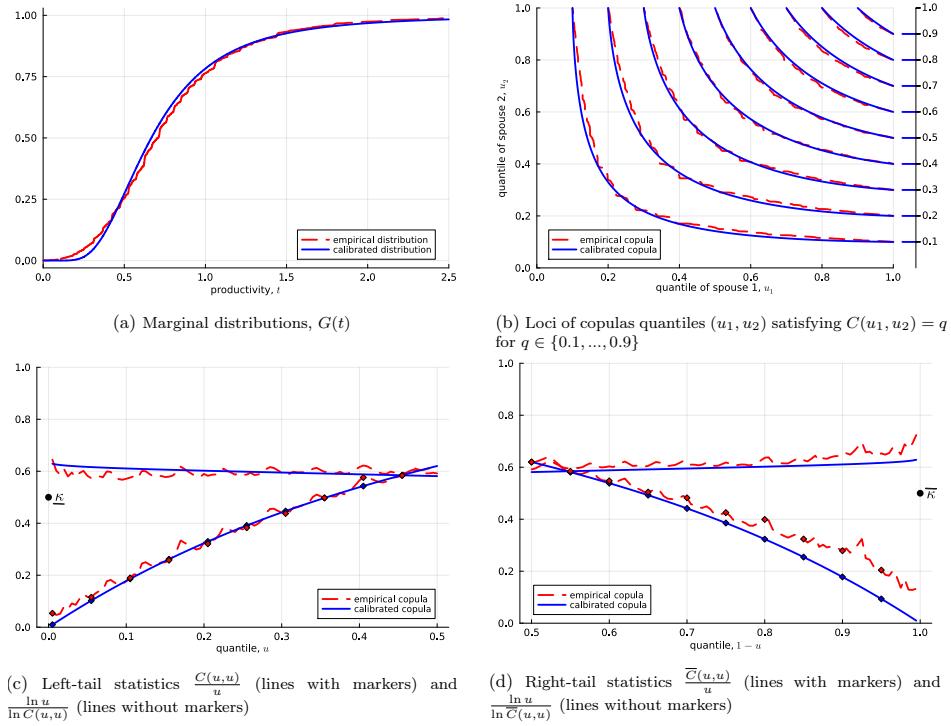
Figure 7: Empirical and calibrated joint distributions of productivities

Figure 8 depicts the optimal taxes computed under the FGM copula in our baseline specification with separable weights and not too redistributive planner. The optimal taxes are much more negatively jointed at the top as compared to the optimal taxes computed under the Gaussian copula (Figure 3). Consistent with Proposition **??** the sign of jointness flips at a threshold, and it is negative (positive) for all large (small) w_i.

Figure 9 illustrates robustness of the optimal taxes. Qualitatively, redistributiveness and modularity of α have identical implications for the optimal taxes under the FGM copula and the Gaussian copula (Figure 4). First, if the planner is more redistrubutive, then the optimal
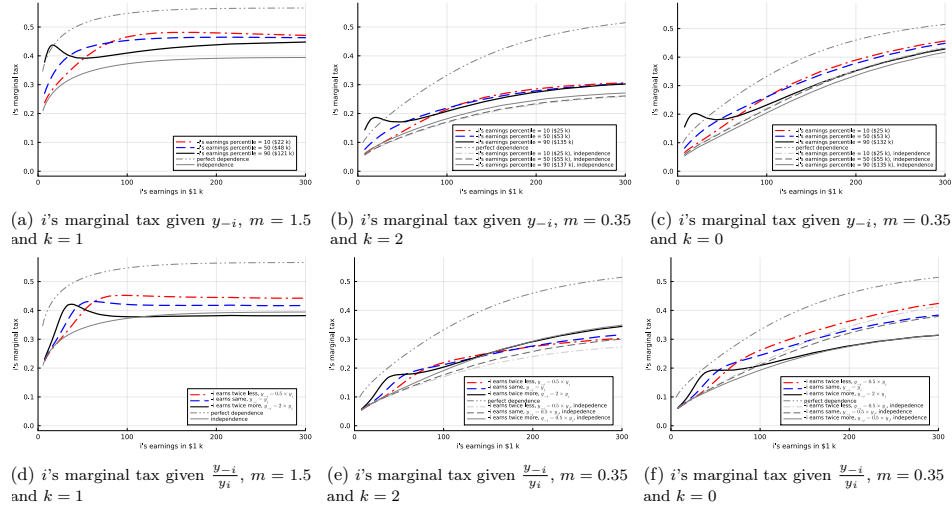
80

(a) i's marginal tax given y_{-i}, $m = 1.5$ and $k = 1$

(b) i's marginal tax given y_{-i}, $m = 0.35$ and $k = 2$

(c) i's marginal tax given y_{-i}, $m = 0.35$ and $k = 0$

(d) i's marginal tax given $\frac{y_{-i}}{y_i}$, $m = 1.5$ and $k = 1$

(e) i's marginal tax given $\frac{y_{-i}}{y_i}$, $m = 0.35$ and $k = 2$

(f) i's marginal tax given $\frac{y_{-i}}{y_i}$, $m = 0.35$ and $k = 0$

Figure 9: Optimal taxes, robustness to m and k

taxes are larger and negative jointness occurs at lower levels. Second, supermodularity amplifies negative jointness at the top, and submodularity amplifies positive jointness at the bottom.

Figure 10 plots the optimal taxes as a function of the total family earnings and the share of the secondary earner. As in the case of the Gaussian copula (Figure 5), the optimal taxes vary substantially with the share of the secondary earner.

(a) U.S. taxes implied by the estimated HSV model

(b) Optimum, $m = 0.35$ and $k = 1$

(c) Optimum, $m = 1.5$ and $k = 1$
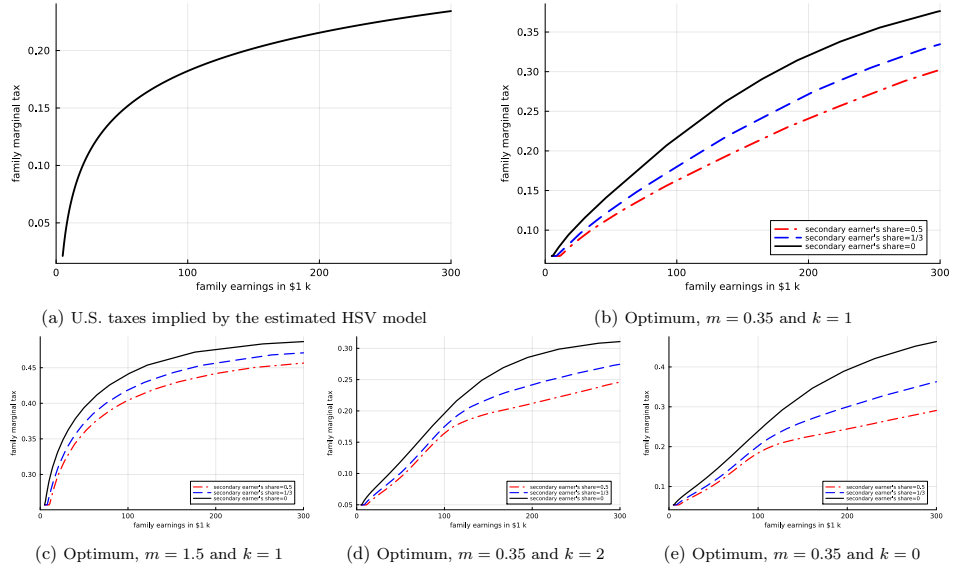
(d) Optimum, $m = 0.35$ and $k = 2$

(e) Optimum, $m = 0.35$ and $k = 0$

Figure 10: Marginal taxes on family earnings