ECONtribute
MARKETS & PUBLIC POLICY

ECONtribute
Discussion Paper No. 349

Pareto-Improvements, Welfare Trade-Offs and the Taxation of Couples

Felix J. Bierbrauer
Andreas Peichl
Pierre C. Boyer
Daniel Weishaar

December 2024

www.econtribute.de

UNIVERSITÄT BONN
UNIVERSITÄT ZU KÖLN

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1-390838866 is gratefully acknowledged.

Pareto-improvements, welfare trade-offs and the taxation of couples*

Felix J. Bierbrauer   Pierre C. Boyer   Andreas Peichl   Daniel Weishaar

December 5, 2024

Abstract

We develop a theory of tax reforms for a setting with multi-dimensional heterogeneity amongst taxpayers and multiple economic decisions that are all subject to fixed and variable costs. The theorems in this paper provide a complete characterization of the conditions under which Pareto- or welfare-improving tax reforms exist. We focus on one application, the taxation of couples, and present a detailed analysis of the behavioral responses to taxation in this setting. Squaring the theorems with this analysis yields sufficient statistics for the existence of Pareto- or welfare-improving tax reforms. In the empirical part, we apply them to US data. Our findings include the following: Tax rates on secondary earnings are inefficiently high when secondary earnings are close to primary earnings. Also, reducing the tax system's degree of jointness is not Pareto-improving. Whether it raises welfare depends on a trade-off between poverty alleviation and gender balance.

Keywords: Taxation of couples, Pareto efficiency, tax reforms, optimal taxation, non-linear income taxation.
JEL classification: C72; D72; D82; H21.

***Bierbrauer**: CMR - Center for Macroeconomic Research, University of Cologne, Germany (bierbrauer@wiso.uni-koeln.de); **Boyer**: CREST, École Polytechnique, Institut polytechnique de Paris, France (pierre.boyer@polytechnique.edu); **Peichl**: ifo Munich, LMU Munich, CESifo, IHS and IZA (peichl@econ.lmu.de); **Weishaar**: LMU Munich (daniel.weishaar@econ.lmu.de). We are extremely grateful to Dan Feenberg for granting us access to the NBER TAXSIM model and providing feedback on our simulations. This paper uses TAXSIM v32 in combination with the Current Population Survey (CPS) data. The authors gratefully acknowledge support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC2126/1 – 390838866 as well as via CRC TRR 190 (project number 280092119) and the Investissements d'Avenir (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047), ANR-19-CE41-0011-01 and ANR-20-CE41-0013-01. Weishaar gratefully acknowledges support from the Joachim Herz Foundation and the Fritz Thyssen Foundation.

1 Introduction

This paper has a pure theory part, an applied theory part, and an empirical part. In the pure theory part, we develop a new approach for the identification of Pareto-improving and welfare-improving tax reforms in a setting with the following features: There is a tax function that potentially has many arguments. Moreover, there is multidimensional heterogeneity amongst taxpayers and every economic decision may give rise to behavioral responses both at the intensive and the extensive margin. Under the given tax system, there may be discontinuities in marginal tax rates. For all these reasons, the optimal choices of taxpayers may not satisfy first-order conditions and behavioral responses may therefore be discontinuous.

We focus on one particular application, the taxation of couples. We take as given a status quo tax system, but we do not make a priori assumptions on its functional form. We also assume that, for each spouse, "making money" comes with both fixed and variable effort costs. Depending on these costs and the properties of the tax system, couples sort themselves into the set of dual-earner couples, the set of single-earner couples, or the set of couples with no earnings. This sorting changes when the tax function is changed. In addition, earnings choices change. In the applied theory part of the paper, we combine a detailed analysis of behavioral responses in this setting with our general theory insights. This approach yields sufficient statistics that can be used to identify Pareto- or welfare-improving tax reforms in the data.

We use these sufficient statistics for an evaluation of the tax treatment of couples in the US. Our analysis is motivated by recurrent debates about reforms of the tax and transfer system with the objective to achieve more gender balance. In particular, there is the concern that current systems excessively discourage the labour supply of secondary earners, most of whom are women, see e.g. Bick and Fuchs-Schündeln (2017). Our analysis (i) identifies the parts of the income distribution with inefficiently high tax rates on secondary earnings, (ii) it shows which welfare functions approve and which ones reject reforms that lower the tax system's degree of jointness and/ or strengthen workfare elements at the bottom of the income distribution. We also present a calibration of optimal marginal tax rates for primary and secondary earners at the top of the income distribution.

Pure Theory. The paper's theory part is based on the following setting: There is a continuum of couples. A couple consists of two spouses with earnings of y_1 and y_2. Couples are confronted with a tax function $(y_1, y_2) \mapsto T(y_1, y_2)$. This tax function may be inherited from the past or it may be an object of theoretical interest such as an optimal tax function.[1] Couples choose y_1 and y_2 optimally given the tax system. Specifically, we assume that the spouses maximize their joint surplus, i.e. the difference between their disposable income and their effort costs.[2]

Methodologically, our key innovation is the use of *conditional revenue functions.* Consider the following thought experiment. Focus on all couples who have primary earnings in a narrow bracket B_1 that starts at an income level y_1. Then focus on the subset of these couples who have secondary earnings in a narrow bracket B_2 that starts at income level y_2. For these couples, consider a small increase of the marginal tax rate on secondary earnings and denote the reform's marginal effect on tax revenue by $\mathcal{R}_2(y_2 \mid y_1)$. The conditional revenue function $y_2 \mapsto \mathcal{R}_2(y_2 \mid y_1)$ documents how this revenue effect varies as we vary the position of B_2 in the range of incomes while holding B_1 fixed. To see how such functions can help to identify inefficiencies in the tax system, suppose that the status quo is a system of progressive and joint taxation, as in France, Germany and the US. Consequently, the larger is y_1, the larger is, for every level of y_2, the marginal tax rate on secondary earnings. For very high y_1, this may discourage secondary earners so much that lower marginal taxes would be actually be self-financing. In this case $\mathcal{R}_2(y_2 \mid y_1)$ takes a negative value which is analogous to tax rates on secondary earnings being above the top of the "Laffer curve", except that the Laffer curve here allows for non-linear taxes and uses primary earnings y_1 as a conditioning variable. A reform that lowers the marginal tax rate on secondary earnings for couples with primary earnings close to y_1 and secondary earnings close to y_2 would then be Pareto-improving. If the status quo has progressive and joint taxation, then the conditioning on y_1 has the potential to make an important difference. For lower values of y_1 marginal tax rates

[1] Stylized alternatives are individual and joint taxation. Under the former, the marginal tax rate on the earnings of one spouse does not depend on the earnings of the other spouse. Under the latter, and with a progressive tax system, the marginal tax rate on the earnings of one spouse is increasing in the earnings of the other spouse.

[2] The maximization of the joint surplus is, under some ancillary conditions, consistent with the spouses engaging in Nash bargaining, see Appendix A.10.

are lower. Hence, there may not be an inefficiency.

The theory part of this paper contains three theorems. They all invoke conditional revenue functions for dual-earner and single-earner couples. Theorems 1 and 2 state, respectively, necessary and sufficient conditions for the Pareto-efficiency of a tax system. Theorem 3 characterizes the welfare implications of tax reforms. It can be used to analyze the extent to which alternative social welfare functions give rise to conflicting views about desirable directions for reform. In particular, Corollary 3 shows that a reduction of marginal tax rates for secondary earners is rejected by a Rawlsian welfare measure if the starting point is a tax system that satisfies the sufficient conditions for Pareto-efficiency in Theorem 2. Moreover, Theorem 3 lends itself to a characterization of welfare-optimal tax systems (Corollary 2). It can also be used to analyze the welfare implications of reforms that alter the tax system's degree of jointness (Corollary 4).

Applied Theory. The general analysis that leads to these Theorems is complemented by a detailed analysis of a couple's behavioral responses to taxation. This analysis leads to a sufficient statistics characterization of the conditional revenue functions for dual- and single-earner couples. With this characterization we can turn to a more detailed analysis of the conditions under which changes of tax rates are desirable. We find conditions for the desirability of negative marginal tax rates that are unprecedented by the previous literature on optimal taxation: When welfare weights are continuous and decreasing in income, then, at the bottom of the income distribution, negative marginal tax rates are desirable even when extensive margin elasticities are zero. We also state conditions under which an optimal tax system is piecewise linear and separable at the top of the income distribution. Thus, under these conditions, small changes in the earnings of one spouse should not affect the marginal tax rate on the earnings of the other spouse.

Empirical Analysis. The empirical part of this paper uses these insights to (i) identify Pareto-improving reforms of the current US tax system, (ii) trace out the welfare implications of changes in the US tax system's degree of jointness, (iii) discuss reforms of the tax-and transfer system that affect low-income couples, and (iv) characterize optimal marginal tax rates for primary and secondary earners at the top of the income distribution.

For this purpose, we combine data on the income distribution within and across married couples from the Current Population Survey (CPS) with information on marginal tax rates from the NBER TAXSIM microsimulation model. In addition, we use tax return data to estimate the properties of the primary and secondary earnings distribution at the top.

We present four main empirical findings. First, for the recent past, the taxation of couples is associated with inefficiently high marginal tax rates for secondary earners. Identifying these inefficiencies requires the use of conditional revenue functions. If primary earnings were averaged out and an unconditional revenue function was used, many of these inefficiencies would remain undetected. The policy implication is that, to reap those efficiency gains, the reduction of marginal tax rates for secondary earners needs to depend also on the level of primary earnings.

Second, reducing the tax system's degree of jointness is not a Pareto-improving reform. It comes with a loss of tax revenue and is therefore rejected by a Rawlsian welfare function. It is approved, however, by a welfare function with weights that are increasing in the women's income share in the couple's total income. Thus, judging the tax system's degree of jointness requires to take a stance on the what the relevant welfare objective looks like. There is a trade-off between poverty alleviation and gender balance.

Third, we identify ranges of incomes at the bottom of the income distribution where reforms towards lower marginal tax rates are "controversial" in the sense that they are approved by welfare functions with continuous weights – among them welfare functions that, say, concentrate weights on the bottom ten percent of the income distribution – while they are rejected by a Rawlsian welfare function.

Finally, there is a substantial gap in optimal top tax rates for primary and secondary earners, even if the elasticities that are capturing the behavioral responses to taxation are the same and the welfare weights for all these people at the top of the income distribution are identical. The differences are driven by the tails of the income distributions. The distribution of primary earnings has a fatter tail so that optimal top tax rates are around 12 percentage points higher for primary earners. In addition, we find that the top tax rates for primary earners in single-earner couples should be larger than for primary earners in dual-earner couples.

Related Literature. There is a rich literature that studies the optimal taxation of couples. Its starting point is the seminal paper by Boskin and Sheshinski (1983) which applied an inverse elasticities logic to the taxation of couples: When secondary earners show behavioral responses to taxation that are stronger than those of primary earners, then it is optimal to tax secondary earnings at a lower rate.[3] The subsequent literature has branched out in numerous ways.[4] Our approach differs in that we analyze reform directions in a neighborhood of a status quo tax system. While we explore some implications of our approach for optimal taxes, we are primarily interested in identifying Pareto- or welfare-improving reforms of an actual tax system.

Our formal analysis is based on a model with multi-dimensional heterogeneity amongst couples. Couples differ in the productive abilities of the primary and the secondary earner, in their fixed costs of labor market participation and, possibly, in their weights in the couple's internal bargaining procedure.[5] This framework is richer than what has previously been considered in papers that approach the optimal taxation of couples as a problem of multi-dimensional screening. Kleven, Kreiner, and Saez (2009) focus on a setting in which a primary earner only makes intensive margin choices and a secondary earner only makes an extensive margin choice. Golosov and Krasikov (2023) focus on spouses who both only make intensive margin choices.[6] Both Kleven, Kreiner, and Saez (2009) and Golosov and Krasikov (2023) present results on optimal jointness, i.e. on the optimal interdependence of marginal tax rates on primary and secondary earnings. We complement these findings with an analysis of reforms that alter a given tax system's degree of jointness. In particular, if the status quo has progressive and joint taxation, then a jointness-reducing reform can be interpreted as bringing the status quo closer to individual taxation. We identify the conditions under which such reforms can be Pareto-improving and the

[3]For a related discussion of gender-based taxation, see Alesina, Ichino, and Karabarbounis (2011).

[4]Non-linear taxes have been considered, labour supply responses at the intensive and the extensive margin have been taken into account, marital status has been treated as endogenous, see e.g. Schroyen (2003), Brett (2009), Kleven, Kreiner, and Saez (2009), Immervoll et al. (2011), Cremer, Lozachmeur, and Pestieau (2012), Gayle and Shephard (2019), Malkov (2021), Alves et al. (2024), Ales and Sleet (2022), and, most recently, Golosov and Krasikov (2023).

[5]See Appendix A.10 for a discussion of Nash bargaining in couples.

[6]In an extension, they show that fixed costs of working are without consequence for optimal taxes when there is random matching on the marriage market.

conditions under which such reforms come with welfare trade-offs.

We use perturbation techniques to identify reform directions that are efficiency-enhancing, or welfare-improving, starting from a given status quo. The perturbation approach is frequently used in analyses of income taxes.[7] In this paper, we extend this approach in a substantial way so that we can explicitly deal with the interdependence of primary and secondary earnings that arises through joint taxation. Problems of optimal multi-dimensional screening are known to give rise to complex patterns of bunching.[8] Our theoretical approach allows for the possibility of bunching and for the possibility of non-continuous behavioral responses to changes in marginal tax rates. We remain agnostic, however, whether these patterns arise because the status quo is an optimal tax system or because the status quo has discontinuities in marginal tax rates for other reasons. Our approach takes account, however, of the possibility that bunching regions or regions with discontinuous behavioral responses change their size in response to a tax perturbation. If the status quo is an optimal tax system, then no such perturbation can have a positive effect on welfare.

Our empirical analysis employs the NBER TAXSIM microsimulation model and CPS micro data.[9] The microsimulation model uses rich data on individual characteristics so that we can elicit, at the level of an individual tax unit, the marginal tax rates and tax liability under the status quo and any alternative tax system. The evaluation of tax reforms rests on empirical estimates of the behavioral responses to taxation. A large literature exists estimating relevant elasticities and our assumptions on behavioral responses are informed by this literature.[10]

Our analysis yields predictions on how tax reforms would affect earnings incentives of the spouses in a couple, aggregate measures of welfare and tax revenue. A

[7]References include Piketty (1997), Saez (2001), Golosov, Tsyvinski, and Werquin (2014), Saez and Stantcheva (2016), Sachs, Tsyvinski, and Werquin (2020), Bierbrauer, Boyer, and Peichl (2021), Jacquet and Lehmann (2021b), Jacquet and Lehmann (2021a), Bergstrom and Dodds (2021), Bierbrauer, Boyer, and Hansen (2023), Ferey, Lockwood, and Taubinsky (2024) and Spiritus et al. (forthcoming).

[8]For recent progress on the characterization of optimal bunching patterns, see Boerma, Tsyvinski, and Zimin (2022), Carlier et al. (2024), see Rochet (2024) for a review of the literature.

[9]Our empirical approach builds on and extends work by Eissa, Kleven, and Kreiner (2008), Bargain et al. (2015) and Bierbrauer, Boyer, and Peichl (2021).

[10]See, e.g., Blundell and Macurdy (1999), Blau and Kahn (2007), Eissa and Hoynes (2004), LaLumia (2008), Bargain, Orsini, and Peichl (2014).

complementary literature in macroeconomics embeds the joint labour supply decisions of couples into quantitative dynamic models. It then traces out the tax and transfer system's implications for the labor market outcomes and the savings decisions of men and women or tax revenue.[11]

In a companion paper, Bierbrauer et al. (2024), we analyze the taxation of couples from a political economy perspective, with a focus on how the support for reforms towards individual taxation has evolved since the 1960s. The paper also contains an analysis of past reforms that affected the distribution of marriage bonuses. It is related to this paper in that it also uses revenue functions for an analysis of tax reforms (as have other papers). There is, however, an important methodological difference: The companion paper is tailored to a status quo with joint taxation and it does not use conditional revenue functions (which are the key innovation from this paper). The analysis in this paper does not restrict the set of tax functions. Therefore, its analysis of Pareto- and welfare-improving reforms is complete in the sense of identifying necessary and sufficient conditions for the existence of such reforms.[12]

Outline. Section 2 contains the theoretical analysis, and Section 3 the empirical results. The last section contains concluding remarks. Formal proofs and supplementary analyses are relegated to the Appendix.

2 Theory

The Status Quo. There is a status quo tax system for married couples $(y_1, y_2) \mapsto T^0(y_1, y_2)$ where y_1 are the earnings of spouse 1 and y_2 are the earnings of spouse 2.[13] To save on notation, we often write $y = (y_1, y_2)$. We assume that T^0 is a continuous function and write $T^0_{y_1}(y)$ and $T^0_{y_2}(y)$ for the marginal tax rates on the earnings of spouse 1 and spouse 2 at y. We write $T^0_{y_1,y_2}(y)$ and $T^0_{y_2,y_1}(y)$ for the tax

[11] See, e.g., Guner, Kaygusuz, and Ventura (2012), Bick and Fuchs-Schündeln (2017), Borella, De Nardi, and Yang (2023) or Holter, Krüger, and Stepancuk (2023).

[12] It also transcends our earlier discussion paper, Bierbrauer et al. (2023) which only covered some sufficient conditions when using unconditional revenue functions.

[13] For now the indices 1 and 2 are labels with no further meaning. In empirical applications, we will assign the index 1 to the spouse with higher earnings in the status quo.

function's cross derivatives.[14]

There is a status quo distribution of incomes. In the status quo y takes values in $\mathcal{Y}_0 = \mathcal{Y}_{0,1} \times \mathcal{Y}_{0,2}$, where, for $i = 1, 2$, $\mathcal{Y}_i = (0, \bar{y}_i]$ is the range of strictly positive earnings and $Y_{0,i} = [0, \bar{y}_i]$ includes earnings of zero. We denote by $F_y : \mathcal{Y}_0 \to [0, 1]$ the *cdf* the represents the joint distribution of incomes in the status quo, and by f_y its density. We use obvious notation for the conditional and marginal distributions that can be derived from F_y. We interpret the status quo distribution of incomes as resulting from the utility-maximizing behavior of couples. Specifically, we assume that the spouses in a married couple solve the following problem: Choose y to maximize $C^0(y) - K(y, \theta)$, where the couple's disposable income is given by

$$C^0(y) \quad = \quad b_m^0 \; + \; y_1 + y_2 \; - \; T^0(y) \; ,$$

and its costs of productive effort are given by

$$K(y, \theta) \quad = \quad k_1(y_1, \omega_1) \; + \; \varphi_1 \, \mathbf{1}(y_1 > 0) \; + \; k_2(y_2, \omega_2) \; + \; \varphi_2 \, \mathbf{1}(y_2 > 0) \; .$$

We assume that $T^0(0, 0) = 0$. Hence, b_m^0 is the intercept of the consumption schedule that couples are facing in the status quo, or, equivalently, the transfer to a couple with no earnings. For each spouse i, the generation of earnings comes with fixed and variable costs of productive effort. We assume that the variable costs are captured by the function k_i which is increasing in the first argument, $k_{i,1} > 0$, satisfies the usual Inada conditions and, moreover, is such that $k_{i,12} < 0$, so that the marginal effort costs decrease in ω_i. We refer to ω_i as a measure of spouse i's productive ability. There also is a fixed cost for each spouse i that we denote by φ_i. The indicator function $\mathbf{1}(y_i > 0)$ takes the value of 1 if $y_i > 0$ and equals zero otherwise. A couple is characterized by a pair $\omega = (\omega_1, \omega_2)$ of productive abilities and a pair of fixed costs $\varphi = (\varphi_1, \varphi_2)$. We will sometimes write for short $\theta = (\omega, \varphi)$ with $\theta_1 = (\omega_1, \varphi_1)$ and $\theta_2 = (\omega_2, \varphi_2)$. The joint distribution of θ_1 and θ_2, denoted by F_θ, is a primitive of the economy. We assume that it is well-behaved in that the marginal and conditional *cdf*s that can be derived from F_θ are differentiable.[15]

[14] If the status quo has *joint taxation* there is a function $\tilde{T}^0$ so that $T^0(y) = \tilde{T}^0(y_1 + y_2)$ for all y. Then, $T^0_{y_1}(y) = T^0_{y_2}(y)$, for all (y_1, y_2).

[15] We discuss alternative modeling choices regarding the behavior of couples and explain the extent to which our analysis is robust to such choices in Appendix A.10.

A Tax Reform. A tax reform replaces the status quo tax system T^0 by a new one T^1. We write $T^1(y) = T^0(y) + \tau\, h(y)$ and refer to the function h as the reform direction and to the scalar τ as the size of the reform. We will focus on marginal effects, i.e. on the effect of raising τ slightly, starting from the status quo, $\tau = 0$. We denote by $y^*(\theta, \tau, h) = (y_1^*(\theta, \tau, h), y_2^*(\theta, \tau, h))$ the utility-maximizing choices of a couple that is confronted with the post-reform budget constraint

$$C^1(y) = b_m^1 \; + \; y_1 + y_2 \; - \; T^0(y) \; - \tau\, h(y) \;,$$

where b_m^1 is the post-reform intercept of the consumption function. We assume that tax revenue is rebated lump sum so that $b_m^1 = b_m^0 + R(\tau, h)$, where $R(\tau, h)$ is the reform-induced change in overall tax revenue,

$$R(\tau, h) \;\; = \;\; \mathbb{E}_\theta \left[T^1(y^*(\theta, \tau, h)) \right] - \mathbb{E}_\theta \left[T^0(y^0(\theta)) \right] \;,$$

and the operator $\mathbb{E}_\theta$ indicates that expectations are take with respect to the distribution F_θ. Finally, $y^0(\theta) := y^*(\theta, 0, h)$ is a shorthand for earnings in the status quo. We write $R_\tau(0, h)$ for a reform's marginal effect on tax revenue. It is an endogenous object that depends inter alia on the behavioral responses to taxation.[16] The spouses' behavioral responses to the tax reform are captured by

$$y_{1,\tau}^0(\theta, h) \;\; := \;\; \frac{\partial}{\partial \tau}\, y_1^*(\theta, \tau, h)_{|\tau=0} \quad \text{and} \quad y_{2,\tau}^0(\theta, h) \;\; := \;\; \frac{\partial}{\partial \tau}\, y_2^*(\theta, \tau, h)_{|\tau=0} \;,$$

whenever these derivatives exist.[17]

Individual and Social Welfare. Using the indirect utility function

$$v(\theta, \tau, h) \;\; := \;\; \max_{y_1, y_2} \quad y_1 + y_2 \; - \; T^0(y) \; - \tau\, h(y) - K(y, \theta) \;,$$

[16]For various reform directions of interest, including the set of all continuous functions h, $R_\tau(0, h)$ will be fully characterized in the course of our analysis.

[17]In the presence of fixed costs, the functions y_1^* and y_2^* jump at "cutoff levels" of φ_1 and φ_2 with the implication that left-sided and right-sided derivatives do not coincide. Discontinuities in marginal tax rates may also imply that left-sided and the right-sided derivative do not coincide. We explicitly deal with these complications below.

we can express the reform induced change in indirect utility for a couple with characteristics θ as $V(\theta, \tau, h) = R(\tau, h) + v(\theta, \tau, h) - v(\theta, 0, h)$. By the envelope theorem, see Milgrom and Segal (2002), the marginal effect on indirect utility, associated with direction h, is given by $V_\tau(\theta, 0, h) = R_\tau(0, h) - h(y^0(\theta))$. A reform in direction h is said to be *Pareto-improving* if, for all θ, $R_\tau(0, h) - h(y^0(\theta)) \geq 0$, with a strict inequality for some θ. Given a specification of welfare weights $\tilde{g} : \theta \mapsto \tilde{g}(\theta)$ with $\mathbb{E}_\theta[\tilde{g}(\theta)] = 1$, a direction h is said to be *welfare-improving* if

$$W_\tau(0, h) \quad := \quad R_\tau(0, h) - \mathbb{E}_\theta\left[\tilde{g}(\theta) h(y^0(\theta))\right] \quad \geq \quad 0\,.$$

2.1 Conditional Revenue and Welfare Functions

Theorems 1 and 2 below state necessary and sufficient conditions for the existence of a Pareto-improving reform direction. Theorem 3 characterizes the revenue and welfare implications of continuous reform directions. These theorems invoke what we refer to as *conditional revenue functions*. Here, as a preliminary step, we define these objects. Conditional revenue functions are derived from a specific class of simple tax reforms. They raise the marginal tax rates on the earnings of spouse i conditional on y_i and y_{-i} belonging to particular brackets of income. For instance, a reform that raises the marginal tax rate on the earnings of spouse 2, conditional on the earnings of spouse 1 belonging to a bracket $B_1(y_{1s}, \ell_1) = [y_{1s}, y_{1s} + \ell_1]$ and the earnings of spouse 2 belonging to a bracket $B_2(y_{2s}, \ell_2) = [y_{2s}, y_{2s} + \ell_2]$ can be represented by a scalar τ and a function $y \mapsto h(y)$ such that

$$h(y) \quad = \quad \begin{cases} y_2 - y_{2s}, & \text{if} \quad y \in B_1(y_{1s}, \ell_1) \times B_2(y_{2s}, \ell_2)\ , \\ \ell_2, & \text{if} \quad y_1 \in B_1(y_{1s}, \ell_1) \quad \text{and} \quad y_2 \geq y_{2s} + \ell_2\ , \\ 0, & \quad\ \text{otherwise}\ . \end{cases}$$

Consequently, the marginal tax rates on the earnings of spouse 2 change when the earnings of the spouses belong to the relevant brackets and do not change otherwise,

$$T^1_{y_2}(y) - T^0_{y_2}(y) = \begin{cases} \tau, & \text{if} \quad y \in B_1(y_{1s}, \ell_1) \times B_2(y_{2s}, \ell_2)\ , \\ 0, & \quad \text{otherwise}\ . \end{cases}$$

Simple reforms so that B_1 and B_2 are brackets with positive length affect the earnings incentives of dual-earner couples. We use different notation for simple reforms that affect the earnings incentives of single-earner couples. For instance, a pair (τ, h^{s1}), with

$$h^{s1}(y) \quad = \quad \begin{cases} y_1 - y_{1s}, & \text{if} \quad y_1 \in B_1(y_{1s}, \ell_1) \quad \text{and} \quad y_2 = 0 \ , \\ \ell_1, & \text{if} \quad y_1 \geq y_{1s} + \ell_1 \quad \text{and} \quad y_2 = 0 \ , \\ 0, & \quad \text{otherwise} \ , \end{cases}$$

raises marginal tax rates on the earnings of spouse 1 conditional on $y \in B_1(y_{1s}, \ell_1)$ and spouse 2 having no earnings, $y_2 = 0$.

For $\tau >> 0$, a simple tax reform creates dominated regions in the space of primary and secondary earnings, i.e. regions where no couple chooses its earnings. Consider Figure 1, a simple tax reform generates a cliff between the purple rectangle with $h(y) = \ell_2$ and the neighboring rectangles with $h(y) = 0$. There also is a discontinuity in marginal tax rates when transiting from the blue rectangle with $h(y) = 0$ to the red rectangle with $h(y) = y_s - y_{2s}$. Typically, there will be bunching at such kink points. By Proposition 5 in part A.5 of the Appendix, this creates no complications for the analysis of such a reform's implications for tax revenue when τ is small. In particular, the impact of behavioral responses on tax revenue is then entirely due to those couples who, in the status quo, choose their earnings y in the region with $h(y) \neq 0$.

Conditional Revenue Functions for Dual-Earner Couples. When we seek to emphasize the parameters that define a simple tax reform we write $R_i(y_{1s}, \ell_1, y_{2s}, \ell_2)$ for its marginal impact on tax revenue, rather than using the more concise notation $R_\tau(0, h)$. The subscript $i \in \{1, 2\}$ indicates whether the marginal tax rates for spouse 1 or spouse 2 are affected by the reform. Of interest is also the derivative of $R_i(y_{1s}, \ell_1, y_{2s}, \ell_2)$ with respect to ℓ_i, evaluated at $\ell_i = 0$, see the graph on the left of Figure 4 for an illustration. We define the shorthand

$$\mathcal{R}_i(y_{is} \mid B_{-i}(y_{-is}, \ell_{-i})) \quad := \quad \frac{\partial}{\partial \ell_i} \, R_i(y_{1s}, \ell_1, y_{2s}, \ell_2)_{|\ell_i = 0} \ .$$

Finally, we define

$$\mathcal{R}_i(y_{is} \mid y_{-is}) \quad := \quad \frac{\partial}{\partial \ell_{-i}} \, \mathcal{R}_i(y_{is} \mid B_{-i}(y_{-is}, \ell_{-i}))_{|\ell_{-i}=0} \; .$$

Henceforth, we refer to $y_i \mapsto \mathcal{R}_i(y_i \mid B_{-i})$ and to $y_i \mapsto \mathcal{R}_i(y_{is} \mid y_{-is})$ as *conditional revenue functions*; $\mathcal{R}_i(y_{is} \mid B_{-i})$ gives the revenue effect from a slight increase of the marginal tax rates that applies to a spouse i with earnings in a narrow bracket starting at the income level y_{is}, conditional on spouse $-i$ having earnings in a (possibly wide) bracket B_{-i}; $\mathcal{R}_i(y_{is} \mid y_{-is})$, by contrast, conditions on spouse $-i$ having earnings in a narrow bracket that starts at income level y_{-is}.[18]

Conditional Revenue Functions for Single Earner Couples. For a simple tax reform that affects single-earner couples with, say, spouse 1 as the single earner, we write $R^s_1(y_{1s}, \ell_1)$ rather than $R_\tau(0, h)$. Of interest is also the derivative of $R^s_1(y_{1s}, \ell_1)$ with respect to ℓ_1, evaluated at $\ell_1 = 0$, , see the graph on the right of Figure 4 for an illustration. We define the shorthand

$$\mathcal{R}^s_1(y_{1s}) \quad := \quad \frac{\partial}{\partial \ell_1} \, R^s_1(y_{1s}, \ell_1)_{|\ell_1=0} \; .$$

Conditional Welfare Functions. Conditional welfare functions give the welfare rather than the revenue implications of simple reforms that affect marginal tax rates over narrow ranges of income. The conditional welfare functions for dual earner couples are formally defined by

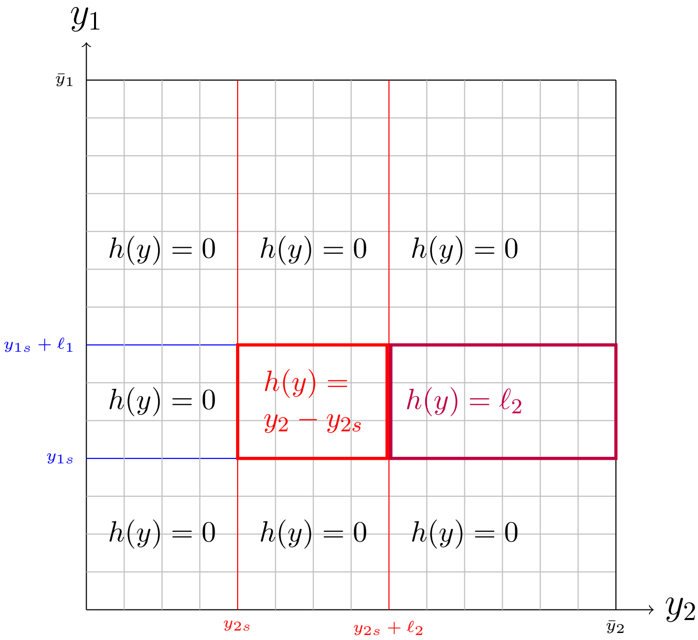
$$\mathcal{W}_i(y_{is} \mid y_{-is}) \;:=\; \mathcal{R}_i(y_{is} \mid y_{-is}) - \frac{\partial}{\partial \ell_{-i}} \left(\frac{\partial}{\partial \ell_i} \, \mathbb{E}_\theta[\tilde{g}(\theta) h(\theta)]_{|\ell_i=0} \right)_{|\ell_{-i}=0} \; .$$

One can show that

$$\mathcal{W}_i(y_{is} \mid y_{-is}) \;:=\; \mathcal{R}_i(y_{is} \mid y_{-is}) - (1 - F_{y_i}(y_{is} \mid y_{-is}))\mathcal{G}(y_{is} \mid y_{-is}) f_{y_{-i}}(y_{-is}) \; .$$

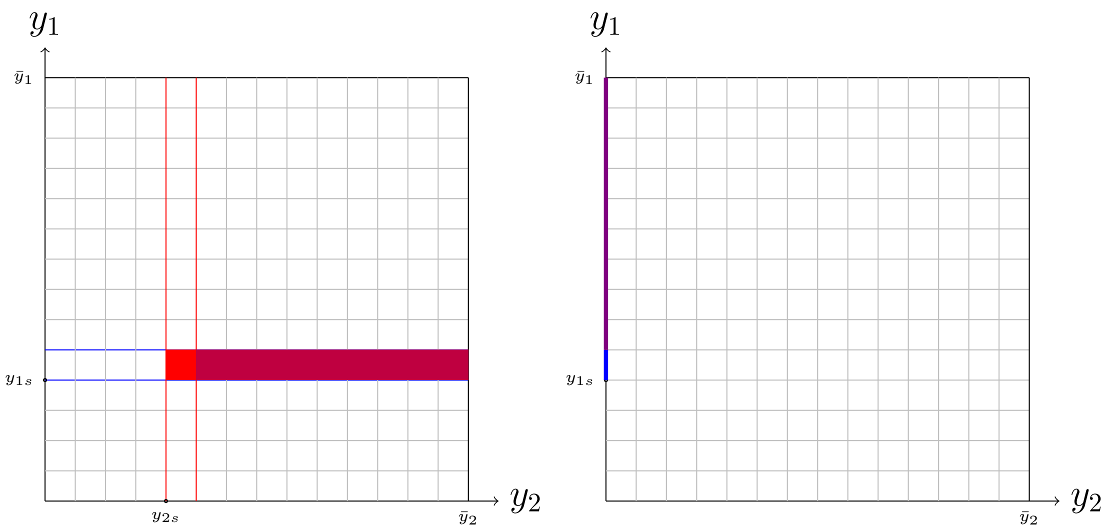
[18]More formally, note that for τ close to zero, a Taylor approximation implies $R(\tau, h) \;\simeq\; \tau \, R_i(y_{1s}, \ell_1, y_{2s}, \ell_2)$. A repeated application of this argument implies that for τ, ℓ_i and ℓ_{-i} close to zero $R(\tau, h) \;\simeq\; \tau \, \ell_i \, \ell_{-i} \, \mathcal{R}_i(y_{is} \mid y_{-is})$.

Figure 1: A simple tax reform



Note: For couples with y in the red rectangle, the marginal tax rate that applies to y_2 is increased by τ. For couples with y in the purple rectangle, the tax burden increases by ℓ_2, while there is no change in marginal tax rates. For τ and ℓ_2 "small", $\tau\,\ell_2\mathcal{R}_2(y_{2s} \mid B_1[y_{1s}, \ell_1])$ is the marginal effect on tax revenue that results when, for couples with y in the red area, the marginal tax rate that applies to y_2 is increased by τ, with the consequence that, for couples with y in the purple area, the tax burden is increased by ℓ_2.

Figure 2: Conditional revenue functions



(a) Dual-earner couples

(b) Single-earner couples

Note: Panel (a): For τ, ℓ_2 and ℓ_1 "small", $\tau\,\ell_2\,\ell_1\,\mathcal{R}_2(y_{2s} \mid y_{1s})$ is the marginal effect on tax revenue that results when, for couples with y in the red area, the marginal tax rate that applies to y_2 is increased by τ, with the consequence that, for couples with y in the purple area, the tax burden is increased by ℓ_2 .

Panel (b): For τ, and ℓ_1 "small", $\tau\,\ell_1\,\mathcal{R}_1^s(y_{1s})$ is the marginal effect on tax revenue that results when, for couples with y on the blue line segment, the marginal tax rate that applies to y_1 is increased by τ, with the consequence that, for couples with y on the violett line segment, the tax burden is increased by ℓ_1.

where $\mathcal{G}(y_{is} \mid y_{-is}) := \mathbb{E}_y(g(s_1, s_2) \mid s_i \geq y_{is} \text{ and } s_{-i} = y_{-is})$ is the average welfare weight of couples where spouse $-i$ has earnings of y_{-is} and spouse i has earnings exceeding y_{is}, and $g(y_1, y_2) := \mathbb{E}_\theta[\tilde{g}(\theta) \mid (y_1^0(\theta), y_2^0(\theta)) = (y_1, y_2)]$ is the average welfare weight of couples with incomes of y_1 and y_2 in the status quo. The obvious adjustments for the case of single earner couples yield conditional welfare functions of the form $\mathcal{W}_i^s(y_{is}) := \mathcal{R}_1^s(y_{1s}) - (1 - F_{y_i}(y_{is} \mid 0))\mathcal{G}(y_{is} \mid 0)F_{y_j}(0)$.

2.2 Theorems

Pareto-efficiency. Let $\mathcal{H}$ by the set of functions $h : y \mapsto h(y)$ that take at least two different values on $y^0(\Theta)$. A tax system is Pareto-efficient only if, for any reform direction $h \in \mathcal{H}$,

$$R_\tau(0, h) - \max_\theta h(y^0(\theta)) \leq 0 . \tag{1}$$

If this condition was violated for some $h \in \mathcal{H}$, there would by a reform (τ, h) that makes some couples strictly better off without making others worse off. Theorem 1 states necessary conditions for the Pareto-efficiency of a tax system. Its proof exploits that, under a Pareto-efficient tax system, condition (1) has to hold for all h in the previously introduced class of simple tax reforms.

Theorem 1 *If a tax system is Pareto-efficient then: For all i and j and for all B_j, the conditional revenue functions $y_i \mapsto \mathcal{R}_i(y_i \mid B_j)$ and $y_i \mapsto \mathcal{R}_i^s(y_i)$ are (i) non-increasing, (ii) bounded from below by 0, and (iii) bounded from above by 1.*

Theorem 1 extends a result from Bierbrauer, Boyer, and Hansen (2023) to a tax function with two arguments. Pareto-efficiency implies that conditional revenue functions have to be non-increasing. If this condition is violated one can Pareto-improve the tax system by means of a self-financing tax cut that applies to all incomes that lie in the income range over which the conditional revenue function is increasing. Pareto-efficiency also implies that conditional revenue functions have to be bounded from below by zero. This is the analogue to the famous Laffer-condition. If there is y so that $\mathcal{R}_i(y_i \mid y_j) < 0$, then lowering the marginal tax rate for spouse i by means of a simple tax reform is self-financing. Finally, there is the "reverse Laffer condition": If $\mathcal{R}_i(y_i \mid y_j)$ gets too large, then increasing the marginal tax rate for spouse i by means of a simple tax reform generates so much additional revenue that even those who now face higher taxes are better off. This is incompatible with Pareto-efficiency in the status quo. Upon letting the length of the conditioning income range vanish, we obtain a Corollary to Theorem 1 that is without precedence in earlier work. As will become clear shortly, the Corollary enables us to transition from the necessary conditions for Pareto-efficiency in Theorem 1 to the sufficient conditions in Theorem 2.

Corollary 1 *If a tax system is Pareto-efficient, then, for all i and j and for all $y \in \mathcal{Y}$, the functions $y_i \mapsto \mathcal{R}_i(y_i \mid y_j)$ are (i) non-increasing, (ii) bounded form below by* 0*, and (iii) bounded from above in the following sense: There is $\bar{\ell}$, so that $\ell_j < \bar{\ell}$ implies*

$$\ell_j \, \mathcal{R}_i(y_i \mid y_j) \quad \leq \quad 1 \, . \tag{2}$$

Theorem 2 *Under the conditions listed in Corollary 1, if h is continuous, then*

$$R_\tau(0, h) \; - \; \max_{y \in \mathcal{Y}} \; h(y) \quad \leq \quad 0 \, . \tag{3}$$

Theorem 2 shows that the conditions in Corollary 1 are sufficient for the local Pareto-efficiency of a tax system: If these conditions are met there is no Pareto-improving direction in the set of continuous functions. Together Corollary 1 and Theorem 2 imply that all conceivable violations of Pareto-efficiency can be found by checking whether the conditional revenue functions have the properties listed in Corollary 1.

The proof of the Theorem proceeds as follows: We take an arbitrary continuous function h as given. We show that any such function can be approximated arbitrarily well by a sum of simple tax reforms. Specifically, we consider functions h_m that are piecewise linear over a partition of $\mathcal{Y}_0$ into m^2 squares and which converge to h as $m \to \infty$. We show that, under the conditions in Corollary 1, there is no function h_m that yields a Pareto-improvement and then extend the argument to the function h. This requires to verify that, for any m, (i) there is no Pareto-improving direction in the set of reforms affecting marginal tax rates only for primary earners, (ii) that there neither is a Pareto-improving direction in the set of reforms affecting marginal tax rates only for secondary earners, and (iii) finally, that there is no Pareto-improving direction that combines changes of marginal tax rates for primary earners and changes of marginal tax rates for secondary earners. The proofs of Theorem 2 and also of Theorem 3 below make intense use of the linearity of Gateaux differentials which implies that the revenue implications of h_m can be written as a linear combination of conditional revenue functions.[19]

Social Welfare. The following theorem provides a characterization of an arbitrary reform direction h's marginal impact on tax revenue and social welfare. Its significance lies in two observations. First, this impact is linear in the changes of marginal tax rates – denoted by h_1 for the earnings of spouse 1 and by h_2 for the earnings of spouse 2 – that are implied by h. Second, the impact associated with a change of marginal tax rates is weighted by the relevant conditional revenue and welfare functions. Thus, the revenue and welfare implications of continuous reforms directions can be written as linear combinations of the revenue and welfare implications of simple reforms.

Theorem 3 *For any continuous reform direction $y \mapsto h(y)$, the marginal effects on tax revenue and social welfare are respectively given by*

$$\begin{aligned} R_\tau(0,h) \;=\; & \int_{\mathcal{Y}} h_1(y)\, \mathcal{R}_1(y_1 \mid y_2)\, dy \;+\; \int_{\mathcal{Y}_1} h_1(y_1,0)\, \mathcal{R}_1^s(y_1)\, dy_1 \\ & + \int_{\mathcal{Y}} h_2(y)\, \mathcal{R}_2(y_2 \mid y_1)\, dy \;+\; \int_{\mathcal{Y}_2} h_2(0,y_2)\, \mathcal{R}_2^s(y_2)\, dy_2 \;, \end{aligned} \tag{4}$$

and

$$\begin{aligned} W_\tau(0,h) \;=\; & \int_{\mathcal{Y}} h_1(y)\, \mathcal{W}_1(y_1 \mid y_2)\, dy \;+\; \int_{\mathcal{Y}_1} h_1(y_1,0)\, \mathcal{W}_1^s(y_1)\, dy_1 \\ & + \int_{\mathcal{Y}} h_2(y)\, \mathcal{W}_2(y_2 \mid y_1)\, dy \;+\; \int_{\mathcal{Y}_2} h_2(0,y_2)\, \mathcal{W}_2^s(y_2)\, dy_2 \;. \end{aligned} \tag{5}$$

The observation that welfare implications are linear in the changes of marginal tax rates has an immediate consequence for welfare-maxima. A welfare maximum has the property that $W_\tau(0,h) = 0$, holds for all h. Given the characterization of $W_\tau(0,h)$ in Theorem 3, this is possible if and only if all conditional welfare functions

[19] Assumptions which guarantee the linearity of the Gateaux differential in the direction h are discussed in Bierbrauer, Boyer, and Hansen (2023).

are equal to zero, everywhere.

Corollary 2 *The following statements are equivalent:*

1. *For all continuous functions* $h : \mathcal{Y}_0 \to \mathbb{R}$, $W_\tau(0, h) = 0$.

2. *For all* $y \in \mathcal{Y}$, $\mathcal{W}_1(y_1 \mid y_2) = \mathcal{W}_2(y_2 \mid y_1) = \mathcal{W}_1^s(y_1) = \mathcal{W}_2^s(y_2) = 0$.

Below we will draw on Corollary 2 for a characterization of optimal top tax rates both on primary and secondary earnings. The following Corollary illustrates the possibility to use Theorem 3 also for an assessment of reforms starting from a possibly suboptimal status quo. Specifically, it shows that the lowering of marginal tax rates on secondary earnings reduces Rawlsian measures of social welfare when the starting point is an efficient tax system.

Corollary 3 *Consider a status quo so that* y *is in the support of* F_y *only if* $y_2 \leq y_1$. *Consider a subset of incomes* $\mathcal{Y}_0'$ *and tax reform so that* $h_2(y) < 0$ *if* $y \in \mathcal{Y}_0' \subset \mathcal{Y}_0$ *and* $h_2(y) = 0$, *otherwise. Let* $h_1(y) = 0$, *for all* $y \in \mathcal{Y}_0$. *Suppose that* $g(y) \leq \bar{g}^d$, *whenever* $y_1 > 0$ *or* $y_2 > 0$. *Suppose that there is no Pareto-improving direction for reform. If* $\bar{g}^d$ *is sufficiently small, then* $W_\tau(0, h) < 0$.

The Corollary is based on the premise that $y_2 \leq y_1$. Thus, the spouse with index 2 is the secondary earner. It then considers a reform so that the marginal tax rates in some set of secondary earnings are reduced while all other marginal tax rates remain unchanged. It shows that this reform is welfare-damaging when two conditions are fulfilled: First, the status quo is an efficient tax system in the sense that the conditions in Theorem 2 are satisfied. Second, the welfare weights of couples with positive earnings are sufficiently low. The Corollary follows from Theorem 3 upon noting that $\mathcal{R}_2(y_2 \mid y_1) > 0$, for all y, when the status quo is an efficient tax system. Thus, for $\bar{g}^d$ close to zero, $\mathcal{W}_2(y_2 \mid y_1) > 0$. Hence, all else equal, a lowering of marginal tax rates for (a subset of) secondary earners, $h_2(y) < 0$, yields a loss of social welfare.

2.3 Using the Theorems to Analyze Changes of Jointness

The term jointness refers to the interdependence of the spouses marginal tax rates. Under a system of progressive and joint taxation, there is positive jointness: the marginal tax rate on the earnings of spouse i is a non-decreasing function of the earnings of spouse $-i$. Under individual taxation, by contrast, there is zero jointness. In the following, we discuss reforms that reduce the tax system's degree of jointness. If the status quo has progressive and joint taxation, then a jointness-reducing reform can be interpreted as bringing the status quo closer to individual taxation.

A reform in direction h alters the tax system's jointness when the cross-derivative h_{12} differs from zero over parts of the income distribution,

$$T^1_{y_1,y_2}(y) \quad = \quad T^0_{y_1,y_2}(y) \ + \ \tau\, h_{12}(y) \ .$$

The following Corollary to Theorem 3 gives the welfare implications associated with a reform in direction h via the implied changes of *jointness.*

Corollary 4 *For any continuous reform direction $y \mapsto h(y)$,*

$$\begin{aligned} W_\tau(0,h) \quad = \quad & \textstyle\int_{\mathcal{Y}} h_{12}(y)\, \mathbf{W}(y) dy \\ & + \textstyle\int_{\mathcal{Y}_1} h_1(y_1,0) \left(\mathcal{W}^s_1(y_1) + \int_{\mathcal{Y}_2} \mathcal{W}_1(y_1 \mid y_2) dy_2 \right) dy_1 \\ & + \textstyle\int_{\mathcal{Y}_2} h_2(0,y_2) \left(\mathcal{W}^s_2(y_2) + \int_{\mathcal{Y}_1} \mathcal{W}_2(y_2 \mid y_1) dy_1 \right) dy_2 \ , \end{aligned}$$

where

$$\mathbf{W}(y) \quad := \quad \int_{y_1}^{\bar{y}_1} \mathcal{W}_2(y_2 \mid y_1) dy_1 \ + \ \int_{y_2}^{\bar{y}_2} \mathcal{W}_1(y_1 \mid y_2) dy_2 \ .$$

An Example. Consider two brackets $B_1(y_{1s}, \ell_1)$ and $B_2(y_{2s}, \ell_2)$ with $y_{1s}, y_{2s} > 0$ and a reform in direction h so that $h(y_1,y_2) \ = \ h^1(y_1)\, h^2(y_2)$ where, for $i = 1, 2$,

$$h^i(y_i) \quad = \quad \begin{cases} 0 & \text{if} \quad y_i \leq y_{is} \ , \\ y_i - y_{is}, & \text{if} \quad y_i \in B_i(y_{is}, \ell_i) \ , \\ \ell_i, & \text{if} \quad y_i \geq y_{is} + \ell_i \ . \end{cases}$$

Note that h is the product of two continuous functions and hence also continuous. Also, note that for $y_1 \in B_1(y_{1s}, \ell_1)$ and $y_2 \in B_2(y_{2s}, \ell_2)$, $h_{12}(y) = 1$. Finally, note that the reform does not involve changes of marginal tax rates for single-earner couples. Therefore, by Corollary 4, the welfare implications of this reform are given by

$$W_\tau(0,h) = \int\limits_{B_1 \times B_2} \mathbf{W}(y,\lambda)\, dy \; . \tag{6}$$

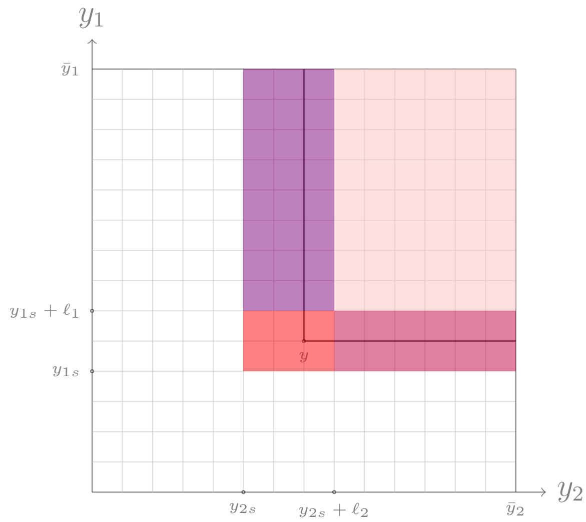
Equation (6) has implications that are noteworthy: First, suppose the status quo is efficient so that all revenue functions are strictly positive. Also suppose that the welfare measure is Rawlsian so that, for all i, and y

$$\mathcal{W}_i(y_i \mid y_j) \quad = \quad \mathcal{R}(y_i \mid y_j) \quad > \quad 0 \; .$$

Then, $\mathbf{W}(y,\lambda) > 0$ for all y and hence $W_\tau(0,h) > 0$. A reduction of jointness ($\tau < 0$) then yields a loss of tax revenue and of Rawlsian welfare. This observation echoes Corollary 2 above. If the status quo is efficient, a move towards individual taxation via a reduction of jointness is rejected by a Rawlsian welfare measure. Second, suppose that a reduction of jointness is instead Pareto-improving. Then, it has to be the case that $W_\tau(0,h) < 0$ for all possible welfare measures, including the Rawlsian measure. Consequently, given a Rawlsian welfare measure, there must be y, so that $\mathbf{W}(y,\lambda) < 0$. But this is possible only if there are i and j so that $\mathcal{R}(y_i \mid y_j) < 0$; i.e. if one of the sufficient conditions in Theorem 2 is violated. More specifically, if a jointness-reducing reform is Pareto-improving, then there is also a simple reform – lowering of the marginal tax rates on the earnings of one spouse only – that is Pareto-improving. Conversely, if the scope for Pareto-improvements via simple reforms has been exhausted, then there is also no possibility to Pareto-improve the tax system via a modification of jointness. To see where all this is coming from, consider Figure 3. The figure illustrates that a reduction of jointness implies a reduction of marginal tax rates in different parts of the income space. The question whether a reduction of jointness is desirable therefore amounts to asking whether these reductions of marginal tax rates are desirable.

Figure 3 is also helpful to clarify the relation between Theorem 3 and Corollary 4. Theorem 3 traces welfare implications back to changes in marginal tax rates. Corollary 4 traces them back to changes in jointness. To describe the connection, we take a closer look at $\mathbf{W}(y)$. Referring to Figure 3, what this expression does is the following: It takes an arbitrary point in the red rectangle and then sums

Figure 3: A change in jointness



Note: The figure illustrates the implications that the reform in the example above has for jointness and marginal tax rates: Jointness changes only in the red rectangle. There, $h_{12}(y) = 1$, and $h_{12}(y) = 0$ elsewhere. Moreover, in the red area the marginal tax rate on the earnings of spouse 1 and also the marginal tax rate on the earnings of spouse 2 change since $h_1(y) = y_2 - y_{2s}$ and $h_2(y) = y_1 - y_{1s}$. In the violet area, the marginal tax rate on the earnings of spouse 1 does not change, $h_1(y) = 0$, but the marginal tax rates on the earnings of spouse 2 changes since $h_2(y) = \ell_1$. Analogously, in the purple area, there is no change in the marginal tax rate for spouse 2, but a change in the marginal rate of spouse 1, $h_1(y) = \ell_2$. In the pink area, there is no change in marginal tax rates, but the total tax burden changes since $h(y) = \ell_1 \ell_2$.

the welfare implications of changes of marginal tax rates in vertical direction (for the spouses with index 1) and the welfare implications of changes of marginal tax rates in horizontal direction (for the spouses with index 2.) The consequence for the overall welfare measure is obtained by "summing over" all the points in the red rectangle.

Corollary 4 gives an intuitive understanding what a change in jointness means: Reducing the jointness for couples where spouse 1 has close to 80,000 USD per year and spouse 2 has close to $60,000$ USD, amounts to a transition into (i) lower marginal tax rates for the earnings of spouse 1 when spouse 1 has more than 80,000 USD, holding spouse 2 fixed at 60,000 USD, and (ii) lower marginal tax rates for the earnings of spouse 2 when spouse 2 has more than 60,000 USD, conditional on spouse 1 making 80,000 USD.

Beyond the Example. Corollary 4 implies that Equation (6) gives the welfare implications for any reform so that (i) $h_{12}(y) = 1$, for $y \in B_1 \times B_2$, and $h_{12}(y) = 0$ otherwise, and (ii) there are no changes of marginal tax rates for single-earner couples. Thus, there is a whole class of different h functions which all have the same welfare implication and the one in the example above is just one function in this class.

In the following, we develop a sufficient statistics characterization of the conditional revenue and welfare functions. This will ultimately enable us to identify the scope for Pareto- or welfare-improving empirically. An intermediate step on this path is relegated to Appendix A.5 where we clarify what behavioral responses look like in our framework and, moreover, what they imply for tax revenue.

2.4 Conditional Revenue Functions as Sufficient Statistics

Conditional revenue functions give the revenue implications of simple tax reforms. Consider an increase of the marginal tax rates on the earnings of spouse 2 conditional on $y \in B_1(y_{1s}, \ell_1) \times B_2(y_{2s}, \ell_2)$ and focus on the limit case $\ell_1 \to 0$ and $\ell_2 \to 0$. The revenue effect of this reform is denoted by $\mathcal{R}_2(y_{2s} \mid y_{1s})$. Proposition 1 below provides a characterization.[20] An analogous characterization of the revenue functions for single-earner couples can be found in Appendix A.7. We first introduce notation that will enable us to state the Proposition in a concise way.

The Extensive Margin. The mass of couples who experience an increase of their tax burden due to an infra-marginal increase of tax rates is given by $1-F_{y_2}(y_{2s} \mid y_{1s})$. To get the reform's impact on the tax revenue that is collected form this set of individuals, this mass is multiplied by $1 - \mathcal{E}^d_x(y_{2s} \mid y_{1s})$, where $\mathcal{E}^d_x(y_{2s} \mid y_{1s})$ is an extensive margin elasticity, reflecting that the reform turns some of these dual-earner couples into single-earner couples.[21]

The Intensive Margin. Conditional on being a dual-earner couple, a couple's optimal earning choices depend only on ω. Intensive margin responses arise when these choices are affected by a tax reform. Henceforth, we denote by $\sigma^{d0}(\omega)$ the mass of dual earner couples among those with productive abilities ω in the status quo.

In Proposition 1 below, the implications of behavioral responses at the intensive

[20] A characterization of $\mathcal{R}_1(y_{1s} \mid y_{2s})$ can be obtained along the same lines.

[21] Formally, $\mathcal{E}^d_x(y_{2s} \mid y_{1s})$ is the percentage change in the mass of dual earner couples that is due to a percentage change of the tax advantage of dual-earner over single-earner couples.

margin for tax revenue are captured by

$$\bar{\beta}^{d}_{I,2}(y_{1s}, y_{2s}) \;=\; \mathbb{E}_{\omega_1}\left[\beta^{d0}_{I,2}(\omega_1, \bar{\omega}_2(y_{2s} \mid \omega_1)) f_{y_2}(y_{2s} \mid \omega_1) \mid y_1^{d0*}(\omega) \;=\; y_{1s}\right],$$

where the expectation is over dual-earner couples with productive abilities ω who, in the status quo, have an income in the region where marginal tax rates on y_2 go up. For such a couple,

$$\beta^{d0}_{I,2}(\omega) \;=\; \sigma^{d0}(\omega)\Big(T^0_{y_1}(y^{d0}(\omega))y^{d0}_{1,\tau}(\cdot) + T^0_{y_2}(y^{d0}(\omega))y^{d0}_{2,\tau}(\cdot)\Big)$$

captures the revenue implications of behavioral responses at the intensive margin.[22] The expectation operator conditions on y_1 being equal to y_{1s} and then goes over the "relevant" combinations of ω_1 and ω_2. The way in which expectations are computed captures, in particular, the possibility that the status quo schedule has a kink at y_{2s}. To see this, for any y_2', and any ω_1, let

$$F_{y_2}(y_2' \mid \omega_1) \quad := \quad F_{\omega_2}(\bar{\omega}_2(y_2' \mid \omega_1) \mid \omega_1)\ ,$$

where, given ω_1, $\bar{\omega}_2(y_2' \mid \omega_1)$ is the largest value of ω_2 so that $y_2^{d0*}(\omega_1, \omega_2) \leq y_2'$. Thus, $F_{y_2}(y_2' \mid \omega_1)$ is a mass of couples who choose $y_2 \leq y_{2s}$. Upon differentiating with respect to y_2', we obtain

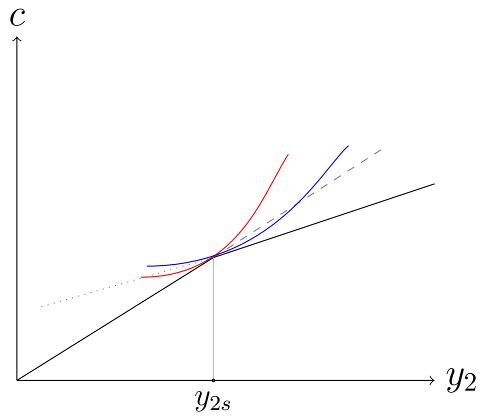
$$f_{y_2}(y_2' \mid \omega_1) \quad := \quad f_{\omega_2}(\bar{\omega}_2(y_2' \mid \omega_1) \mid \omega_1)\frac{\partial}{\partial y_2'}\bar{\omega}_2(y_2' \mid \omega_1)\ .$$

If couples with $y_2 = y_{2s}$ satisfy a first order condition, then $f_{y_2}(y_{2s} \mid \omega_1)$ is a measure of the mass of couples who choose y_2 very close to y_{s2}. If, by contrast, there is a kink because marginal tax rates jump upwards, see the left part of Figure 4, then there is a mass point at y_{2s} and $f_{y_2}(y_{2s} \mid \omega_1)$ indicates how this mass changes when the marginal tax rate changes at y_{2s}. Finally, if there is a kink because marginal tax rates jump downwards, see the right part of Figure 4, then $f_{y_2}(y_{2s} \mid \omega_1)$ indicates how the mass of people who choose y_2 strictly smaller than y_{2s} changes.

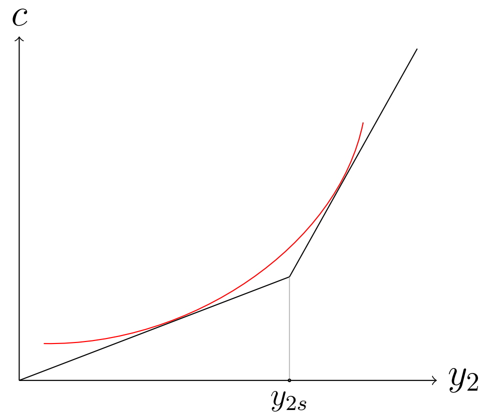
Proposition 1 *Consider an increase of the marginal tax rates on the earnings of spouse 2 conditional on $y \in B_1(y_{1s}, \ell_1) \times B_2(y_{2s}, \ell_2)$. For the limit case $\ell_1 \to 0$ and*

[22] See Appendix A.8.3 for a detailed characterization of $\beta^{d0}_{I,2}(\omega)$.

Figure 4: Intensive margin responses with discontinuities in marginal tax rates



(a) Bunching



(b) Jumping

Note: The black line is the frontier of the couple's budget set conditional on $y_1 = y_{1s}$.
Panel (a): There is a kink at y_{2s} because marginal tax rates jump upwards. Holding fixed the productive ability of spouse 1, ω_1, and the fixed costs φ_1 and φ_2, there is a mass of couples who choose $y_2 = y_{2s}$. These couples differ in the value of ω_2. The red line is the indifference curve of the couple with the minimal value of ω_2. The blue line is the indifference curve of the couple with the maximal value of ω_2. When the marginal tax rate changes at y_{2s} then this maximal value changes and hence the mass of couples who bunch at the kink.
Panel (b) There is a kink at y_{2s} because marginal tax rates jump downwards. Holding fixed the productive ability of spouse 1, ω_1, and the fixed costs φ_1 and φ_2, there is a mass of couples who choose $y_2 < y_{2s}$ and a mass of couples who choose $y_2 > y_{2s}$. These couples differ in the value of ω_2. The red line is the indifference curve of the couple with the critical value of ω_2 that separates these two sets. Note that the earnings function $y_2^*(\cdot)$ jumps upwards at this critical level When the marginal tax rate changes at y_{2s} then this critical value changes and hence the mass of couples who choose $y_2 < y_{2s}$.

$\ell_2 \to 0$, *the revenue effect of this reform is given by*

$$\mathcal{R}(y_{2s} \mid y_{1s}) \quad = \quad f_{y_1}(y_{1s}) \left(\bar{\beta}^d_{I,2}(y_{1s}, y_{2s}) + (1 - F_{y_2}(y_{2s} \mid y_{1s})) \left(1 - \mathcal{E}^d_x(y_{2s} \mid y_{1s})\right)\right) \ ,$$

where $y_1 \mapsto f^{y_1}(y_1)$ is the density associated with the marginal distribution of y_1 and $y_2 \mapsto F^{y_2}(y_{2s} \mid y_{1s})$ is the cdf of y_2, conditional on y_1 being equal to y_{1s}.

The revenue function simplifies further under an additional assumption: When the tax system is piecewise linear, effort-cost functions are iso-elastic

$$k_2(y_2, \omega_2) \quad = \quad \frac{1}{1 + \frac{1}{\varepsilon_2}} \left(\frac{y_2}{\omega_2}\right)^{1+\frac{1}{\varepsilon_2}} \ ,$$

and earnings choices are characterized by first order conditions, then[23]

$$\bar{\beta}^d_{I,2}(y_{1s}, y_{2s}) \quad = \quad -f_{y_2}(y_{2s} \mid y_{1s}) \, \frac{T^0_{y_2}(y_{1s}, y_{2s})}{1 - T^0_{y_2}(y_{1s}, y_{2s})} \, y_{2s} \, \varepsilon_2 \ . \qquad (7)$$

[23]The derivation can be found in part A.8 of the Appendix.

For later reference, we note that Proposition 1 can be extended so as to obtain a sufficient statistics characterization also of the corresponding conditional welfare functions. This yields

$$\begin{aligned} \mathcal{W}_2(y_{2s} \mid y_{1s}) \;=\; & f_{y_1}(y_{1s})\Big(\bar{\beta}^d_{I,2}(y_{1s}, y_{2s}) + (1 - F_{y_2}(y_{2s} \mid y_{1s})) \times \\ & \big(1 - \mathcal{G}(y_{2s} \mid y_{1s}) - \mathcal{E}^d_x(y_{2s} \mid y_{1s})\big)\Big) \,. \end{aligned} \tag{8}$$

2.5 Tax Rates at the Bottom and Optimal Tax Rates at the Top

In the following, we will use the sufficient statistics characterization of conditional revenue and welfare functions to discuss two specific policy design questions: (a) The desirability of workfare elements in the tax and transfer system, and (b) optimal tax rates at the top of the income distribution.

Marginal Tax Rates at the Bottom. The following assumption is empirically plausible at the bottom of the income distribution.

Assumption 1 *Suppose that there is a range of incomes $Y' = Y'_1 \times Y'_2$ over which the following properties all hold true: (i) $\mathcal{G}(y_2 \mid y_1) \geq 1$ and $0 \leq \mathcal{E}^d_x(y_{2s} \mid y_{1s}) < 1$, with at least one of the weak inequalities being strict. (ii) The choices of dual earner couples with $y^{d0}(\theta) \in Y'$ satisfy first order conditions. (iii) Effort cost functions are iso-elastic. (iv) The tax system in is piecewise linear. (v) Marginal tax rates are (weakly) positive.*

To see that part (i) is plausible for low incomes, suppose that $(y_1, y_2) \mapsto \mathcal{G}(y_2 \mid y_1)$ is continuous and non-increasing both in y_1 and y_2. Since $\mathbb{E}_y[g(y)] = 1$ this implies that $1-\mathcal{G}(y_{2s} \mid y_{1s}) \leq 0$ for y_{1s} and y_{2s} both sufficiently close to zero. This inequality is strict if the function $(y_1, y_2) \mapsto \mathcal{G}(y_2 \mid y_1)$ is strictly decreasing. Part (i) allows both for extensive margin elasticities of zero and for positive ones. Empirically, it has been shown that extensive marginal elasticities are largest at low levels of income, while the assumption that extensive margin elasticities are bounded from above by 1 is also in line with the empirical literature. Part (ii) is a regularity condition that simplifies the analysis. Part (iii) is another simplification which, moreover, relates the welfare implications of tax changes to an elasticity that is frequently estimated, the Frisch elasticity. Since any continuous tax function can be well approximated by piecewise linear tax system, part (iv) does not seem to be strong restriction.[24]

Proposition 2 *Under Assumption 1,*

$$\mathcal{W}(y_{2s} \mid y_{1s}) \quad < \quad 0 \quad < \quad \mathcal{R}(y_{2s} \mid y_{1s})\,.$$

We omit a formal proof of Proposition 2. It follows immediately from squaring Proposition 1 with Equations (7) and (8). The Proposition gives conditions under which it is desirable to lower marginal tax rates on the earnings of spouse 2.[25] Moreover, when $\mathcal{W}(y_{2s} \mid y_{1s}) < 0$, and $0 < \mathcal{R}(y_{2s} \mid y_{1s})$, then the lowering of marginal tax rates has two effects: First, it reduces the intercept of the consumption schedule, and therefore is harmful to those with zero earnings both before and after the reform. A Rawlsian welfare function therefore rejects the reform. Second, it makes the secondary earner's budget line in a (c, y)-space steeper, so that the couple's disposable income rises more strongly in secondary earnings. This resembles introducing or strengthening elements of workfare in the tax and transfer system: Less unconditional transfers and more high-powered earnings incentives. It is remarkable that such reforms are desirable for a large class of social welfare functions, including ones with significant inequality aversion. Proposition 2 also implies that optimal marginal tax rates are negative at the bottom of the income distribution: If parts (i) - (v) of Assumption 1 hold, then the condition $\mathcal{W}(y_{2s} \mid y_{1s}) = 0$, yields negative marginal tax rates. Note in particular, that with $\mathcal{G}(y_{2s} \mid y_{1s}) > 1$, negative marginal tax rates are desirable even with an extensive margin elasticity of zero. This contrasts with results from Mirrleesian models with a single-decision maker. In those models, extensive margin responses are necessary for the possibility to rationalize negative marginal tax rates when welfare weights are monotonic in income.[26]

[24] The following paragraph, moreover, presents conditions under which the conditions for welfare optimality can be satisfied fully – as opposed to approximately – with a piecewise linear tax system.

[25] A symmetric argument applies to the earnings of spouse 1.

[26] The seminal reference is Saez (2002), see Bierbrauer, Boyer, and Hansen (2023) for a more extensive discussion of this literature.

Top Tax Rates. We impose a set of assumptions that are analogous to the ones used by Diamond (1998) and Saez (2001) in their characterizations of top tax rates in the basic Mirrless model that does not have couples.

Assumption 2 *Suppose that there is a range of incomes $Y' = Y_1' \times Y_2'$ over which the following properties all hold true: (i) There are numbers $\bar{g}_1$ and $\bar{g}_2$ so that $\mathcal{G}(y_1 \mid y_2) = \bar{g}_1$ and $\mathcal{G}(y_2 \mid y_1) = \bar{g}_2$, for all $(y_1, y_2) \in Y'$. (ii) There are numbers $\bar{\alpha}_1$ and $\bar{\alpha}_2$ so that*

$$\frac{f_{y_1}(y_1 \mid y_2)\, y_1}{1 - F_{y_1}(y_1 \mid y_2)} \;=\; \bar{\alpha}_1 \quad \textit{and} \quad \frac{f_{y_2}(y_2 \mid y_1)\, y_2}{1 - F_{y_2}(y_2 \mid y_1)} \;=\; \bar{\alpha}_2 \;,$$

for all $(y_1, y_2) \in Y'$. (iii) The choices of dual earner couples with $y^{d0}(\theta) \in Y'$ satisfy first order conditions. Moreover, their extensive margin elasticities are zero. (iv) Effort cost functions are iso-elastic.

Again, these assumptions do not formally refer to the top of the income distribution. Empirically, however, they are most plausible near the top. For Y' near the top, part (i) says that welfare weights do not discriminate between "the rich" and "the super rich". They may discriminate, however, between rich primary and rich secondary earners which is the case when $\bar{g}_1 \neq \bar{g}_2$. Part (ii) says that, in the relevant income range, conditional income distributions are (well approximated by) Pareto-distributions, and moreover that the Pareto-parameter for the conditional income distribution of spouse i is, locally, independent of the earnings of spouse $-i$. Parts (iii) and (iv) are as in the previous section on marginal tax rates at the bottom of the income distribution.

Proposition 3 *Suppose that $g(y) > 0$ for all $y \in \mathcal{Y}$. Define t_i^* implicitly as the solution to*

$$\frac{t_i^*}{1 - t_i^*} \;=\; \Big(1 - \bar{g}_i\Big)\, \frac{1}{\bar{\alpha}_i}\, \frac{1}{\varepsilon_i}.$$

Under Assumption 2, $T^0_{y_1}(y) = t_1^*$ and $T^0_{y_2}(y) = t_2^*$, solves $\mathcal{W}_1(y_1 \mid y_2) = 0$ and $\mathcal{W}_2(y_2 \mid y_1) = 0$, for all $(y_1, y_2) \in Y'$.

By Corollary 2, a welfare-maximization implies $\mathcal{W}_1(y_1 \mid y_2) = \mathcal{W}_2(y_2 \mid y_1) = 0$. Assumption 2 makes it possible to find an analytical solution to this system of equations. Near the top the solution takes the form of two optimal tax formulae, one applying to y_1, one applying to y_2, that are both akin to the famous ABC formula developed in Diamond (1998). In particular, under Assumption 2 the optimal tax function is separable and piecewise linear with zero jointness for $y \in Y'$.

3 Empirical Analysis

We now build on the theoretical framework to empirically estimate conditional revenue functions. We use these revenue functions to illustrate inefficiencies in the tax treatment of married couples (Section 3.2), and to describe the welfare consequences of reducing jointness (Section 3.3). Section 3.4 analyzes the welfare implications of lowering marginal tax rates at the bottom of the income distribution while we estimate optimal marginal top tax rates for primary and secondary earners in Section 3.5.

3.1 Conditional Revenue Functions

To estimate conditional revenue functions, we draw on detailed household micro data from the Annual Social and Economic Supplement of the CPS (CPS-ASEC).[27] The detailed demographic characteristics in the CPS allow us to obtain information on the (conditional) income distributions of primary and secondary earners in married couples. Information on tax rates is obtained through the NBER TAXSIM (v32) microsimulation model.[28] By combining CPS data with TAXSIM, we estimate tax rates and tax liabilities under the status quo at the household level. This allows us to capture complexities in the tax code that go beyond the information embedded in statutory marginal tax rates. We focus on the tax system as of 2019.

[27]See Flood et al. (2021) and `https://cps.ipums.org` for a detailed description of CPS data. Appendix B.1 provides details on the data preparation. We use CPS data because it provides separate demographic and earnings information for both spouses. In contrast, the tax return micro data (SOI-PUF) from the Internal Revenue Service (IRS) used in Bierbrauer, Boyer, and Peichl (2021) does not contain this information (except for the year 1974; see Figure B4 for a comparison).

[28]See Feenberg and Coutts (1993) and `https://users.nber.org/~taxsim/` for detailed information on the TAXSIM microsimulation model, and Appendix B.1 for details on its combination with the CPS data.

From Theory to Data. In our empirical application we use conditional revenue functions, where we condition on the other spouse's earnings y_{-i} belonging to a bracket B_{-i}. The relation between these conditional revenue functions characterized in Proposition 1 and the conditional revenue functions used in the empirical application arises via[29]

$$\mathcal{R}_i(y_i \mid B_{-i}) := \int_{y_{-i}\in B_{-i}} \mathcal{R}_i(y_i \mid y_{-i})\, dy_{-i}\ . \tag{10}$$

In our empirical application we deal with the US tax system which is piecewise linear. We, moreover, assume that variable effort costs are captured by an iso-elastic function. Finally, to simplify the analysis, we assume that choices in the status quo satisfy first-order conditions. From Proposition 1 and Equations (10) and (7), it then follows that

$$\mathcal{R}_2(y_{2s} \mid B_1) := s_1(B_1)\Big(I_2^d(y_{2s}, B_1)\ +\ M_2^{xd}(y_{2s}, B_1)\Big)\ , \tag{11}$$

where $s_1(B_1) = F_{y_1}(\bar{b}_1) - F_{y_1}(\underline{b}_1)$ is the mass of couples with $y_1 \in B_1 = [\underline{b}_1, \bar{b}_1]$.

$$I_2^d(y_{2s}, B_1) \quad := \quad -\mathbb{E}_{y_1}\left[f_{y_2}(y_{2s} \mid y_1)\frac{T^0_{y_2}(y_1, y_{2s})}{1 - T^0_{y_2}(y_1, y_{2s})}\ y_{2s}\,\varepsilon_2 \mid y_1 \in B_1\right]$$

gives the revenue effect of behavioral responses at the intensive margin, where

$$M_2^{xd}(y_{2s}, B_1) \quad := \quad \mathbb{E}_{y_1}\left[(1 - F_{y_2}(y_{2s} \mid y_1))(1 - \mathcal{E}_x^d(y_{2s} \mid y_1)) \mid y_1 \in B_1\right]$$

is a term that compounds a simple tax reform's mechanical effect on tax revenue and the effect due to behavioral responses at the extensive margin.[30]

[29] For later reference, we also define the unconditional revenue function

$$\mathcal{R}_i(y_i) := \int_{\mathcal{Y}_{-i}} \mathcal{R}_i(y_i \mid y_{-i})\, dy_{-i}\ , \tag{9}$$

which gives the revenue implication of raising the marginal tax rate at y_i, irrespective of the partner's earnings.

[30] In Appendix B.2, we describe in detail, under which assumptions we estimate the ingredients of the conditional revenue function empirically. There, we also discuss the estimation of unconditional revenue functions. In particular, we assume that conditional revenue functions are constant over the brackets of interest. This constant value then coincides with the bracket average.

Conditioning Brackets. Our analysis in the main text is based on a specification of brackets based on quantiles of the income distribution: We condition on primary earnings belonging to specific deciles of the distribution of primary earnings. This gives us the possible values of B_1. Similarly, when constructing revenue functions conditional on secondary earnings, we consider the deciles of the secondary earner income distribution that will give us B_2.[31]

Behavioral Responses. We consider a baseline scenario with an intensive margin elasticity for primary earners (resp. secondary earners) of 0.25 (resp. 0.75). We also consider alternatives that are documented in Table B2. We consider an extensive margin elasticity of 0.2. We show that our main findings also hold without extensive margin responses, i.e., an extensive margin elasticity of 0.

3.2 Inefficiencies

Figure 5 displays conditional revenue functions for primary and secondary earners.[32] Figure 5a shows where marginal tax rates on secondary earnings are inefficiently high: The inefficiencies are concentrated in the upper deciles of the distribution of secondary earnings. This is most pronounced when primary earnings are in a middle range. For instance, consider primary earnings above the median (Q6). The associated conditional revenue function is positive and monotonically decreasing until secondary earnings exceed $70,000. Then, the function turns negative indicating inefficiently high marginal tax rates on secondary earnings. Hence, a reduction of secondary earners marginal tax rates for all couples with secondary earnings above $70,000 and primary earnings in the sixth income decile would be self-financing.

An inspection of the unconditional revenue function shows that these inefficiencies can only be eliminated with tax cuts that are tailored to the relevant decile of the primary earnings distribution. To see this, note that the unconditional revenue function at secondary earnings of $70,000 is positive. Thus, cutting marginal tax rates on secondary earnings for all couples with y_2 close to $70,000 would not be self-financing.

[31] See Table B3 for the thresholds of the respective income deciles.

[32] Empirical ingredients to these revenue functions, such as marginal tax rates, conditional income distributions, and intensive margin responses are reported in the Appendix, see Figures B7 – B16 (B17–B26).

When applied to primary earnings, this test does not identify inefficiencies (see Figure 5b). Thus, it is not the case that marginal tax rates are inefficiently high across the board and by implication also for secondary earnings. They are too high only for secondary earnings which are close to primary earnings.

In the Appendix, we replicate the analysis by conditioning on earnings quintiles (Figure B30), vingtiles (Figure B31), and on statutory income tax brackets (Figure B32). In line with the main figure, the magnitude of inefficiencies for secondary earners are largest in the middle of the primary earnings income distribution or, alternatively, for couples with primary earners in "middle" tax brackets.

3.3 Reducing Jointness

By Corollary 4, if the status quo tax system is efficient, a Rawlsian welfare measure objects to a reform that reduces jointness. However, as shown in the previous section, the status quo tax system is not efficient. It involves inefficiently high marginal tax rates in parts of the income space. Whether Rawlsian welfare is improved by a jointness-reducing reform is thus an open question that we discuss in this section. We also provide answers for alternative welfare measures.

We consider reductions of jointness tailored to specific combinations of income brackets for primary and secondary earnings. By Corollary 4, for τ small, the welfare implications of a jointness reducing reform are well approximated by

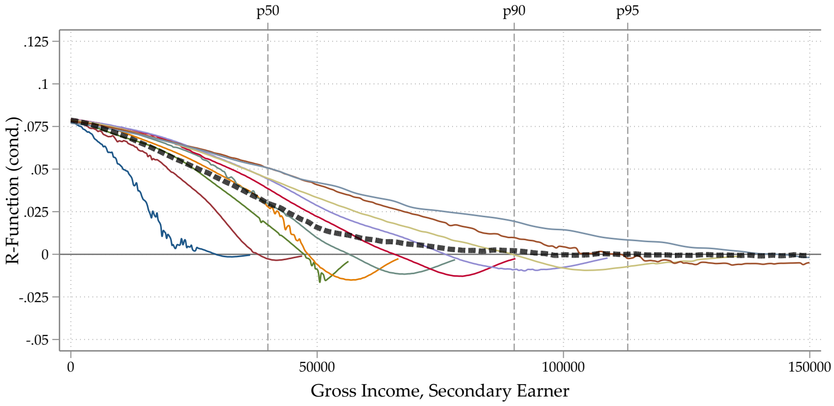
$$\begin{aligned} W &= \tau \int\limits_{B_1 \times B_2} \left[\int_{\underline{y}_1}^{\bar{y}_1} \mathcal{W}_2(y_2 \mid y_1) dy_1 \; + \int_{\underline{y}_2}^{\bar{y}_2} \mathcal{W}_1(y_1 \mid y_2) dy_2 \; \right] dy \\ &= \tau \int\limits_{B_1 \times B_2} \left[\mathcal{R}_2 \left(y_2 \mid [\underline{y}_1, \bar{y}_1] \right) + \mathcal{R}_1 \left(y_1 \mid [\underline{y}_2, \bar{y}_2] \right) - 2 \, s\left(y_1, y_2 \right) \mathcal{G}\left(y_1, y_2 \right) \right] dy \; , \end{aligned} \tag{12}$$

where $s\left(y_1, y_2\right)$ is the share of couples with primary earnings above y_1 and secondary earnings above y_2, and $\mathcal{G}(y_1, y_2) = \mathbb{E}[g(s_1, s_2) | s_1 \geq y_1, s_2 \geq y_2]$ is the average welfare weight for couples in this group. The second line in (12) expresses the welfare implications of a jointness-reducing reform using conditional revenue functions, which capture the loss in overall revenue and cumulative welfare weights which capture the gains of those who benefit from a lower tax burden.

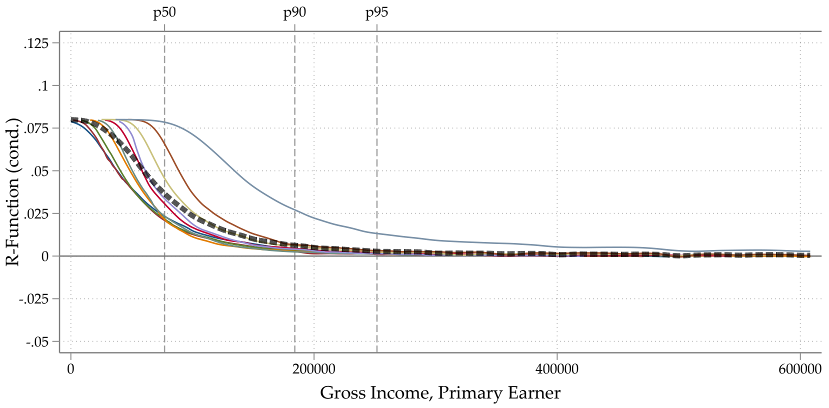
If the relevant brackets are not too wide, we can approximate the above expres-

Figure 5: Conditional revenue functions, deciles (2019)

(a) Secondary earners



(b) Primary earners



Note: This figure shows conditional revenue functions for secondary (resp. primary) earners in married dual earner couples conditional on primary (resp. secondary) earnings income deciles in Panel a (resp. b) as of 2019. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 (resp. 0.75) for primary (resp. secondary) earners and an extensive margin elasticity of 0.2. The figure also displays modified unconditional revenue functions for secondary and primary earners where unconditional revenue functions have been scaled by 0.1 to facilitate comparability with the conditional revenue functions. Results without extensive margin responses are shown in Figure B27. Results for alternative elasticities (Table B2) are shown in Figures B28 and B29.
Source: Own calculations based on CPS-ASEC (2019).

sion for W by

$$W = \tau\ \ell_1\ \ell_2 \left[\mathcal{R}_2\left(y_{2s} \mid [y_{1s}, \bar{y}_1]\right) + \mathcal{R}_1\left(y_{1s} \mid [y_{2s}, \bar{y}_2]\right) - 2\ s\left(y_{1s}, y_{2s}\right) \mathcal{G}\left(y_{1s}, y_{2s}\right) \right], \quad (13)$$

where y_{1s} and y_{s2} are, respectively, the income levels at which the relevant brackets start, whereas ℓ_1 and ℓ_2 are their respective lengths.

In our application of this formula, we consider brackets of lengths $l_1 = l_2 = 500$ and we set $\tau = -0.01$, corresponding to a one percentage point reduction of marginal tax rates. We select the thresholds y_{1s} for primary earnings based on the deciles of the primary earner income distribution, and the thresholds for secondary earnings y_{2s} based on the deciles of the distribution of secondary earnings conditional on the respective primary earner income decile. The relevant bracket combinations are shown in Table B6. We, moreover, focus on specific welfare functions that enable us to highlight the trade-offs associated with a reduction of jointness. Specifically, we consider a welfare function with equal weights for all couples. With preferences that are quasi-linear in consumption, welfare is then simply a measure of the total economic surplus. Changes in this welfare measure indicate how the distortions due to the behavioral responses to taxation change. We also consider welfare functions that, for lack of a better term, we call quasi-Rawlsian. Those welfare functions assign positive weights only to couples with joint income $y_1 + y_2$ in the bottom decile of the joint income distribution of dual earner couples. This welfare function indicates whether a tax reform is in the interest of "the poor". Finally, we consider "feminist" welfare functions with weights that increase in the women's share in a couples joint income. This welfare function indicates whether a reform is in the interest of "working women".[33]

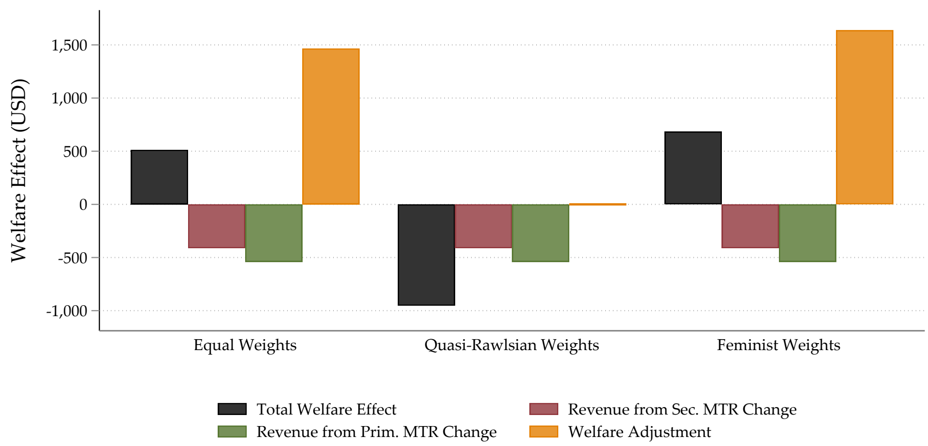
Reducing Jointness Around the Median. Figure 6 illustrates the welfare effect for a reform that reduces jointness in $B_1 \times B_2$, where B_1 starts at the median level of primary earnings and y_2 starts at the median of secondary earnings conditional on $y_1 \in B_1$. It provides a decomposition into the various components that appear on the right hand side of Equation (13). The left hand side corresponds to the black bar. The Figure shows that the reform is rejected by a quasi-Rawlsian wel-

[33] For the detailed specification of welfare weights, see Table B5. Appendix B.3 further provides detailed information on the required ingredients and construction of empirical welfare effects.

fare function and approved both by a Feminist one and by one with equal weights.

Note that the revenue effects coming from the respective changes in secondary and primary earner marginal tax rates do not necessarily point into the same direction. For instance, when marginal tax rates are inefficiently high, e.g. under strong behavioral responses and for relatively rich secondary earners (see Figure B34), the (conditional) cut of secondary earner marginal tax rates might lead to a revenue gain. The aggregate revenue and welfare effects then depend on the relative magnitude of revenue effects coming from the tax rate changes on primary earnings and on secondary earnings.

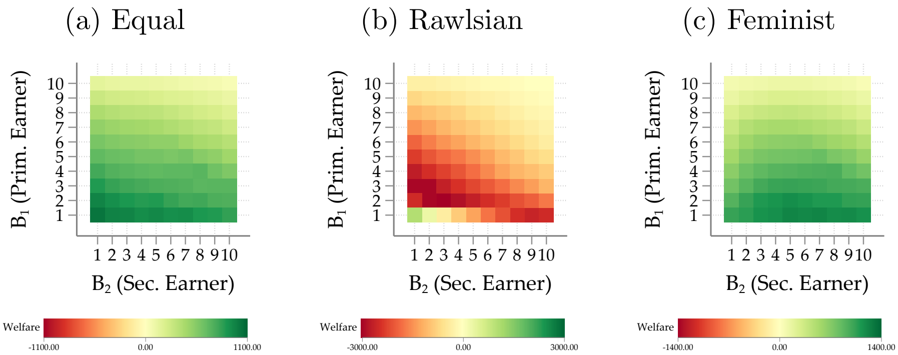
Figure 6: Welfare effect, reducing jointness at the median (2019)



Note: This figure visualizes the welfare effects implied by a reduction in jointness as of 2019 for $(y_1, y_2) \in B_1 \times B_2$, where B_1 is the 6th decile of the distribution of primary earnings and B_2 is the 6th decile in the distribution of secondary earnings conditional on $y_1 \in B_1$. The aggregate welfare effect and its components is shown for three different welfare weights (see Table B5) and is based on the bracket lengths $\ell_1 = \ell_2 = 500$ and $\tau = -0.01$. We take intensive margin responses according to the baseline elasticity scenario (see Table B2) into account and also include extensive margin responses. The black bar indicates the aggregate welfare effect. The red (resp. green) bar visualizes the revenue change coming from the decrease of secondary (resp. primary) earnings. The orange bar illustrates the welfare adjustment corresponding to the last term in equation (13). For details on the decomposition of the welfare effect, see also Appendix B.3. Figure B33 illustrates the conditional revenue functions, the two revenue effects are based on.
Source: Own calculations based on CPS-ASEC (2019).

Different Combinations of Brackets. We can repeat the exercise across all combinations of primary and secondary earnings shown in Table B6 and evaluate the welfare consequences of jointness-reducing reforms in the respective brackets. Figure 7 summarizes the results for all three welfare measures and all bracket combinations. Under equal weights, a reduction in jointness is always considered beneficial, no matter at which parts of the income distribution it is carried out. The positive welfare impact is larger if behavioral responses are stronger (see Figure B35). The feminist welfare function gives a similar picture, but with more pronounced welfare gains for reform in the range of high secondary earnings. The quasi-Rawlsian welfare measure rejects reductions in jointness, except at the very bottom. The opposition is weak however for reductions of jointness in brackets with high secondary earnings as those reforms are close to self-financing (see Section 3.1).

Figure 7: Welfare implications of decreasing jointness (2019)



Note: This figure displays the welfare implications of decreasing jointness as of 2019 by decreasing secondary and primary marginal tax rates conditional on particular brackets of primary earnings and secondary earnings. Bracket thresholds are shown in Table B6. The reform applies to a bracket of length $l(B_1) = l(B_2) = 500$ and has the magnitude $\tau = -0.01$. The figure distinguishes between three different forms of welfare weights (see Table B5) and is based on intensive margin responses according to the baseline elasticity scenario (see Table B2). All results are shown including extensive margin responses. Results for different elasticity scenarios and without extensive margin responses are shown in Figures B35 and B36.
Source: Own calculations based on CPS-ASEC (2019).

3.4 Lowering Marginal Tax Rates at the Bottom

Proposition 2 points to the possibility that welfare functions with continuous weights and a Rawlsian welfare functions may give rise to different conclusions regarding the desirability of reforms that lower marginal tax rates at the bottom of the income distribution. In this section, we demonstrate that this situation is empirically relevant. Specifically, we will show that a strictly Rawlsian and a quasi-Rawlsian welfare function reach opposite conclusions.

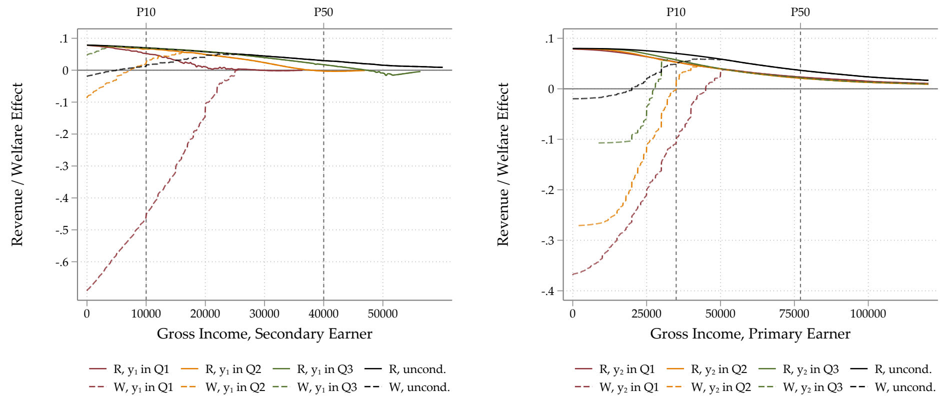
Our analysis makes use of the empirical conditional revenue functions that were already used in our analysis of inefficiencies in Section 3.1. To get to the welfare implications, we extend the approach that was used to construct these functions to the conditional welfare functions that were introduced in Section 2.[34]

[34]Our empirical implementation of the conditional welfare functions uses equation (8) and then

Figure 8 shows results for a reduction of marginal tax rates at a specific level of secondary earnings conditional on primary earnings belonging to the three bottom deciles of the primary earnings distribution. It also presents analogous results for a lowering of marginal tax rates on primary earnings. The solid lines are the conditional revenue functions from Section 3.1 which describe the revenue effects of the tax change. The dashed lines indicate the welfare effect of the tax change. The figure also displays the revenue and welfare effects of reforms that lower marginal tax rates on secondary or primary earnings unconditionally (black lines).

Figure 8: Welfare effects of reforms at the bottom (2019)

(a) Secondary earnings (b) Primary earnings



Note: This figure shows the revenue and welfare effects for secondary (resp. primary) earners in married dual earner couples conditional on primary (resp. secondary) earnings income deciles at the bottom of the income distribution in Panel a (resp. b) as of 2019. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 (resp. 0.75) for primary (resp. secondary) earners and an extensive margin elasticity of 0.2. The figure displays conditional revenue functions (solid lines, also displayed in detail in Figure 5) and the welfare effect based on quasi-Rawlsian welfare weights according to Table B5 (dashed lines). Black lines indicate unconditional revenue and welfare functions. Results without extensive margin responses are shown in Figure B38. Results for different elasticity scenarios in line with Table B2 are shown in Figure B39.
Source: Own calculations based on CPS-ASEC (2019).

We can extract three insights about tax reforms at low incomes. First, in line with the theoretical reasoning in Proposition 2, the quasi-Rawlsain welfare function takes negative values at low secondary earnings conditional on low primary earnings (Figure 8a). Thus, reducing marginal tax rates is desirable according to this welfare measure. This is also true at low primary earnings conditional on low secondary earnings (Figure 8b). Second, the conditioning on the other spouse's earnings makes

follows our empirical implementation of the conditional revenue functions.

an important difference. An unconditional quasi-Rawlsian welfare function takes negative value over a much more narrow income range at the very bottom.[35] Third, lowering marginal tax rates is costly in terms of tax revenue and therefore rejected by a strictly Rawlsian welfare function. This holds true both for conditional and for unconditional revenue functions. Lowering marginal tax rates would thus be akin to an introduction of workfare elements into the tax-and-transfer system: The intercept of the consumption schedule goes down but its slope goes up as the net-of-tax-rate increases.

3.5 Optimal Top Tax Rates

Recall that under the premisses of Proposition 3 top tax rates for primary and secondary earners are given by

$$\frac{t_i^*}{1-t_i^*} = (1-\bar{g}_i)\,\frac{1}{\bar{\alpha}_i}\frac{1}{\varepsilon_i} \qquad i = 1, 2\ . \tag{14}$$

We seek to trace out the implications of differences in the Pareto-coefficients for primary and secondary earnings for top tax rates. We therefore assume, for ease of exposition, that "rich people" have welfare weights of zero, $g_1 = g_2 = 0$. As in the previous sections, in our baseline, we assume that $\varepsilon_1 = 0.25$ and $\varepsilon_2 = 0.75$.[36]

Pareto Coefficients. We estimate Pareto coefficients for primary and secondary earners among jointly filing tax units using tabulated income tax data from the Statistics of Income (SOI) program of the Internal Revenue Service (IRS).[37] We use the tabulated tax return data and apply generalized Pareto interpolation techniques by Blanchet, Fournier, and Piketty (2022) to estimate properties of the wage income distribution. In the main text, we focus on 2018, the most recent available year for this data. Panel A of Table 1 shows the estimated Pareto coefficients at the 95th and 99th percentile of the wage income distribution. At both percentiles, the Pareto

[35] When excluding extensive margin responses, the unconditional quasi-Rawlsian welfare function is positive throughout, while the unconditional ones take negative values (Figure B38)— a fact driven by the behavior of (un)conditional welfare weights (Figure B37).

[36] We proceed under Assumption 2. In particular, we take elasticities to be constant over the relevant part of the income distribution. An implication is that the Pareto coefficient does not vary with the tax rates at the top.

[37] For details on the data source and preparation, see Appendix B.1.2.

coefficient of the primary earner wage income distribution is around 40 percent lower than the respective coefficient for secondary earners.

Results. Based on the estimated Pareto coefficients, Panel B of Table 1 displays optimal top tax rates for different elasticities. Even under the assumption that $\varepsilon_1 = \varepsilon_2$ at the top of the income distribution, the optimal tax rates for primary earners are around 12 percentage points higher than the one for secondary earners.[38]

As shown by Figure B40, for the time period where tabulated tax data for primary and secondary earners is available (2008-2018), the optimal top tax rate gap stayed relatively constant. In Appendix B.5, we also provide heterogeneity analyses that estimate top tax rates for primary earners separately for single earner and dual earner couples. The results show that marginal tax rates for primary earners in single earner couples should be higher than for primary earners in dual earner couples, since the Pareto coefficients of the former group are much lower. In particular, optimal top tax rates for primary earners in dual earner couples should be closer to the secondary earners optimal top tax rate.

Table 1: Pareto coefficients and optimal tax rates (2018)

	P95		P99	
	Primary	Secondary	Primary	Secondary
Panel A: Pareto coefficients				
Wages on W2 form	1.83	3.06	1.75	3.12
Panel B: Optimal tax rates				
Elasticity = .25	69%	57%	70%	56%
Elasticity = .5	52%	40%	53%	39%
Elasticity = .75	42%	30%	43%	30%

Note: This table shows Pareto coefficients and optimal top tax rates. Panel A displays Pareto coefficients for primary and secondary earners based on a generalized Pareto interpolation using tabulated data on wages on W2 forms for joint return taxpayers with wage income. We distinguish between interpolations at the 95th and the 99th percentile. Panel B displays optimal top tax rates associated with these Pareto coefficients. We distinguish between different elasticities.
Source: Own calculations based on SOI Tax Stats - Individual Information Return Form W-2 Statistics (2018).

[38] In Appendix B.5, we show that this gap is also present when using CPS data to estimate Pareto coefficients and optimal top tax rates.

4 Concluding Remarks

In this paper, we have introduced a theory of tax reforms for a setting with (i) multi-dimensional heterogeneity amongst taxpayers, (ii) multiple economic decisions that are all subject to both fixed and variable costs, and therefore involve multiple discontinuities in the behavioral responses to taxation, and (iii) a tax function that treats all these decisions as separate arguments. For such a setting, the theorems in this paper provide a complete characterization for the existence of Pareto- or welfare-improving tax reforms.

We then explored the implications of these Theorems for a particular policy design problem, the taxation of couples. For this application, we developed sufficient statistics that can be used to check (i) whether and where marginal tax rates on secondary earnings are inefficiently high, (ii) whether and where changes to the tax system's degree of jointness would be welfare-improving, (iii) which welfare functions approve and which welfare functions reject reforms that introduce or strengthen workfare elements at the bottom of the income distribution, and (iv) to characterize optimal marginal tax rates at the top of the income distribution.

Finally, we provided answers to these questions against the background of the tax treatment of couples in the US. We showed that (i) marginal tax rates on secondary earnings are inefficiently high in parts of the income distribution, (ii) that a reduction of jointness is approved by Feminist welfare measures, but rejected by Rawlsian ones, (iii) that "more" workfare is desirable for a large class of social welfare functions, including ones that concentrate welfare weights at the bottom, but rejected by a Rawlsian welfare function, and (iv) at the top, marginal tax rates on secondary earners should be lower than marginal tax rates on primary earnings (even if "rich men" and "rich women" receive the same welfare weights and show the same behavioral responses to taxation) because the distribution of primary earnings has fatter tails than the distribution of secondary earnings.

References

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

Alesina, A., A. Ichino, and L. Karabarbounis. 2011. "Gender-Based Taxation and the Division of Family Chores." *American Economic Journal: Economic Policy* 3 (2): 1–40.

Alves, C.B., C.E. da Costa, F. Lobel, and H. Moreira. 2024. "Intrahousehold inequality and the joint taxation of household earnings." *Journal of Public Economics* 239:105208.

Bargain, O., M. Dolls, H. Immervoll, D. Neumann, A. Peichl, N. Pestel, and S. Siegloch. 2015. "Tax policy and income inequality in the United States, 1979–2007." *Economic Inquiry* 53 (2): 1061–1085.

Bargain, O., K. Orsini, and A. Peichl. 2014. "Comparing labor supply elasticities in europe and the united states new results." *Journal of Human Resources* 49 (3): 723–838.

Bergstrom, K., and W. Dodds. 2021. "Optimal taxation with multiple dimensions of heterogeneity." *Journal of Public Economics* 200:104442.

Bick, A., and N. Fuchs-Schündeln. 2017. "Taxation and Labour Supply of Married Couples across Countries: A Macroeconomic Analysis." *Review of Economic Studies* 85 (3): 1543–1576.

Bierbrauer, F.J., P.C. Boyer, and E. Hansen. 2023. "Pareto-improving tax reforms and the Earned Income Tax Credit." *Econometrica* 91 (3): 1077–1103.

Bierbrauer, F.J., P.C. Boyer, and A. Peichl. 2021. "Politically feasible reforms of nonlinear tax systems." *American Economic Review* 111 (1): 153–91.

Bierbrauer, F.J., P.C. Boyer, A. Peichl, and D. Weishaar. 2023. *The taxation of couples.* CEPR Discussion Paper 18138.

———. 2024. *The political economy of joint taxation.* Mimeo.

Blanchet, T., J. Fournier, and T. Piketty. 2022. "Generalized Pareto Curves: Theory and Applications." *Review of Income and Wealth* 68 (1): 263–288.

Blau, F. D., and L. M. Kahn. 2007. "Changes in the labor supply behavior of married women: 1980–2000." *Journal of Labor Economics* 25 (3): 393–438.

Blundell, R., and T. Macurdy. 1999. "Chapter 27 - Labor Supply: A Review of Alternative Approaches," edited by Orley C. Ashenfelter and David Card, 3:1559–1695. Handbook of Labor Economics. Elsevier.

Boerma, J., A Tsyvinski, and A.P. Zimin. 2022. *Bunching and Taxing Multidimensional Skills.* NBER Working Paper 30015.

Borella, M., M. De Nardi, and F. Yang. 2023. "Are Marriage-Related Taxes and Social Security Benefits Holding Back Female Labour Supply?" *Review of Economic Studies* 90 (1): 102–131.

Boskin, M. J., and E. Sheshinski. 1983. "Optimal tax treatment of the family: Married couples." *Journal of Public Economics* 20 (3): 281–297.

Brett, C. 2009. "Optimal nonlinear taxes for families." *International Tax and Public Finance* 14 (3): 225–261.

Carlier, G., X. Dupuis, J.-C. Rochet, and J. Thanassoulis. 2024. "A general solution to the quasi linear screening problem." *Journal of Mathematical Economics* 114:103025.

Cremer, H., J.-M. Lozachmeur, and P. Pestieau. 2012. "Income taxation of couples and the tax unit choice." *Journal of Population Economics* 25 (2): 763–778.

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

Eissa, N., and H. W. Hoynes. 2004. "Taxes and the labor market participation of married couples: the Earned Income Tax Credit." *Journal of Public Economics* 88 (9-10): 1931–1958.

Eissa, N., H.J. Kleven, and C.T. Kreiner. 2008. "Evaluation of four tax reforms in the United States: Labor supply and welfare effects for single mothers." *Journal of Public Economics* 92 (3): 795–816.

Feenberg, D. R., and E. Coutts. 1993. "An Introduction to the TAXSIM Model." *Journal of Policy Analysis and Management* 12 (1): 189–194.

Ferey, A., B.B. Lockwood, and D. Taubinsky. 2024. "Sufficient statistics for nonlinear tax systems with general across-income heterogeneity." *American Economic Review* 114 (10): 3206–3249.

Flood, S., M. King, R. Rodgers, S. Ruggles, J. R. Warren, and M. Westberry. 2021. "Integrated Public Use Microdata Series, Current Population Survey: Version 9.0 [dataset]."

Gayle, G.-L., and A. Shephard. 2019. "Optimal taxation, marriage, home production, and family labor supply." *Econometrica* 87 (1): 291–326.

Golosov, M., and I. Krasikov. 2023. *The Optimal Taxation of Couples.* NBER Working Paper 31140.

Golosov, M., A. Tsyvinski, and N. Werquin. 2014. *A Variational Approach to the Analysis of Tax Systems.* NBER Working Paper 20780.

Guner, N., R. Kaygusuz, and G. Ventura. 2012. "Taxation and household labour supply." *Review of Economic Studies* 79 (3): 1113–1149.

Holter, H., D. Krüger, and S. Stepancuk. 2023. *Until the IRS Do Us Part: (Optimal) Taxation of Households.* Mimeo. University of Pennsylvania.

Immervoll, H., H. J. Kleven, C. T. Kreiner, and N. Verdelin. 2011. "Optimal tax and transfer programs for couples with extensive labor supply responses." *Journal of Public Economics* 95 (11): 1485–1500.

Jacquet, L., and E. Lehmann. 2021a. *How to Tax Different Incomes?* CEPR Discussion Paper 16571.

———. 2021b. "Optimal income taxation with composition effects." *Journal of the European Economic Association* 19 (2): 1299–1341.

Kleven, H. J., C. T. Kreiner, and E. Saez. 2009. "The Optimal Income Taxation of Couples." *Econometrica* 77 (2): 537–560.

LaLumia, S. 2008. "The effects of joint taxation of married couples on labor supply and non-wage income." *Journal of Public Economics* 92 (7): 1698–1719.

Malkov, E. 2021. *Optimal Income Taxation of Singles and Couples.* Unpublished Manuscript.

Milgrom, P., and I. Segal. 2002. "Envelope Theorems for Arbitrary Choice Sets." *Econometrica* 70 (2): 583–601.

Piketty, T. 1997. "La redistribution fiscale face au chômage." *Revue française d'économie* 12:157–201.

Rochet, J.-C. 2024. "Multidimensional screening after 37 years." *Journal of Mathematical Economics:* 103010.

Sachs, D., A. Tsyvinski, and N. Werquin. 2020. "Nonlinear tax incidence and optimal taxation in general equilibrium." *Econometrica* 88 (2): 469–493.

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

———. 2002. "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses." *Quarterly Journal of Economics* 117 (3): 1039–1073.

Saez, E., and S. Stantcheva. 2016. "Generalized Social Marginal Welfare Weights for Optimal Tax Theory." *American Economic Review* 106 (1): 24–45.

Schroyen, F. 2003. "Redistributive taxation and the household: the case of individual filings." *Journal of Public Economics* 87 (11): 2527–2547.

Spiritus, K., E. Lehmann, S. Renes, and F. Zoutman. Forthcoming. "Optimal taxation with multiple incomes and types." *Theoretical Economics.*

Supplemental Appendix

A Appendix for the theory part / Section 2 **2**

A.1 Proofs of Theorem 1 and Corollary 1 2

A.2 Proof of Theorem 2 5

A.3 Proof of Theorem 3 15

A.4 Proof of Corollary 4 17

A.5 Behavioral responses and tax revenue 18

A.6 Proof of Proposition 1 27

A.7 Conditional revenue functions: Single-earner couples 38

A.8 Intensive margin responses of dual earner couples 47

A.9 Proof of Proposition 3 51

A.10 Alternative modelling choices 54

B Appendix for the empirical part / Section 3 **57**

B.1 Data 57

B.2 Conditional revenue functions 64

B.3 Reducing jointness 85

B.4 Tax reforms at the bottom 88

B.5 Optimal top tax rates 94

A Appendix for the theory part / Section 2

A.1 Proofs of Theorem 1 and Corollary 1

We focus on the statements in Theorem 1 that involve conditional revenue functions for dual earner couples. The extension to single-earner couples follows from the same arguments and is therefore omitted. Recall that a reform in direction $h \in \mathcal{H}$ is Pareto-improving if

$$R_\tau(0,h) \;-\; \max_{y\in\mathcal{Y}} \; h(y) \quad \geq \quad 0\;. \tag{15}$$

A reform in direction $-h \in \mathcal{H}$ is Pareto-improving if

$$R_\tau(0,h) \;-\; \min_{y\in\mathcal{Y}} \; h(y) \quad \leq \quad 0\;. \tag{16}$$

Let $(i,j) \in \{(1,2),(2,1)\}$. Consider a simple reform h so that

$$h(y) \quad = \quad h_i(y_i)\;\mathbf{1}(y_j \in [y_{js}, y_{js} + \ell_j])\;,$$

where

$$h_i(y_i) \quad = \quad \begin{cases} 0, & \text{if} \;\; y_i \leq y_{is} \\ y_i - y_{is} & \text{if} \;\; y_{is} \leq y \leq y_{is} + \ell_i \\ \ell_i, & \text{if} \;\; y_{is} \geq y_{is} + \ell_1 \end{cases}$$

It follows from (15) and (16) that a reform in direction h is Pareto-improving if

$$R_\tau(0,h) \quad > \quad \ell_i\;. \tag{17}$$

A reform in direction $-h$ is Pareto-improving if

$$R_\tau(0,h) \quad < \quad 0\;. \tag{18}$$

Lemma 1 *If, for some B_j, $\mathcal{R}_i(y_i \mid B_j) > 1$, there exist y_{is} and $\ell_i > 0$ so that (17) holds. If $\mathcal{R}_i(y_i \mid B_j) < 0$, there exist y_{is} and $\ell_i > 0$ so that (18) holds.*

Proof. For ℓ_i close to zero, we have that[39]

$$R_\tau(0,h) \quad \simeq \quad \ell_i\,\mathcal{R}_i(y_i \mid B_j)\;.$$

Therefore, for ℓ_i close to zero, (17) holds if

$$\ell_i\,\mathcal{R}_i(y_i \mid B_j) \quad > \quad \ell_i\;,$$

or, equivalently, if $\mathcal{R}_i(y_i \mid B_j) > 1$. Analogously, for ℓ_i close to zero, (18) holds if

$$\ell_i\,\mathcal{R}_i(y_i \mid B_j) \quad < \quad 0\;,$$

or, equivalently, if $\mathcal{R}_i(y_i \mid B_j) < 0$. □

Lemma 2 *If, for some* B_j, $\underline{y}_i < \overline{y}_i$ *and* $\mathcal{R}_i(\overline{y}_i \mid B_j) > \mathcal{R}_i(\underline{y}_i \mid B_j)$, *there is a Reform-direction* h *so that (15) holds.*

Proof. We define

$$h \quad = \quad \underline{\tau}_i\,\underline{h}_i(y_i)\,\mathbf{1}(y_j \in B_j) + \overline{\tau}_i\,\overline{h}_i(y_i)\,\mathbf{1}(y_j \in B_j)\;,$$

where

$$\underline{h}_i(y_i) \quad = \quad \begin{cases} 0, & \text{if } \; y_i \leq \underline{y}_i \\ y_i - \underline{y}_i & \text{if } \; \underline{y}_i \leq y \leq \underline{y}_i + \underline{\ell}_i \\ \underline{\ell}_i, & \text{if } \; y_{is} \geq \underline{y}_i + \ell_i \end{cases}$$

[39] With $\ell_i = 0$, we have that $T^0(y) = T^1(y)$, for all y and hence $y^*(\theta,\tau,h) = y^0(\theta)$. There is neither a change of the tax system nor of behavior, hence there is no change of tax revenue so that

$$R_\tau(\tau,h) = R_i(y_{is},0,y_{js},\ell_j) = 0\;,$$

for all $\tau \geq 0$. In particular, $R_i(0,y_{is},y_{js},\ell_j) = 0$. A first order Taylor approximation, moreover yields

$$R_\tau(0,h) \quad \simeq \quad R_i(y_{is},0,y_{js},\ell_j) \; + \; \ell_i\,\mathcal{R}_i(y_i \mid B_j) \quad = \quad \ell_i\,\mathcal{R}_i(y_i \mid B_j)\;.$$

and

$$\overline{h}_i(y_i) \quad = \quad \begin{cases} 0, & \text{if} \;\; y_i \leq \overline{y}_i - \overline{\ell}_i \\ y_i - \overline{y}_i & \text{if} \;\; \overline{y}_i - \overline{\ell}_i \leq y \leq \overline{y}_i \\ \ell_i, & \text{if} \;\; y_{is} \geq \underline{y}_i \; . \end{cases}$$

We also choose $\underline{\tau}_i < 0$, $\overline{\tau}_i > 0$ and $\underline{\tau}_i + \overline{\tau}_i = 0$. We finally let $\underline{\ell}_i = \overline{\ell}_i =: \ell_i$. This construction ensures that, conditional on $y_j \in B_j$,

$$h(y) \begin{cases} = \; 0, & \text{if} \;\; y_1 \leq \underline{y}_1 \\ < \; 0, & \text{if} \;\; \underline{y}_1 < y < \overline{y}_1 \\ = \; 0, & \text{if} \;\; y_1 \geq \overline{y}_1 \; . \end{cases}$$

so that $\max_y h(y) = 0$. Thus, to verify that (15) holds, it remains to be shown that $R_\tau(0,h) > 0$. To see this, note that by the linearity of the Gateaux differential, for ℓ_i close to zero, we have

$$\begin{aligned} R_\tau(0,h) \;\; &= \;\; \underline{\tau}_i \; \ell_i \; \mathcal{R}_i(\underline{y}_i \mid B_j) + \overline{\tau}_i \; \ell_i \; \mathcal{R}_i(\overline{y}_i \mid B_j) \\ &> \;\; \underline{\tau}_i \; \ell_i \; \mathcal{R}_i(\underline{y}_i \mid B_j) + \overline{\tau}_i \; \ell_i \; \mathcal{R}_i(\underline{y}_i \mid B_j) \\ &= \;\; 0 \; . \end{aligned}$$

□

Lemmas 1 and 2 imply Theorem 1. If a given tax system is Pareto-efficient, then there must not exist a Pareto-improving reform direction. The following Lemma which establishes Corollary 1 exploits this property to give necessary conditions for the Pareto-efficiency of a tax system.

Lemma 3 *If there is no Pareto-improving reform direction, then the following holds, for all* $(i,j) \in \{(1,2),(2,1)\}$ *and for all* $y \in \mathcal{Y}$*: (i) There is a number* $\bar{\ell}_j$*, so that, for all* $\ell_j < \bar{\ell}_j$*,*

$$\ell_j \; \mathcal{R}_i(y_i \mid y_j) \quad \leq \quad 1 \; . \tag{19}$$

(ii)

$$\mathcal{R}_i(y_i \mid y_j) \quad \geq \quad 0 \; . \tag{20}$$

(iii) The function $y_i \mapsto \mathcal{R}_i(y_i \mid y_j)$ is non-increasing.

Proof. All statements follow from the observation that, for ℓ_j close to zero,

$$\mathcal{R}_i(y_i \mid B_j[y_{js}, \ell_j]) \quad \simeq \quad \ell_j \, \mathcal{R}_i(y_i \mid y_j) \, .$$

Therefore, if (19) was violated, there would exists ℓ_j so that

$$\mathcal{R}_i(y_i \mid B_j[y_{js}, \ell_j]) \quad > \quad 1$$

thereby contradicting the non-existence of a Parteo-improving reform direction, see Proposition 1. If (20) was violated, there would exist ℓ_j so that

$$\mathcal{R}_i(y_i \mid B_j[y_{js}, \ell_j]) \quad < \quad 0$$

again contradicting the non-existence of a Pareto-improving reform direction, see Proposition 1. If the function $y_i \mapsto \mathcal{R}_i(y_i \mid y_j)$ was increasing, this would imply the existence of ℓ_j and $\underline{y}_i < \overline{y}_i$ with $\mathcal{R}(\overline{y}_i \mid B_j[y_{js}, \ell_j]) > \mathcal{R}_i(\underline{y}_i \mid B_j[y_{js}, \ell_j])$, once more contradicting the non-existence of a Pareto-improving reform direction, see Proposition 2. □

A.2 Proof of Theorem 2

Outline. Let $h \in \mathcal{H}$ be a continuous function. We approximate h by a combination of simple tax reforms h_m and show that

$$R_\tau(0, h) \; - \; \max_{y \in \mathcal{Y}} \; h_m(y) \quad \leq \quad 0 \, . \tag{21}$$

We then extend the argument to the function h.

A.2.1 Constructing h_m.

Consider a partition of $\mathcal{Y}_1$ in m adjacent brackets of equal length ℓ_1, indexed by $k_1 \in K_1 = \{1, 2, \ldots, m\}$, with bracket $B^1_{k_1}$ starting at $(k_1 - 1)\ell_1$ and ending at $k_1 \ell_1$. A partition of $\mathcal{Y}_2$ in m adjacent intervals of equal length ℓ_2 is defined analogously, with brackets indexed by $k_2 \in K_2 = \{1, 2, \ldots, m\}$.

A simple tax reform that changes marginal tax rates in the interior of $\mathcal{Y}_1 \times \mathcal{Y}_2$ can now be identified by $k \in K = K_1 \times K_2$ and by the changes in marginal tax rates τ_k^1 and τ_k^2 applying, respectively, to the earnings of spouse 1 and spouse 2. For such a reform we write

$$h_k(y_1, y_2) \quad = \quad \tau_k^1 \, h_{k_1+1}^1(y_1) \, \mathbf{1}(y_2 \in B_{k_2+1}^2) \; + \; \tau_k^2 \, h_{k_2+1}^2(y_2) \, \mathbf{1}(y_1 \in B_{k_1+1}^1) \; ,$$

where

$$h_{k_1+1}^1(y_1) = \begin{cases} 0, & \text{if} \quad y_1 \leq k_1 \, \ell_1 \; , \\ y_1 - k_1 \, \ell_1, & \text{if} \quad k_1 \, \ell_1 \leq y_1 \leq (k_1+1)\ell_1 \; , \\ \ell_1, & \text{if} \quad y_1 \geq (k_1+1)\ell_1 \; , \end{cases}$$

and analogously for $h_{k_2+1}^2$. We write

$$\begin{aligned} h^{int}(y_1, y_2) \quad &:= \quad \textstyle\sum_{k \in K} h^k(y_1, y_2) \\ &= \quad \textstyle\sum_{k \in K} \tau_k^1 \, \bar{h}_1^k(y_1, y_2) \; + \; \sum_{k \in K} \tau_k^1 \, \bar{h}_2^k(y_1, y_2) \end{aligned}$$

for a reform that combines all these simple tax reforms, where

$$\bar{h}_k^1(y_1, y_2) \quad := \quad h_{k_1+1}^1(y_1) \, \mathbf{1}(y_2 \in B_{k_2+1}^2)$$

and

$$\bar{h}_k^2(y_1, y_2) \quad := \quad h_{k_2+1}^2(y_2) \, \mathbf{1}(y_1 \in B_{k_1+1}^1) \; .$$

Simple tax reforms that change marginal tax rates on the boundary of $\mathcal{Y}_1 \times \mathcal{Y}_2$ are identified by the index k_1 (for the segment with $y_2 = 0$) or by the index k_2 (for the segment with $y_1 = 0$). The change in the marginal tax rate applying to the spouse with positive earnings is denoted either by $\tau_{k_1}^1$ or $\tau_{k_2}^2$. We write

$$h^{b1}(y_1, y_2) \quad := \quad \sum_{k_1 \in K_1} \tau_{k_1}^1 \bar{h}_{k_1}^{b1}(y_1, y_2)$$

with

$$\bar{h}^{b1}_{k_1}(y_1, y_2) \quad := \quad h^1_{k_1+1}(y_1)\mathbf{1}(y_2 = 0)$$

for a combination of simple tax reforms that affect the marginal tax rates of spouse 1, conditional on spouse 1 being the couple's single earner. We define the functions $\bar{h}^{b2}_{k_2}$ analogously. We now combine all these simple reforms into a reform direction $h_m : (y_1, y_2) \to h_m(y_1, y_2)$ with

$$h_m(y_1, y_2) \quad = \quad h^{int}(y_1, y_2) + h^{b1}(y_1, y_2) + h^{b2}(y_1, y_2)$$

Note that the reform direction h_m depends on how fine the partition of $\mathcal{Y}_1 \times \mathcal{Y}_2$ into m^2 squares is. Finally, to complete the construction, for $k = (k_1, k_2)$, we define the changes of marginal tax rates stipulated by h_m with respect to the function h that h_m approximates,

$$\frac{h(y_1^{k_1+1}, y_2^{k_2}) - h(y_1^{k_1}, y_2^{k_2})}{\ell_1} \quad =: \quad \tau^1_k$$

and

$$\frac{h(y_1^{k_1}, y_2^{k_2+1}) - h(y_1^{k_1}, y_2^{k_2})}{\ell_2} \quad =: \quad \tau^2_k \ ,$$

where $y_1^{k_1} := k_1 \ \ell_1$ and $y_2^{k_2} := k_2 \ \ell_2$. For later reference, note that this implies, in particular, that

$$\lim_{\ell_1 \to 0} \frac{h(y_1^{k_1+1}, y_2^{k_2}) - h(y_1^{k_1}, y_2^{k_2})}{\ell_1} \quad = \quad h_1(y_1^{k_1}, y_2^{k2}) \quad = \quad \tau^1_k$$

and

$$\lim_{\ell_2 \to 0} \frac{h(y_1^{k_1}, y_2^{k_2+1}) - h(y_1^{k_1}, y_2^{k_2})}{\ell_2} \quad = \quad h_2(y_1^{k_1}, y_2^{k2}) \quad = \quad \tau^2_k \ .$$

The relation between h and h_m. Consider the partition of the set $\mathcal{Y}_1 \times \mathcal{Y}_2$ into m^2 squares as described above. Thus, square $k = (k_1, k_2)$ has endpoints $(y_1^{k_1}, y_2^{k_2})$, $(y_1^{k_1+1}, y_2^{k_2})$, $(y_1^{k_1}, y_2^{k_2+1})$, and $(y_1^{k_1+1}, y_2^{k_2+1})$ where $y_1^{k_1} = k_1\ell_1$, $y_1^{k_1+1} = (k_1 + 1)\ell_1$, $y_2^{k_2} = k_2\ell_2$, $y_2^{k_2+1} = (k_2 + 1)\ell_2$. We now argue that if $h(0, 0) = h_m(0, 0)$, then the functions h and h_m coincide at the corners of any square. To see this, let $h(y_1^{k_1}, y_2^{k_2}) = h_m(y_1^{k_1}, y_2^{k_2})$. We now show that this implies that also

$$h_m(y_1^{k_1+1}, y_2^{k_2+1}) = h(y_1^{k_1+1}, y_2^{k_2+1})$$

Note that

$$\begin{aligned} h_m(y_1^{k_1+1}\ell_1, y_2^{k_2+1}) &= h_m(y_1^{k_1}, y_2^{k_2}) \;+\; h_m(y_1^{k_1+1}, y_2^{k_2+1}) - h_m(y_1^{k_1+1}, y_2^{k_2}) \\ &\quad + \; h_m(y_1^{k_1+1}, y_2^{k_2}) - h_m(y_1^{k_1}, y_2^{k_2}) \\ &= h(y_1^{k_1}, y_2^{k_2}) \;+\; \tau_2^{k_1+1,k_2}\,\ell_2 \;+\; \tau_1^{k}\,\ell_1 \\ &= h(y_1^{k_1}, y_2^{k_2}) \;+\; h(y_1^{k_1+1}, y_2^{k_2+1}) - h(y_1^{k_1+1}, y_2^{k_2}) \\ &\quad + \; h(y_1^{k_1+1}, y_2^{k_2}) - h(y_1^{k_1}, y_2^{k_2}) \\ &= h(y_1^{k_1+1}, y_2^{k_2+1})\,. \end{aligned}$$

A.2.2 Revenue implications of h_m.

We seek to characterize the Gateaux differential of tax revenue in direction h_m, $R_\tau(0, h_m)$. Using that

$$\begin{aligned} h_m(y_1, y_2) &= h^{int}(y_1, y_2) + h^{b1}(y_1, y_2) + h^{b2}(y_1, y_2) \\ &= \textstyle\sum_{k\in K} \tau_k^1\, \bar{h}_k^1(y_1, y_2) \;+\; \sum_{k\in K} \tau_k^2\, \bar{h}_k^2(y_1, y_2) \\ &\quad + \textstyle\sum_{k_1\in K_1} \tau_{k_1}^1 \bar{h}_{k_1}^{b1}(y_1, y_2) \quad + \sum_{k_2\in K_2} \tau_2^{k_2} \bar{h}_{k_2}^{b2}(y_1, y_2) \end{aligned}$$

and the linearity of the Gateaux differential, we obtain

$$\begin{aligned} R_\tau(0, h_m) &= \textstyle\sum_{k\in K} \tau_k^1\; R_\tau(0, \bar{h}_1^k) + \sum_{k\in K} \tau_k^2\; R_\tau(0, \bar{h}_2^k) \\ &\quad + \textstyle\sum_{k_1\in K_1} \tau_{k_1}^1 R_\tau(0, \bar{h}_{k_1}^{b1}) \quad + \sum_{k_2\in K_2} \tau_{k_2}^2 R_\tau(0, \bar{h}_{k_2}^{b2})\,. \end{aligned}$$

Large m. We have partitioned $\mathcal{Y}_1 \times \mathcal{Y}_2$ into m^2 squares. For $\ell_1 = \frac{\bar{y}_1}{m}$ and $\ell_2 = \frac{\bar{y}_2}{m}$ close to zero it follows from first order Taylor approximations that

$$R_\tau(0, \bar{h}^1_k) \quad \simeq \quad \ell_1\, \ell_2\, \mathcal{R}_1(y_1^{k_1} \mid y_2^{k_2})\ ,$$

where $y_1^{k_1} = \ell_1\ k_1$ and $y_2^{k_2} = \ell_2\ k_2$. The approximation is perfect in the limit as $m \to \infty$. Analogously, we have that

$$R_\tau(0, \bar{h}^2_k) \quad \simeq \quad \ell_1\, \ell_2\, \mathcal{R}_2(y_2^{k_2} \mid y_1^{k_1})\ ,$$

$$R_\tau(0, \bar{h}^{b1}_{k_1}) \quad \simeq \quad \ell_1\, \mathcal{R}^s_1(y_1^{k_1})\ ,$$

and

$$R_\tau(0, \bar{h}^{b2}_{k_2}) \quad \simeq \quad \ell_2\, \mathcal{R}^s_2(y_2^{k_2})\ .$$

Thus, for $\ell_1 = \frac{\bar{y}_1}{m}$ and $\ell_2 = \frac{\bar{y}_2}{m}$ close to zero and by the linearity of the Gateaux differential, we obtain

$$\begin{aligned} R_\tau(0, h_m) \ \simeq\ & \textstyle\sum_{k\in K} \tau^1_k\, \ell_1\, \ell_2\, \mathcal{R}_1(y_1^{k_1} \mid y_2^{k_2}) \quad + \sum_{k\in K} \tau^2_k\, \ell_1\, \ell_2\, \mathcal{R}_2(y_2^{k_2} \mid y_1^{k_1}) \\ & + \textstyle\sum_{k_1\in K_1} \tau^1_{k_1} \ell_1\, \mathcal{R}^s_1(y_1^{k_1}) \quad + \sum_{k_2\in K_2} \tau^2_{k_2} \ell_2\, \mathcal{R}^s_2(y_2^{k_2})\ . \end{aligned} \tag{22}$$

A.2.3 Proof that h_m is not a Parteo-improving direction

Rewriting h_m. For later reference note that

$$\textstyle\sum_{k\in K} \tau^1_k\, \bar{h}^1_k(y_1, y_2) \quad = \quad \displaystyle\sum_{k_2=0}^{K_2} \sum_{k_1=0}^{K_1} \tau^1_{k_1,k_2}\, \bar{h}^1_{k_1,k_2}(y_1, y_2)$$

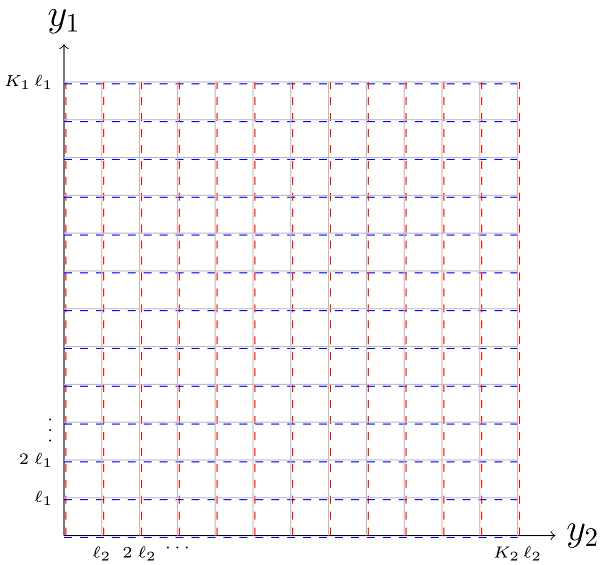
$$\textstyle\sum_{k\in K} \tau^k_2\, \bar{h}_k(y_1, y_2)^2 \quad = \quad \displaystyle\sum_{k_1=0}^{K_1} \sum_{k_2=0}^{K_2} \tau^2_{k_1,k_2}\, \bar{h}^2_{k_1,k_2}(y_1, y_2)$$

so that we can write

$$h_m(y_1, y_2) \;=\; \sum_{k_2 \in K_2} \sum_{k_1 \in K_1} \tau^1_{k_1,k_2}\, \bar{h}^1_{k_1,k_2}(y_1, y_2) \;+ \sum_{k_1 \in K_1} \sum_{k_2 \in K_2} \tau^2_{k_1,k_2}\, \bar{h}^2_{k_1,k_2}(y_1, y_2)$$

$$+ \sum_{k_1 \in K_1} \tau^1_{k_1}\, \bar{h}^{b1}_{k_1}(y_1, y_2) \;+ \sum_{k_2 \in K_2} \tau^2_{k_2}\, \bar{h}^{b2}_{k_2}(y_1, y_2)\,. \tag{23}$$

Figure A1: Summing over all k



Note: A point y_k on the grid has coordinates $y_{k1} = k_1\ \ell_1$ and $y_{k2} = k_2\ \ell_2$. In summing over all k, we can either go from one red line to the next and sum over all points, or go from one blue line to the next and sum over all points. The summation in (23) does the former for simple reforms that affect the marginal tax rates for spouse 1 and the latter for simple reforms that affect the marginal tax rates for spouse 2.

Focussing on corners. We seek to show that

$$R_\tau(0, h_m) \;-\; \max_{y \in \mathcal{Y}}\; h_m(y) \;\;\leq\;\; 0\,. \tag{24}$$

In doing so we can, without loss of generality, limit attention to the values that the function h_m achieves at the corners of the squares in Figure A1: The function h_m is piecewise linear and therefore achieves its maximum at a corner. Thus, to establish that (24) holds, we show that for all $k_1 \in K_1$ and all $k_2 \in K_2$

$$R_\tau(0, h_m) \;-h_m(y_1^{k_1}, y_2^{k_2}) \;\;\leq\;\; 0\,. \tag{25}$$

Using (22) and (23) inequality (25) can also be written as

$$A \;+\; B \;+\; C \;+\; D \quad \leq \quad 0\,, \tag{26}$$

where

$$A \;=\; \sum_{k_2\in K_2}\sum_{k_1\in K_1} \tau^1_{k_1,k_2}\;\ell_1\;\ell_2\;\mathcal{R}_1(y_1^{k_1}\mid y_2^{k_2})$$

$$-\sum_{k_2\in K_2}\sum_{k_1\in K_1} \tau^1_{k_1,k_2}\;\bar{h}^1_{k_1,k_2}(y_1^{k_1},y_2^{k_2})\;,$$

$$B \;=\; \sum_{k_1\in K_1}\sum_{k_2\in K_2} \tau^2_{k_1,k_2}\;\ell_1\;\ell_2\;\mathcal{R}_2(y_2^{k_2}\mid y_1^{k_1})$$

$$-\sum_{k_1\in K_1}\sum_{k_2\in K_2} \tau^2_{k_1,k_2}\;\bar{h}^2_{k_1,k_2}(y_1^{k_1},y_2^{k_2})\;,$$

$$C \;=\; \textstyle\sum_{k_1\in K_1}\tau^1_{k_1}\;\ell_1\;\mathcal{R}^s_1(y_1^{k_1}) \quad - \displaystyle\sum_{k_1\in K_1}\tau^1_{k_1}\;h^1_{k_1+1}(y_1^{k_1})\mathbf{1}(y_2=0)\;,$$

and

$$D \;=\; \textstyle\sum_{k_2\in K_2}\tau^2_{k_2}\ell_2\;\mathcal{R}^s_2(y_2^{k_2}) \quad - \displaystyle\sum_{k_2\in K_2}\tau^2_{k_2}\;h^2_{k_2+1}(y_2^{k_2})\mathbf{1}(y_1=0)\;.$$

The inequality in (25) follows form the observations that, under the conditions listed in Theorem 1,

$$A\;\leq\;0,\quad B\;\leq\;0,\quad C\;\leq\;0\quad\text{and}\quad D\;\leq\;0\,.$$

In the following we show that $A \;\leq\; 0$. $B,C,D\leq 0$ can be shown by a straightforward adaptation of the arguments used to establish that $A\leq 0$.

Lemma 4 *Under the conditions listed in Corollary 1,*

$$\sum_{k_2=0}^{K_2}\sum_{k_1=0}^{K_1}\tau_1^{k_1,k_2}\;\ell_1\;\ell_2\;\mathcal{R}(y_1^{k_1}\mid y_2^{k_2}) \;\;\leq\;\; \sum_{k_2=0}^{K_2}\sum_{k_1=0}^{K_1}\tau_1^{k_1,k_2}\;h_1^{k_1,k_2}(y_1^{k_1},y_2^{k_2})\;.$$

Proof. Fix some $k_2 \in \{0, 1, \ldots, K_2\}$. It suffices to show that for any such k_2,

$$\sum_{k_1=0}^{K_1} \tau^1_{k_1,k_2}\ \ell_1\ \ell_2\ \mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \ \leq\ \sum_{k_1=0}^{K_1} \tau^1_{k_1,k_2}\ h_1^{k_1,k_2}(y_1^{k_1}, y_2^{k_2})\ . \tag{27}$$

Step 1. Define the auxiliary function H_1

$$H_1(j_1) \quad := \quad \sum_{k_1=0}^{j_1} \tau^1_{k_1,k_2}\ h^1_{k_1,k_2}(y_1^{k_1}, y_2^{k_2})$$

and let $j_1^* \in$ argmax $H_1(j_1)$. Note that

$$\sum_{k_1=j_1^*}^{j^*} \tau^1_{k_1,k_2}\ h^1_{k_1,k_2}(y_1^{k_1}, y_2^{k_2}) \quad = \quad \tau^1_{j_1^*,k_2}\ \ell_1$$

and that

$$\tau^1_{j_1^*,k_2}\ell_1\ \ell_2\ \mathcal{R}(y_1^{j_1^*} \mid y_2^{k_2}) \quad \leq \quad \tau^1_{j_1^*,k_2}\ \ell_1$$

since, by assumption,

$$\ell_2\ \mathcal{R}(y_1^{j_1^*} \mid y_2^{k_2}) \quad \leq 1\ .$$

Thus we have that

$$\sum_{k_1=j_1^*}^{j^*} \tau^1_{k_1,k_2}\ell_1\ \ell_2\ \mathcal{R}(y_1^{j_1^*} \mid y_2^{k_2}) \quad \leq \quad \sum_{k_1=j_1^*}^{j^*} \tau^1_{k_1,k_2}\ h^1_{k_1,k_2}(y_1^{k_1}, y_2^{k_2})\ . \tag{28}$$

Step 2. We now argue that also

$$\sum_{k_1=j_1^*-1}^{j^*} \tau^1_{k_1,k_2}\ell_1\ \ell_2\ \mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \quad \leq \quad \sum_{k_1=j_1^*-1}^{j_1^*} \tau_1^{k_1,k_2}\ h^1_{k_1,k_2}(y_1^{k_1}, y_2^{k_2})\ . \tag{29}$$

Note that

$$\sum_{k_1=j_1^*-1}^{j_1^*} \tau^1_{k_1,k_2}\ h_1^{k_1,k_2}(y_1^{k_1}, y_2^{k_2}) \quad = \quad \sum_{k_1=j_1^*-1}^{j_1^*} \tau_1^{k_1,k_2}\ \ell_1 \quad \geq \quad 0$$

otherwise j_1^* would not be a maximizer of H_1. Also, by assumption,

$$\mathcal{R}(y_1^{j_1^*} \mid y_2^{k_2}) \quad \leq \quad \mathcal{R}(y_1^{j_1^*-1} \mid y_2^{k_2}) \ ,$$

so that

$$\sum_{k_1=j_1^*-1}^{j^*} \tau^1_{k_1,k_2} \ell_1 \ \ell_2 \ \mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \quad \leq \quad \mathcal{R}(y_1^{j_1^*} \mid y_2^{k_2}) \ell_2 \sum_{k_1=j_1^*-1}^{j_1^*} \tau^1_{k_1,k_2} \ \ell_1 \ .$$

Since, by assumption,

$$\ell_2 \ \mathcal{R}(y_1^{j_1^*} \mid y_2^{k_2}) \quad \leq 1 \ ,$$

(29) follows.

Step 3. A repeated application of this argument yields

$$\sum_{k_1=0}^{j^*} \tau^1_{k_1,k_2} \ell_1 \ \ell_2 \ \mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \quad \leq \quad \sum_{k_1=0}^{j_1^*} \tau^1_{k_1,k_2} \ h_1^{k_1,k_2}(y_1^{k_1}, y_2^{k_2}) \ . \tag{30}$$

Step 4. We now claim that, for any $z \geq j_1^* + 1$,

$$\sum_{k_1=j_1^*+1}^{z} \tau^1_{k_1,k_2} \ell_1 \ \ell_2 \ \mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \quad \leq \quad 0 \ . \tag{31}$$

Note that

$$\sum_{k_1=j_1^*+1}^{z} \tau^1_{k_1,k_2} \ h_1^{k_1,k_2}(y_1^{k_1}, y_2^{k_2}) \quad = \quad \sum_{k_1=j_1^*+1}^{z} \tau^1_{k_1,k_2} \ \ell_1 \quad \leq \quad 0 \tag{32}$$

otherwise j_1^* would not be a maximizer of H_1. In particular, this requires that $\tau_1^{j_1^*+1,k_2} \leq 0$ so that

$$\sum_{k_1=j_1^*+1}^{j_1^*+1} \tau_1^{k_1,k_2} \ell_1 \ \ell_2 \ \mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \quad \leq \quad 0 \ ,$$

using that, by assumption, $\mathcal{R}(y_1^{j_1^*+1} \mid y_2^{k_2}) \geq 0$. Now suppose that also $\tau_1^{j_1^*+2,k_2} \leq 0$, then, using that $\mathcal{R}(y_1^{j_1^*+2} \mid y_2^{k_2}) \geq 0$, it follows that also

$$\sum_{k_1=j_1^*+1}^{j_1^*+2} \tau_1^{k_1,k_2} \ell_1 \, \ell_2 \, \mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \quad \leq \quad 0$$

If, by contrast, $\tau^1_{j_1^*+2,k_2} \geq 0$, then, using $\mathcal{R}(y_1^{j_1^*+2} \mid y_2^{k_2}) \leq \mathcal{R}(y_1^{j_1^*+1} \mid y_2^{k_2})$

$$\sum_{k_1=j_1^*+1}^{j_1^*+2} \tau_1^{k_1,k_2} \ell_1 \, \ell_2 \, \mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \;\; \leq \;\; \mathcal{R}(y_1^{j_1^*+1} \mid y_2^{k_2}) \ell_2 \sum_{k_1=j_1^*+1}^{j_1^*+2} \tau^1_{k_1,k_2} \, \ell_1 \; ,$$

where the right-hand side expression is negative by (32). Again, it follows that

$$\sum_{k_1=j_1^*+1}^{j_1^*+2} \tau^1_{k_1,k_2} \, \ell_1 \, \ell_2 \, \mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \quad \leq \quad 0 \; .$$

A repeated application of this argument establishes (31). In particular, this implies that

$$\sum_{k_1=j_1^*+1}^{K_1} \tau^1_{k_1,k_2} \ell_1 \, \ell_2 \, \mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \quad \leq \quad 0 \; . \tag{33}$$

Step 5. The inequality in (27) now follows from combining (30) and (33). □

A.2.4 Proof that h is not a Pareto-improving direction

Suppose that, in contrast to what we seek to show,

$$R_\tau(0, h) \quad - \quad \max_{y \in \mathcal{Y}} \; h(y) \quad > \quad 0 \; .$$

Since

$$R_\tau(0, h) \quad = \quad \lim_{m \to \infty} \quad R_\tau(0, h_m)$$

and since for every $\epsilon > 0$, there is $\hat{m}(\epsilon)$ so that $m > \hat{m}(\epsilon)$ implies that, for all $y \in \mathcal{Y}$,

$$\mid h_m(y) - h(y) \mid \quad \leq \quad \epsilon \; ,$$

this implies that, there exists $m > \hat{m}(\epsilon)$ so that

$$R_\tau(0, h_m) \quad - \quad \max_{y \in \mathcal{Y}} \ h_m(y) \quad > \quad 0 \ .$$

This contradicts (24). Thus, the assumption that h is a Pareto-improving direction has led to a contradiction and must be false.

A.3 Proof of Theorem 3

We only show that (4) holds. The proof that also (5) holds can be obtained along the same lines.

Let h be a function that is continuous on the bounded domain $\mathcal{Y}$ and which is hence differentiable almost everywhere. We approximate reform direction h by a collection of simple tax reforms as described in the proof of Theorem 2, see section A.2. As a first step, we seek to characterize the Gateaux differential of tax revenue in direction h_m, $R_\tau(0, h_m)$. Using that

$$\begin{aligned} h_m(y_1, y_2) &= h^{int}(y_1, y_2) + h^{b1}(y_1, y_2) + h^{b2}(y_1, y_2) \\ &= \textstyle\sum_{k \in K} \tau_1^k \ h_1^k(y_1, y_2) \ + \ \sum_{k \in K} \tau_2^k \ h_2^k(y_1, y_2) \\ &\quad + \textstyle\sum_{k_1 \in K_1} \tau_1^{k_1} h^{bk_1}(y_1) \mathbf{1}(y_2 = 0) \\ &\quad + \textstyle\sum_{k_2 \in K_2} \tau_2^{k_2} h^{bk_2}(y_2) \mathbf{1}(y_1 = 0) \end{aligned}$$

and the linearity of the Gateaux differential, we obtain

$$\begin{aligned} R_\tau(0, h_m) &= \textstyle\sum_{k \in K} \tau_1^k \ R_\tau(0, h_1^k) + \sum_{k \in K} \tau_2^k \ R_\tau(0, h_2^k) \\ &\quad + \textstyle\sum_{k_1 \in K_1} \tau_1^{k_1} R_\tau(0, h^{bk_1} \mathbf{1}(y_2 = 0)) \\ &\quad + \textstyle\sum_{k_2 \in K_2} \tau_2^{k_2} R_\tau(0, h^{bk_2} \mathbf{1}(y_1 = 0)) \ . \end{aligned}$$

Taking limits. We have partitioned $\mathcal{Y}_1 \times \mathcal{Y}_2$ into m^2 squares. We now take limits and let $m \to \infty$.

Note that for $\ell_1 = \frac{\bar{y}_1}{m}$ and $\ell_2 = \frac{\bar{y}_2}{m}$ close to zero it follows from first order Taylor approximations that

$$R_\tau(0, h_1^k) \quad \simeq \quad \ell_1 \ \ell_2 \ \mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \ ,$$

where $y_1^{k_1} = \ell_1\ k_1$ and $y_2^{k_2} = \ell_2\ k_2$. The approximation is perfect in the limit as $m \to \infty$. Analogously, we have that

$$R_\tau(0, h_2^k) \quad \simeq \quad \ell_1\ \ell_2\ \mathcal{R}(y_2^{k_2} \mid y_1^{k_1})\ ,$$

$$R_\tau(0, h^{bk_1}\mathbf{1}(y_2 = 0)) \quad \simeq \quad \ell_1\ \mathcal{R}(y_1^{k_1} \mid 0)\ ,$$

and

$$R_\tau(0, h^{bk_2}\mathbf{1}(y_1 = 0)) \quad \simeq \quad \ell_2\ \mathcal{R}(y_2^{k_2} \mid 0)\ .$$

Thus, for $\ell_1 = \frac{\bar{y}_1}{m}$ and $\ell_2 = \frac{\bar{y}_2}{m}$ close to zero and by the linearity of the Gateaux differential, we obtain

$$\begin{aligned} R_\tau(0, h_m) \ \simeq\ & \textstyle\sum_{k \in K} \tau_1^k\ \ell_1\ \ell_2\ \mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) + \sum_{k \in K} \tau_2^k\ \ell_1\ \ell_2\ \mathcal{R}(y_2^{k_2} \mid y_1^{k_1}) \\ & + \textstyle\sum_{k_1 \in K_1} \tau_1^{k_1} \ell_1\ \mathcal{R}(y_1^{k_1} \mid 0) \\ & + \textstyle\sum_{k_2 \in K_2} \tau_2^{k_2} \ell_2\ \mathcal{R}(y_2^{k_2} \mid 0)\ . \end{aligned}$$

and

$$\begin{aligned} & \lim_{m \to \infty} \ R_\tau(0, h_m) = \\ & \lim_{m \to \infty} \ \textstyle\sum_{k \in K} \tau_1^k\ (y_1^{k_1+1} - y_1^{k_1})\ (y_2^{k_2+1} - y_2^{k_2})\ \mathcal{R}(y_1^{k_1} \mid y_2^{k_2}) \\ & + \lim_{m \to \infty} \ \textstyle\sum_{k \in K} \tau_2^k\ \ (y_1^{k_1+1} - y_1^{k_1})\ (y_2^{k_2+1} - y_2^{k_2})\ \mathcal{R}(y_2^{k_2} \mid y_1^{k_1}) \\ & + \lim_{m \to \infty} \ \textstyle\sum_{k_1 \in K_1} \tau_1^{k_1}\ (y_1^{k_1+1} - y_1^{k_1})\ \mathcal{R}(y_1^{k_1} \mid 0) \\ & + \lim_{m \to \infty} \ \textstyle\sum_{k_2 \in K_2} \tau_2^{k_2}\ (y_2^{k_2+1} - y_2^{k_2})\ \mathcal{R}(y_2^{k_2} \mid 0)\ . \end{aligned}$$

All the expressions on the right-hand side are Riemanns sums that converge to the corresponding Riemann integral as $m \to \infty$. Therefore,

$$\begin{aligned} R_\tau(0, h) \ &= \ \lim_{m \to \infty} \ R_\tau(0, h_m) \\ &= \ \textstyle\int_{\mathcal{Y}_1 \times \mathcal{Y}_1} h_1(y_1, y_2)\ \mathcal{R}(y_1 \mid y_2)\ dy_2\ dy_1 \\ & \quad + \textstyle\int_{\mathcal{Y}_1 \times \mathcal{Y}_1} h_2(y_1, y_2)\ \mathcal{R}(y_2 \mid y_1)\ dy_2\ dy_1 \\ & \quad + \textstyle\int_{\mathcal{Y}_1} h_1(y_1, 0)\ \mathcal{R}(y_1 \mid 0)\ dy_1 \\ & \quad + \textstyle\int_{\mathcal{Y}_2} h_2(0, y_2)\ \mathcal{R}(y_2 \mid 0)\ dy_2\ . \end{aligned}$$

A.4 Proof of Corollary 4

Starting from the characterization of $W_\tau(0,h)$ in Theorem 3, we can substitute

$$h_1(y_1,0)+\int_0^{y_2} h_{12}(y_1,s_2)\,ds_2$$

for $h_1(y_1,y_2)$ and

$$h_2(0,y_2)+\int_0^{y_1} h_{21}(s_1,y_2)\,ds_1$$

for $h_2(y_1,y_2)$. This yields

$$\begin{aligned} W_\tau(0,h) &= \textstyle\int_{\mathcal{Y}}\Big(\int_0^{y_2} h_{12}(y_1,s_2)\,ds_2\Big)\mathcal{W}_1(y_1\mid y_2)dy \\ &\quad+\textstyle\int_{\mathcal{Y}}\Big(\int_0^{y_1} h_{21}(s_1,y_2)\,ds_1\Big)\mathcal{W}_2(y_2\mid y_1)dy \\ &\quad+\textstyle\int_{\mathcal{Y}_1} h_1(y_1,0)\,\Big(\mathcal{W}_1^s(y_1)+\int_{\mathcal{Y}_2}\mathcal{W}_1(y_1\mid y_2)dy_2\Big)dy_1 \\ &\quad+\textstyle\int_{\mathcal{Y}_2} h_2(0,y_2)\,\Big(\mathcal{W}_2^s(y_2)+\int_{\mathcal{Y}_1}\mathcal{W}_2(y_2\mid y_1)dy_1\Big)dy_2\ , \end{aligned}$$

One can use integration by parts to verify that

$$\textstyle\int_{\mathcal{Y}_2}\Big(\int_0^{y_2} h_{12}(y_1,s_2)\,ds_2\Big)\mathcal{W}_1(y_1\mid y_2)dy_2 = \int_{\mathcal{Y}_2} h_{12}(y_1,y_2)\Big(\int_{y_2}^{\bar{y}_2}\mathcal{W}_1(y_1\mid s_2)ds_2\Big)dy_2$$

and

$$\textstyle\int_{\mathcal{Y}_1}\Big(\int_0^{y_1} h_{21}(s_1,y_2)\,ds_1\Big)\mathcal{W}_2(y_2\mid y_1)dy_1 = \int_{\mathcal{Y}_1} h_{21}(y_1,y_2)\Big(\int_{y_1}^{\bar{y}_1}\mathcal{W}_2(y_2\mid s_1)ds_1\Big)dy_1$$

Thus,

$$\begin{aligned} W_\tau(0,h) &= \textstyle\int_{\mathcal{Y}} h_{12}(y_1,y_2)\Big(\int_{y_2}^{\bar{y}_2}\mathcal{W}_1(y_1\mid s_2)ds_2\Big)dy \\ &\quad+\textstyle\int_{\mathcal{Y}} h_{21}(y_1,y_2)\Big(\int_{y_1}^{\bar{y}_1}\mathcal{W}_2(y_2\mid s_1)ds_1\Big)dy \\ &\quad+\textstyle\int_{\mathcal{Y}_1} h_1(y_1,0)\,\Big(\mathcal{W}_1^s(y_1)+\int_{\mathcal{Y}_2}\mathcal{W}_1(y_1\mid y_2)dy_2\Big)dy_1 \\ &\quad+\textstyle\int_{\mathcal{Y}_2} h_2(0,y_2)\,\Big(\mathcal{W}_2^s(y_2)+\int_{\mathcal{Y}_1}\mathcal{W}_2(y_2\mid y_1)dy_1\Big)dy_2\ , \end{aligned}$$

Using that $h_{12}(y)=h_{21}(y)$, for all $y\in\mathcal{Y}$ yields the expression in Corollary 4.

A.5 Behavioral responses and tax revenue

A couple's choice problem. Given a tax reform (τ, h), we denote by

$$y^{*d}(\omega, \tau, h) \quad \in \quad \text{argmax}_y \, C^1(y) - k_1(y_1, \omega_1) \; - \; k_2(y_2, \omega_2) \; , \tag{34}$$

the couples' optimal earnings choices conditional on being a dual-earner couple. We denote by $v^d(\omega, \tau, h)$ the indirect utility associated with a solution to the problem in (34). Analogously, we write

$$y_i^{*s}(\omega_i, \tau, h) \quad \in \quad \text{argmax}_y \, C^1(y_i, 0) - k_i(y_i, \omega_i) \tag{35}$$

for the optimal choice conditional on the spouse with index $i \in \{1, 2\}$ being the single earner and $v_i^s(\omega_i, \tau, h)$ for the indirect utility associated with a solution to the problem in (35). A couple will choose to be a dual earner couple when the resulting payoff of $v^d(\omega, \tau, h) - \varphi_1 - \varphi_2$. exceeds the payoff from its alternative options. Alternatives are that only spouse 1 or only spouse 2 exerts productive effort with payoffs that are, respectively, given by $v_1^s(\omega_1, \tau, h) - \varphi_1$ and $v_2^s(\omega_2, \tau, h) - \varphi_2$. Thus, for dual earner couples,

$$v^d(\omega, \tau, h) - v_1^s(\omega_1, \tau, h) \; := \; \hat{\varphi}_2(\omega, \tau, h) \quad \geq \quad \varphi_2 \; , \quad \text{and}$$

$$v^d(\omega, \tau, h) - v_2^s(\omega_2, \tau, h) \; := \; \hat{\varphi}_1(\omega, \tau, h) \quad \geq \quad \varphi_1 \; .$$

It also has to be true that $v^d(\omega, \tau, h) \geq \varphi_1 + \varphi_2$. Otherwise, the couple would be better off with no earnings at all. A couple will choose spouse i as its single-earner when this yields a payoff that dominates what is achievable as a dual earner couple or with spouse $-i$ as the single earner. The latter implies

$$v_i^s(\omega_i, \tau, h) - \varphi_i \quad \geq \quad v_{-i}^s(\omega_{-i}, \tau, h) - \varphi_{-i} \; .$$

The payoff with i as the single earner must also dominate having no earnings at all,

$$v_i^s(\omega_i, \tau, h) \; := \; \hat{\hat{\varphi}}_i(\omega_i, \tau, h) \quad \geq \quad \varphi_i \; .$$

Finally, for couples with no earnings, it must be true that, for both i, $\varphi_i \geq \hat{\hat{\varphi}}(\omega_i, \tau, h)$ and $v^d(\omega, \tau, h) \leq \varphi_1 + \varphi_2$. The following Proposition states implications of these conditions under a set of additional assumptions, see Figure A2 for an illustration.

Proposition 4 *Suppose that the following premises are satisfied:*

i) For all i, $y_i^{*s}(\omega_i, \tau, h) > y_i^{*d}(\omega, \tau, h)$.

ii) For all i, $\lim_{\omega_i \to 0} v_i^s(\omega_i, \tau, h) = 0$ and $\lim_{\omega_{-i} \to 0} v^d(\omega, \tau, h) = v_i^s(\omega_i, \tau, h)$.

Then, for any $\omega >> 0$, and for any $i = 1, 2$,

$$\hat{\hat{\varphi}}(\omega_i, \tau, h) \quad > \quad \hat{\varphi}_i(\omega, \tau, h) \; . \tag{36}$$

Hence, $y^*(\omega, \varphi, \tau, h) =$

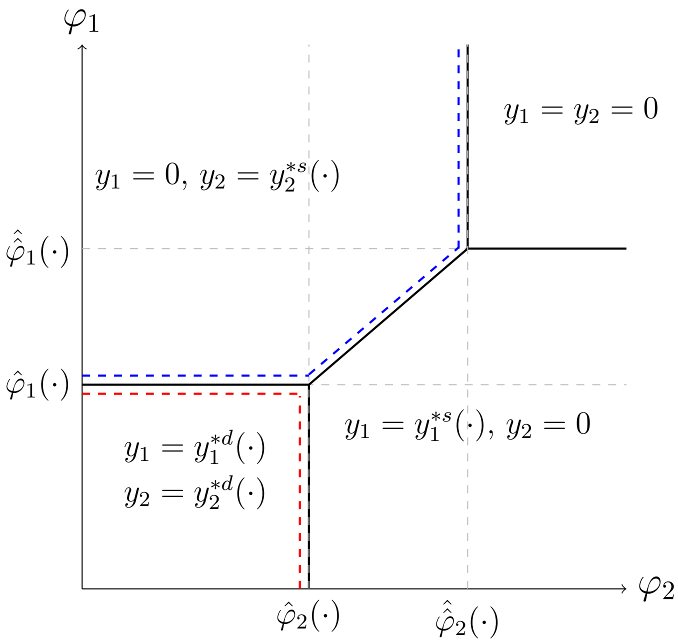
$$\begin{cases} y^{d*}(\cdot), & \text{if } \; \varphi_1 < \hat{\varphi}_1(\cdot) \; \text{ and } \; \varphi_2 < \hat{\varphi}_2(\cdot) \; , \\ (y_1^{*s}(\cdot), 0), & \text{if } \; \varphi_1 < \min\{\Delta v^s(\cdot) + \varphi_2, \hat{\hat{\varphi}}_1(\cdot)\} \; \text{ and } \; \varphi_2 > \hat{\varphi}_2(\cdot) \; , \\ (0, y_2^{*s}(\cdot)), & \text{if } \; \varphi_1 > \hat{\varphi}_1(\cdot) \; \text{ and } \; \varphi_2 < \min\{-\Delta v^s(\cdot) + \varphi_1, \hat{\hat{\varphi}}_2(\cdot)\} \; , \\ (0, 0), & \text{if } \; \varphi_1 > \hat{\hat{\varphi}}_1(\cdot) \; \text{ and } \; \varphi_2 > \hat{\hat{\varphi}}_2(\cdot) \; . \end{cases} \tag{37}$$

where $\Delta v^s(\omega, \tau, h) := v_1^s(\omega_1, \tau, h) - v_2^s(\omega_2, \tau, h)$.

A formal proof of Proposition 4 can be found at the end of this Section. The first premise of Proposition 4 is fulfilled under a progressive tax system that gives rise to positive jointness. Under such a system, the marginal tax rate on the earnings of one spouse is increasing in the earnings of the other spouse. Thus, when spouse i is a single earner her marginal tax rate is lower as compared to any other constellation in which she is part of a dual earner couple. Therefore, y_i^* is decreasing in y_{-i}. The second premise appears natural on the assumption that y_i^* goes to zero as ω_i goes to zero and that, in this case, also i's contribution to the couple's payoff goes to zero.[40] Under these assumptions, the cutoff types $\hat{\hat{\varphi}}(\omega_i, \tau, h)$ and $\hat{\varphi}_i(\omega, \tau, h)$ are ordered as follows: Spouse i has positive earnings – whatever the earnings choice for spouse $-i$ – when $\varphi < \hat{\varphi}$. Likewise, spouse i has no earnings when $\varphi > \hat{\hat{\varphi}}$, again, irrespectively

[40]It follows form the Spence-Mirrlees single-crossing condition that y_i^* is a decreasing function of ω_i. To avoid lenghty formal arguments, we refrain from stating further assumptions on the effort cost function and the tax function that imply premise ii).

Figure A2: A couple's discrete choice problem



Note: The figure shows how the earnings choices of the spouses in a couple with productivities $\omega = (\omega_1, \omega_2)$ depend on their fixed costs $\varphi = (\varphi_1, \varphi_2)$. For fixed costs in the lower left rectangle both spouses generate income. For fixed costs in the upper right area, the couple has no earnings at all. For fixed costs in the upper left area, spouse 1 is the single earner. For fixed costs in the lower right area, spouse 2 is the single earner.

of whether $y_{-i} = 0$ or $y_{-i} > 0$. For intermediate fixed costs, however, whether or not spouse i has positive earnings depends on whether the other spouse has $y_{-i} = 0$ or $y_{-i} > 0$.

In the sequel, we will characterize behavioral responses to taxation and, ultimately, derive sufficient statistics formulas for the conditional revenue functions. In doing so we will proceed under the assumption that, in the status quo, i.e. for $\tau = 0$, $\hat{\hat{\varphi}}(\omega_i, \tau, h) \quad > \quad \hat{\varphi}_i(\omega, \tau, h)$ holds for all i and all ω. Imposing this assumption eases the exposition. Otherwise we would have to get into a distinction of various subcases.

Behavioral responses to a simple tax reform: dual earner couples. Consider a simple tax reform (τ, h) so that, conditional on $y_1 \in B_1(y_{1s}, \ell_1)$, the couples' tax burden increases by $\tau h_2(y_2)$, where

$$h_2(y_2) \quad = \quad \begin{cases} y_2 - y_{2s}, & \text{if} \quad y_2 \in B_2(y_{2s}, \ell_2)\ , \\ \ell_2, & \text{if} \quad y_2 \geq y_{2s} + \ell_2\ , \\ 0, & \qquad \text{otherwise}\ . \end{cases}$$

Let $\mathcal{Y}_{h_2} = \{y \mid h_2(y) \neq 0\}$ be the "treated" range of incomes and denote by

$$\Omega^d(\tau, h_2) = \{\omega \mid y^{*d}(\omega, \tau, h_2) \in \mathcal{Y}_{h_2}\}$$

the corresponding set of productive abilities.[41] Obviously, there will be behavioral responses to infinitesimal changes of τ only for couples with $\omega \in \Omega(\tau, h_2)$. For all other couples, $\mathcal{Y}(h_2)$ is a dominated region and what happens there has no consequence for their choices.

Now consider a couple with $\omega \in \Omega(\tau, h_2)$. The couples' behavioral response at the intensive margin is given by

$$y^{*d}_{\tau}(\omega, \tau, h_2) \quad = \quad (y^{*d}_{1,\tau}(\omega, \tau, h), y^{*d}_{2,\tau}(\omega, \tau, h)) \ ,$$

i.e. by the derivative of its optimal earnings choice conditional on being a dual-earner couple.[42] Since indirect utility $v^d(\omega, \tau, h)$ goes down, some couples give up on having two earners. This is the behavioral response at the extensive margin. It is illustrated by the dashed red lines in Figure A2. They indicate that both $\hat{\varphi}_1(\omega, \tau, h)$ and $\hat{\varphi}_2(\omega, \tau, h)$ shrink in response to the reform. Couples with fixed costs between the black and red dashed lines move out of the set of dual-earner couples and into the set of single-earner couples.

Single earner couples. Consider a simple tax reform so that, conditional on $y_2 = 0$, the couples' tax burden increases by $\tau h_1(y_1)$, where

$$h_1(y_1) \quad = \quad \begin{cases} y_1 - y_{1s}, & \text{if} \quad y_1 \in B_1(y_{1s}, \ell_1) \ , \\ \ell_1, & \text{if} \quad y_1 \geq y_{1s} + \ell_1 \ , \\ 0, & \qquad \text{otherwise} \ . \end{cases}$$

[41]We allow for the possibility that optimal choices are not unique. The statement $y^{*d}(\omega, \tau, h_2) \in \mathcal{Y}_{h_2}$ is a shorthand for: There is an element y of $\text{argmax}_y \ C^1(y) - k_1(y_1, \omega_1) - k_2(y_2, \omega_2)$ that belongs to $\mathcal{Y}_{h_2}$.

[42]Left-sided and right-sided derivatives of y^{*d} may not be identical when there are multiple optimal choices conditional on being a dual earner couple. Clarifying whether the left-sided or the right-sided derivative captures the behavioral responses to a tax reform then requires a more detailed analysis.

Again, let $\mathcal{Y}_{h_1} = \{y \mid h_1(y) \neq 0\}$ be the "treated" range of incomes and denote by

$$\Omega_i^s(\tau, h_i) = \{\omega \mid y_i^{*s}(\omega, \tau, h) \in \mathcal{Y}_{h_1}\}$$

the corresponding set of productive abilities. The behavioral response at the intensive margin is given by $y_{1,\tau}^{*s}(\omega_i, \tau, h_1)$. i.e. by the derivative of the optimal earnings choice conditional on spouse 1 being the couple's single earner. When indirect utility $v_1^s(\omega_i, \tau, h)$ goes down some couples give up on having spouse 1 as the single earner. Depending on their fixed costs, some move into being a dual earner couple, others assign the role of the single earner now to spouse $-i$ and some give up on having earnings at all. These behavioral responses at the extensive margin are illustrated by the dashed blue lines in Figure A2.

Implications for tax revenue. We now clarify the repercussion that behavioral responses have for tax revenue. As will become clear, even though the earnings function that has been characterized in Proposition 4 exhibits discontinuous jumps at the cutoff levels for fixed costs, tax revenue is a differentiable function of the reform intensity τ if the distribution of fixed costs is "well-behaved." Consider a tax reform h and denote by $r(\omega, \tau, h)$ the change in tax revenue that is due to the set of couples who all have productive abilities ω, but may differ in their fixed costs of generating income. Thus,

$$\begin{aligned} r(\omega, \tau, h) + r^0(\omega) \quad &= \quad \sigma^d(\omega, \tau, h)\ T^1(y^{d*}(\omega, \tau, h), \tau, h) \\ &\quad + \sigma_1^s(\omega, \tau, h)\ T^1(y_1^{s*}(\omega_1, \tau, h), 0, \tau, h) \\ &\quad + \sigma_2^s(\omega, \tau, h)\ T^1(0, y_2^{s*}(\omega_2, \tau, h), \tau, h) \end{aligned}$$

where, for any level of y, $T^1(y, \tau, h) := T^0(y) + \tau h(y)$, $r^0(\omega)$ denotes revenues in the status quo and

$$\sigma^d(\omega, \tau, h) \quad = \quad \int_0^{\hat{\varphi}_1(\omega,\tau,h)} \int_0^{\hat{\varphi}_2(\omega,\tau,h)} f^{\varphi}(\varphi_1, \varphi_2 \mid \omega)\ d\varphi_2\ d\varphi_1\ ,$$

is the fraction of dual earner couples among those with productive ability ω, and the density f^{φ} characterizes the joint distribution of φ_1 and φ_2, conditional on the couples' productivities being ω. The fractions of couples where only spouse 1 or only spouse 2 has positive earnings are analogously defined. The derivative of $r(\omega, \tau, h)$ with respect to τ shows how these behavioral responses affect tax revenue. Specifically,

$$r_\tau(\omega, \tau, h) = r_\tau^I(\omega, \tau, h) + r_\tau^E(\omega, \tau, h) \ , \tag{38}$$

where $r_\tau^I(\omega, \tau, h)$ captures the revenue implications of intensive margin responses and $r_\tau^E(\omega, \tau, h)$ captures the revenue implications of extensive margin responses.[43] An aggregation over all possible combinations of ω gives us the change in tax revenue due to the reform (τ, h):

$$R(\tau, h) = \mathbb{E}_\omega[r(\omega, \tau, h) \ \mathbf{1}(\omega \in \Omega^d(\tau, h))]$$

Upon taking a derivative with respect to τ and evaluating the resulting expression at the status quo, we obtain the marginal effect on tax revenue that is associated with reform direction h. Proposition 5 provides a characterization.

Proposition 5

$$R_\tau(0, h) = \mathbb{E}_\omega[r_\tau(\omega, 0, h) \ \mathbf{1}(\omega \in \Omega^d(0, h))]$$

The Proposition – see the end of this section for a formal proof – says that a reform direction's marginal effect on tax revenue is only due those couples who, in the status quo, make choices in the "treated region" i.e. in the range of incomes with $h(y) \neq 0$. This observation greatly simplifies the characterization of conditional

[43]More specifically,

$$\begin{aligned} r_\tau^I(\omega, \tau, h) \ &= \ \sigma^d(\omega, \tau, h)\Big(T^1_{y_1}(y^{d*}(\cdot), \tau, h) y^{d*}_{1,\tau}(\cdot) + T^1_{y_2}(y^{d*}(\cdot), \tau, h) y^{d*}_{2,\tau}(\cdot) + h(y^{d*}(\cdot))\Big) \\ &+\sigma^s_1(\omega, \tau, h)\Big(T^1_{y_1}(y^{s*}_1(\cdot), 0, \tau, h) y^{s*}_{1,\tau}(\cdot) + h(y^{s*}_1(\cdot))\Big) \\ &+\sigma^s_2(\omega, \tau, h)\Big(T^1_{y_2}(0, y^{s*}_2(\cdot), \tau, h) y^{s*}_{2,\tau}(\cdot) + h(y^{s*}_2(\cdot))\Big) \qquad \text{and} \end{aligned} \tag{39}$$

$$\begin{aligned} r_\tau^E(\omega, \tau, h) \ &= \ \sigma^d_\tau(\omega, \tau, h) \ T^1(y^{d*}(\omega, \tau, h), \tau, h) \\ &+\sigma^s_{1,\tau}(\omega, \tau, h) \ T^1(y^{s*}_1(\omega_1, \tau, h), 0, \tau, h) \\ &+\sigma^s_{2,\tau}(\omega, \tau, h) \ T^1(0, y^{s*}_2(\omega_2, \tau, h), \tau, h) \ . \end{aligned} \tag{40}$$

revenue functions in the subsequent section.[44]

A.5.1 Proofs

Proof of Proposition 4. We show that under the premises of the Proposition, we have, for all $\omega >> 0$.

$$\hat{\varphi}_1(\omega, \tau, h) \; < \; \hat{\hat{\varphi}}_1(\omega_1, \tau, h) \quad \text{and} \quad \hat{\varphi}_2(\omega, \tau, h) \; < \; \hat{\hat{\varphi}}_2(\omega_2, \tau, h) \; .$$

We only show that $\hat{\varphi}_1(\cdot) \; < \; \hat{\hat{\varphi}}_1$, or, equivalently, that

$$v^d(\omega, \tau, h) - v_2^s(\omega_2, \tau, h) \quad < \quad v_1^s(\omega_1, \tau, h) \; .$$

By assumption,

$$\lim_{\omega_2 \to 0} v^d(\omega, \tau, h) - v_2^s(\omega_2, \tau, h) \quad = \quad v_1^s(\omega_1, \tau, h) \; .$$

It therefore suffices to show that $v^d(\omega, \tau, h) - v_2^s(\omega_2, \tau, h)$ is a decreasing function of ω_2. An application of the Envelope Theorem reveals that

$$\frac{\partial}{\partial \omega_2} v^d(\omega, \tau, h) \; = \; -k_{2,2}(y_2^{*d}(\cdot), \omega_2)$$

and

$$\frac{\partial}{\partial \omega_2} v_2^s(\omega_2, \tau, h) \; = \; -k_{2,2}(y_2^{*s}(\cdot), \omega_2)$$

where, by assumption, $y_2^{*s}(\cdot) > y_2^{*d}(\cdot)$. Hence,

$$\begin{aligned} \tfrac{\partial}{\partial \omega_2} \Big(v^d(\omega, \tau, h) - v_2^s(\omega_2, \tau, h) \Big) \quad &= \quad k_{2,2}(y_2^{*s}(\cdot), \omega_2) - k_{2,2}(y_2^{*d}(\cdot), \omega_2) \\ &= \quad \textstyle\int_{y_2^{*d}(\cdot)}^{y_2^{*s}(\cdot)} k_{2,12}(y, \omega_2) \; dy \; , \end{aligned}$$

where $k_{2,12} < 0$ by the Spence-Mirrless single crossing property.

[44] To be clear, the set $\Omega^d(\tau, h)$ will generally depend on τ. Possibly a reform creates dominated regions in space of incomes and those regions may become larger as the reform intensifies, so that $\tau' > \tau$ implies that $\Omega^d(\tau', h)$ is a strict subset of $\Omega^d(\tau, h)$. As formally demonstrated in the proof of Proposition 5, this force is, however, inconsequential, at $\tau = 0$

Proof of Proposition 5. Consider a reform direction h with $\mathcal{Y}_h = \{y \mid h(y) \neq 0\}$ in the interior of $\mathcal{Y}$ and with

$$\Omega^d(\tau, h) \quad = \quad \{\omega \mid y^{d*}(\omega, \tau, h) \in \mathcal{Y}_h\}$$

Let

$$\underline{\omega}_1(\tau, h) \quad = \quad \min\{\omega_1 \mid \exists \omega_2 \text{ s.t } (\omega_1, \omega_2) \in \Omega^d(\tau, h)\}$$

and

$$\overline{\omega}_1(\tau, h) \quad = \quad \max\{\omega_1 \mid \exists \omega_2 \text{ s.t } (\omega_1, \omega_2) \in \Omega^d(\tau, h)\} \ .$$

For given $\omega_1 \in \Omega_1^d(\tau, h) := [\underline{\omega}_1(\tau, h), \overline{\omega}_1(\tau, h)]$, let

$$\underline{\omega}_2(\omega_1, \tau, h) \quad = \quad \min\{\omega_2 \mid (\omega_1, \omega_2) \in \Omega^d(\tau, h)\}$$

and

$$\overline{\omega}_2(\omega_1, \tau, h) \quad = \quad \max\{\omega_2 \mid (\omega_1, \omega_2) \in \Omega^d(\tau, h)\} \ .$$

Armed with this notation we can write

$$R(\tau, h) \quad = \quad \mathbb{E}_\omega[r(\omega, \tau, h)\mathbf{1}(\omega \in \Omega^d(\tau, h))]$$

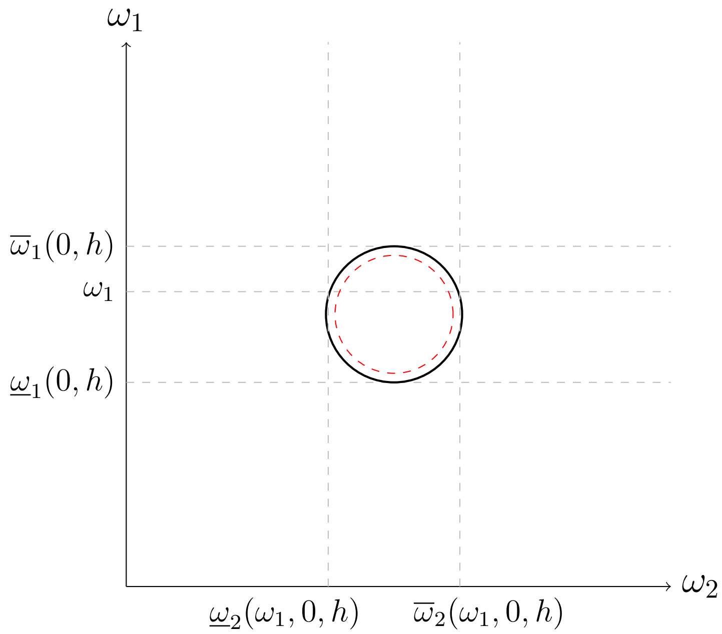
$$\int_{\underline{\omega}_1(\tau,h)}^{\overline{\omega}_1(\tau,h)} \int_{\underline{\omega}_2(\omega_1,\tau,h)}^{\overline{\omega}_2(\omega_1,\tau,h)} r(\omega, \tau, h) f_{\omega_2}(\omega_2 \mid \omega_1) d\omega_2 \ f_{\omega_1}(\omega_1) d\omega_1 \ ,$$

where f_{ω_1} is the marginal distribution of ω_1 associated with F_θ and $f_{\omega_2}(\cdot \mid \omega_1)$ is the distribution of ω_2 conditional on ω_1. Differentiating with respect to τ and repeatedly applying of Leibnitz' rule yields

$$R_\tau(\tau, h) \quad = \quad \int_{\underline{\omega}_1(\tau,h)}^{\overline{\omega}_1(\tau,h)} \int_{\underline{\omega}_2(\omega_1,\tau,h)}^{\overline{\omega}_2(\omega_1,\tau,h)} r_\tau(\omega, \tau, h) \ f_{\omega_2}(\omega_2 \mid \omega_1) d\omega_2 \ f_{\omega_1}(\omega_1) d\omega_1 \ ,$$

$$+A(\tau, h) \ + \ B(\tau, h) \ + \ C(\tau, h) \ + \ D(\tau, h) \ ,$$

Figure A3: A couple's discrete choice problem



Note: The area in the black circle is the set $\Omega^d(0,h)$. (The circle is a generic shape. There is no presumption that the set takes this specific geometric form.) If for all $y \in \mathcal{Y}_h$, $h(y) > 0$, the treated area becomes less attractive and the set $\Omega^d(\tau,h)$ is therefore decreasing in τ, as indicated by the dashed red circle, which represents $\Omega^d(\tau,h)$ for some positive level of τ. At $\tau = 0$, the revenue loss from the fact that fewer couples make choices in the treated area is equal to zero. The reason is that every loss of a couple with a type ω on the black circle is multiplied with the change in their tax payment. At $\tau = 0$, this change is zero, $r(\omega, 0, h) = 0$.

where

$$A(\tau,h) = \left(\int_{\underline{\omega}_2(\overline{\omega}_1(\cdot),\tau,h)}^{\overline{\omega}_2(\overline{\omega}_1(\cdot),\tau,h)} r(\overline{\omega}_1(\cdot),\omega_2,\tau,h) f_{\omega_2}(\omega_2 \mid \omega_1) d\omega_2 \right) \frac{\partial}{\partial \tau}\overline{\omega}_1(\tau,h) \ ,$$

$$B(\tau,h) = -\left(\int_{\underline{\omega}_2(\underline{\omega}_1(\cdot),\tau,h)}^{\overline{\omega}_2(\underline{\omega}_1(\cdot),\tau,h)} r(\underline{\omega}_1(\cdot),\omega_2,\tau,h) f_{\omega_2}(\omega_2 \mid \omega_1) d\omega_2 \right) \frac{\partial}{\partial \tau}\underline{\omega}_1(\tau,h) \ ,$$

$$C(\tau,h) = \int_{\underline{\omega}_1(\tau,h)}^{\overline{\omega}_1(\tau,h)} r(\omega_1,\overline{\omega}_2(\cdot),\tau,h) \ \frac{\partial}{\partial \tau}\underline{\omega}_2(\omega_1,\tau,h) \ f_{\omega_1}(\omega_1) d\omega_1 \ ,$$

and

$$D(\tau,h) \quad = \quad -\int\limits_{\underline{\omega}_1(\tau,h)}^{\overline{\omega}_1(\tau,h)} r(\omega_1,\underline{\omega}_2(\cdot),\tau,h)\ \frac{\partial}{\partial\tau}\underline{\omega}_2(\omega_1,\tau,h)\ f_{\omega_1}(\omega_1)d\omega_1\ .$$

Note that $\tau = 0$ implies that $T^1(y) = T^0(y)$, for all y. Hence $r(\omega, 0, h) = 0$, for all ω and all h. This implies, moreover, that

$$A(0,h) \quad = \quad B(0,h) \quad = \quad C(0,h) \quad = \quad D(0,h) \quad = \quad 0\ .$$

Therefore,

$$R_\tau(0,h) \quad = \quad \int\limits_{\underline{\omega}_1(0,h)}^{\overline{\omega}_1(0,h)} \int\limits_{\underline{\omega}_2(\omega_1,0,h)}^{\overline{\omega}_2(\omega_1,0,h)} r_\tau(\omega,0,h)\ f_{\omega_2}(\omega_2 \mid \omega_1)d\omega_2\ f_{\omega_1}(\omega_1)d\omega_1\ ,$$

$$= \quad \mathbb{E}_\omega[r_\tau(\omega,0,h)\ \mathbf{1}(\omega \in \Omega^d(0,h))]\ .$$

A.6 Proof of Proposition 1

Notation. The joint distribution of productive abilities types is represented by a density $(\omega_1, \omega_2) \mapsto f_\omega(\omega_1, \omega_2)$, with marginal densities $\omega_1 \mapsto f_{\omega_1}(\omega_1)$ and $\omega_2 \mapsto f_{\omega_2}(\omega_2)$. We write $\omega_2 \mapsto f_{\omega_2}(\omega_2 \mid \omega_1)$ for the density of ω_2, conditional on ω_1, where

$$f_{\omega_2}(\omega_2 \mid \omega_1) \quad := \quad \frac{f_\omega(\omega_1,\omega_2)}{f_{\omega_1}(\omega_1)}\ .$$

The density $(\omega_1, \omega_2) \mapsto f_\omega(\omega_1, \omega_2)$ is an unobserved primitive of the economy. What is observed, by contrast, is the status quo distribution of earnings. We denote by $(y_1, y_2) \mapsto f_y(y_1, y_2)$ the density that represents the joint distribution of earnings, i.e. the cross-section distribution of $y_1^0(\theta)$ and $y_2^0(\theta)$. The marginal densities are denoted by f_{y_1} and f_{y_2} and the marginal *cdf*s by F_{y_1} and F_{y_2}. We write $y_2 \mapsto f_{y_2}(y_2 \mid y_1)$ and $y_2 \mapsto F_{y_2}(y_2 \mid y_1)$ for the conditional densities and *cdf*s.

Let Θ' be a measurable subset of $\Theta = \Theta_1 \times \Theta_2$. We use $P(\Theta')$ as a shorthand for

$$\int_{\Theta} \mathbf{1}(\theta \in \Theta') f(\theta) \, d\theta \ ,$$

where $\mathbf{1}$ is the indicator function, and the density f is associated with the joint distribution of θ_1 and θ_2.

A.6.1 Proof of Proposition 1

Recall that by Proposition 5

$$R_\tau(0,h) = \mathbb{E}_\omega[r_\tau(\omega,0,h)\ \mathbf{1}(\omega \in \Omega^d(0,h))] \ .$$

Characterizing $\Omega^d(0,h)$. Fix $B_1 = [y_{1s}, y_{1s}+\ell_1]$ and $B_2 = [y_{2s}, y_{2s}+\ell_2]$ so that $\exists \omega : y^{d0}(\omega) \in B := B_1 \times B_2$, where $y^{d0}(\omega) = (y_1^{d0}(\omega), y_2^{d0}(\omega))$ is a shorthand for $y^{d*}(\omega, 0, h)$. Let

$$\underline{\omega}_1(y_{1s}) \quad := \quad \min \{\omega_1 \in \Omega_1 \mid \exists\, \omega_2 \text{ s.t. } y_1^{d0}(\omega_1,\omega_2) \geq y_{1s}\}$$

be the lowest value of ω_1 consistent with earnings of individual 1 in the bracket ranging from y_{1s} to $y_{1s}+\ell_1$. Analogously, define the largest value of ω_1 by

$$\overline{\omega}_1(y_{1s}+\ell_1) \quad := \quad \max \{\omega_1 \in \Omega_1 \mid \exists\, \omega_2 \text{ s.t. } y_1^{d0}(\omega_1,\omega_2) \leq y_{1s}+\ell_1\} \ .$$

We will sometimes write $\Omega_1(B) = [\underline{\omega}_1(y_{1s}), \overline{\omega}_1(y_{1s}+\ell_1)]$.

Henceforth, consider some fixed $\omega_1 \in [\underline{\omega}_1(y_{1s}), \overline{\omega}_1(y_{1s}+\ell_1)]$. There is then a non-empty subset $\Omega_2[B,\omega_1]$ of Ω_2 so that $\omega_2 \in \Omega_2[B,\omega_1]$ implies $y_1^{d0}(\omega_1,\omega_2) \in [y_{1s}, y_{1s}+\ell_1]$.

There is then a, possibly empty, subset $\Omega_2^{in}[B,\omega_1]$ of $\Omega_2[B,\omega_1]$ so that $\omega_2 \in \Omega_2^{in}[B,\omega_1]$ implies $y_2^{0}(\omega_1,\omega_2) \in [y_{2s}, y_{2s}+\ell_2]$. If the set is non-empty, denote by $\underline{\omega}_2^{in}(y_{2s} \mid \omega_1)$ and $\overline{\omega}_2^{in}(y_{2s}+\ell_2 \mid \omega_1)$ the minimal and the maximal element of this set.

There also is a, possibly empty, subset $\Omega_2^{+}[B,\omega_1]$ of $\Omega_2[B,\omega_1]$ so that $\omega_2 \in \Omega_2^{+}[B,\omega_1]$ implies $y_2^{0}(\omega_1,\omega_2) \geq y_{2s}+\ell_2$. If the set is non-empty, denote by $\underline{\omega}_2^{+}(y_{2s}+\ell_2 \mid \omega_1)$ and $\overline{\omega}_2^{+}(B \mid \omega_1)$ the minimal and the maximal element of this set.

Characterizing $r_\tau(\omega, 0, h)$**.** From equations (38) - (40) in Section A.5 it follows that

$$\begin{aligned} r_\tau(\omega,0,h) &= \sigma^{d0}(\omega)\Big(T^0_{y_1}(y^{d0}(\omega))y^{d0}_{1,\tau}(\cdot) + T^0_{y_2}(y^{d0}(\omega))y^{d0}_{2,\tau}(\cdot) + h(y^{d0}(\omega))\Big) \\ &\quad +\sigma^{s0}_1(\omega)\Big(T^0_{y_1}(y^{s0}_1(\omega_1),0)y^{s0}_{1,\tau}(\cdot) + h(y^{s0}_1(\omega_1),0)\Big) \\ &\quad +\sigma^{s0}_2(\omega)\Big(T^0_{y_2}(0,y^{s0}_2(\omega_2))y^{s0}_{2,\tau}(\cdot) + h(0,y^{s0}_2(\omega_2))\Big) \\ &\quad +\sigma^{d0}_\tau(\omega,h)\; T^0(y^{d0}(\omega)) \\ &\quad +\sigma^{s0}_{1,\tau}(\omega,h)\; T^0(y^{s0}_1(\omega_1),0) \\ &\quad +\sigma^{s0}_{2,\tau}(\omega,h)\; T^0(0,y^{s0}_2(\omega_2))\;. \end{aligned} \tag{41}$$

where $\sigma^{d0}(\omega)$, $\sigma^{s0}_1(\omega)$ and $\sigma^{s0}_2(\omega)$ are, respectively, the fractions of dual and single earner couples amongst those with productive abilities ω in the status quo, and the expressions involving a subscript of τ and a superscript of 0 refer to derivatives with respect to τ that are evaluated at $\tau = 0$. We introduce shorthands – capturing, respectively, behavioral responses at the intensive and the extensive margin, β^{d0}_I and β^{d0}_E, and mechanical effects on tax revenue due to changes in the tax functions for dual or single-earner couples, μ^{d0}, μ^{0s}_1, μ^{0s}_2 – so that we can write more compactly

$$\begin{aligned} r_\tau(\omega,0,h) &= \beta^{d0}_I(\omega) \quad + \quad \mu^0(\omega) \\ &= \beta^{d0}_I(\omega) + \beta^{d0}_E(\omega) \quad + \quad \mu^{d0}(\omega) + \mu^{0s}_1(\omega) + \mu^{0s}_2(\omega) \end{aligned}$$

where

$$\begin{aligned} \beta^{d0}_I(\omega) &:= \sigma^{d0}(\omega)\Big(T^0_{y_1}(y^{d0}(\omega))y^{d0}_{1,\tau}(\cdot) + T^0_{y_2}(y^{d0}(\omega))y^{d0}_{2,\tau}(\cdot)\Big) \\ &\quad +\sigma^{s0}_1(\omega)\; T^0_{y_1}(y^{s0}_1(\omega_1),0)y^{s0}_{1,\tau}(\cdot) \\ &\quad +\sigma^{s0}_2(\omega)\; T^0_{y_2}(0,y^{s0}_2(\omega_2))y^{s0}_{2,\tau}(\cdot)\;, \end{aligned}$$

$$\begin{aligned} \beta^{d0}_E(\omega) &:= +\sigma^{d0}_\tau(\omega,h)\; T^0(y^{d0}(\omega)) \\ &\quad +\sigma^{s0}_{1,\tau}(\omega,h)\; T^0(y^{s0}_1(\omega_1),0) \\ &\quad +\sigma^{s0}_{2,\tau}(\omega,h)\; T^0(0,y^{s0}_2(\omega_2))\;, \end{aligned}$$

$$\mu^{d0}(\omega) := \sigma^{d0}(\omega)\; h(y^{d0}(\omega))\;,$$

$$\mu^{0s}_1(\omega) := \sigma^{s0}_1(\omega)\; h(y^{s0}_1(\omega_1),0) \quad \text{and} \quad \mu^{0s}_2(\omega) := \sigma^{s0}_2(\omega)\; h(0,y^{s0}_2(\omega_1))\;.$$

A simple reform affecting only dual earner couples. We are considering a simple reform that affects marginal tax rates only in the interior of $\mathcal{Y}$. Hence $h(y) = 0$, whenever $y_1 = 0$ or $y_2 = 0$ and therefore also $y^{s0}_{1,\tau}(\cdot) = y^{s0}_{2,\tau}(\cdot) = 0$. Thus, for $\omega \in \Omega^d(0, h)$ we have

$$r_\tau(\omega, 0, h) \;=\; \beta^{d0}_I(\omega) + \beta^{d0}_E(\omega) \quad + \quad \mu^{d0}(\omega) \tag{42}$$

where

$$\beta^{d0}_I(\omega) \;=\; \sigma^{d0}(\omega)\Big(T^0_{y_1}(y^{d0}(\omega))y^{d0}_{1,\tau}(\cdot) + T^0_{y_2}(y^{d0}(\omega))y^{d0}_{2,\tau}(\cdot)\Big)\;. \tag{43}$$

Behavioral responses at the extensive margin for a simple reform affecting only dual earner couples: Characterization. For ease of notation, in this paragraph we take the reform direction h as given and write cutoff types and other endogenous objects (indirect utilities, earnings) as functions of ω and τ, as opposed to writing them as a function of ω, τ and h. Thus, we write $\hat{\varphi}(\omega, \tau) = (\hat{\varphi}_1(\omega, \tau), \hat{\varphi}_2(\omega, \tau))$. Henceforth we denote by $\sigma^d(\hat{\varphi}(\omega, \tau), \omega)$ the fraction of type ω couples so that both are generating income. Analogously, we denote by $\sigma^s_1(\hat{\varphi}(\omega, \tau), \omega)$ and $\sigma^s_2(\hat{\varphi}(\omega, \tau), \omega)$ the fractions of type ω couples so that only spouse 1 or only spouse 2 is generating income. Note that

$$\sigma^d(\hat{\varphi}(\omega, \tau), \omega) \;=\; \int_0^{\hat{\varphi}_1(\omega,\tau)} \int_0^{\hat{\varphi}_2(\omega,\tau)} f^{\varphi}(\varphi_1, \varphi_2 \mid \omega)\; d\varphi_2\; d\varphi_1\;,$$

where the density f^{φ} characterizes the joint distribution of φ_1 and φ_2, conditional on the couples' productivities ω. Moreover,

$$\begin{aligned}\sigma^s_1(\hat{\varphi}(\omega, \tau), \omega) \;=\;& \textstyle\int_0^{\hat{\varphi}_1(\omega,\tau)} \int_{\hat{\varphi}_2(\omega,\tau)}^{\infty} f^{\varphi}(\varphi_1, \varphi_2 \mid \omega)\; d\varphi_2\; d\varphi_1 \\ &+ \textstyle\int_{\hat{\varphi}_1(\omega,\tau)}^{\hat{\hat{\varphi}}_1} \int_{\varphi_1 - \Delta v^s(\omega)}^{\infty} f^{\varphi}(\varphi_1, \varphi_2 \mid \omega)\; d\varphi_2\; d\varphi_1\;.\end{aligned}$$

Analogously,

$$\begin{aligned}\sigma^s_2(\hat{\varphi}(\omega, \tau), \omega) \;=\;& \textstyle\int_{\hat{\varphi}_1(\omega,\tau)}^{\infty} \int_0^{\hat{\varphi}_2(\omega,\tau)} f^{\varphi}(\varphi_1, \varphi_2 \mid \omega)\; d\varphi_2\; d\varphi_1 \\ &+ \textstyle\int_{\hat{\varphi}_2(\omega,\tau)}^{\hat{\hat{\varphi}}_2} \int_{\Delta v^s(\omega)\,+\,\varphi_2}^{\infty} f^{\varphi}(\varphi_1, \varphi_2 \mid \omega)\; d\varphi_1\; d\varphi_2\;.\end{aligned}$$

Note that a reform that affects only dual earner couples, does neither affect Δv^s nor $\hat{\hat{\varphi}}_1$ and $\hat{\hat{\varphi}}_2$, see Figure A3 for an illustration. Hence, for the purpose of this paragraph, we treat these objects only as functions of ω, but not as functions of τ. (This will be different below, when we consider reforms that affect single-earner couples.)

For later use, we compute the derivatives of σ^d, σ^s_1 and σ^s_2 with respect to τ. For any of these quantities we indicate a derivative with respect to the first argument $\hat{\varphi}_1$ with a subscript $\hat{\varphi}_1$, and a derivative with respect to the second argument $\hat{\varphi}_2$ with a subscript $\hat{\varphi}_2$. Repeated applications of both the chain rule and Leibnitz' rule yield:

$$\sigma^d_\tau(\hat{\varphi}(\tau,\omega),\omega) \;\; := \;\; \hat{\varphi}_{2,\tau}(\omega,\tau)\;\sigma^d_{\hat{\varphi}_2}(\hat{\varphi}(\omega,\tau),\omega) + \hat{\varphi}_{1,\tau}(\omega,\tau)\;\sigma^d_{\hat{\varphi}_1}(\hat{\varphi}(\omega,\tau),\omega)$$

where

$$\sigma^d_{\hat{\varphi}_2}(\hat{\varphi}(\omega,\tau),\omega) := \int_0^{\hat{\varphi}_1(\omega,\tau)} f^\varphi(\varphi_1,\hat{\varphi}_2(\omega,\tau)\mid\omega)\;\; d\varphi_1$$

and

$$\sigma^d_{\hat{\varphi}_1}(\hat{\varphi}(\omega,\tau),\omega) := \int_0^{\hat{\varphi}_2(\omega,\tau)} f^\varphi(\hat{\varphi}_1(\omega,\tau),\varphi_2\mid\omega)\; d\varphi_2 \;\; .$$

Analogously,

$$\sigma^s_{1,\tau}(\hat{\varphi}(\omega,\tau),\omega) \;\; := \;\; -\hat{\varphi}_{2\tau}(\omega,\tau)\;\sigma^d_{\hat{\varphi}_2}(\hat{\varphi}(\omega,\tau),\omega)$$

and

$$\sigma^s_{2,\tau}(\hat{\varphi}(\omega,\tau),\omega) \;\; := \;\; -\hat{\varphi}_{1\tau}(\omega,\tau)\;\sigma^d_{\hat{\varphi}_1}(\hat{\varphi}(\omega,\tau),\omega) \;\; .$$

Note in particular that

$$\sigma^d_\tau(\hat{\varphi}(\tau,\omega),\omega) \;\; = \;\; -\sigma^s_{1,\tau}(\hat{\varphi}(\tau,\omega),\omega) - \sigma^s_{2,\tau}(\hat{\varphi}(\tau,\omega),\omega)\;.$$

From the definitions $\hat{\varphi}_1(\omega,\tau) := v^d(\omega,\tau) - v^s_2(\omega_2,\tau)$ and $\hat{\varphi}_2(\omega,\tau) := v^d(\omega,\tau) -$

$v_1^s(\omega_1, \tau)$ and the envelope theorem it follows that

$$-\hat{\varphi}_{1,\tau}(\omega,\tau) \;=\; -\hat{\varphi}_{2,\tau}(\omega,\tau) \;=\; h(y^*(\tau,\omega)) \;=$$

$$\begin{cases} y_2^*(\tau,\omega) - y_{2s}, & \text{if} \quad y_1^*(\tau,\omega) \in B_1 \text{ and } y_2^*(\tau,\omega) \in B_2 \\ \ell_2, & \text{if} \quad y_1^*(\tau,\omega) \in B_1 \text{ and } y_2^*(\tau,\omega) \geq y_{2s} + \ell_2 \\ 0, & \quad\text{otherwise,} \end{cases}$$

where, for ease of notation, we have suppressed the dependence of y^* on the reform direction h. If we now evaluate all these terms in the status quo, i.e. at $\tau = 0$, we obtain the expressions that are collected in the following Lemma.

Lemma 5

$$\sigma_\tau^{d0}(\omega) = -h(y^0(\omega))\Big(\sigma^d_{\hat{\varphi}_2}(\cdot) + \sigma^d_{\hat{\varphi}_1}(\cdot)\Big)_{|\tau=0}\,, \tag{44}$$

$$\sigma_{1,\tau}^{s0}(\omega) = h(y^0(\omega))\;\sigma^d_{\hat{\varphi}_2}(\cdot)_{|\tau=0}\,, \tag{45}$$

and

$$\sigma_{2,\tau}^{s0}(\omega) = h(y^0(\omega))\;\sigma^d_{\hat{\varphi}_1}(\cdot)_{|\tau=0}\,. \tag{46}$$

Upon substituting these expressions in to the above definition of $\beta_E^{d0}(\omega)$ we obtain the following Corollary.

Corollary 5

$$\beta_E^{d0}(\omega) = -h(y^{d0}(\omega))\,\Big(\sigma^d_{\hat{\varphi}_1}(\cdot)_{|\tau=0}\Delta^{T0}_{-1}(\omega) \;+\; \sigma^d_{\hat{\varphi}_2}(\cdot)_{|\tau=0}\Delta^{T0}_{-2}(\omega)\Big)\,, \tag{47}$$

where

$$\Delta T^0_{-1}(\omega) \quad := \quad T^0(y^{d0}(\omega)) - T^0(0, y_2^{s0}(\omega_2))$$

and

$$\Delta T^0_{-2}(\omega) \quad = \quad T^0(y^{d0}(\omega)) - T^0(y_1^{s0}(\omega_1), 0) \ .$$

Ceteris paribus, $\hat{\varphi}^0_2$ is one-by-one decreasing in $\Delta^{T0}_{-2}(\omega)$ and that $\hat{\varphi}^0_1$ is one-by-one decreasing in $\Delta^{T0}_{-1}(\omega)$.[45] Therefore, we can as well interpret $\sigma^d_{\hat{\varphi}_1}(\cdot)$ as a derivative of $\sigma^d(\cdot)$ with respect to $\Delta T^0_{-1}(\omega)$ and $\sigma^d_{\hat{\varphi}_2}(\cdot)$ as a derivative of $\sigma^d(\cdot)$ with respect to $\Delta T^0_{-2}(\omega)$. We write

$$\chi^{d0}_{-1}(\omega) \quad := \quad \frac{\sigma^d_{\hat{\varphi}_1}(\cdot)_{\tau=0}}{\sigma^{d0}(\omega)} \ \Delta T^0_{-1}(\omega) \ .$$

for the elasticity of σ^d with respect to $\Delta^{T0}_{-1}(\omega)$ in the status quo. Analogously, we define

$$\chi^{d0}_{-2}(\omega) \quad := \quad \frac{\sigma^d_{\hat{\varphi}_2}(\cdot)_{\tau=0}}{\sigma^{d0}(\omega)} \ \Delta T^0_{-2}(\omega) \ .$$

Armed with this notation we rewrite the expression for $\beta^{d0}_E(\omega)$ obtained in Corollary 5

$$\beta^{d0}_E(\omega) \ := \ -h(y^{d0}(\omega)) \ \sigma^{d0}(\omega) \Big(\chi^{d0}_{-1}(\omega) + \chi^{d0}_{-2}(\omega) \Big) \ .$$

[45] To see this, note that, $\hat{\varphi}_1(\omega, \tau) := v^d(\omega, \tau) - v^s_2(\omega_2, \tau)$ can be written as

$$\begin{aligned} \hat{\varphi}_1(\omega, \tau) \quad &= \quad \Delta \mathbf{s}_{-1}(\omega, \tau) \ - \ \Big(T^0(y^d(\omega, \tau)) - T^0(0, y^s_2(\omega_2, \tau)) + \tau \ h(y^d(\omega, \tau)) \Big) \\ &= \quad \Delta \mathbf{s}_{-1}(\omega, \tau) - \Big(\Delta T_{-1}(\tau, \omega) + \tau \ h(y^d(\omega, \tau)) \Big) \end{aligned} \qquad (48)$$

where $\mathbf{s}_{-i}(\omega, \tau) := \mathbf{s}^d(\omega, \tau) - \mathbf{s}^s_i(\omega, \tau)$ and $\mathbf{s}^d(\omega, \tau) := y^d(\omega, \tau) - k_1(y^d_1(\omega, \tau), \omega_1) - k_2(y^d_2(\omega, \tau), \omega_2)$ is the surplus (i.e. output minus variable cost) generated by a dual-earner couple with productive abilities ω, whereas $\mathbf{s}^s_i(\omega, \tau) := y^s_i(\omega_i, \tau) - k_i(y^s_i(\omega_i, \tau), \omega_i)$ is the surplus generated by a couple with spouse i as the single-earner. Thus, $\mathbf{s}_{-i}(\omega, \tau)$ is the surplus that is lost when spouse $-i$ no longer opts for positive earnings. In the status quo, (48) becomes

$$\hat{\varphi}^0_1(\omega) \quad = \quad \Delta \mathbf{s}^0_{-1}(\omega) \ - \ \Delta T^0_{-1}(\omega) \ . \qquad (49)$$

Upon letting

$$\bar{\chi}^{d0}(\omega) \quad := \quad \chi_{-1}^{d0}(\omega) + \mu_{-2}^{d0}(\omega)$$

and interpreting this object as the exit rate – the percentage change of $\sigma_{d0}(\omega)$ that results when joint income is taxed more – we can write this even more concisely, see the following Lemma which summarizes the analysis in this paragraph.

Lemma 6

$$\beta_E^{d0}(\omega) := -h(y^{d0}(\omega)) \;\; \sigma^{d0}(\omega) \; \bar{\chi}^{d0}(\omega) \; . \tag{50}$$

Corollary 6

$$\beta_E^{d0}(\omega) \; + \; \mu^{d0}(\omega) \quad = \quad \sigma^{d0}(\omega)(1 - \bar{\chi}^{d0}(\omega))h(y^{d0}(\omega)) \; . \tag{51}$$

The corollary illuminates the relation between a setup that has only behavioral responses at the intensive margin and one that has behavioral responses both at intensive and the extensive margin: Behavioral responses at the extensive margin response are as if there were only intensive margin responses but weaker mechanical effects of a tax reform. For ease of reference, we write henceforth

$$r_\tau(\omega, 0, h) \;\; = \;\; \beta_I^{d0}(\omega) \quad + \quad \mu_E^{d0}(\omega) \; h(y^{d0}(\omega)) \tag{52}$$

where

$$\beta_I^{d0}(\omega) \;\; = \;\; \sigma^{d0}(\omega)\Big(T_{y_1}^0(y^{d0}(\omega))y_{1,\tau}^{d0}(\cdot) + T_{y_2}^0(y^{d0}(\omega))y_{2,\tau}^{d0}(\cdot)\Big) \; . \tag{53}$$

and

$$\mu_E^{d0}(\omega) \;\; = \;\; \sigma^{d0}(\omega)(1 - \bar{\chi}^{d0}(\omega)) \tag{54}$$

is our notation for extensive-margin inclusive mechanical effects, per dollar of tax increase.

Aggregation. Note that $\beta_I^{d0}(\omega)$ is zero for $\omega_2 \in \Omega_2^+[B,\omega_1]$ In this case, marginal tax rates do not change so that there is no behavioral response at the intensive margin. We can now write

$$\mathbb{E}_\omega[r_\tau(\omega,0,h)\,\mathbf{1}(\omega \in \Omega^d(0,h))] \;\; =: \;\; \mathbf{B}(y_{1s},\ell_1,y_{2s},\ell_2) + \mathbf{M}(y_{1s},\ell_1,y_{2s},\ell_2)\;,$$

where

$$\mathbf{B}_I(y_{1s},\ell_1,y_{2s},\ell_2) \;\; = \;\; \int_{\Omega_1(B)} \int_{\Omega_2^{in}(B,\omega_1)} \beta_I^{d0}(\omega)\,dF(\omega)\;,$$

and

$$\mathbf{M}(y_{1s},\ell_1,y_{2s},\ell_2) \;\; = \;\; \int_{\Omega_1(B)} \int_{\Omega_2^{in}(B,\omega_1)} \mu_E^{d0}(\omega)\Big\{y_2^0(\omega) - y_{2s}\Big\}dF(\omega)$$

$$+\,\ell_2 \int_{\Omega_1(B)} \int_{\Omega_2^+(B,\omega_1)} \mu_E^{d0}(\omega)dF(\omega)\;.$$

and $dF(\omega)$ is a shorthand for $f_{\omega_2}(\omega_2 \mid \omega_1)d\omega_2\; f_{\omega_1}(\omega_1)d\omega_1$.

Revenue implications for $\ell_2 \to 0$. We now compute the derivative of

$$\mathbf{R}(y_{1s},\ell_1,y_{2s},\ell_2) \quad := \quad \mathbf{B}(y_{1s},\ell_1,y_{2s},\ell_2) + \mathbf{M}(y_{1s},\ell_1,y_{2s},\ell_2)$$

with respect to ℓ_2 and evaluate the resulting expressions at $\ell_2 = 0$. Note that in changing ℓ_2 we are changing the length of the bracket B_2 and therefore also the types $\overline{\omega}_2^{in}(y_{2s}+\ell_2 \mid \omega_1)$ and $\underline{\omega}_2^{+}(y_{2s}+\ell_2 \mid \omega_1)$ on the boundaries of $\Omega_2^{in}(B,\omega_1)$ and $\Omega_2^+(B,\omega_1)$. Note that the set $\Omega_1(B)$ does not change as ℓ_2 varies. This yields

$$\mathbf{R}_{\ell_2}(y_{1s},\ell_1,y_{2s},0) \;\; = \;\; \mathbf{B}_{\ell_2}(y_{1s},\ell_1,y_{2s},0) \;+\; \mathbf{M}_{\ell_2}(y_{1s},\ell_1,y_{2s},0)\;.$$

where

$$\mathbf{M}_{\ell_2}(y_{1s},\ell_1,y_{2s},0) \; = \int_{\Omega_1(B)} \int_{\Omega_2^+(y_{2s},\omega_1)} \mu_E^{d0}(\omega)dF(\omega)\;.$$

and

$$\Omega_2^+(y_{2s},\omega_1) \quad := \quad \{\omega_2 \mid y_2^{d0}(\omega) \geq y_{2s}\ , \text{for } \omega_1 \in \Omega_1(B)\}\ .$$

Moreover,

$$\mathbf{B}_{\ell_2}(y_{1s},\ell_1,y_{2s},0) \ =$$

$$\int\limits_{\underline{\omega}_1(y_{1s})}^{\overline{\omega}_1(y_{1s}+\ell_1)} \beta_I^{d0}(\omega_1,\hat{\omega}_2) f_{\omega_2}(\hat{\omega}_2 \mid \omega_1)\ \left(\frac{\overline{\omega}_2^{in}(y_{2s}+\ell_2|\omega_1)}{\partial \ell_2}\right) |_{\ell_2=0}\ f_{\omega_1}(\omega_1) d\omega_1\ .$$

for $\hat{\omega}_2 = \overline{\omega}_2^{in}(y_{2s} \mid \omega_1)$.

Rewriting $\mathbf{M}_{\ell_2}(y_{1s},\ell_1,y_{2s},0)$. Note that $\mathbf{M}_{\ell_2}(y_{1s},\ell_1,y_{2s},0)$ can equivalently be written as

$$\mathbf{M}_{\ell_2}(y_{1s},\ell_1,y_{2s},0) = P(\Theta')\,\mathbb{E}_\omega\left[\mu_E^{d0}(\omega) \mid \theta \in \Theta'\right]\ ,$$

where

$$\begin{aligned}
P(\Theta') &= P\left(\{\theta \mid y_1^0(\theta)\ \in\ [y_{1s}, y_{1s}+\ell_1]\ \text{ and }\ y_2^0(\theta) \geq y_{2s}\}\right) \\
&= P\left(\{\theta \mid y_2^0(\theta) \geq y_{2s}\} \mid y_1^0(\theta)\ \in\ [y_{1s}, y_{1s}+\ell_1]\right)\Big(F_{y_1}(y_{1s}+\ell_1) - F_{y_1}(y_{1s})\Big) \\
&= \Big(1 - F_{y_2}(y_{2s} \mid y_1^0(\theta)\ \in\ [y_{1s}, y_{1s}+\ell_1])\Big)\Big(F_{y_1}(y_{1s}+\ell_1) - F_{y_1}(y_{1s})\Big)\ .
\end{aligned}$$

Rewriting $\mathbf{B}_{\ell_2}(y_{1s},\ell_1,y_{2s},0)$. Recall that

$$\mathbf{B}_{\ell_2}(y_{1s},\ell_1,y_{2s},0)$$

$$= \int\limits_{\underline{\omega}_1(y_{1s})}^{\overline{\omega}_1(y_{1s}+\ell_1)} \beta_I^{d0}(\omega_1,\overline{\omega}_2^{in}(y_{2s} \mid \cdot)) f\omega_2(\overline{\omega}_2^{in}(y_{2s} \mid \cdot) \mid \omega_1)\ \left(\frac{\overline{\omega}_2^{in}(y_{2s}+\ell_2|\cdot)}{\partial \ell_2}\right) |_{\ell_2=0}\ f_{\omega_1}(\omega_1) d\omega_1\ .$$

For any y_2', and any ω_1, let

$$F_{y_2}(y_2' \mid \omega_1) \quad := \quad F_{\omega_2}(\bar{\omega}_2(y_2' \mid \omega_1) \mid \omega_1)\ ,$$

where $\bar{\omega}_2(y_2' \mid \omega_1)$ is the largest value of ω_2 so that $y_2^{d0}(\omega_1, \omega_2) \leq y_2'$. Upon differentiating with respect to y_2', we obtain

$$f_{y_2}(y_2' \mid \omega_1) \quad := \quad f_2(\bar{\omega}_2(y_2' \mid \omega_1) \mid \omega_1) \frac{\partial}{\partial y_2'} \bar{\omega}_2(y_2' \mid \omega_1) \ .$$

We can therefore write

$$\mathbf{B}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0)$$

$$= \int_{\underline{\omega}_1(y_{1s})}^{\overline{\omega}_1(y_{1s}+\ell_1)} \beta_I^{d0}(\omega_1, \bar{\omega}_2(y_{2s} \mid \omega_1)) f_{y_2}(y_{2s} \mid \omega_1) \ f_{\omega_1}(\omega_1) d\omega_1 \ ,$$

$$= \Big(F^{y_1}(y_{1s} + \ell_1) - F^{y_1}(y_{1s}) \Big) \mathbb{E}_{\omega_1} \Big[\beta_I^{d0}(\omega_1, \bar{\omega}_2(y_{2s} \mid \omega_1)) f^{y_2}(y_{2s} \mid \omega_1) \mid y_1^0(\theta) \ \in \ [y_{1s}, y_{1s} + \ell_1] \Big] \ .$$

Rewriting $\mathbf{R}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0)$. Recall that

$$\mathbf{R}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0) \ = \ \mathbf{B}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0) \ + \ \mathbf{M}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0) \ .$$

Using the results above this can be written as

$$\mathbf{R}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0) \quad = \quad \Big(F^{y_1}(y_{1s} + \ell_1) - F^{y_1}(y_{1s}) \Big) \Big(\mathcal{B}(y_{1s}, \ell_1, y_{2s}) + \mathcal{M}(y_{1s}, \ell_1, y_{2s}) \Big) \ ,$$

where

$$\mathcal{B}(y_{1s}, \ell_1, y_{2s}) \ := \ \mathbb{E}_{\omega_1} \Big[\beta_I^{d0}(\omega_1, \bar{\omega}_2(y_{2s} \mid \omega_1)) f_{y_2}(y_{2s} \mid \omega_1) \mid y_1^{d0}(\omega) \ \in \ B_1 \Big]$$

and

$$\mathcal{M}(y_{1s}, \ell_1, y_{2s}) \ := \ \Big(1 - F_{y_2}(y_{2s} \mid y_1^{d0}(\omega) \ \in \ B_1) \Big) \mathbb{E}_{\omega} \left[\mu_E^{d0}(\omega) \mid \omega \in \Omega' \right]$$

Henceforth, for brevity, we will write $\mathcal{R}(y_{1s}, \ell_1, y_{2s})$ rather than $\mathbf{R}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0)$.

Taking the derivative with respect to ℓ_1 and evaluating at $\ell_1 = 0$. This yields

$$\mathcal{R}_{\ell_1}(y_{1s}, 0, y_{2s}) \quad = \quad f^{y_1}(y_{1s})\Big(\mathcal{B}(y_{1s}, 0, y_{2s}) + \mathcal{M}(y_{1s}, 0, y_{2s})\Big) \ ,$$

where

$$\mathcal{B}(y_{1s}, 0, y_{2s})) \quad = \quad \mathbb{E}_{\omega_1}\Big[\beta_I^{d0}(\omega_1, \bar{\omega}_2(y_{2s} \mid \omega_1))f_{y_2}(y_{2s} \mid \omega_1) \mid y_1^0(\theta) \ = \ y_{1s}\Big]$$

and

$$\mathcal{M}(y_{1s}, 0, y_{2s})) \quad = \quad \Big(1 - F_{y_2}(y_{2s} \mid y_{1s})\Big)\mathbb{E}_{\omega}\left[\mu_E^{d0}(\omega) \mid \theta \in \Theta''\right]$$

where

$$\Theta'' \quad := \quad \{\theta \mid y_1^0(\theta) = y_{1s} \text{ and } y_2^0(\theta) \geq y_{2s}\} \ .$$

Upon writing, more compactly,

$$\mathcal{B}(y_{1s}, 0, y_{2s}) \quad =: \quad \bar{\beta}_I^{d0}(y_{1s}, y_{2s}) \ ,$$

$$\mathbb{E}_{\omega}\left[\mu_E^{d0}(\omega) \mid \theta \in \Theta''\right] \quad =: \quad \bar{\mu}_E^{d0}(y_{1s}, y_{2s}) \ ,$$

and

$$\mathcal{R}(y_{2s} \mid y_{1s}) \quad := \quad \mathcal{R}_{\ell_1}(y_{1s}, 0, y_{2s}) \ ,$$

we ultimately obtain

$$\mathcal{R}(y_{2s} \mid y_{1s}) \quad = \quad f^{y_1}(y_{1s})\left(\bar{\beta}_I^{d0}(y_{1s}, y_{2s}) + \Big(1 - F^{y_2}(y_{2s} \mid y_{1s})\Big)\bar{\mu}_E^{d0}(y_{1s}, y_{2s})\right) \ .$$

A.7 Conditional revenue functions: Single-earner couples

The revenue functions for single earner couples are shaped by similar forces as the ones for dual earner couples. For concreteness, we characterize the revenue function $y_2 \mapsto \mathcal{R}_2^s(y_2)$. Recall that $\mathcal{R}_2^s(y_{2s})$ is the marginal effect on tax revenue associated with an increase on the earnings of spouse 2, conditional on spouse 2 having earnings in narrow bracket that starts at y_{2s} and conditional on spouse 2 being a single earner.[46]

The intensive margin. Specifically, the revenue implications of intensive margin behavioral response are captured by

$$\beta_{2,I}^{s0}(\omega) \quad := \quad \sigma_2^{s0}(\omega)\; T_{y_2}^0(0, y_2^{s0}(\omega_2))\; y_{2,\tau}^{s0}(\cdot)\ , \quad \text{and}$$

$$\bar{\beta}_{2,I}^{s0}(y_{2s}) \quad := \quad \mathbb{E}_{\omega_1}\Big[\beta_{2,I}^{s0}(\omega_1, \bar{\omega}_2(y_{2s})) f_{y_2}(y_{2s} \mid \omega_1) \mid y_1 = 0\Big]\ .$$

The term $\beta_{2,I}^{s0}(\omega)$ looks at the set of couples with productive abilities ω. Of these couples, a fraction $\sigma_2^{s0}(\omega)$ are single earner couples with $y_1 = 0$ and y_2 in the range where marginal tax rates change. Then, $\bar{\beta}_{2,I}^{s0}(y_{2s})$ is obtained as the average of $\beta_{2,I}^{s0}(\omega)$ among all such couples, again, taking account of the possibility that the status quo tax schedule has a kink at $y = (0, y_{2s})$.

Mechanical effects and the extensive margin. The reform comes with an increase of the tax burden for single earner couples with $y_1 = 0$ and $y_2 \geq y_{2s}$. Consider couples with productive ability ω, who conditional on being a single earner couple with $y_1 = 0$, choose $y_2 \geq y_{2s}$ in the status quo. The reform implies that some of these couples become dual-earner couples, some are turned into couples with spouse 1 as the single earner and some are turned into couples with no earnings, see Figure A3. The percentage change in the mass of these single earner couples is denoted by $\mathcal{E}_{x2}^s(y_{2s})$.

Proposition 6

$$\mathcal{R}_2^s(y_{2s}) \quad = \quad F_{y_1}(0)\Big(\bar{\beta}_{2,I}^{s0}(y_{2s}) + (1 - F_{y_2}(y_{2s} \mid 0))\ (1 - \mathcal{E}_{x2}^s(y_{2s}))\Big)\ ,$$

where $y_1 \mapsto F^{y_1}(y_1)$ is the cdf associated with the marginal distribution of y_1 and $y_2 \mapsto F^{y_2}(y_{2s} \mid 0)$ is the cdf of y_2, conditional on y_1 being equal to 0.

[46] A characterization of $y_1 \mapsto \mathcal{R}_1^s(y_1)$ can be obtained along the same lines.

A.7.1 Proof of Proposition 6

We are now considering a reform (τ, h), with

$$h(y) = \begin{cases} y_2 - y_{2s}, & \text{if} \quad y \in \{0\} \times B_2(y_{2s}, \ell_2) \ , \\ \ell_2, & \text{if} \quad y_2 \geq y_{2s} + \ell_2 \quad \text{and} \quad y_1 = 0 \ , \\ 0, & \quad \text{otherwise} \ , \end{cases}$$

that raises marginal tax rates on the earnings of spouse 2 conditional on $y_2 \in B_2(y_{2s}, \ell_2)$ and spouse 1 having no earnings, $y_1 = 0$.

Consider the set of couples with productivity $\omega = (\omega_1, \omega_2)$ that includes a subset with $y_1^0(\theta) = 0$ and $y_2^0(\theta) \geq y_{2s}$ in the status quo. Figure A3 shows how the reform affects the solution to the couple's discrete choice problem: In particular, the blue line indicates how the mass of couples with spouse 2 as a single earner shrinks, this is the extensive margin response. In addition, there is an intensive margin behavioral response as the earnings incentives of spouse 2, conditional on being a single earner, change. Equations (41) therefore simplifies to

$$r_\tau(\omega, 0, h) = \beta_{2,I}^{s0}(\omega) + \mu_2^{s0}(\omega) + \beta_{2,E}^{s0}(\omega) \ , \tag{55}$$

where

$$\beta_{2,I}^{s0}(\omega) := \sigma_2^{s0}(\omega)\Big(T_{y_2}^0(0, y_2^{s0}(\omega_2))y_{2,\tau}^{s0}(\cdot)\Big) \ ,$$

$$\mu_2^{s0}(\omega) := \sigma_2^{s0}(\omega)h(0, y_2^{s0}(\omega_2)) \ ,$$

and

$$\begin{aligned} \beta_{2,E}^{s0}(\omega) := \ & \sigma_\tau^{d0}(\omega, h) \ T^0(y^{d0}(\omega)) \\ & + \sigma_{1,\tau}^{s0}(\omega, h) \ T^0(y_1^{s0}(\omega_1), 0) \\ & + \sigma_{2,\tau}^{s0}(\omega, h) \ T^0(0, y_2^{s0}(\omega_2)) \ . \end{aligned} \tag{56}$$

Comparative statics at the extensive margin. Henceforth, we denote by

$$\begin{aligned} & \sigma_2^s(\hat{\varphi}_1(\tau, \omega), \hat{\hat{\varphi}}_2(\tau, \omega), \Delta v^s(\tau, \omega), \omega) \\ & = \int_{\hat{\varphi}_1(\tau,\omega)}^{\infty} \int_0^{\hat{\varphi}_2} f^\varphi(\varphi_1, \varphi_2 \mid \omega) \ d\varphi_2 \ d\varphi_1 \\ & \quad + \int_{\hat{\varphi}_2}^{\hat{\hat{\varphi}}_2(\tau,\omega)} \int_{\Delta v^s(\tau,\omega) \, + \, \varphi_2}^{\infty} f^\varphi(\varphi_1, \varphi_2 \mid \omega) \ d\varphi_1 \ d\varphi_2 \ . \end{aligned}$$

the mass of couples where only spouse 2 generates income. Let $\sigma^s_{2,\hat{\varphi}_1}(\cdot)$, $\sigma^s_{2,\hat{\hat{\varphi}}_2}(\cdot)$ and $\sigma^s_{2,\Delta v^s}(\cdot)$ denote, respectively, the derivative of $\sigma^s_2(\cdot)$ with respect to its first, second and third argument. An application of Leibnitz' rule yields

$$\begin{aligned}&\sigma^s_{2\tau}(\hat{\varphi}_1(\tau,\omega),\hat{\hat{\varphi}}_2(\tau,\omega),\Delta v^s(\tau,\omega),\omega)\\ &= \sigma^s_{2,\hat{\varphi}_1}(\cdot)\ \hat{\varphi}_{1\tau}(\tau,\omega) \quad + \quad \sigma^s_{2,\hat{\hat{\varphi}}_2}(\cdot)\ \hat{\hat{\varphi}}_{2\tau}(\tau,\omega) \quad + \quad \sigma^s_{2,\Delta v^s}(\cdot)\ \Delta v^s_\tau(\tau,\omega)\ .\end{aligned}$$

where

$$\sigma^s_{2,\hat{\varphi}_1}(\cdot) \quad = \quad -\int_0^{\hat{\varphi}_2} f^\varphi(\varphi_1(\tau,\omega),\varphi_2 \mid \omega)d\varphi_2\ ,$$

$$\sigma^s_{2,\hat{\hat{\varphi}}_2}(\cdot) \quad = \quad \int_{\hat{\hat{\varphi}}_1}^{\infty} f^\varphi(\varphi_1,\hat{\hat{\varphi}}_2(\tau,\omega) \mid \omega)d\varphi_1\ ,$$

and

$$\sigma^s_{2,\Delta v^s}(\cdot) \quad = \quad -\int_{\hat{\varphi}_2}^{\hat{\hat{\varphi}}_2(\tau,\omega)} f^\varphi(\Delta v^s(\tau,\omega)+\varphi_2,\varphi_2 \mid \omega)d\varphi_2\ .$$

We denote by $\sigma^s_1(\hat{\varphi}_1(\tau,\omega),\hat{\hat{\varphi}}_2(\tau,\omega),\Delta v^s(\tau,\omega),\omega)$ and $\sigma^d(\hat{\varphi}_1(\tau,\omega),\hat{\hat{\varphi}}_2(\tau,\omega),\Delta v^s(\tau,\omega),\omega)$, respectively, the mass of couples where only spouse 1 generates income and the mass where both spouses generate income. We note that

$$\sigma^d_\tau(\hat{\varphi}_1(\tau,\omega),\hat{\hat{\varphi}}_2(\tau,\omega),\Delta_s(\tau,\omega),\omega) \quad = \quad -\sigma^s_{2,\hat{\varphi}_1}(\cdot)\ \hat{\varphi}_{1\tau}(\tau,\omega)$$

and that

$$\sigma^s_{1\tau}(\hat{\varphi}_1(\tau,\omega),\hat{\hat{\varphi}}_2(\tau,\omega),\Delta v^s(\tau,\omega),\omega) \quad = \quad -\sigma^s_{2,\Delta v^s}(\cdot)\ \Delta v^s_\tau(\tau,\omega)$$

Moreover, from the envelope theorem it follows that

$$\hat{\varphi}_{1\tau}(\tau,\omega) \;=\Delta v^s_\tau(\tau,\omega) \;=\; -\hat{\hat{\varphi}}_{2\tau}(\tau,\omega) \;=\; h(0,y^{s*}_2(\tau,\omega)) \;=$$

$$\begin{cases} 0, & \text{if} \quad y^{s*}_2(\tau,\omega) \leq y_{2s} \\ y^{s*}_2(\tau,\omega) - y_{2s}, & \text{if} \quad y^{s*}_2(\tau,\omega) \in [y_{2s}, y_{2s}+\ell_2] \\ \ell_2, & \text{if} \quad y^{s*}_2(\tau,\omega) \geq y_{2s}+\ell_2 \end{cases}$$

where, for ease of notation, we have suppressed the dependence of y_2^{s*} on the reform direction h. If we now evaluate all this expressions in the status quo, i.e. at $\tau = 0$, we obtain the following Lemma.

Lemma 7

$$\sigma_\tau^{d0}(\omega) = -h(0, y_2^{s0}(\omega_2))\ \sigma^s_{2,\hat{\varphi}_1}(\cdot)_{|\tau=0}\ , \tag{57}$$

$$\sigma_{1,\tau}^{s0}(\omega) = -h(0, y_2^{s0}(\omega_2))\ \sigma^s_{2,\Delta v^s}(\cdot)_{|\tau=0}\ , \tag{58}$$

and

$$\sigma_{2,\tau}^{s0}(\omega) = h(0, y_2^{s0}(\omega_2))\Big(\sigma^s_{2,\hat{\varphi}_1}(\cdot) + \sigma^s_{2,\Delta v^s}(\cdot) + \sigma^s_{2,\hat{\hat{\varphi}}_2}(\cdot)\Big)_{|\tau=0}\ . \tag{59}$$

Corollary 7

$$\begin{aligned} \beta_E^{d0}(\omega) \;&=\; -h(0, y_2^{s0}(\omega_2))\Big(\sigma^s_{2,\hat{\varphi}_1}(\cdot)_{|\tau=0}\Delta^{T0}_{-1}(\omega)\ +\ \sigma^s_{2,\Delta v^s}(\cdot)_{|\tau=0}\Delta^{T0}_{2-1}(\omega) \\ &\qquad -\sigma^s_{2,\hat{\hat{\varphi}}_2}(\cdot)_{|\tau=0}T^0(0, y_2^{s0}(\omega_2))\Big)\ , \end{aligned} \tag{60}$$

where

$$\Delta T^0_{-1}(\omega) \quad := \quad T^0(y^{d0}(\omega)) - T^0(0, y_2^{s0}(\omega_2))\ , \quad \text{and}$$

$$\Delta T^0_{2-1}(\omega) \quad := \quad T^0(y_1^{s0}(\omega_1), 0) - T^0(0, y_2^{s0}(\omega_2))\ .$$

Note that, ceteris paribus, $\hat{\varphi}_1^0$ and $\Delta v^s(\tau,\omega)$ are one-by-one increasing in $T_0(0, y_2^{s0}(\omega_2))$ and that $\hat{\hat{\varphi}}_2^0$ is one-by-one decreasing in $T_0(0, y_2^{s0}(\omega_2))$. This follows from the definitions of these cutoff types. We can therefore as well interpret $\sigma^s_{2,\hat{\varphi}_1}(\cdot)$ as a derivative of $\sigma^s_2(\cdot)$ with respect to $-\Delta^{T0}_{-1}(\omega)$, $\sigma^s_{2,\Delta v^s}(\cdot)$ as a derivative of σ^s_2 with respect to $-\Delta^{T0}_{2-1}(\omega)$ and, finally, $\sigma^s_{2,\hat{\hat{\varphi}}_2}(\cdot)$ as a derivative with respect to $T_0(0, y_2^{s0}(\omega_2))$, see also Footnote 45.We write

$$\chi^{s0}_{2,-1}(\omega) \quad := \quad \frac{\sigma^s_{2,\hat{\varphi}_1}(\cdot)_{\tau=0}}{\sigma^{s0}_2(\omega)}\ \Delta T^0_{-1}(\omega)\ ,$$

$$\chi^{s0}_{2,2-1}(\omega) \quad := \quad \frac{\sigma^s_{2,\Delta v^s}(\cdot)_{\tau=0}}{\sigma^{s0}_2(\omega)} \; \Delta T^0_{2-1}(\omega) \; ,$$

and

$$\chi^{s0}_{2,-2}(\omega) \quad := \quad \frac{\sigma^s_{2,\hat{\varphi}_2}(\cdot)_{|\tau=0}}{\sigma^{s0}_2(\omega)} \; T_0(0, y^{s0}_2(\omega_2)) \; .$$

Upon collection terms and defining

$$\bar{\chi}^{s0}_2(\omega) \quad := \quad \chi^{s0}_{2,-1}(\omega) + \chi^{s0}_{2,2-1}(\omega) - \chi^{s0}_{2,-2}(\omega)$$

we summarize the analysis in this paragraph with the following Lemma.

Lemma 8

$$\beta^{s0}_{2,E}(\omega) := -h(0, y^{s0}_2(\omega_2)) \;\; \sigma^{s0}_2(\omega) \; \bar{\chi}^{s0}_2(\omega) \; . \tag{61}$$

Corollary 8

$$\beta^{s0}_{2,E}(\omega) \; + \; \mu^{s0}_2(\omega) \quad = \quad \sigma^{s0}_2(\omega)(1 - \bar{\chi}^{s0}_2(\omega))h(0, y^{s0}_2(\omega_2)) \; . \tag{62}$$

The corollary illuminates the relation between a setup that has only behavioral responses at the intensive margin and one that has behavioral responses both at intensive and the extensive margin: Behavioral responses at the extensive margin response are as if there were only intensive margin responses but weaker mechanical effects of a tax reform. For ease of reference, we write henceforth

$$r_\tau(\omega, 0, h) \quad = \quad \beta^{s0}_{2,I}(\omega) \quad + \quad \mu^{s0}_{2,E}(\omega) \; h(0, y^{s0}_2(\omega_2)) \tag{63}$$

where

$$\beta^{s0}_{2,I}(\omega) \quad = \quad \sigma^{s0}_2(\omega) \; T^0_{y_2}(0, y^{s0}_2(\omega_2)) \; y^{s0}_{2,\tau}(\cdot) \; , \tag{64}$$

and

$$\mu^{s0}_{2,E}(\omega) \quad = \quad \sigma^{s0}_2(\omega)(1 - \bar{\chi}^{s0}_2(\omega)) \tag{65}$$

is our notation for extensive-margin inclusive mechanical effects, per dollar of tax increase.

Aggregation. Let $B_2 = [y_{2s}, y_{2s} + \ell_2]$. We can now write

$$\mathbb{E}_\omega[r_\tau(\omega, 0, h)\, \mathbf{1}(\omega \in \Omega_2^s(0, h))] \quad =: \quad \mathbf{B}(\{0\}, y_{2s}, \ell_2) + \mathbf{M}(\{0\}, y_{2s}, \ell_2) \; ,$$

where

$$\mathbf{B}(\{0\}, y_{2s}, \ell_2) \quad = \int\limits_{\Omega_1} \int\limits_{\Omega_2^{in}(B_2)} \beta_{2,I}^{s0}(\omega)\, dF(\omega) \; ,$$

and

$$\mathbf{M}(\{0\}, y_{2s}, \ell_2) \quad = \int\limits_{\Omega_1} \int\limits_{\Omega_2^{in}(B_2)} \mu_{2,E}^{s0}(\omega) \Big\{ y_2^0(\omega) - y_{2s} \Big\} dF(\omega)$$

$$+ \, \ell_2 \int\limits_{\Omega_1} \int\limits_{\Omega_2^{+}(B_2)} \mu_{2,E}^{s0}(\omega) dF(\omega) \; .$$

Revenue implications for $\ell_2 \to 0$. We now compute the derivative of

$$\mathbf{R}(\{0\}, y_{2s}, \ell_2) \quad := \quad \mathbf{B}(\{0\}, y_{2s}, \ell_2) + \mathbf{M}(\{0\}, y_{2s}, \ell_2)$$

with respect to ℓ_2 and evaluate the resulting expressions at $\ell_2 = 0$. Note that in changing ℓ_2 we are changing the length of the bracket B_2 and therefore also the types $\overline{\omega}_2^{in}(y_{2s} + \ell_2)$ and $\underline{\omega}_2^{+}(y_{2s} + \ell_2)$ on the boundaries of $\Omega_2^{in}(B_2, \omega)$ and $\Omega_2^{+}(B_2)$. This yields

$$\mathbf{R}_{\ell_2}(\{0\}, y_{2s}, 0) \quad = \quad \mathbf{B}_{\ell_2}(\{0\}, y_{2s}, 0) \; + \; \mathbf{M}_{\ell_2}(\{0\}, y_{2s}, 0) \; ,$$

where

$$\mathbf{M}_{\ell_2}(\{0\}, y_{2s}, 0) \; = \int\limits_{\Omega_1} \int\limits_{\Omega_2^{+}(y_{2s})} \mu_{2,E}^{s0}(\omega) dF(\omega) \; .$$

and

$$\Omega_2^{+}(y_{2s}) \quad := \quad \{\omega_2 \mid y_2^{s0}(\omega_2) \geq y_{2s}\} \; .$$

Moreover,

$$\mathbf{B}_{\ell_2}(\{0\}, y_{2s}, 0) \ =$$

$$\int_{\Omega_1} \beta^{s0}_{2,I}(\omega_1, \hat{\omega}_2) f_{\omega_2}(\hat{\omega}_2 \mid \omega_1) \ \left(\tfrac{\overline{\omega}^{in}_2(y_{2s}+\ell_2)}{\partial \ell_2} \right) |_{\ell_2=0} \ f_{\omega_1}(\omega_1) d\omega_1 \ .$$

for $\hat{\omega}_2 = \overline{\omega}^{in}_2(y_{2s})$.

Rewriting $\mathbf{M}_{\ell_2}(\{0\}, y_{2s}, 0)$. Note that $\mathbf{M}_{\ell_2}(\{0\}, y_{2s}, 0)$ can equivalently be written as

$$\mathbf{M}_{\ell_2}(\{0\}, y_{2s}, 0) = P\left(\Theta'\right) \mathbb{E}_{\omega} \left[\mu^{s0}_{2,E}(\omega) \mid \theta' \in \Theta' \right] \ ,$$

where

$$\begin{aligned} P\left(\Theta'\right) \quad &= \quad P\left(\{\theta \mid y^0_1(\theta) = 0 \ \text{ and } \ y^0_2(\theta) \geq y_{2s}\}\right) \\ &= \quad P\left(\{\theta \mid y^0_2(\theta) \geq y_{2s}\} \mid y^0_1(\theta) = 0\right) F_{y_1}(0) \\ &= \quad \Big(1 - F_{y_2}(y_{2s} \mid y^0_1(\theta) = 0)\Big) F_{y_1}(0) \ . \end{aligned}$$

Rewriting $\mathbf{B}_{\ell_2}((\{0\}, y_{2s}, 0)$. Recall that

$$\mathbf{B}_{\ell_2}(\{0\}, y_{2s}, 0)$$

$$= \int_{\Omega_1} \beta^{s0}_{2,I}(\omega_1, \overline{\omega}^{in}_2(y_{2s})) f_{\omega_2}(\overline{\omega}^{in}_2(y_{2s}) \mid \omega_1) \ \left(\tfrac{\overline{\omega}^{in}_2(y_{2s}+\ell_2)}{\partial \ell_2} \right) |_{\ell_2=0} \ f_{\omega_1}(\omega_1) d\omega_1 \ .$$

For any y'_2, and any ω_1, let

$$F_{y_2}(y'_2 \mid \omega_1) \quad := \quad F_{\omega_2}(\bar{\omega}_2(y'_2) \mid \omega_1) \ ,$$

where $\bar{\omega}_2(y'_2)$ is the largest value of ω_2 so that $y^{s0}_2(\omega_2) \leq y'_2$. Upon differentiating with respect to y'_2, we obtain

$$f_{y_2}(y'_2 \mid \omega_1) \quad := \quad f_2(\bar{\omega}_2(y'_2) \mid \omega_1) \frac{\partial}{\partial y'_2} \bar{\omega}_2(y'_2) \ .$$

We can therefore write

$$\mathbf{B}_{\ell_2}(\{0\}, y_{2s}, 0)$$

$$= \int\limits_{\Omega_1} \beta^{s0}_{2,I}(\omega_1, \bar{\omega}_2(y_{2s} \mid \omega_1)) f_{y_2}(y_{2s} \mid \omega_1)\ f_{\omega_1}(\omega_1) d\omega_1\ ,$$

$$= F^{y_1}(0)\ \mathbb{E}_{\omega_1}\Big[\beta^{s0}_{2,I}(\omega_1, \bar{\omega}_2(y_{2s})) f^{y_2}(y_{2s} \mid \omega_1) \mid y^0_1(\theta) = 0\Big]\ .$$

Rewriting $\mathbf{R}_{\ell_2}(y_{1s}, \ell_1, y_{2s}, 0)$. Recall that

$$\mathbf{R}_{\ell_2}(\{0\}, y_{2s}, 0) \ \ = \ \ \mathbf{B}_{\ell_2}(\{0\}, y_{2s}, 0)\ +\ \mathbf{M}_{\ell_2}(\{0\}, y_{2s}, 0)\ .$$

Using the results above this can be written as

$$\mathbf{R}_{\ell_2}(\{0\}, y_{2s}, 0) \ \ = \ \ F^{y_1}(0)\Big(\mathcal{B}(\{0\}, y_{2s}) + \mathcal{M}(\{0\}, y_{2s})\Big)\ ,$$

where

$$\mathcal{B}(\{0\}, y_{2s}) \ \ := \ \ \mathbb{E}_{\omega_1}\Big[\beta^{s0}_{2,I}(\omega_1, \bar{\omega}_2(y_{2s})) f_{y_2}(y_{2s} \mid \omega_1) \mid y^0_1(\theta) = 0\Big]$$

and

$$\mathcal{M}(\{0\}, y_{2s}) \ \ := \ \ \Big(1 - F_{y_2}(y_{2s} \mid y^0_1(\theta) = 0)\Big)\mathbb{E}_{\omega}\left[\mu^{d0}_{E}(\omega) \mid \theta \in \Theta'\right]$$

Upon writing

$$\bar{\beta}^{s0}_{2,I}(y_{2s}) \quad := \quad \mathbb{E}_{\omega_1}\Big[\beta^{s0}_{2,I}(\omega_1, \bar{\omega}_2(y_{2s})) f_{y_2}(y_{2s} \mid \omega_1) \mid y^0_1(\theta) = 0\Big]$$

and

$$\bar{\mu}^{s0}_{2,E}(y_{2s}) \quad := \quad \mathbb{E}_{\omega}\left[\mu^{s0}_{2,E}(\omega) \mid \theta \in \Theta'\right]$$

we ultimately, obtain

$$\begin{aligned}\mathcal{R}_2^s(y_{2s}) &:= \mathbf{R}_{\ell_2}(\{0\}, y_{2s}, 0) \\ &= F_{y_1}(0)\Big(\bar{\beta}_{2,I}^{s0}(y_{2s}) + (1 - F_{y_2}(y_{2s} \mid y_1^0(\theta) = 0))\ \bar{\mu}_{2,E}^{s0}(y_{2s})\Big)\end{aligned}$$

A.8 Intensive margin responses of dual earner couples

A.8.1 First order conditions

We consider a dual earner couple whose choices in the status quo satisfy the first order conditions

$$1 - T_{y_1}^0(y_1^*, y_2^*) = k_{1,y_1}(y_1, \omega_1) \tag{66}$$

and

$$1 - T_{y_2}^0(y_1^*, y_2^*) = k_{2,y_2}(y_1, \omega_1) \tag{67}$$

We now consider the possibility that the marginal tax rate on the earnings of spouse 1 increases by τ_1 or that the the marginal tax rate on the earnings of spouse 1 increases by τ_2. To capture this possibility, we let $\tau \in \{\tau_1, \tau_2\}$ and interpret both y_1^* and y_2^* as functions of τ. Specifically, $y_1^*(\tau)$ and $y_2^*(\tau)$ solve

$$1 - T_{y_1}^0(y_1^*(\tau), y_2^*(\tau)) - \tau\ \mathbf{1}(\tau = \tau_1) = k_{1,y_1}(y_1^*(\tau), \omega_1) \tag{68}$$

and

$$1 - T_{y_2}^0(y_1^*(\tau), y_2^*(\tau)) - \tau\ \mathbf{1}(\tau = \tau_2) = k_{2,y_2}(y_2^*(\tau), \omega_1)\ . \tag{69}$$

Differentiating these two equations with respect to τ, evaluating the resulting expressions in the status quo, i.e. for $\tau = 0$, and using (66) and (67), yields the following system of equations

$$\rho_{y_1}^1(y^*)\ \delta y_1^*\ +\ \rho_{y_2}^1(y^*)\ \delta y_2^*\ - \mathbf{1}(\tau = \tau_1) = \kappa^1(y_1^*, \omega_1)\ \delta y_1^*\ , \tag{70}$$

where

$$\delta y_1^* \ := \ \frac{y_{1,\tau}^*}{y_1^*} \quad \text{and} \quad \delta y_2^* \ := \ \frac{y_{2,\tau}^*}{y_2^*}$$

are the changes of y_1^* and y_2^* in percent and

$$\rho_{y_1}^1(y) \quad := \quad \frac{-T_{y_1 y_1}^0(y)}{1 - T_{y_1}^0(y)} \; y_1 \ ,$$

$$\rho_{y_2}^1(y) \quad := \quad \frac{-T_{y_1 y_2}^0(y)}{1 - T_{y_1}^0(y)} \; y_2 \ ,$$

are the elasticities of the net of tax or retention rate on the earnings of spouse 1 with respect to y_1 and y_2, respectively. Note that a tax system that is progressive in that marginal tax rates on y_1 are non-decreasing in y_1 has $\rho_{y_1}^1(y) \leq 0$. A tax system that has positive jointness so that the marginal tax rates on y_1 are non-decreasing in y_2 has $\rho_{y_2}^1(y) \leq 0$. Finally,

$$\kappa^1(y_1, \omega_1) \quad := \quad \frac{k_{1,y_1 y_1}(y_1, \omega_1)}{k_{1,y_1}(y_1, \omega_1)} \; y_1$$

is the elasticity of spouse 1's marginal effort costs with respect to the earnings of spouse 1. Note that, with an iso-elastic effort cost function,

$$k_1(y_1, \omega_1) = \frac{1}{1 + \frac{1}{\varepsilon_1}} \left(\frac{y_1}{\omega_1} \right)^{1 + \frac{1}{\varepsilon_1}} \ ,$$

$$\kappa^1(y_1, \omega_1) \quad = \quad \frac{1}{\varepsilon_1} \ .$$

Analogously, we obtain

$$\rho_{y_1}^2(y^*) \; \delta y_1^* \ + \ \rho_{y_2}^2(y^*) \; \delta y_2^* \ - \mathbf{1}(\tau = \tau_2) = \kappa^2(y_2^*, \omega_2) \; \delta y_2^* \ , \tag{71}$$

Equations (70) and (71) can be used to solve for δy_1^* and δy_2^* this yields

$$\delta y_1^* = \frac{1}{D(y^*)} \left(\left(\rho_{y_1}^2(y^*) - \kappa^2(y_2^*, \omega_1) \right) \frac{\mathbf{1}(\tau = \tau_1)}{1 - T_{y_1}^0(y)} - \rho_{y_2}^1(y^*) \frac{\mathbf{1}(\tau = \tau_2)}{1 - T_{y_2}^0(y)} \right) \tag{72}$$

and

$$\delta y_2^* = \frac{1}{D(y^*)}\left(\left(\rho_{y_2}^1(y^*) - \kappa^1(y_1^*,\omega_1)\right)\frac{\mathbf{1}(\tau=\tau_2)}{1-T_{y_2}^0(y)} - \rho_{y_1}^2(y^*)\frac{\mathbf{1}(\tau=\tau_1)}{1-T_{y_1}^0(y)}\right) \tag{73}$$

where

$$D(y^*) \quad := \quad \left(\rho_{y_1}^1(y^*) - \kappa^1(y_1^*,\omega_1)\right)\left(\rho_{y_2}^2(y^*) - \kappa^2(y_2^*,\omega_1)\right) \; - \; \rho_{y_2}^1(y^*)\,\rho_{y_1}^2(y^*)\;.$$

Piecewise-linear tax system. If the tax system is piecewise linear, then, locally, $\rho_{y_1}^1(y^*) = \rho_{y_1}^2(y^*) = \rho_{y_2}^2(y^*) = \rho_{y_1}^2(y^*) = 0$. In this case, (72) and (73) imply

$$\delta y_1^* = -\frac{1}{\kappa^1(y_1^*,\omega_1)}\,\frac{\mathbf{1}(\tau=\tau_1)}{1-T_{y_1}^0(y)} \tag{74}$$

and

$$\delta y_2^* = -\frac{1}{\kappa^2(y_2^*,\omega_1)}\,\frac{\mathbf{1}(\tau=\tau_2)}{1-T_{y_2}^0(y)}\;. \tag{75}$$

With iso-elastic effort costs this becomes, moreover,

$$\delta y_1^* = -\varepsilon_1\,\frac{\mathbf{1}(\tau=\tau_1)}{1-T_{y_1}^0(y)} \tag{76}$$

and

$$\delta y_2^* = -\varepsilon_2\,\frac{\mathbf{1}(\tau=\tau_2)}{1-T_{y_2}^0(y)}\;. \tag{77}$$

Family Taxation. If there is a system of family taxation, then, for every y there is a number $p(y)$ so that

$$\frac{\rho_{y_1}^1(y)}{y_1} \;=\; \frac{\rho_{y_1}^2(y)}{y_1} \;=\; \frac{\rho_{y_2}^2(y)}{y_2} \;=\; \frac{\rho_{y_2}^1(y)}{y_2} \quad=\quad p(y)\;.$$

A.8.2 Implications for tax revenue

Recall that behavioral responses of a dual earner couples affect marginal tax rates via

$$\beta_I^0(\omega) \;=\; \sigma^{d0}(\omega)\Big(T^0_{y_1}(y^{d0}(\omega))y^{d0}_{1,\tau}(\cdot) + T^0_{y_2}(y^{d0}(\omega))y^{d0}_{2,\tau}(\cdot)\Big)$$

We now use the characterization of $y^{d0}_{1,\tau}(\cdot)$ and $y^{d0}_{2,\tau}(\cdot)$ in the previous section – in particular, equations (72) and (73) – for an analysis of

$$\begin{aligned} r^d_\tau(y^*) \;&:=\; T^0_{y_1}(y^*)\, y^*_{1,\tau}(\cdot) \quad + \quad T^0_{y_2}(y^*)\, y^*_{2,\tau}(\cdot) \\ &=\; T^0_{y_1}(y^*)\, \delta y^*_1\, y^*_1 \quad + \quad T^0_{y_2}(y^*)\, \delta y^*_2\, y^*_2\ . \end{aligned}$$

Straightforward computations yield

$$\begin{aligned} r^d_\tau(y^*) \;:=\; & \tfrac{\mathbf{1}(\tau=\tau_1)}{(1-T^0_{y_1}(y^*))D(y^*)}\Big(T^0_{y_1}(y^*)\, y^*_1\, (\rho^2_{y_1}(y^*) - \kappa^2(y^*_2,\omega_1)) \;-\; T^0_{y_2}(y^*)\, y^*_2\, \rho^2_{y_1}(y^*)\Big) \\ & + \tfrac{\mathbf{1}(\tau=\tau_2)}{(1-T^0_{y_2}(y^*))D(y^*)}\Big(T^0_{y_2}(y^*)\, y^*_2\, (\rho^1_{y_2}(y^*) - \kappa^1(y^*_2,\omega_1)) \;-\; T^0_{y_1}(y^*)\, y^*_1\, \rho^1_{y_2}(y^*)\Big)\ . \end{aligned}$$

Piecewise-linear tax system and iso-elastic effort costs. In this case, this simplifies to

$$r^d_\tau(y^*) \;:=\; -\mathbf{1}(\tau=\tau_1)\, \tfrac{T^0_{y_1}(y^*)}{1-T^0_{y_1}(y^*)}\, y^*_1\, \varepsilon_1 \quad - \quad \mathbf{1}(\tau=\tau_2)\, \tfrac{T^0_{y_2}(y^*)}{1-T^0_{y_2}(y^*)}\, y^*_2\, \varepsilon_2\ .$$

Family taxation. In this case, for every y, there is a number $T'(y)$ so that $T^0_{y_1}(y) = T^0_{y_2}(y) = T'(y)$. Therefore

$$\begin{aligned} r^d_\tau(y^*) \;:=\; & \tfrac{\mathbf{1}(\tau=\tau_1)T'(y^*)}{(1-T'(y^*))D(y^*)}\Big(y^*_1\, (\rho^2_{y_1}(y^*) - \kappa^2(y^*_2,\omega_1)) \;-\; y^*_2\, \rho^2_{y_1}(y^*)\Big) \\ & \tfrac{\mathbf{1}(\tau=\tau_2)T'(y^*)}{(1-T'(y^*))D(y^*)}\Big(y^*_2\, (\rho^1_{y_2}(y^*) - \kappa^1(y^*_2,\omega_1)) \;-\; y^*_1\, \rho^1_{y_2}(y^*)\Big)\ . \end{aligned}$$

A.8.3 Characterization of $\beta^{d0}_{I,2}(\omega)$.

Recall that

$$\beta^{d0}_{I,2}(\omega) \quad = \quad T^0_{y_1}(y^{d0}(\omega))y^{d0}_{1,\tau}(\cdot) + T^0_{y_2}(y^{d0}(\omega))y^{d0}_{2,\tau}(\cdot) \; . \tag{78}$$

For $\mathbf{1}(\tau = \tau_2) = 1$, it follows from the analysis in the previous section that

$$y^{d0}_{1,\tau}(\omega) \quad = \quad -y^{d0}_1(\omega)\, \frac{\rho^1_{y_2}(y^{d0}(\omega))}{(1 - T^0_{y_2}(y^{d0}(\omega))D(y^{d0}(\omega))} \tag{79}$$

and

$$y^{d0}_{2,\tau}(\omega) \quad = \quad y^{d0}_2(\omega)\, \frac{\rho^1_{y_2}(y^{d0}(\omega)) - \kappa^1(y^{d0}_1(\omega), \omega_1)}{(1 - T^0_{y_2}(y^{d0}(\omega))D(y^{d0}(\omega))} \; . \tag{80}$$

By substituting (79) and (80) into (78) we obtain the reform's marginal impact on the couple's tax payment. The resulting expression simplifies when (i) there is a system of family taxation so that, in the status quo, there is a joint marginal tax rate for the earnings of spouse 1 and spouse 2, henceforth denoted by $T^{0\prime}(y^{d0}(\omega)) = T^0_{y_1}(y^{d0}(\omega)) = T^0_{y_2}(y^{d0}(\omega))$, (ii) the status quo tax system is piece-wise-linear so that

$$\rho^1_{y_1}(y) \quad = \quad \rho^1_{y_2}(y) \quad = \quad \rho^2_{y_2}(y) \quad = \quad \rho^2_{y_1}(y) \quad = \quad 0 \; ,$$

for all y where marginal tax rates are well-defined, and (iii) effort costs are iso-elastic. Then,

$$T^0_{y_1}(y^{d0}(\omega))y^{d0}_{1,\tau}(\cdot) + T^0_{y_2}(y^{d0}(\omega))y^{d0}_{2,\tau}(\cdot) \quad = \quad -\frac{T^{0\prime}(y^{d0}(\omega))}{1 - T^{0\prime}(y^{d0}(\omega))}\; y^{d0}_2(\omega)\; \varepsilon_2 \; .$$

Hence, only the earnings and the Frisch elasticity of spouse 2 matter for the loss of tax revenue due to behavioral responses at the intensive margin.

A.9 Proof of Proposition 3

We use the conditions $\mathcal{W}_1(y_1 \mid y_2) = 0$ and $\mathcal{W}_2(y_2 \mid y_1) = 0$ to derive a candidate solution for optimal tax rates. We then verify that, under the candidate solution it is indeed the case, that $\mathcal{W}_1(y_1 \mid y_2) = 0$ and $\mathcal{W}_2(y_2 \mid y_1) = 0$.

Fix $(y_1, y_2) \in Y'$. With extensive margin elasticities of zero, and using part i) of Assumption 2, the two optimality conditions $\mathcal{W}_1(y_1 \mid y_2) = 0$ and $\mathcal{W}_2(y_2 \mid y_1) = 0$ can be written as

$$\bar{\beta}^d_{I,1}(y_1, y_2) + \Big(1 - \lambda_1^* \, \bar{g}_1\Big)\Big(1 - F_{y_1}(y_1 \mid y_2)\Big) = 0 \; , \tag{81}$$

and

$$\bar{\beta}^d_{I,2}(y_1, y_2) + \Big(1 - \lambda_2^* \, \bar{g}_2\Big)\Big(1 - F_{y_2}(y_2 \mid y_1)\Big) = 0 \; . \tag{82}$$

Part iii) of Assumption 2, implies that there is only one ω with $y^{d0}(\omega) = (y_1, y_2)$. Thus,

$$\bar{\beta}^d_{I,1}(y_1, y_2) = \beta^{d0}_{I,1}(\omega) f_{y_1}(y_1 \mid y_2) \; , \tag{83}$$

and

$$\bar{\beta}^d_{I,2}(y_1, y_2) = \beta^{d0}_{I,1}(\omega) f_{y_1}(y_1 \mid y_2) \tag{84}$$

Substituting (83) and (84) into (81) and (82), using part (ii) of Assumption yields

$$\beta^{d0}_{I,1}(\omega)\frac{1}{y_1} = \Big(1 - \lambda_1^* \, \bar{g}_1\Big)\frac{1}{\bar{\alpha}_1} \; , \tag{85}$$

and

$$\beta^{d0}_{I,2}(\omega)\frac{1}{y_2} = \Big(1 - \lambda_2^* \, \bar{g}_2\Big)\frac{1}{\bar{\alpha}_2} \; . \tag{86}$$

Recall from part A.8 of the Appendix that

$$\beta^{d0}_{I,1}(\omega) = \Big(T^0_{y_1}(y^{d0}(\omega)) y^{d0}_{1,\tau_1}(\cdot) + T^0_{y_2}(y^{d0}(\omega)) y^{d0}_{2,\tau_1}(\cdot)\Big) \; . \tag{87}$$

where

$$y^{d0}_{1,\tau_1}(\cdot) = y_1 \, \frac{1}{D(y^*)}\Big(\rho^2_{y_1}(y^*) - \kappa^2(y_2^*, \omega_2)\Big)\frac{1}{1 - T^0_{y_1}(y)} \tag{88}$$

and

$$y^{d0}_{2,\tau_1}(\cdot) = -y_2 \ \frac{1}{D(y^*)}\rho^2_{y_1}(y^*)\frac{1}{1-T^0_{y_1}(y)} \tag{89}$$

Analogously,

$$\beta^{d0}_{I,2}(\omega) = \Big(T^0_{y_1}(y^{d0}(\omega))y^{d0}_{1,\tau_2}(\cdot) + T^0_{y_2}(y^{d0}(\omega))y^{d0}_{2,\tau_2}(\cdot) \Big) \ . \tag{90}$$

where

$$y^{d0}_{1,\tau_2}(\cdot) = -y_1 \ \frac{1}{D(y^*)}\rho^1_{y_2}(y^*)\frac{1}{1-T^0_{y_2}(y)} \tag{91}$$

and

$$y^{d0}_{2,\tau_2}(\cdot) = y_2 \ \frac{1}{D(y^*)}\Big(\rho^1_{y_2}(y^*) - \kappa^1(y^*_1,\omega_1)\Big)\frac{1}{1-T^0_{y_2}(y)} \tag{92}$$

Substituting all these expressions into equations (85) and (86) yields a system of partial differential equations (recall that the second derivatives of the tax function appear in all the ρ-terms.) Rather than working with that system, we hypothesize that, under Assumption 2, there is a solution that is piecewise linear. Under this hypothesis all the ρ-terms are zero with the implications that

$$y^{d0}_{2,\tau_1}(\cdot) \quad = \quad y^{d0}_{1,\tau_2}(\cdot) \quad = \quad 0$$

$$y^{d0}_{1,\tau_1}(\cdot) \quad = \quad -y_1 \ \frac{\varepsilon_1}{1-T^0_{y_1}(y)}$$

and

$$y^{d0}_{2,\tau_2}(\cdot) \quad = \quad -y_2 \ \frac{\varepsilon_2}{1-T^0_{y_2}(y)} \ .$$

Therefore,

$$\beta^{d0}_{I,1}(\omega) = -\frac{T^0_{y_1}(y)}{1-T^0_{y_1}(y)} \ y_1 \ \varepsilon_1 \ , \tag{93}$$

and

$$\beta^{d0}_{I,2}(\omega) = -\frac{T^0_{y_2}(y)}{1 - T^0_{y_2}(y)}\; y_2\; \varepsilon_2 \; . \tag{94}$$

Substituting (93) and (94) into (85) and (86) yields

$$\frac{T^0_{y_1}(y)}{1 - T^0_{y_1}(y)} \quad = \quad \left(1 - \lambda^*_1 \,\bar{g}_1\right) \frac{1}{\bar{\alpha}_1} \, \frac{1}{\varepsilon_1} \; , \tag{95}$$

and

$$\frac{T^0_{y_2}(y)}{1 - T^0_{y_2}(y)} \quad = \quad \left(1 - \lambda^2_1 \,\bar{g}_2\right) \frac{1}{\bar{\alpha}_2} \, \frac{1}{\varepsilon_2} \; . \tag{96}$$

and hence

$$T^0_{y_1}(y) = t^*_1 \quad \text{and} \quad T^0_{y_2}(y) = t^*_2 \; . \tag{97}$$

To complete the proof, we finally note that if we choose for all $y \in Y'$, $T^0_{y_1}(y) = t^*_1$ and $T^0_{y_1}(y) = t^*_1$, then this implies that the tax system is indeed piecewise linear: When marginal tax rates are constant, second derivatives are zero. Moreover, equations (85) and (86) hold. Under Assumption 2 this implies that $\mathcal{W}_1(y_1 \mid y_2) = 0$ and $\mathcal{W}_2(y_2 \mid y_1) = 0$, for all $y \in Y'$, as had to be shown.

A.10 Alternative modelling choices

Alternative assumptions on the preferences of married individuals. The couples' preferences specified above are such that the couples' disposable income is a public good for the spouses in a couple. Moreover, the spouses' preferences are quasi-linear in disposable income. Furthermore, in the couples' objective function, each spouse's effort costs receive the same weight. Theorems 1 - 3 below could be proven under much weaker assumptions: There could be a separate utility function for each spouse that, moreover, is not quasi-linear in disposable income. The disposable income could be treated as resource that is split between the couples – as opposed to being a public good. Who receives how much of that private good, and also who contributes how much to the couple's pre-tax income could be determined via Nash bargaining. Finally, the Nash bargaining might also be over the allocation of household duties, see Bierbrauer et al. 2023 for a formal treatment under these alternative assumptions. More specific assumptions are needed, however, in the context of an empirical application that uses sufficient statistics for the purpose of checking whether the conditions in the Theorems hold in the data. To ease the exposition, we base the whole analysis in this paper on the most simple setup which is prominent also in the related literature. In the empirical application, we will, moreover, invoke the assumption that the variable costs of productive effort are iso-elastic so that, for $i = 1, 2$,

$$k_i(y_i, \omega_i) = \frac{1}{1 + \frac{1}{\varepsilon_i}} \left(\frac{y_i}{\omega_i} \right)^{1+\frac{1}{\varepsilon_i}} ,$$

where ε_i can be interpreted as the Frisch elasticity of spouse i's labour supply.

Nash bargaining. The formalism above nests the possibility that couples engage in Nash bargaining and, moreover, that there is heterogeneity in the bargaining weights of the spouses with index 1 and the spouses with index 2. To see this, suppose that spouse i has a utility function

$$u(c_m, y_i, \theta_i) \quad = \quad c_m - k_i(y_i, \omega_i) - \varphi_i \, \mathbf{1}(y_i > 0) \; .$$

and that the outcome of Nash bargaining is as if the couples were maximizing

$$\begin{aligned} &\gamma_1 \; u(c_m, y_1, \theta_1) + \gamma_2 \; u(c_m, y_2, \theta_2) = \\ &c_m - \gamma_1 \; k_1(y_1, \omega_1) - \gamma_1 \; \varphi_2 \, \mathbf{1}(y_1 > 0) - \gamma_2 \; k_2(y_2, \omega_2) - \gamma_2 \; \varphi_2 \, \mathbf{1}(y_2 > 0) \; , \end{aligned}$$

where γ_1 and $\gamma_2 = 1 - \gamma_2$ are, respectively, the bargaining weights of spouse 1 and spouse 2. With iso-elastic cost functions the resulting behavior is as if $\gamma_1 = \gamma_2 = 1$, while, the productive ability and fixed costs of spouse i are respectively equal to $\gamma_i^{1+\frac{1}{\varepsilon_i}} \omega_i$ and $\gamma_i \, \varphi_i$. Thus, working with an arbitrary distribution of θ, F_θ encompasses the possibility that couples differ in three dimensions: bargaining weights, productive abilities and fixed costs.

Singles and couples. We look at married couples in isolation. In particular, we do not consider tax reforms that jointly alter the tax treatment of singles and the tax treatment of couples, see Bierbrauer et al. 2023 for an extensive discussion of this possibility. For the analysis of Pareto-improving reform directions this is without loss of generality: A Pareto-improving reform of an overall tax system – consisting of a tax function that applies to singles and one that applies to married couples – is possible if and only if one of the following conditions is met: (i) there is a Pareto-improving reform of the tax function for singles, while holding the tax function for couples fixed, and (ii) there is a Pareto-improving reform of the tax function for couples, while holding the tax function for singles fixed. Welfare-improving reforms of an overall system involve tradeoffs between singles and married individuals; e.g. between single mothers and married women. In line with the literature on optimal welfare-maximizing taxes, we focus on welfare-improvements amongst couples, while holding the tax function for singles fixed.[47]

[47] There is the possibility that both the tax function for singles and the tax function for couples are welfare-optimal on a stand-alone basis, while there are still welfare gains, say, from extracting more tax revenues from couples so as to raise the intercept of the single's consumption function.

B Appendix for the empirical part / Section 3

This section of the appendix contains additional explanations, supplementary graphs and tables on the empirical part of the paper in Section 3. We first present details on the data preparation (Section B.1). Subsequently, supplementary insights are presented separately for each of the respective subset of results, i.e., for results on empirical conditional revenue functions (Section B.2), welfare effects of jointness-reducing reforms (Section B.3), tax reforms at the bottom of the income distribution (Section B.4), and optimal top tax rates (Section B.5).

B.1 Data

In Section 3 of the main text, we use household micro data from the Current Population Survey (CPS) and tabulated income tax return data from the Statistics of Income (SOI) program. CPS data is used for the estimation of conditional revenue functions (Section 3.1), and for analyzing the welfare effects of reducing jointness (Section 3.3), as well as tax reforms at the bottom of the income distribution (Section 3.4). For the calibration of optimal top tax rates in Section 3.5, we use tabulated income tax return data. In the following, we describe the details of the two data sources and respective data preparation.

B.1.1 Current Population Survey (CPS)

The Current Population Survey (CPS) is conducted by the US Census Bureau and the Bureau of Labor Statistics and contains nationally representative cross-sectional survey data from 1962 onward. We use data from the Annual Social and Economic Supplement of the Current Population Survey (CPS-ASEC).[48] The sample size of CPS-ASEC increased from around 30,000 households in 1962 to more than 90,000 in the most recent wave. In contrast to tax return micro data such as the public use files (IRS-SOI PUF) from the Statistics of Income (SOI) division of the Internal Revenue Service (IRS), as, e.g., used by Bargain et al. (2015) or Bierbrauer, Boyer, and Peichl (2021), the CPS data contain exact information about the incomes of

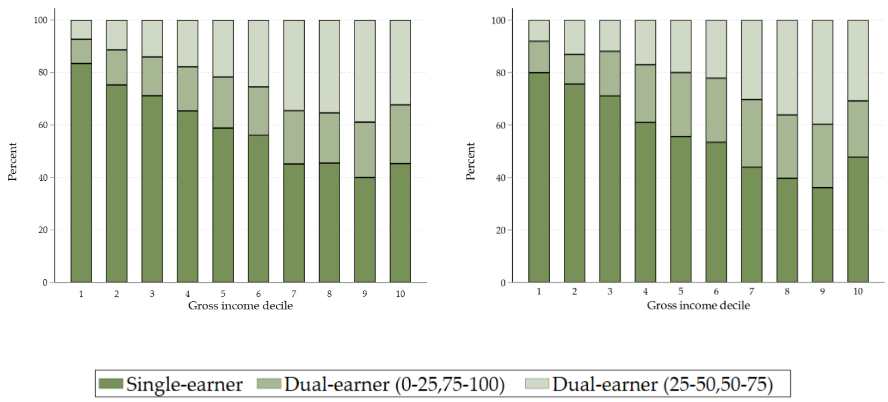
[48] See Flood et al. (2021) and `https://cps.ipums.org` for a detailed description of CPS data.

primary and secondary earners of the tax unit.[49]

To adapt the CPS to the input requirements of the microsimulation model, we transform the CPS from a household-level data set to a tax unit level data set. For this purpose, we form tax units by joining all married spouses with their dependent children. Single individuals and unmarried spouses form separate tax units. Children of single individuals are in most cases allocated to the household head. Adult individuals with a total income below the year-specific personal exemption threshold are assumed to reflect dependents of the household head. Table B1 illustrates in detail the correspondence between variables utilized in NBER TAXSIM and variables in the CPS data.

Figure B4: Comparison of CPS and SOI data (1974), couple types

(a) CPS (b) SOI



Note: This figure displays for the tax year of 1974 the distribution of married couple types across deciles of the per capita income distribution. The figure compares the distribution based on the CPS data (Figure B4a) to the distribution based on the IRS-SOI PUF tax return micro data (Figure B4b). All estimates are based on tax units with strictly positive gross income in which both spouses are between 25 and 55 years old.
Source: Our calculations based on CPS-ASEC and SOI PUF.

Treatment of top incomes. In the CPS data, information on top incomes is limited by (i) public topcoding, and (ii) internal censoring. We address both limitations by harmonizing the treatment of top incomes across the different survey years

[49] In the IRS-SOI PUF, the relevant information on salaries and wages from the W2-form of the primary and secondary earner is only available for the year 1974 and imputed for all other years using an undocumented procedure. For 1974, in which reliable information is available, the distribution of different couple types across per capita income distribution is very similar to the CPS data (see Figure B4). Moreover, Bargain et al. (2015) compare inequality measures as well as the direct effects of tax policies based on CPS and SOI-PUF data and show that results are very similar (except for the very top of the distribution).

and by following Piketty and Saez (2003) and Piketty (2003) in assuming that top incomes are well represented by a Pareto distribution.

In a first step, we address the challenge that public topcoding methods vary over time. In most recent years (since 2011), the Census Bureau uses a rank proximity swapping procedure to preserve the privacy for top income earners while maintaining the internal distribution of top incomes. In this procedure, values at or above a specific swap threshold are switched against other top income values within a bounded interval. For previous years, however, the CPS data originally contains top income values that are based on different procedures, in particular traditional topcoding (1962-1995), and a replacement value system procedure (1996-2010). To be able to consistently analyze the data, we apply the most recent method of rank proximity swapping also to previous years using supplementary files provided by IPUMS.[50] Thereby, we preserve the internally used distribution of top incomes whenever possible.

In a second step, we address the challenge that top incomes are also internally censored based on the value range limits of the income variables. As shown by Larrimore et al. (2008), since these censoring thresholds have changed discretely at specific points in time, the share of individuals affected by censoring varies and can reach up to one percent in specific years. To address the unequal representation of censored incomes, we replace censored incomes by random draws from a Pareto distribution. In particular, we first identify for every year and every income type the highest possible income T assigned in a given year. Based on this censoring threshold, we generate for every year and every income type the parameter α of a Pareto distribution with density $f(Y) = \alpha * T^{\alpha} * Y^{-\alpha-1}$. We thereby assume that incomes above the 99th percentile follow a Pareto distribution and thus estimate the shape parameter α as

$$\alpha = \frac{ln\left(\frac{N_{Y \geq p99}}{N_{Y=T}}\right)}{ln\left(\frac{Y_T}{Y_{p99}}\right)}$$

where $N_{Y \geq p99}$ is the number of individuals with an income above the 99th percentile

[50]For details on the treatment of top incomes in general and the data used for rank proximity swapping, see `https://cps.ipums.org/cps/topcodes_tables.shtml` and `https://cps.ipums.org/cps/income_cell_means.shtml`.

of the income distribution, $N_{Y=T}$ is the number of individuals at the highest income, and Y_T and Y_{p99} are the top income and the income at the 99th percentile respectively.[51] Finally, we use the distribution to replace the top incomes T by random draws from this calibrated distribution.[52]

Sample restrictions. In this paper, we focus on married couples.[53] We assume that married couples always file jointly. While married couples can also file separately, this filing status is usually not beneficial (see Figure B6) and is chosen by less than two percent of all tax units (see Figure B5).[54] Similarly, we abstract from the qualifying widow(er) filing status that gives widowed individuals a preferential tax treatment in the two years following the spouses' death. Given our sample restriction, the occurrence of widow(er)s is negligible (see also Figure B5). If not indicated otherwise, we restrict the sample to tax units in which primary and secondary taxpayer are between 25 and 55 years old and have non-negative gross income. This sample restriction is guided by (i) our model that considers neither education nor retirement decisions, and (ii) the assumptions on labor supply responses to taxation that are not valid for young and old people with weak labor force attachment.

Throughout the analysis, we calculate tax payments as well as average and marginal tax rates based on the federal income tax and abstract from state income tax and social security payroll taxes. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income.

[51]Discussions of different estimation methods for the shape parameter of the Pareto distribution can be found in Armour, Burkhauser, and Larrimore (2016) and Blanchet, Fournier, and Piketty (2022).

[52]To reduce the impact of random sampling on our results, we use quantiles of the distribution. The number of quantiles utilized depends on the number of individuals at the top income. For instance, if we observe 25 individuals at the top income, we assign these individuals income levels that correspond to the 25 quantiles of the randomly drawn values from the calibrated Pareto distribution. Thereby, we preserve the properties of the distribution while limiting the influence of random draws.

[53]Our companion paper Bierbrauer et al. (2024) discusses the relationship between singles and married couples explicitely.

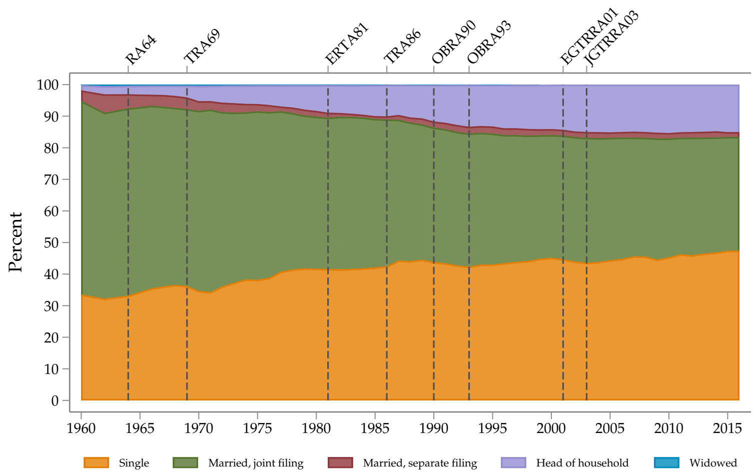
[54]Filing separately can be beneficial in very particular circumstances that we do not observe, i.e., in the case of substantial itemizable deductions (e.g. high medical expenses or student loan repayments).

Table B1: TAXSIM variables and CPS application

TAXSIM Variable	Explanation	CPS Application
taxsimid	Case ID	N/A
year	Tax year	ASEC income reference year
state	State	State of residence
mstat	Marital Status	Marital status (married vs. unmarried)
page	Age of primary taxpayer	Age of husband
sage	Age of spouse	Age of spouse
depx	Number of dependents	Number of children below and of age 18 + additional dependents
dep13	Number of children under 13	Number of children under 13
dep17	Number of children under 17	Number of children under 17
dep18	Number of qualifying children for EITC.	Number of children below and of age 18
pwages	Wage and salary income of Primary Taxpayer	Wage income + business income + farm income of husband
swages	Wage and salary income of Spouse	Wage income + business income + farm income of spouse
dividends	Dividend income	Income from dividends
intrec	Interest Received	Income from interest
stcg	Short Term Capital Gains or losses	N/A
ltcg	Long Term Capital Gains or losses.	Capital gains - capital losses
otherprop	Other property income	Income from rent
nonprop	Other non-property income	Income from other Source not specified + income from alimony
pensions	Taxable Pensions and IRA distributions	Retirement income
gssi	Gross Social Security Benefits	Social Security income
ui	Unemployment compensation received	Income from unemployment benefits
transfers	Other non-taxable transfer Income	Welfare (public assistance) income + income from worker's compensation + income from veteran benefits + income from survivor benefits + income from disability benefits + income from child support + income from educational assistance + income from SSI + income from assistance
rentpaid	Rent Paid	N/A
proptax	Real Estate taxes paid	Annual property taxes
otheritem	Other Itemized deductions	Indirect calculation via difference between adjusted gross income and taxable income calculated by the Census Bureau's taxy model.
childcare	Child care expenses	N/A
mortgage	Deductions not included in otheritem	N/A

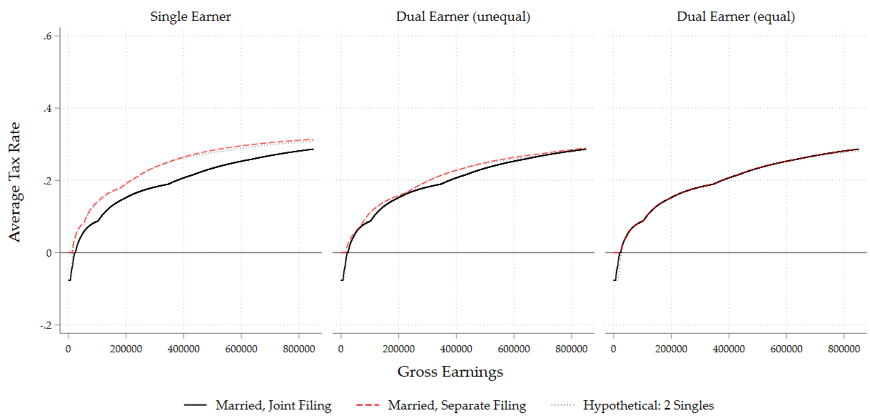
Note: This table displays the variables utilized as part of the tax calculation via the NBER TAXSIM (v32) microsimulation model and the corresponding information from the CPS used for the respective variables. For details on TAXSIM (v32) see Feenberg and Coutts (1993) and `https://users.nber.org/~taxsim/`.
Source: NBER TAXSIM and CPS-ASEC

Figure B5: Filing status according to SOI data



Note: This figure shows the distribution of filing status from 1960 to 2016. Filing statuses are based on the IRS-SOI PUF administrative tax return micro data.
Source: Bierbrauer et al. (2024), calculations based on SOI PUF

Figure B6: Married couples filing jointly and separately (2019)



Note: This figure shows how the average tax rate of a couple with specific gross earnings differs between whether this couple files separately or jointly. In addition, the figure also shows the average tax rate of two singles with the same joint income. The figure differentiates further by the type of couple: single earner couples (95% / 5%), unequal dual earner couples (75% / 25%) and dual earner couples with equal incomes (50% / 50%).
Source: Bierbrauer et al. (2024), calculations based on NBER TAXSIM and CPS-ASEC

B.1.2 IRS income tax tabulations

For the estimation of Pareto coefficients, we use tabular data from the Statistics of Income (SOI) program.[55] The data contains detailed tabular information on joint return tax payers with wage income, separated by the size of wage income, gender, and the share of total wage income. It includes all primary and secondary taxpayers with wages, tips and other compensation as reported in the wage and tax statement of income tax return form W-2.

Data preparation. We prepare the tables for 2008-2018 in two steps. First, since some cells have been merged by the IRS to avoid information disclosure for specific taxpayers, we need to impute some information. Such an imputation is not necessary to estimate Pareto coefficients for a particular group. However, to aggregate information across groups, e.g., to obtain data on all secondary earners within a specific income range, the income ranges need to be harmonized. To fill in missing values for unmerged cells, we distribute the merged number of tax payers and wages according to the distribution of all tax payers of a specific gender. For instance, for the tax year 2018, among women with a wage share between 25 and 50 percent, there are 127 tax payers who earn between 1.5 million and 10 million (merged cell). To allocate these tax payers to the more detailed income brackets of 1.5-2 / 2-5 / 5-10 million (unmerged cells), we use information on how tax payers are distributed within the merged income cell among all women. This allocation procedure is applied for both the number of tax payers and the sum of wages, separately for men and women.

Second, we combine the cell information for men and women with different wage shares to data on primary and secondary earners, i.e. we combine the number of taxpayers and the sum of wages for men and women with wage shares between 0 and 50 percent (secondary earners) and for men and women with wage shares between 50 and 100 percent (primary earners). The final table contains separate primary and secondary earner data on the number of tax payers and the sum of wages for different income ranges. We further produce tables for single earner couples and equal earning dual earner couples, for which the primary (secondary) earner's income share lies between 50 (25) and 75 (50) percent. In case of inconsistencies,

[55]See IRS, Statistics of Income Division, Form W-2 study, June 2019.

for instance when the imputation yields an average income that is lower than the bracket threshold, we abstract from this data point for the Pareto interpolation.

Pareto interpolation. We use the tabular information on the number of taxpayers and the sum of wage income in a particular wage income bracket to generate empirical quantiles associated with the lower threshold and average wage income of a specific wage income bracket. We then follow the generalized Pareto interpolation technique by Blanchet, Fournier, and Piketty (2022) to obtain information on specific percentiles of the distribution and the associated Pareto coefficient. In contrast to standard Pareto interpolation, generalized Pareto interpolation in principle allows the Pareto coefficient to vary across the distribution and thereby provide a more flexible way to describe distributions based on tabular data.

B.2 Conditional revenue functions

In Section 3.1 of the main text, we compute empirical conditional revenue functions for the analysis of Pareto-improving tax reforms in the domain of couples taxation. In this section of the appendix, we provide more detailed explanations on how empirical revenue functions are constructed (Section B.2.1). We also show supplementary results on the empirical ingredients of these conditional revenue functions (Section B.2.2), robustness tests regarding behavioral responses at the extensive and intensive margin (Section B.2.3), and results for alternative choices of the conditioning brackets (Section B.2.4).

B.2.1 From theory to data

As discussed in the main text, to bring conditional revenue functions to the data, we condition on income ranges as described in equation (10). In the main text, we showed in equation (11) that the adjusted equation for secondary earners based on Proposition 1 reads

$$\mathcal{R}_2(y_{2s} \mid B_1) := s_1(B_1)\Big(I_2^d(y_{2s}, B_1) \ + \ M_2^{xd}(y_{2s}, B_1)\Big) \tag{98}$$

where $s_1(B_1) = F_{y_1}(\bar{b}_1) - F_{y_1}(\underline{b}_1)$ is the mass of couples with $y_1 \in B_1 = [\underline{b}_1, \bar{b}_1]$,

$$I_2^d(y_{2s}, B_1) \quad := \quad -\mathbb{E}_{y_1}\left[f_{y_2}(y_{2s} \mid y_1)\frac{T^0_{y_2}(y_1, y_{2s})}{1 - T^0_{y_2}(y_1, y_{2s})}\ y_{2s}\,\varepsilon_2 \mid y_1 \in B_1\right]$$

gives the revenue effect of behavioral responses at the intensive margin, where

$$M_2^{xd}(y_{2s}, B_1) \quad := \quad \mathbb{E}_{y_1}\left[(1 - F_{y_2}(y_{2s} \mid y_1))(1 - \mathcal{E}_x^d(y_{2s} \mid y_1)) \mid y_1 \in B_1\right] .$$

For the empirical application, we assume that the conditional distribution of secondary earnings does not vary within the specific bracket of primary earnings that we are looking at. This assumption is necessary, since we require enough data points to estimate the secondary earner income distribution. We also assume that the extensive margin responses do not vary across the condtioning earnings variable. Consequently, the empirically estimated revenue function simplifies to

$$\begin{aligned}\mathcal{R}\left(y_{2s} \mid \underline{y}_1^Q \leq y_{1s} < \overline{y}_1^Q\right) = \underbrace{s_1^Q}_{\text{Share Inc. Range}} & \Bigg(\underbrace{- \bar{\beta}^d_{I,2}\left(y_{1s}, y_{2s}\right) f^{y_2}(y_{2s} \mid \underline{y}_1^Q \leq y_{1s} < \overline{y}_1^Q)}_{\text{Intensive Margin}} \\ & + \Big(1 - \underbrace{F^{y_2}\left(y_{2s} \mid \underline{y}_1^Q \leq y_{1s} < \overline{y}_1^Q\right)}_{\text{Sec. Earnings Cond. CDF}}\Big) \underbrace{\left(1 - \mathcal{E}_x^d\right)}_{\text{Extensive Margin}} \Bigg) \\ \bar{\beta}^d_{I,2} = \mathbb{E}_{y_{1s}} & \left[\frac{T^{0'}\left(y^{d0}\right)}{1 - T^{0'}\left(y^{d0}\right)} y_{2s}^{d0}\varepsilon_2 \mid \underline{y}_1^Q \leq y_{1s} < \overline{y}_1^Q\right]\end{aligned} \tag{99}$$

where $\underline{y}_1^Q$ and $\overline{y}_1^Q$ indicate lower and upper thresholds for the respective primary earner income range Q, upon which the revenue functions are conditioned.

The equation provides an intuitive understanding of the mechanics behind the conditional revenue function. The term s_1^Q reflects the share of dual earner couples whose primary earnings fall in the respective income range, and are thus affected by a tax reform. The first part of the equation captures the change in tax revenue through behavioral responses at the intensive margin. The intensive margin behavioral response affects all secondary earners at the income level y_{2s} with primary earnings in the respective income range. $\bar{\beta}^d_{I,2} = \mathbb{E}_{y_{1s}}\left[\frac{T^{0'}\left(y^{d0}\right)}{1-T^{0'}\left(y^{d0}\right)} y_{2s}^{d0}\varepsilon_2 \mid \underline{y}_1^Q \leq y_{1s} < \overline{y}_1^Q\right]$

describes the reaction of these secondary earners, and is based on an average at every secondary earnings level y_{2s}, because marginal tax rates faced by secondary earners, as well as elasticities, might vary across primary earnings in the bracket. The second part of the equation captures the mechanical change in tax revenue. This mechanical change in tax revenue is downscaled by the behavioral response at the extensive margin. As a whole, the value and shape of the empirical conditional revenue function characterize the potential for Pareto improving tax reforms for secondary earners with a spouse in specific income ranges (see Theorem 1).

We also compare *conditional* revenue functions to the respective *unconditional* revenue function. The unconditional revenue function illustrates the potential for Pareto improvements through tax rate changes that are independent from the earnings of the other spouse. For secondary earners in dual earner couples, this unconditional revenue function reads

$$\mathcal{R}\left(y_{2s}\right) = \underbrace{-\mathbb{E}_{y_{1s}}\left[\frac{T^{0'}\left(y^{d0}\right)}{1-T^{0'}\left(y^{d0}\right)} y_{2s}^{d0}\varepsilon_2\right] f^{y_2}(y_{2s})}_{\text{Intensive Margin}} + \underbrace{\left(1-F^{y_2}\left(y_{2s}\right)\right)}_{\text{Mechanical Effect}} \underbrace{\left(1-\mathcal{E}_x^d\right)}_{\text{Extensive Margin}} . \quad (100)$$

A comparison between conditional and unconditional revenue functions reveals whether inefficiencies in the tax code can be cured by changing marginal tax rates for *all* secondary earners or whether curing inefficiencies requires the joint consideration of secondary and primary earnings.

B.2.2 Ingredients

The estimation of revenue functions requires assumptions about labor supply elasticities and conditioning brackets for the reform. In addition, we need estimates on the (conditional) income distributions and intensive margin response terms.

Table B2 display three different scenarios that we use throughout the paper and that are informed by the literature listed in the table note. Table B3 presents the upper thresholds of the income deciles for the primary and secondary earner income distributions that are used as conditioning brackets.

We estimate all income distributions by means of a kernel density estimation using a Gaussian kernel and 20,000 grid points. The income grid ranges from zero to one million dollars.

The intensive margin response terms in equation (99) and (100) demand knowledge of a conditional average. For instance, the secondary earner conditional revenue function contains an average of the intensive margin response at every secondary earnings income level conditional on primary earnings in a specific decile. We estimate these averages by means of fitting a local polynomial regression of degree zero (local mean smoothing) with a Gaussian kernel as a weighting function.

Figures B7 - B16 show the empirical ingredients to compute the conditional revenue functions for secondary earners, in particular marginal tax rates, conditional income distributions, and behavioral responses to taxation. Figures B17-B26 show the respective ingredients for the conditional revenue functions for primary earners.

Table B2: Assumptions about labor supply elasticities

	Primary earner	Secondary earner
Low elasticity scenario	0.15	0.35
Baseline elasticity scenario	0.25	0.75
High elasticity scenario	0.5	1.5

Note: This table displays our assumptions about the labor supply elasticities for primary and secondary earners in married couples. Assumptions are guided by the range of estimates found in the literature, e.g. Blundell and Macurdy (1999), Blau and Kahn (2007), Eissa and Hoynes (2004), LaLumia (2008), Saez, Slemrod, and Giertz (2012), Bargain, Orsini, and Peichl (2014), and Neisser (2021).

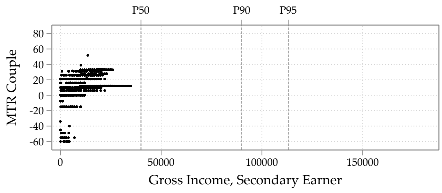
Table B3: Income deciles (2019)

Decile	Primary earnings	Secondary earnings
Q1	35000	10001
Q2	46002	19000
Q3	55604	25015
Q4	65228	32005
Q5	77056	40000
Q6	91032	47300
Q7	108966	55009
Q8	135030	67785
Q9	184155	90001
Q10	18009756	1358632

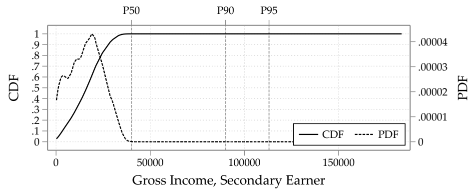
Note: This table shows the upper thresholds of the deciles for the primary and secondary earner income distribution among dual earner couples as of 2019. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income.
Source: Own calculations based on CPS-ASEC (2019)

Figure B7: Conditional revenue functions, sec. earners, PE Q1

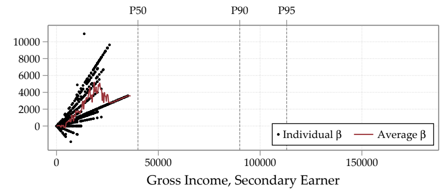
(a) Marginal tax rate (%)



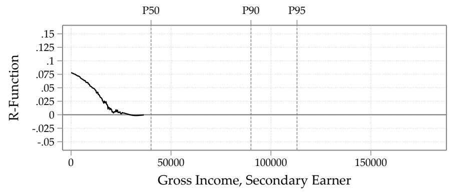
(b) Conditional distribution



(c) β



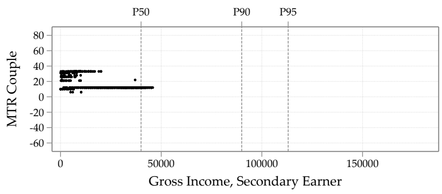
(d) Revenue function



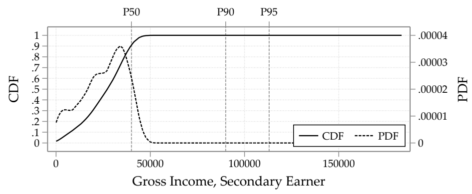
Note: This figure shows the ingredients necessary to estimate equation (99) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for primary earnings in decile Q1. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.75 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B8: Conditional revenue functions, sec. earners, PE Q2

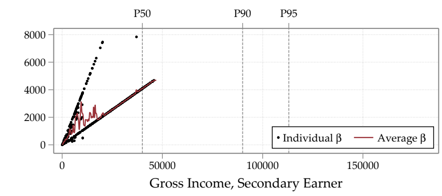
(a) Marginal tax rate (%)



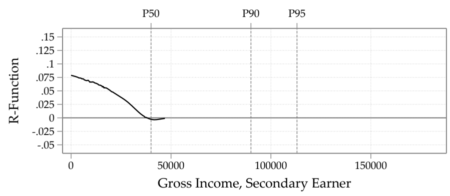
(b) Conditional distribution



(c) β



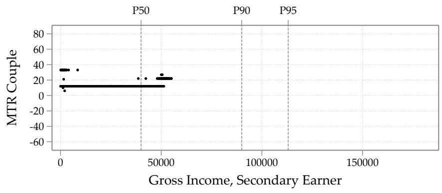
(d) Revenue function



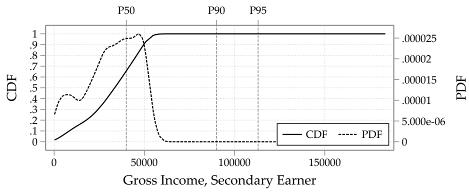
Note: This figure shows the ingredients necessary to estimate equation (99) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for primary earnings in decile Q2. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.75 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B9: Conditional revenue functions, sec. earners, PE Q3

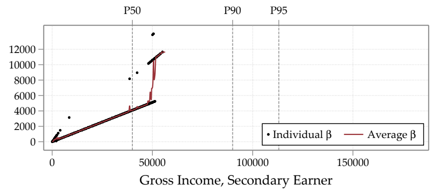
(a) Marginal tax rate (%)



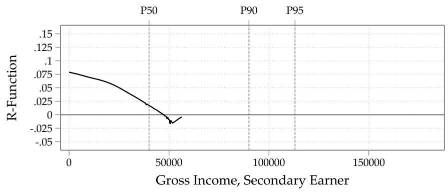
(b) Conditional distribution



(c) β



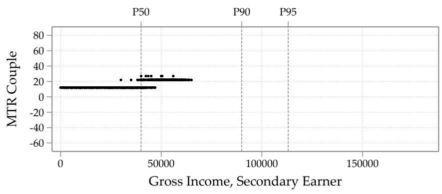
(d) Revenue function



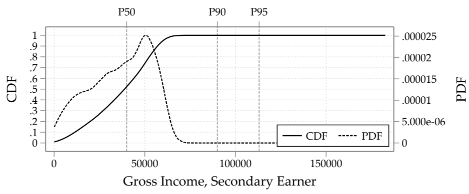
Note: This figure shows the ingredients necessary to estimate equation (99) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for primary earnings in decile Q3. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.75 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B10: Conditional revenue functions, sec. earners, PE Q4

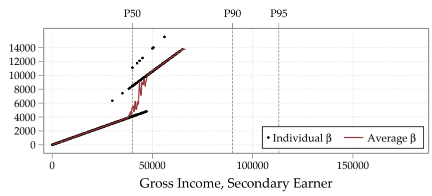
(a) Marginal tax rate (%)



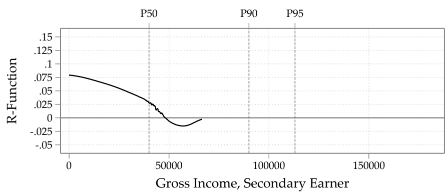
(b) Conditional distribution



(c) β



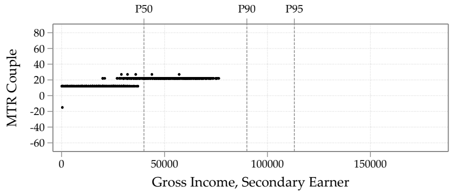
(d) Revenue function



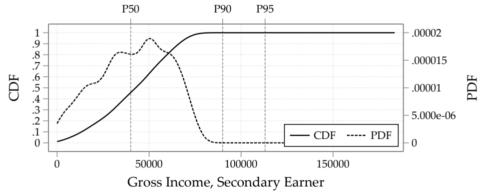
Note: This figure shows the ingredients necessary to estimate equation (99) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for primary earnings in decile Q4. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.75 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B11: Conditional revenue functions, Q5 PE, Ingredients

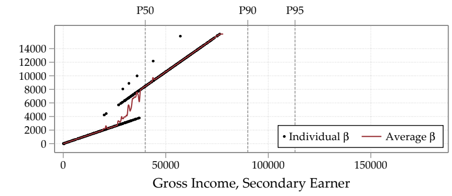
(a) Marginal tax rate (%)



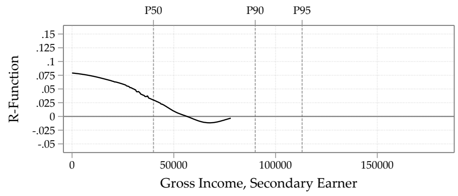
(b) Conditional distribution



(c) β



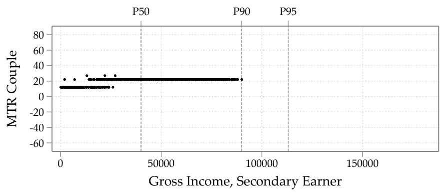
(d) Revenue function



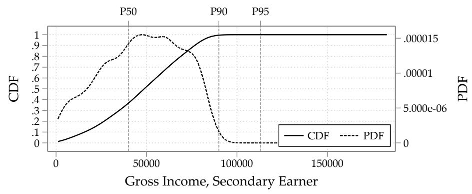
Note: This figure shows the ingredients necessary to estimate equation (99) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for primary earnings in decile Q5. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.75 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B12: Conditional revenue functions, sec. earners, PE Q6

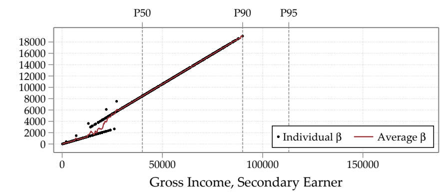
(a) Marginal tax rate (%)



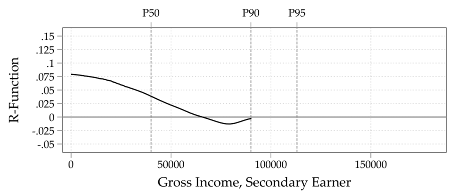
(b) Conditional distribution



(c) β

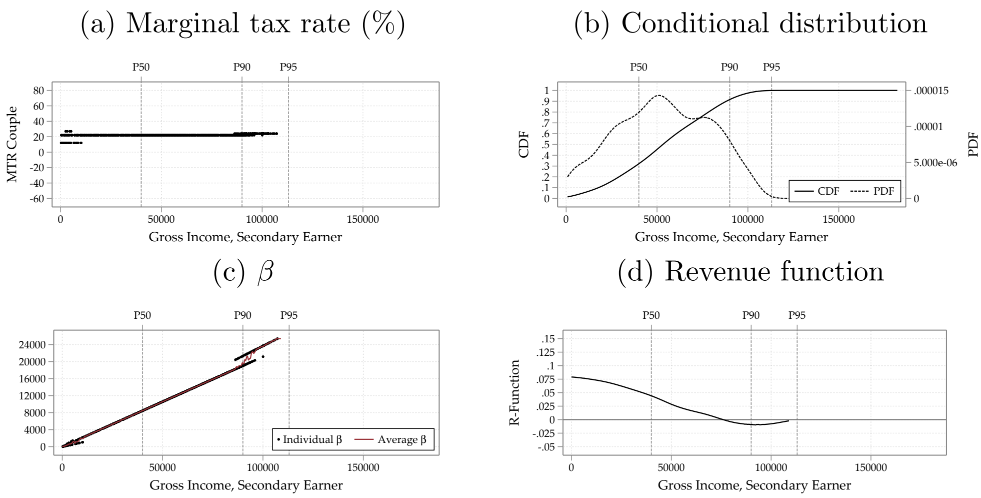


(d) Revenue function



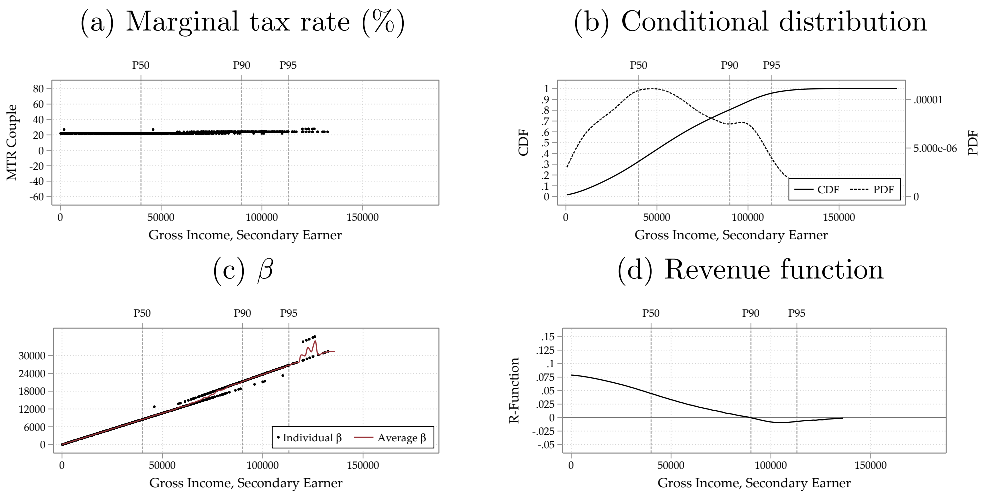
Note: This figure shows the ingredients necessary to estimate equation (99) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for primary earnings in decile Q6. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.75 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B13: Conditional revenue functions, sec. earners, PE Q7



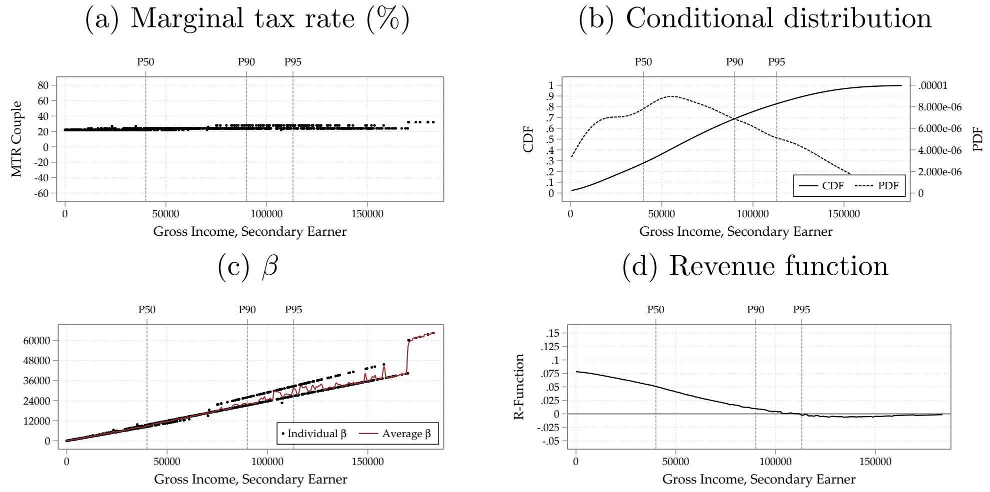
Note: This figure shows the ingredients necessary to estimate equation (99) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for primary earnings in decile Q7. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.75 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B14: Conditional revenue functions, sec. earners, PE Q8



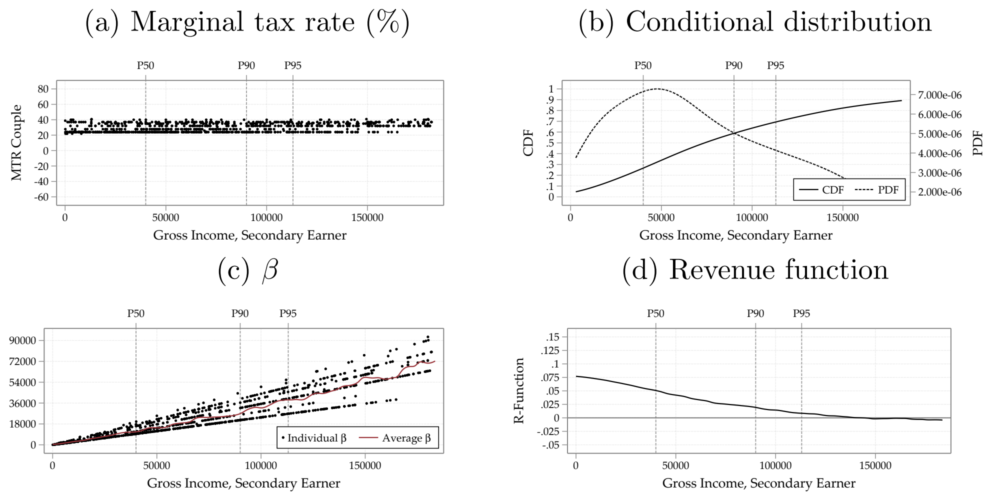
Note: This figure shows the ingredients necessary to estimate equation (99) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for primary earnings in decile Q8. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.75 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B15: Conditional revenue functions, sec. earners, PE Q9



Note: This figure shows the ingredients necessary to estimate equation (99) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for primary earnings in decile Q9. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.75 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

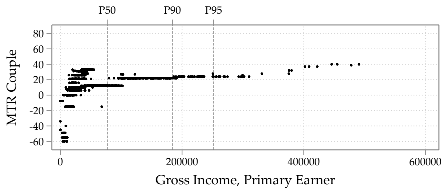
Figure B16: Conditional revenue functions, sec. earners, PE Q10



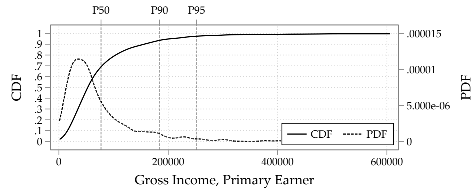
Note: This figure shows the ingredients necessary to estimate equation (99) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for primary earnings in decile Q10. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.75 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B17: Conditional revenue functions, prim. earners, SE Q1

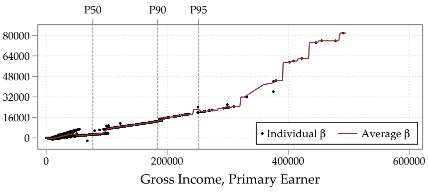
(a) Marginal tax rate (%)



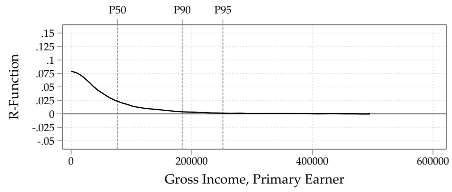
(b) Conditional distribution



(c) β



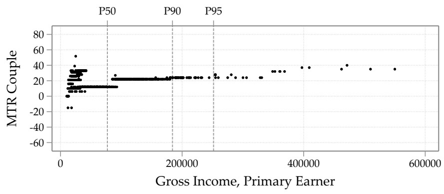
(d) Revenue function



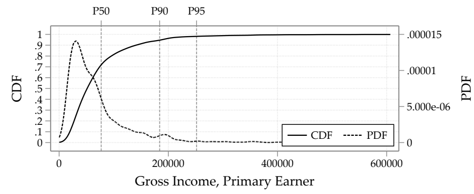
Note: This figure shows the ingredients necessary to estimate equation (99) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for secondary earnings in decile Q1. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B18: Conditional revenue functions, prim. earners, SE Q2

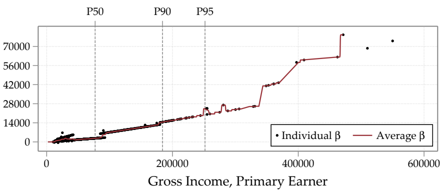
(a) Marginal tax rate (%)



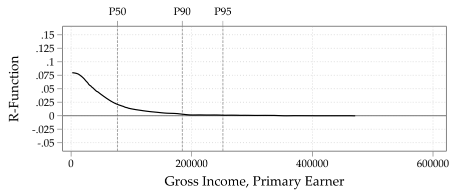
(b) Conditional distribution



(c) β



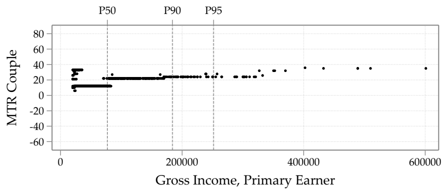
(d) Revenue function



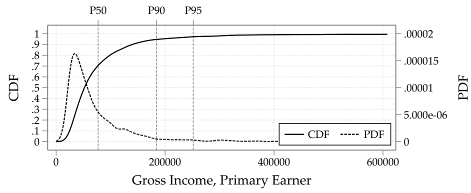
Note: This figure shows the ingredients necessary to estimate equation (99) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for secondary earnings in decile Q2. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B19: Conditional revenue functions, prim. earners, SE Q3

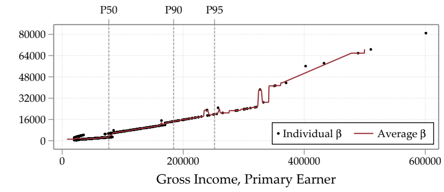
(a) Marginal tax rate (%)



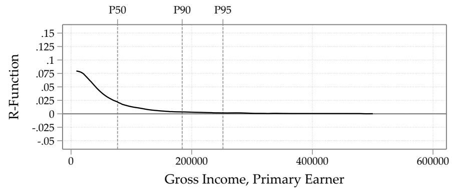
(b) Conditional distribution



(c) β



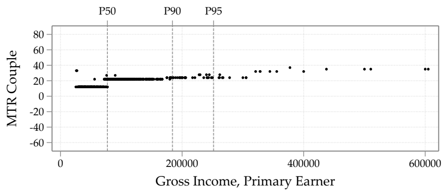
(d) Revenue function



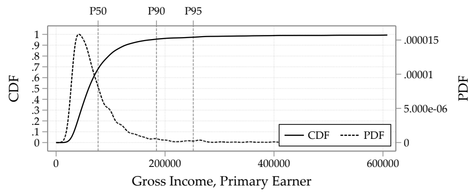
Note: This figure shows the ingredients necessary to estimate equation (99) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for secondary earnings in decile Q3. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B20: Conditional revenue functions, prim. earners, SE Q4

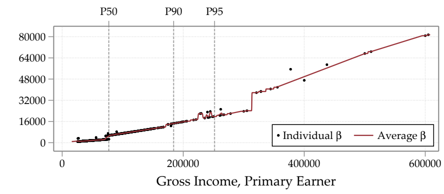
(a) Marginal tax rate (%)



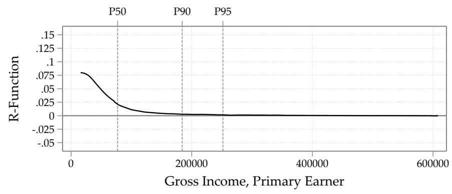
(b) Conditional distribution



(c) β

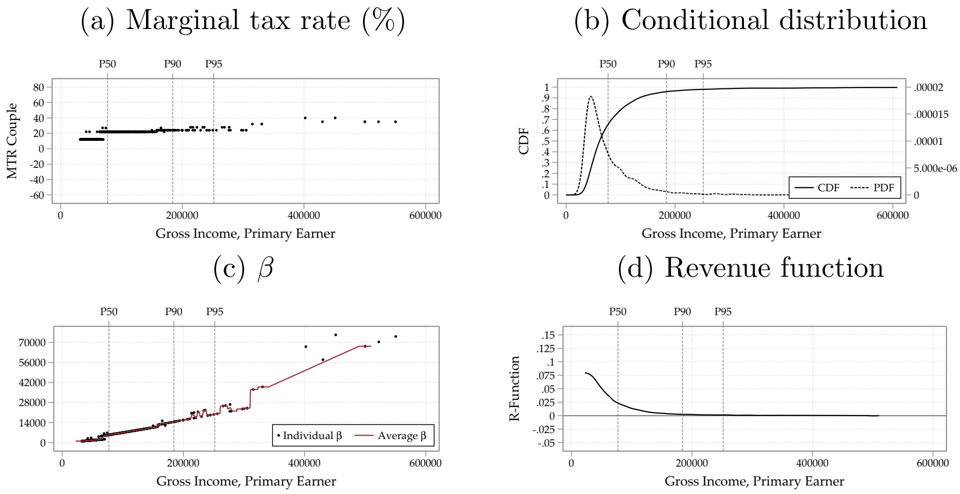


(d) Revenue function



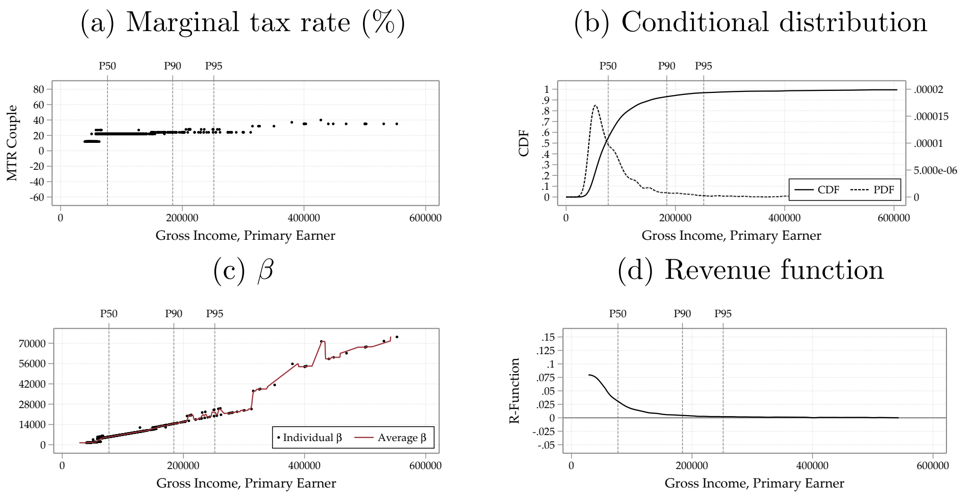
Note: This figure shows the ingredients necessary to estimate equation (99) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for secondary earnings in decile Q4. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B21: Conditional revenue functions, prim. earners, SE Q5



Note: This figure shows the ingredients necessary to estimate equation (99) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for secondary earnings in decile Q5. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

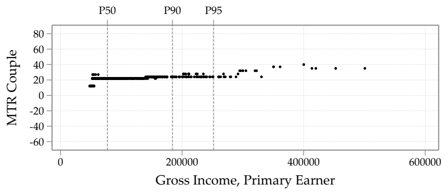
Figure B22: Conditional revenue functions, prim. earners, SE Q6



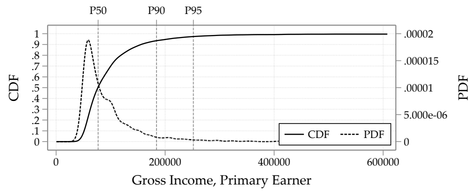
Note: This figure shows the ingredients necessary to estimate equation (99) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for secondary earnings in decile Q6. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B23: Conditional revenue functions, prim. earners, SE Q7

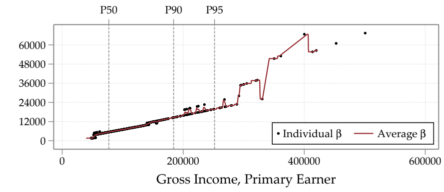
(a) Marginal tax rate (%)



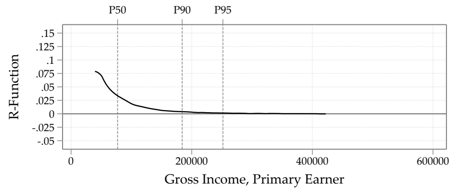
(b) Conditional distribution



(c) β



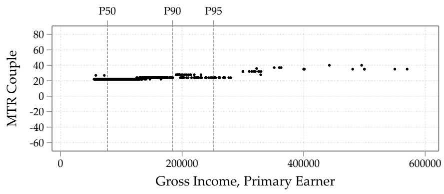
(d) Revenue function



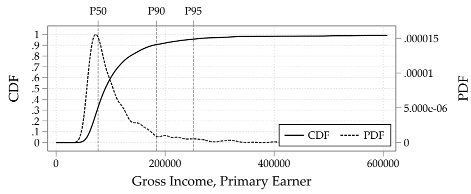
Note: This figure shows the ingredients necessary to estimate equation (99) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for secondary earnings in decile Q7. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B24: Conditional revenue functions, prim. earners, SE Q8

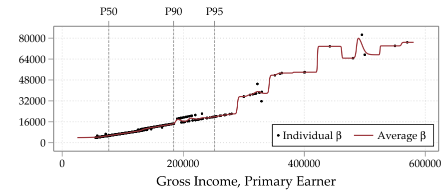
(a) Marginal tax rate (%)



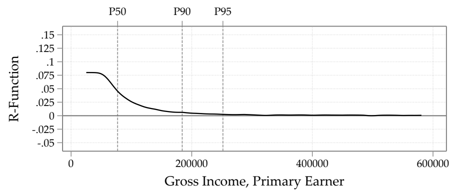
(b) Conditional distribution



(c) β



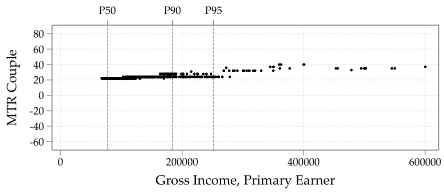
(d) Revenue function



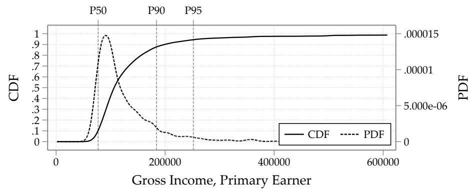
Note: This figure shows the ingredients necessary to estimate equation (99) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for secondary earnings in decile Q8. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B25: Conditional revenue functions, prim. earners, SE Q9

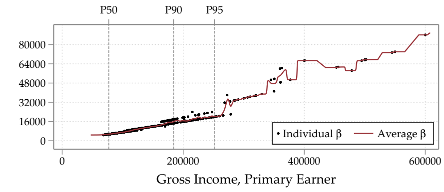
(a) Marginal tax rate (%)



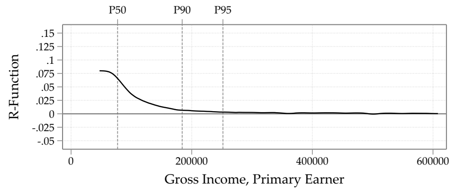
(b) Conditional distribution



(c) β



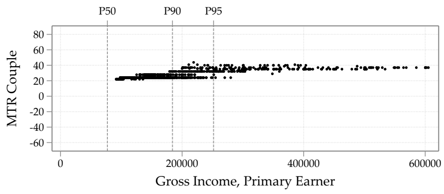
(d) Revenue function



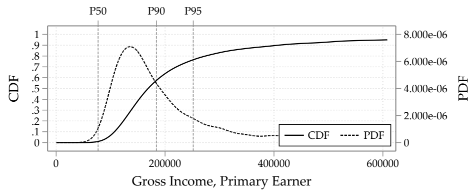
Note: This figure shows the ingredients necessary to estimate equation (99) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for secondary earnings in decile Q9. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B26: Conditional revenue functions, prim. earners, SE Q10

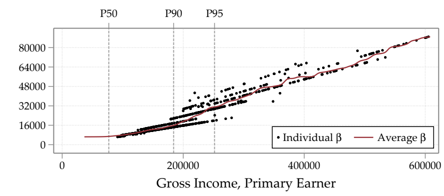
(a) Marginal tax rate (%)



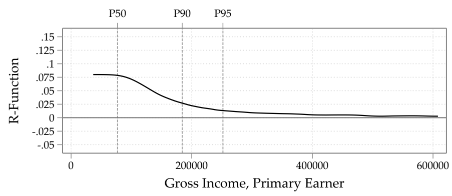
(b) Conditional distribution



(c) β



(d) Revenue function



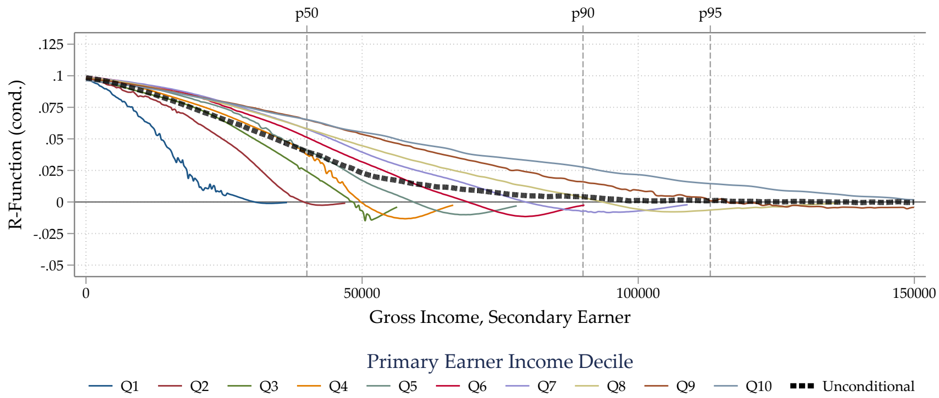
Note: This figure shows the ingredients necessary to estimate equation (99) – i.e. the marginal tax rate (Panel a), the conditional income distribution (Panel b), the intensive margin response β (Panel c), and the resulting conditional revenue function (Panel d) – among dual earner couples as of 2019 for secondary earnings in decile Q10. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

B.2.3 Robustness

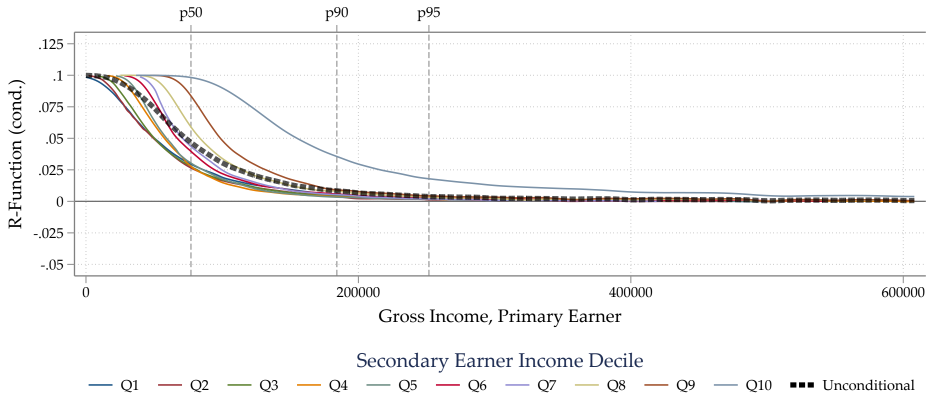
In the baseline, conditional revenue functions are computed under the assumption of extensive margin and moderate intensive margin responses. We provide robustness tests regarding the exclusion of margin responses (Figure B27) and for the use of low (high) elasticity scenarios from Table B2 in Figure B28 (B29).

Figure B27: Cond. revenue functions, deciles (2019), no extensive margin

(a) Secondary earners



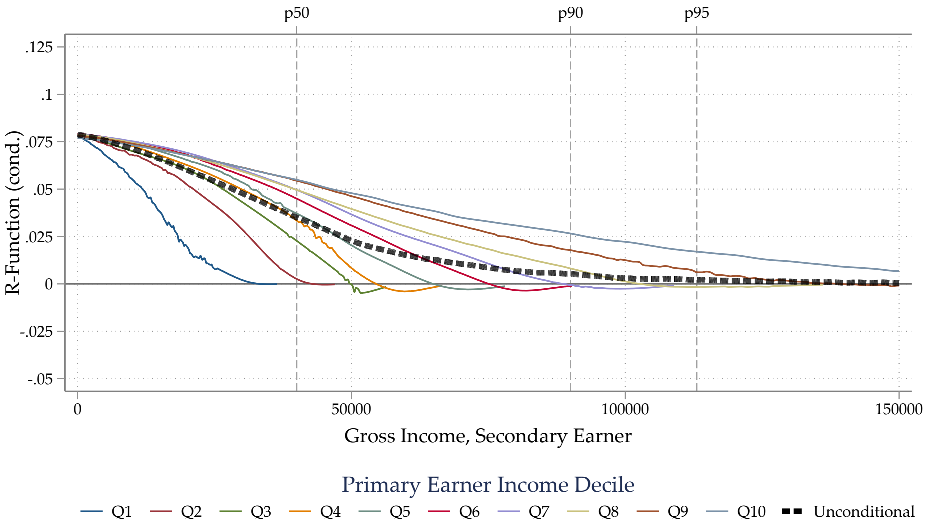
(b) Primary earners



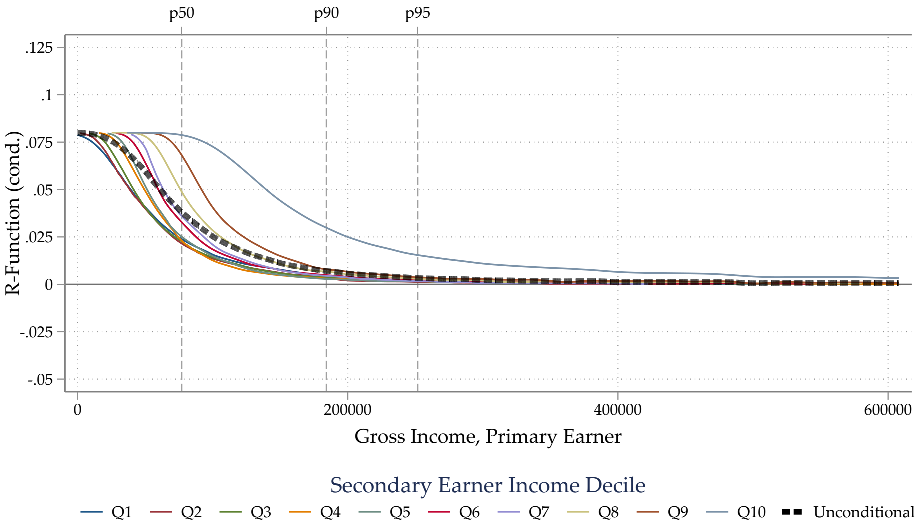
Note: This figure shows conditional revenue functions for secondary (primary) earners in married dual earner couples conditional on primary (secondary) earnings income deciles in Panel a (b) as of 2019. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive margin into account. We assume intensive margin elasticities of 0.25 (0.75) for primary (secondary) earners. The figure also displays modified unconditional revenue functions for secondary and primary earners where unconditional revenue functions have been scaled by 0.1 to facilitate comparability with the conditional revenue functions.
Source: Own calculations based on CPS-ASEC (2019)

Figure B28: Cond. revenue functions, deciles (2019), low elasticity

(a) Secondary earners



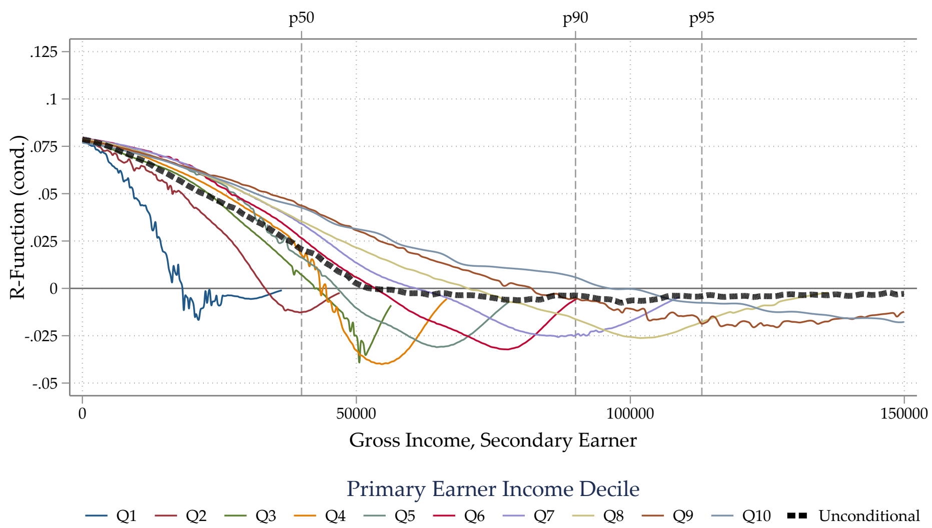
(b) Primary earners



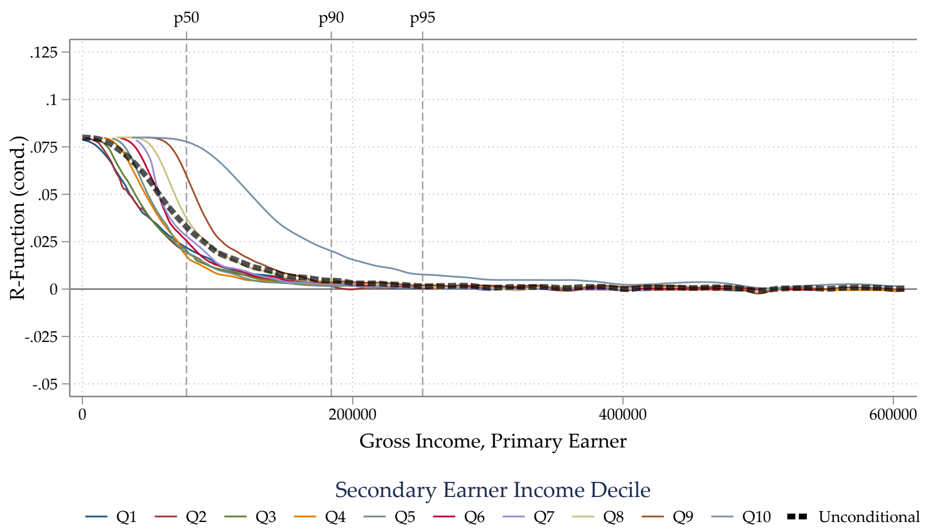
Note: This figure shows conditional revenue functions for secondary (primary) earners in married dual earner couples conditional on primary (secondary) earnings income deciles in Panel a (b) as of 2019. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.15 (0.35) for primary (secondary) earners and an extensive margin elasticity of 0.2. The figure also displays modified unconditional revenue functions for secondary and primary earners where unconditional revenue functions have been scaled by 0.1 to facilitate comparability with the conditional revenue functions. Results for the baseline elasticity scenario are shown in Figure 5 of the main text.
Source: Own calculations based on CPS-ASEC (2019)

Figure B29: Cond. revenue functions, deciles (2019), high elasticity

(a) Secondary earners



(b) Primary earners



Note: This figure shows conditional revenue functions for secondary (primary) earners in married dual earner couples conditional on primary (secondary) earnings income deciles in Panel a (b) as of 2019. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.5 (1.5) for primary (secondary) earners and an extensive margin elasticity of 0.2. The figure also displays modified unconditional revenue functions for secondary and primary earners where unconditional revenue functions have been scaled by 0.1 to facilitate comparability with the conditional revenue functions. Results for the baseline elasticity scenario are shown in Figure 5 of the main text.
Source: Own calculations based on CPS-ASEC (2019)

B.2.4 Alternative conditioning brackets

In the main text, we construct conditioning brackets by looking at deciles of the primary and secondary earner income distribution. Our framework, however, is flexible enough to capture any construction of conditioning brackets. Empirically, we are limited by data constraints, since we need enough observations to estimate conditional distributions. If we condition on wider brackets, we receive more precise results; however we potentially also average out more inefficiencies. If we condition on narrower brackets, we require less averaging, but also have less observations to estimate distributions. In this section, we provide results for wider brackets (quintiles in Figure B30), and more narrow brackets (vingtiles in Figure B31).

An alternative way of conditioning on primary or secondary earnings is to select brackets based on the statutory income tax schedule. Table B4 displays the corresponding income thresholds and the share of primary and secondary earners with incomes in the respective tax bracket. Figure B32 displays the estimated conditional revenue functions for secondary (primary) earners conditional on primary (secondary) earnings being in one of the seven income brackets in 2019. In contrast to the conditioning based on deciles, revenue functions for different tax brackets do not have a common intercept with the vertical axis, because the share of secondary earners in a particular income range is no longer the same. However, conditional revenue functions again show inefficiently high marginal tax rates for some "middle tax bracket" income levels (see Figure B32a).

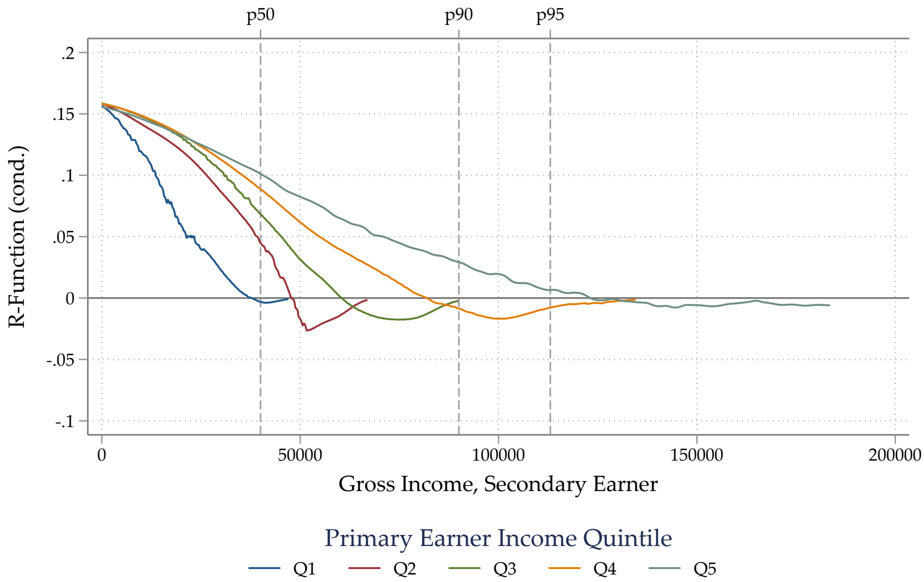
Table B4: Taxable income brackets (2019)

Bracket	Upper threshold	MTR	Share prim. earner	Share sec. earner
1	19400	10	1.83%	20.41%
2	78950	12	48.89%	65.41%
3	168400	22	36.83%	12.75%
4	321450	24	9.46%	1.29%
5	408200	32	1.01%	0.02%
6	612350	35	1.01%	0.05%
7	.	37	0.98%	0.07%

Note: This table shows the upper thresholds (column 2) and marginal tax rates (column 3) for the statutory tax schedule for married couples filing jointly as of 2019. All threshold values refer to taxable income. Column 5 (6) displays the share of primary (secondary) earners whose primary (secondary) earnings lie in a respective bracket.
Source: Own calculations based on CPS-ASEC (2019) and U.S. Federal Individual Income Tax Rates and Brackets (2019)

Figure B30: Conditional revenue functions, quintiles (2019)

(a) Secondary earners

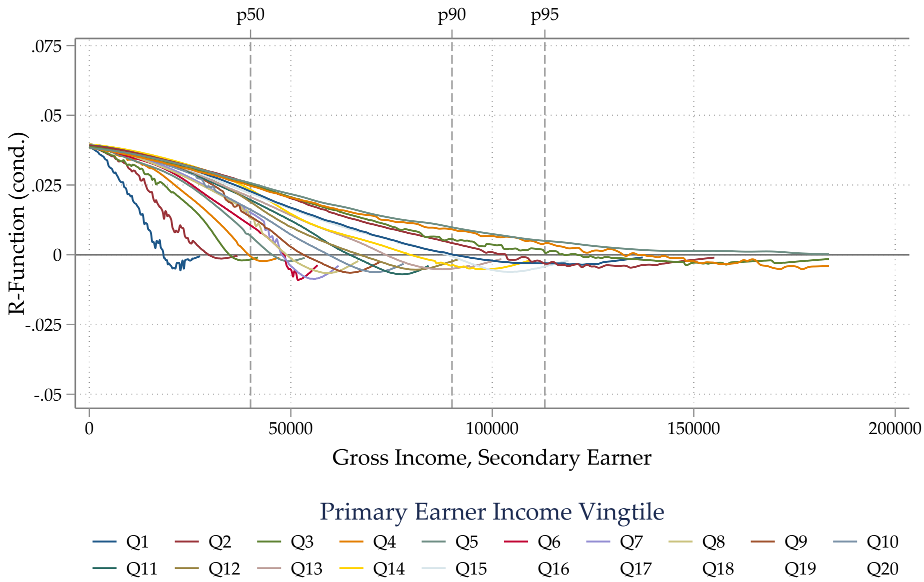


(b) Primary earners

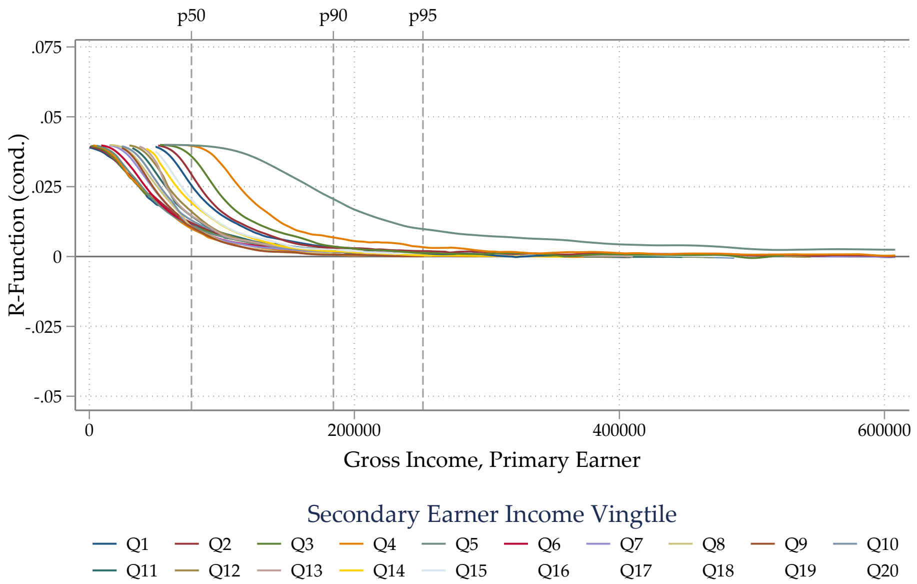
Note: This figure shows conditional revenue functions for secondary (primary) earners in married dual earner couples conditional on primary (secondary) earnings income quintiles in Panel a (b) as of 2019. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 (0.75) for primary (secondary) earners and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B31: Conditional revenue functions, vingtiles (2019)

(a) Secondary earners



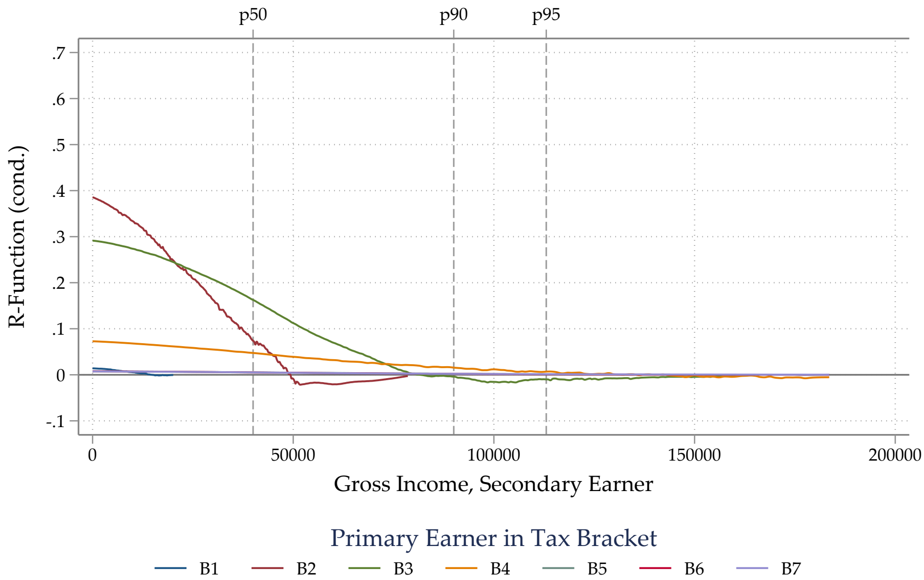
(b) Primary earners



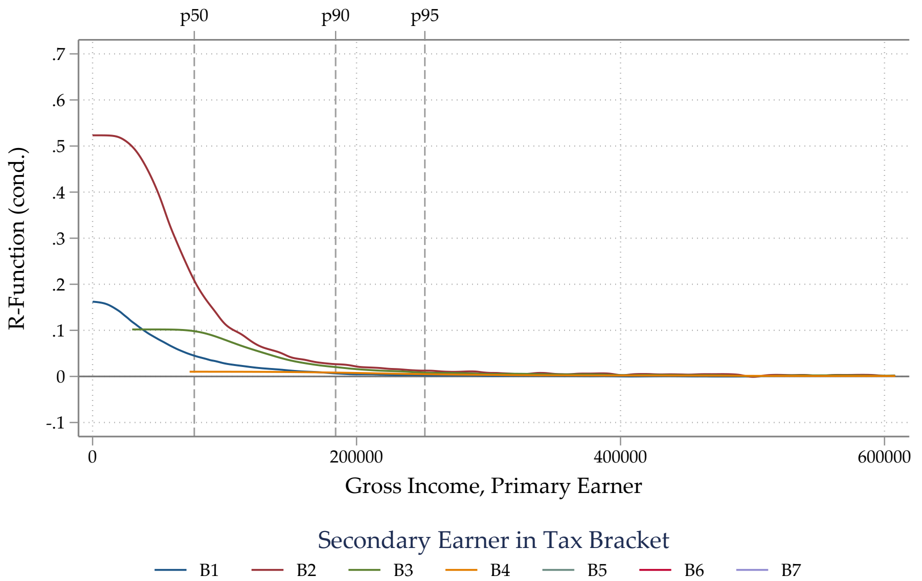
Note: This figure shows conditional revenue functions for secondary (primary) earners in married dual earner couples conditional on primary (secondary) earnings income vingtiles in Panel a (b) as of 2019. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 (0.75) for primary (secondary) earners and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

Figure B32: Conditional revenue functions, tax brackets (2019)

(a) Secondary earners



(b) Primary earners



Note: This figure shows conditional revenue functions for secondary (primary) earners in married dual earner couples conditional on primary (secondary) earnings in specific income tax brackets in Panel a (b) as of 2019. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. We assume intensive margin elasticities of 0.25 (0.75) for primary (secondary) earners and an extensive margin elasticity of 0.2.
Source: Own calculations based on CPS-ASEC (2019)

B.3 Reducing jointness

In Section 3.3 of the main text, we estimate the welfare effects of jointness-reducing reforms at different combinations of primary and secondary earnings. In this section of the appendix, we provide detailed information on the construction of empirical welfare assessments (Section B.3.1), and show supplementary results and robustness checks (Section B.3.2).

B.3.1 From theory to data

Following equation (13), the welfare effect of a reform can be approximated as

$$W = \tau\ \ell_1\ \ell_2 \Bigg[\mathcal{R}_2\left(y_{2s} \mid [y_{1s}, \bar{y}_1]\right) + \mathcal{R}_1\left(y_{1s} \mid [y_{2s}, \bar{y}_2]\right) - 2\ s\left(y_{1s}, y_{2s}\right) \mathcal{G}\left(y_{1s}, y_{2s}\right) \Bigg] \tag{101}$$

where the first two terms are revenue effects coming from the reduction in marginal tax rates for secondary (primary) earners conditional on primary (secondary) earnings ranges while the third term captures the adjustment required by specific welfare objectives. The conditional revenue functions are estimated as

$$\begin{aligned} \mathcal{R}_i\left(y_i \mid [\underline{b}_j, \bar{y}_j]\right) = & \underbrace{s(y_j \geq \underline{b}_j)}_{\text{Share in Inc. Range}} \left(\underbrace{-\overline{\beta}^d_{I,i}\left(y_i, y_j\right) f^{y_i}(y_i \mid y_j \geq \underline{b}_j)}_{\text{Intensive Margin}} \right) \\ & + \underbrace{s(y_j \geq \underline{b}_j)}_{\text{Share in Inc. Range}} \left(1 - \underbrace{F^{y_i}\left(y_i \mid y_j \geq \underline{b}_j\right)}_{\text{Cond. CDF}}\right) \underbrace{\left(1 - \mathcal{E}^d_x\right)}_{\text{Extensive Margin}} \end{aligned} \tag{102}$$

$$\overline{\beta}^d_{I,i} = \mathbb{E}_{y_i}\left[\frac{T^{0'}\left(y^{d0}\right)}{1 - T^{0'}\left(y^{d0}\right)} y_i^{d0} \varepsilon_i \mid y_j \geq \underline{b}_j \right] \quad , \quad (i,j) \in \{(1,2),(2,1)\} \quad . \tag{103}$$

Specifications. For consistency reasons, assumptions about behavioral responses follow the previous section (see Table B2). We use three different specifications of welfare weights displayed in Table B5. Reforms are carried out at bracket combinations defined by the primary earnings deciles and the respective conditional secondary earnings deciles (see Table B6).

Table B5: Welfare weights

Welfare Weight Type	Welfare Weight
Equal Weights	$g(y_1, y_2) = 1$
Quasi-Rawlsian Weights	$g(y_1, y_2) = \begin{cases} 1, & \text{for } y_1 + y_2 \leq \mathrm{P}_{10} \\ 0, & \text{for } y_1 + y_2 \geq \mathrm{P}_{10} \end{cases}$
Feminist Weights	$g(y_1, y_2) = \begin{cases} \frac{y_2}{y_1+y_2}, & \text{for } y_2 = y_{\text{woman}} \\ \frac{y_1}{y_1+y_2}, & \text{for } y_1 = y_{\text{woman}} \end{cases}$

Note: This table shows the specifications of different welfare weights used in the empirical analysis. For quasi-Rawlsian weights, P_{10} refers to the tenth percentile of the joint income distribution of dual earner couples. Welfare weights are normalized to mean one.

Table B6: Reforms on jointness, selected reform brackets (2019)

Primary earnings bracket	$\underline{b}_1$	Secondary earnings bracket, $\underline{b}_2$: $B_2(1)$	$B_2(2)$	$B_2(3)$	$B_2(4)$	$B_2(5)$	$B_2(6)$	$B_2(7)$	$B_2(8)$	$B_2(9)$	$B_2(10)$
$B_1(1)$	1	1	3200	7488	10167	14001	16001	19001	20010	24001	27750
$B_1(2)$	35001	1	7500	15002	20001	24000	28002	30115	34001	36001	40001
$B_1(3)$	46003	1	8484	18006	25001	30001	34011	38001	41802	45603	50001
$B_1(4)$	55625	1	12001	20001	27002	33501	40001	45001	49792	52002	56004
$B_1(5)$	65255	1	14288	23150	30086	36040	42202	49550	53700	60006	65353
$B_1(6)$	77060	1	16102	26523	35012	42005	50001	55271	61100	70005	76401
$B_1(7)$	91036	1	18200	30005	39200	46225	52800	60124	70010	79501	88060
$B_1(8)$	109000	1	16200	28307	37757	47000	55028	65125	76150	90001	101051
$B_1(9)$	135031	1	16001	30005	42510	55001	65218	77252	91088	107345	128001
$B_1(10)$	184202	1	15106	30014	45200	59100	73500	92301	114400	140223	186300

Note: This table shows the lower bracket thresholds used for the analysis of different jointness-reducing reforms. Primary earner brackets (vertical) are chosen based on deciles of the primary earner income distribution. Secondary earner brackets (horizontal) are chosen based on the secondary earner deciles conditional on the respective primary earner decile. The red (blue) cells indicate the bracket combination, for which reforms are discussed in detail (Figure B33 and Figure B34).
Source: Own calculations based on CPS-ASEC (2019)

Mechanics and intuition. To understand the mechanics behind the welfare evaluation based on the selection of welfare weights in Table B5, it can be helpful to rewrite and combine equations (101) - (103) and decompose the welfare effect as

$$\begin{aligned} W = \tau\, l(B_1)\, l(B_2) \Bigg[& \underbrace{s(y_1 \geq \underline{b}_1) f^{y_2}(\underline{b}_2 \mid y_1 \geq \underline{b}_1)}_{\text{Share Affected, Behavioral Effect}} \Bigg(\underbrace{-\overline{\beta}^d_{I,2}\,(y_1, \underline{b}_2)}_{\text{Intensive Margin}} \Bigg) \\ & + \underbrace{s(y_1 \geq \underline{b}_1, y_2 \geq \underline{b}_2)}_{\text{Share Affected, Mechanical Effect}} \Bigg(1 - \underbrace{\mathcal{E}^d_x}_{\text{Ext. margin}} - \underbrace{\mathcal{G}\,(\underline{b}_1, \underline{b}_2)}_{\text{Welfare}} \Bigg) \\ & + \underbrace{s(y_2 \geq \underline{b}_2) f^{y_1}(\underline{b}_1 \mid y_2 \geq \underline{b}_2)}_{\text{Share Affected, Behavioral Effect}} \Bigg(\underbrace{-\overline{\beta}^d_{I,1}\,(y_2, \underline{b}_1)}_{\text{Intensive Margin}} \Bigg) \\ & + \underbrace{s(y_2 \geq \underline{b}_2, y_1 \geq \underline{b}_1)}_{\text{Share Affected, Mechanical Effect}} \Bigg(1 - \underbrace{\mathcal{E}^d_x}_{\text{Extensive Margin}} - \underbrace{\mathcal{G}\,(\underline{b}_2, \underline{b}_1)}_{\text{Welfare}} \Bigg) \Bigg]. \end{aligned} \tag{104}$$

Equation (104) first decomposes the welfare effect into the welfare effect based on (i) changes of secondary marginal tax rates conditional on primary earnings (lines 1+2), and the welfare effect based on changes of primary marginal tax rates conditional on secondary earnings (lines 3+4). Each of the two effects can be further decomposed into the change in tax revenue coming from intensive margin behavioral responses (lines 1 and 3), and from the mechanical revenue effect corrected for both extensive margin responses and welfare evaluations (lines 2 and 4). For instance, an increase in secondary earner marginal tax rates triggers intensive margin behavioral responses for all couples with primary earnings above $\underline{b_1}$ and secondary earnings at $\underline{b_2}$. In contrast, all couples with primary earnings above $\underline{b_1}$ and secondary earnings above $\underline{b_2}$ pay now more taxes due to the reform. This mechanical effect is potentially downscaled by the extensive margin elasticity representing dual earner couples that decide to become single earner couples as a response to the increased tax liability. In addition, the mechanical effect is further downscaled by how much one cares about those couples that pay for the increase in revenue in terms of a higher tax liability.

Revisiting the social welfare objectives specified in Table B5 can provide some additional intuition. If the social planner operates under a Rawlsian welfare objective, the welfare terms are zero if the reform is done at income levels that are above the tenth decile of the joint income distribution. Welfare evaluations then boil down to how much tax revenue is raised through the reform. Under equal welfare weights, in contrast, one would care about all couples equally, i.e. a one dollar higher mechanical tax revenue increase is completely counteracted by lower welfare of those that are paying for this revenue with a higher tax liability. Welfare evaluations then boil down to the intensive and extensive margin behavioral responses that are triggered by the reform.

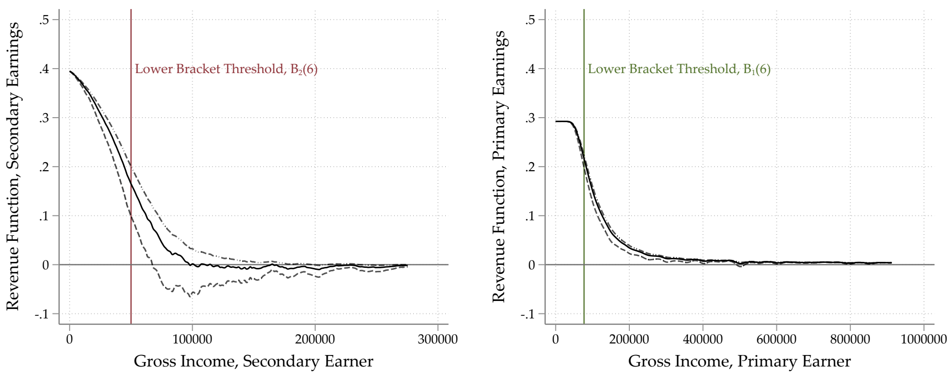
B.3.2 Supplementary graphs and robustness

Figure B33 provides supplementary details on the empirical conditional revenue functions that are used to estimate the welfare effect for a reduction of jointness around the median in Figure 6. We also provide an illustration of the welfare effect for a bracket combination of mid levels of primary earnings and high secondary earnings that is discussed in the main text (Figure B34). The effects of a reform at this bracket combination illustrates that the revenue effects from jointness-reducing reforms can move in opposite directions for the primary and secondary earner marginal tax rate changes. In the particular case, due to the inefficiencies in the tax treatment of secondary earnings, the tax cut for secondary earner marginal tax rates yields a revenue increase. However, the corresponding revenue loss coming from the tax cut of primary earner marginal tax rates is higher. Consequently, despite the inefficiencies, a quasi-Rawlsian welfare function would still not agree to a jointness-reducing reform in this area.

Figure B35 replicates the results on the welfare implications of changing jointness at different bracket combinations and illustrates, how results change under different elasticity scenarios. Figure B36 provides the results without extensive margin responses.

Figure B33: Welfare effect, reducing jointness at $B_1(6)$, $B_2(6)$

(a) $\mathcal{R}_2\left(y_2 \mid [\underline{b}_1, \bar{y}_1]\right)$ (b) $\mathcal{R}_1\left(y_1 \mid [\underline{b}_2, \bar{y}_2]\right)$



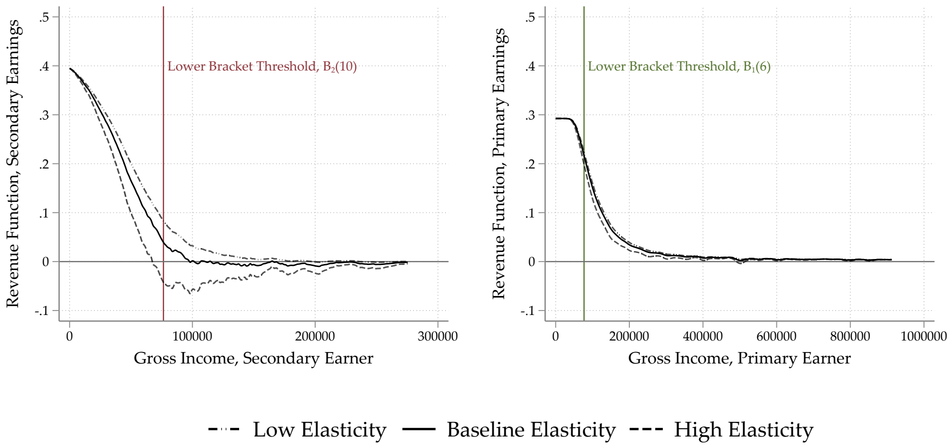
Note: This figure visualizes the estimation of the two revenue effects required for estimating the welfare effect implied by a reduction in jointness for $(y_1, y_2) \in B_1 times B_2$, where B_1 is the 6th-decile of the distribution of primary earnings and B_2 is the sixth decile in the distribution of secondary earnings conditional on $y_1 \in B_1$ (see Figure 6). Figure B33b shows the relevant conditional revenue function describing the effect of increasing marginal tax rates for primary earners above a certain threshold conditional on secondary earnings being above the bracket threshold b_2 of the bracket $B_2(6)$. The solid red (green) lines indicate the lower bracket thresholds of secondary (primary) earnings that are relevant for the welfare evaluation, i.e. b_2 (b_1). Both revenue functions distinguish between different intensive margin elasticity scenarios and include extensive margin behavioral responses.
Source: Own calculations based on CPS-ASEC (2019)

B.4 Tax reforms at the bottom

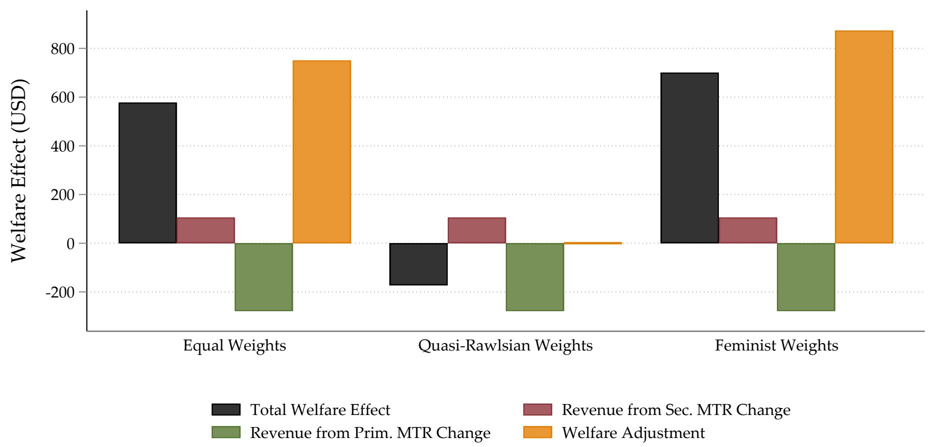
In Section 3.4 of the main text, we estimate welfare effects of tax reforms at the bottom of the income distribution. This section presents graphs on the shape of

Figure B34: Welfare effect, reducing jointness at $B_1(6)$, $B_2(10)$

(a) $\mathcal{R}_2\left(y_2 \mid [\underline{b}_1, \bar{y}_1]\right)$ (b) $\mathcal{R}_1\left(y_1 \mid [\underline{b}_2, \bar{y}_2]\right)$



(c) Welfare effect

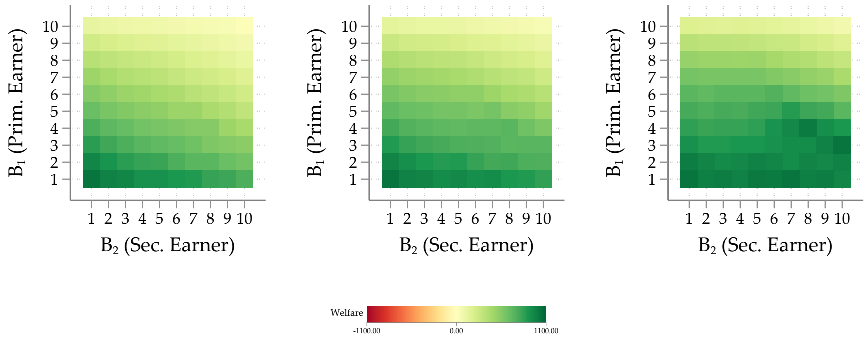


Note: This figure visualizes the welfare effects implied through a change in jointness at the bracket combination $B_1(6)$ and $B_2(10)$. Figure B34a shows the relevant conditional revenue function describing the effect of increasing marginal tax rates for secondary earners above a certain threshold conditional on primary earnings being above the bracket threshold b_1 of the bracket $B_1(6)$. Figure B34b shows the relevant conditional revenue function describing the effect of increasing marginal tax rates for primary earners above a certain threshold conditional on secondary earnings being above the bracket threshold b_2 of the bracket $B_2(10)$. The solid red (green) lines indicate the lower bracket thresholds of secondary (primary) earnings that are relevant for the welfare evaluation, i.e. b_2 (b_1). Both revenue functions distinguish between different intensive margin elasticity scenarios and include extensive margin behavioral responses. Figure B34c displays the aggregate welfare effect and its components based on the bracket lengths $l(B_1) = l(B_2) = 500$ and $\tau = -0.01$, and for the high intensive margin elasticity scenario, and including extensive margin responses. The black bar indicates the aggregate welfare effect. The red (green) bar visualizes the revenue change coming from the decrease of secondary (primary) earnings. The orange bar illustrates the welfare adjustment.
Source: Own calculations based on CPS-ASEC (2019)

Figure B35: Decreasing jointness, varying elasticities

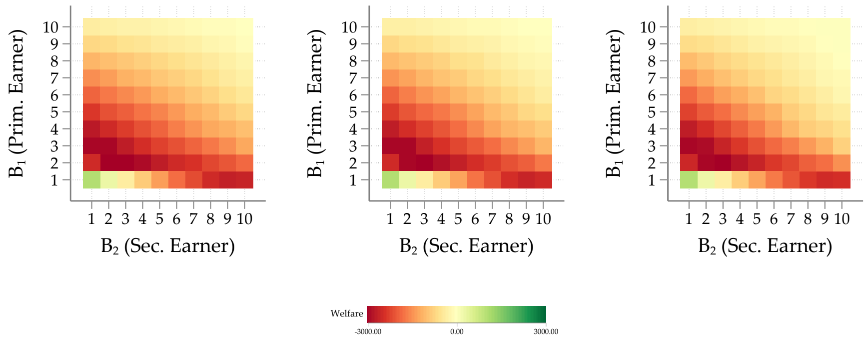
Equal weights

(a) Low elasticity (b) Baseline elasticity (c) High elasticity



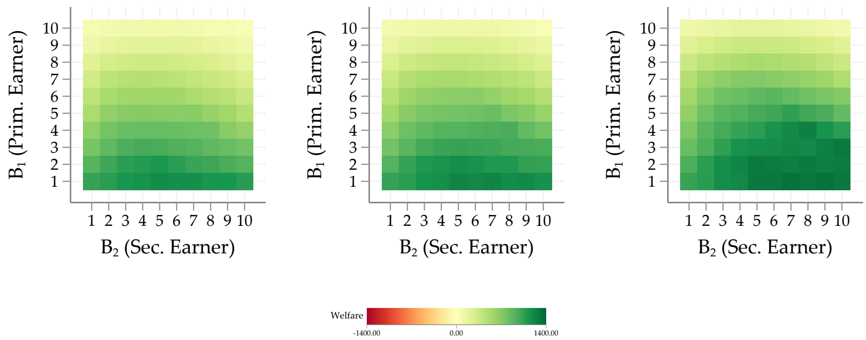
Quasi-Rawlsian weights

(d) Low Elasticity (e) Baseline Elasticity (f) High Elasticity



Feminist weights

(g) Low elasticity (h) Baseline elasticity (i) High elasticity

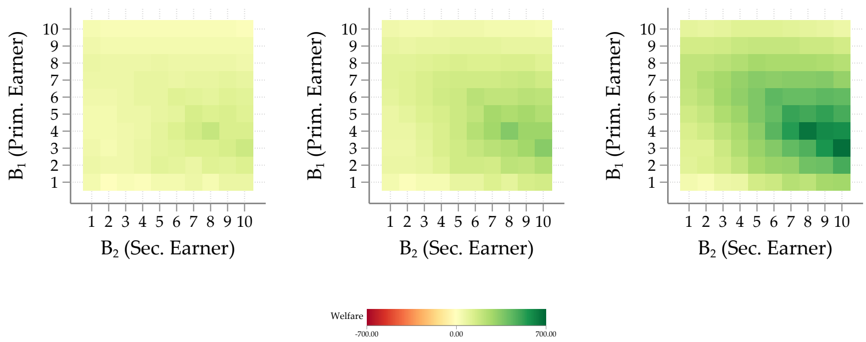


Note: This figure displays the welfare implications of decreasing jointness by decreasing secondary and primary marginal tax rates conditional on particular brackets of primary earnings and secondary earnings. Bracket thresholds are shown in Table B6. The reform applies to a bracket of length $l(B_1) = l(B_2) = 500$ and has the magnitude $\tau = -0.01$. The figure distinguishes between three different forms of welfare weights (see Table B5) and three elasticity scenarios (see Table B2). All results are shown including extensive margin responses. Results without extensive margin responses are shown in Table B36.
Source: Own calculations based on CPS-ASEC (2019)

Figure B36: Decreasing jointness, no extensive margin

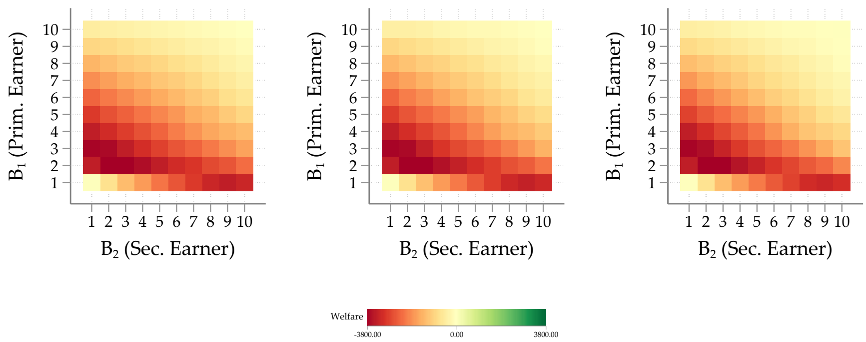
Equal weights

(a) Low elasticity (b) Baseline elasticity (c) High elasticity



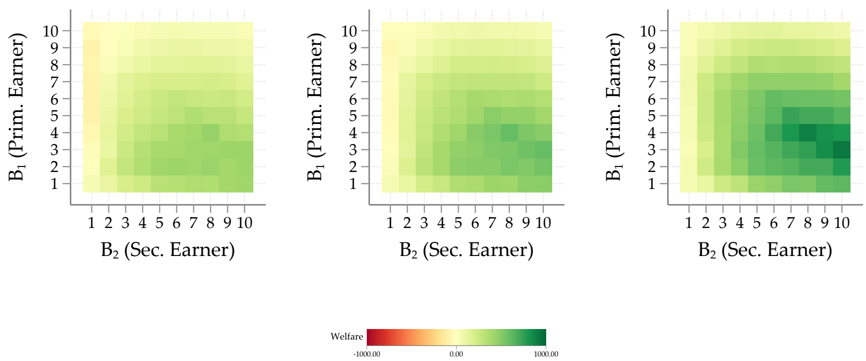
Quasi-Rawlsian weights

(d) Low elasticity (e) Baseline elasticity (f) High elasticity



Feminist weights

(g) Low wlasticity (h) Baseline elasticity (i) High elasticity

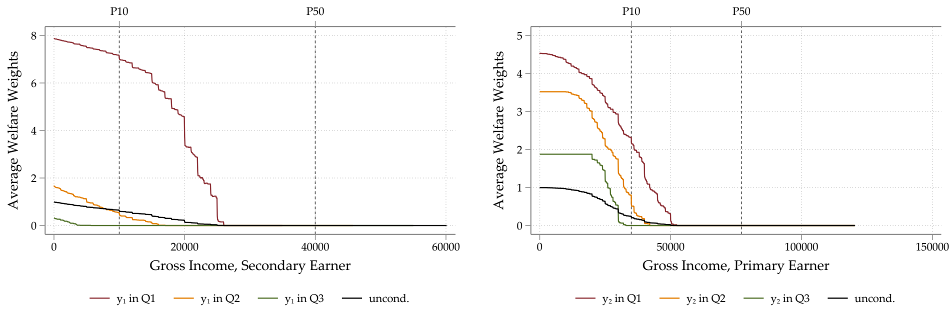


Note: This figure displays the welfare implications of decreasing jointness by decreasing secondary and primary marginal tax rates conditional on particular brackets of primary earnings and secondary earnings. Bracket thresholds are shown in Table B6. The reform applies to a bracket of length $l(B_1) = l(B_2) = 500$ and has the magnitude $\tau = -0.01$. The figure distinguishes between three different forms of welfare weights (see Table B5) and three elasticity scenarios (see Table B2). All results are shown excluding extensive margin responses. Results with extensive margin responses are shown in Figure B35.
Source: Own calculations based on CPS-ASEC (2019)

welfare weights (Figure B37) and robustness tests on the exclusion of extensive margin responses (Figure B38) and the choice of elasticities (Figure B39).

Figure B37: Cond. reforms at bottom (2019), welfare weights

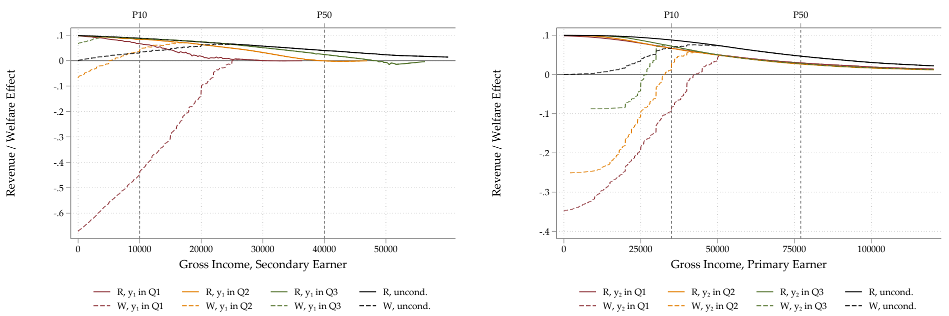
(a) Secondary earnings (b) Primary earnings



Note: This figure shows average welfare weights used for the calculation of welfare effects in Figure 8. Weights for secondary (primary) earners above a specific value conditional on primary (secondary) earnings income deciles at the bottom are shown in Panel a (b). Weights are based on a quasi-Rawlsian welfare function (see Table B5) which concentrates weights on couples below the 10th decile of the joint income distribution. Black lines indicate unconditional welfare weights.
Source: Own calculations based on CPS-ASEC (2019)

Figure B38: Cond. reforms at bottom (2019), no extensive margin

(a) Secondary earnings (b) Primary earnings

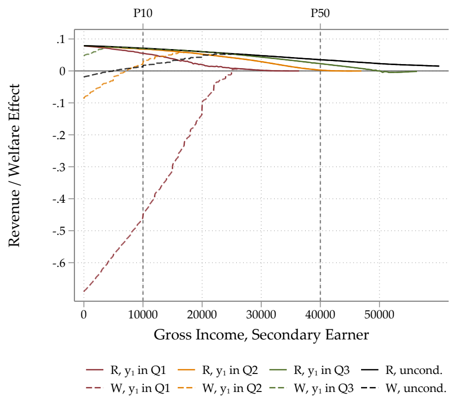


Note: This figure replicates Figure 8 without extensive margin responses. The figure shows the revenue and welfare effects for secondary (primary) earners in married dual earner couples conditional on primary (secondary) earnings income deciles at the bottom of the income distribution in Panel a (b) as of 2019. We assume intensive margin elasticities of 0.25 (0.75) for primary (secondary) earners. The figure displays the revenue effect in terms of conditional revenue functions (solid lines) and the welfare effect based on quasi-Rawlsian welfare weights (dashed lines).
Source: Own calculations based on CPS-ASEC (2019)

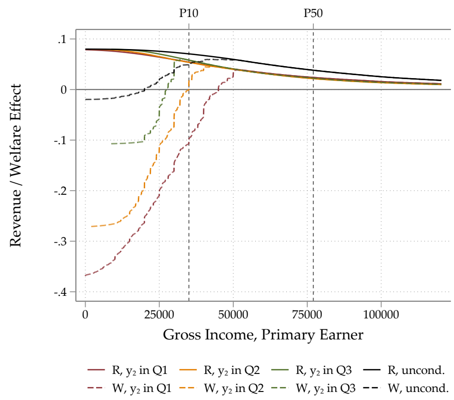
Figure B39: Cond. reforms at bottom (2019), different elasticities

Low elasticity scenario

(a) Secondary earnings

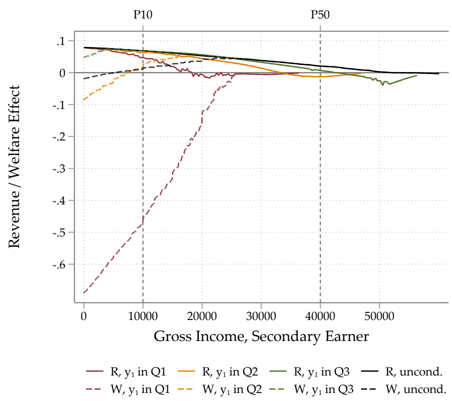


(b) Primary earnings

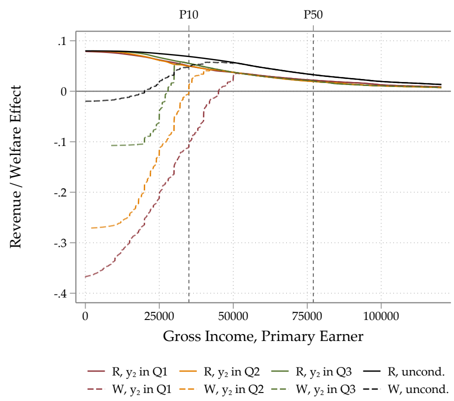


High elasticity scenario

(c) Secondary earnings



(d) Primary earnings



Note: This figure replicates Figure 8 under the low elasticity scenario from Table B2. The figure shows the revenue and welfare effects for secondary (primary) earners in married dual earner couples conditional on primary (secondary) earnings income deciles at the bottom of the income distribution in Panel a (b) as of 2019. All estimates are based on tax units with non-negative gross income in which both spouses are between 25 and 55 years old. Our pre-tax gross income variable of interest contains wage income, farm income, business income, income from dividends, income from interest, income from rent, and retirement income. We take behavioral responses at the intensive and extensive margin into account. In the low elasticity scenario, we assume intensive margin elasticities of 0.15 (0.35) for primary (secondary) earners. In the high elasticity scenario, we assume intensive margin elasticities of 0.5 (1.5) for primary (secondary) earners. The extensive margin scaling factor is held constant at 0.9. The figure displays the revenue effect in terms of conditional revenue functions (solid lines) and the welfare effect based on quasi-Rawlsian welfare weights according to Table B5 (dashed lines). Black lines indicate unconditional revenue and welfare functions.
Source: Own calculations based on CPS-ASEC (2019)

B.5 Optimal top tax rates

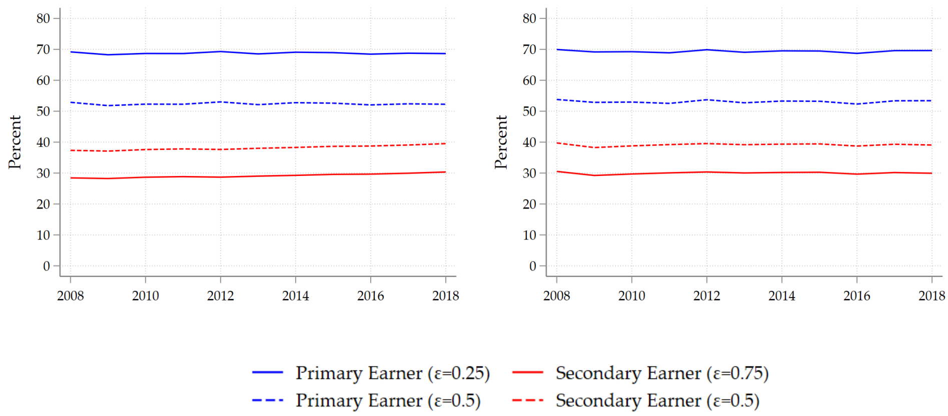
In Section 3.5 of the main text, we use tabular income tax information on all married couples for the estimation of Pareto coefficients and associated optimal tax rates in the most recent available year. In the following, we provide insights on the stability of optimal tax rates over time (Section B.5.1), heterogeneity analyses that estimate Pareto coefficients and optimal tax rates separately for single earner and dual earner couples (Section B.5.2), and robustness tests that estimate Pareto coefficients from CPS data (Section B.5.3).

B.5.1 Optimal tax rates over time

Figure B40 presents optimal tax rates based on Pareto coefficients that have been calculated at the 95th and 99th percentile for all available years of the tabulated income tax return data. Since the distributional properties did not change over this relatively recent time horizon, optimal top tax rates stay rather constant.

Figure B40: Optimal top tax rates over time

(a) Top tax rate (P95 Pareto coef.) (b) Top tax rate (P99 Pareto coef.)



Note: This figure displays optimal top tax rates for different values of the Frisch elasticity based on the Pareto coefficients estimated from tabulated data on wages on W2 forms for joint return taxpayers with wage income.
Source: Own calculations based on SOI Tax Stats - Individual Information Return Form W-2 (2018)

B.5.2 Heterogeneity analyses

Tables B7 (B8) display Pareto coefficients and respective optimal tax rates for single earner couples (dual earner couples with relatively equal income shares). The tables provide two insights. First, the Pareto coefficient for single earner couples (primary earners) is much lower than for dual earner couples (primary earner and secondary earner) pushing towards higher optimal tax rates for single earner couples compared to dual earner couples. Second, within dual earner couples, the optimal tax rate gap between primary and secondary earners almost vanishes, since Pareto coefficients are very similar.

Table B7: Pareto coefficients and optimal tax rates, single earner couples (2018)

	P95		P99	
	Primary	Secondary	Primary	Secondary
Panel A: Pareto coefficients				
Wages on W2 form	1.65	-	1.61	-
Panel B: Optimal tax rates				
Elasticity = .25	71%	-	71%	-
Elasticity = .5	55%	-	55%	-
Elasticity = .75	45%	-	45%	-

Note: This table shows Pareto coefficients and optimal top tax rates. Panel A displays Pareto coefficients for primary and secondary earners based on a generalized Pareto interpolation using tabulated data on wages on W2 forms for joint return single earner couples with wage income. We distinguish between interpolations at the 95th and the 99th percentile. Panel B displays optimal top tax rates associated with these Pareto coefficients. We distinguish between different elasticities.
Source: Own calculations based on SOI Tax Stats - Individual Information Return Form W-2 Statistics (2018)

Table B8: Pareto coefficients and optimal tax rates, dual earner couples (2018)

	P95		P99	
	Primary	Secondary	Primary	Secondary
Panel A: Pareto coefficients				
Wages on W2 form	2.95	3.07	3.16	3.22
Panel B: Optimal tax rates				
Elasticity = .25	58%	57%	56%	55%
Elasticity = .5	40%	39%	39%	38%
Elasticity = .75	31%	30%	30%	29%

Note: This table shows Pareto coefficients and optimal top tax rates. Panel A displays Pareto coefficients for primary and secondary earners based on a generalized Pareto interpolation using tabulated data on wages on W2 forms for joint return dual earner couples with wage income. Dual earner couples have a secondary earner income share of at least 25 percent. We distinguish between interpolations at the 95th and the 99th percentile. Panel B displays optimal top tax rates associated with these Pareto coefficients. We distinguish between different elasticities.
Source: Own calculations based on SOI Tax Stats - Individual Information Return Form W-2 Statistics (2018)

B.5.3 Robustness

In the main text, we estimate Pareto coefficients with tabular income tax return data that provides high quality at the top of the distribution. However, we can also calculate Pareto coefficients directly using CPS data. For this purpose, we estimate the empirical cumulative distribution function (ECDF) and plot the log-transformed ECDF against the log-transformed income level for specific percentiles (see Figure B41). The slope of the regression lines indicate the Pareto coefficient. In line with main results from tabulated tax return data, the Pareto coefficient for primary earners is larger than the one for secondary earners. This holds irrespective of the type of income used in the estimation and the sample restriction (see Table B9). As shown in Table B10, the optimal tax rate gap between primary and secondary earners is even larger than what is suggested by the tabulated data.

Table B9: Pareto coefficients, CPS (2018)

Sample restr.	Income concept	Top 5%		Top 1%	
		Prim. earner	Sec. earner	Prim. earner	Sec. earner
Baseline	Wages	1.63 (N=933)	3.35 (N=953)	1.37 (N=206)	3.27 (N=194)
Baseline	Wage, Business, Farm Income	1.61 (N=942)	3.33 (N=1022)	1.37 (N=194)	3.33 (N=216)
Baseline	Gross Income	1.64 (N=925)	3.28 (N=933)	1.42 (N=187)	3.28 (N=185)
Alternative	Wages	1.72 (N=1700)	3.25 (N=1743)	1.42 (N=378)	3.21 (N=358)
Alternative	Wage, Business, Farm Income	1.69 (N=1789)	3.13 (N=1871)	1.4 (N=350)	3.04 (N=351)
Alternative	Gross Income	1.74 (N=1705)	3.05 (N=1729)	1.47 (N=348)	3 (N=333)

Note: This table shows the Pareto coefficients estimated for different sample restrictions, different income concepts, and at different parts of the income distribution. The baseline sample is restricted to all married couples where both spouses are between 25 and 55 years old. Under an alternative sample restriction, we include all adult married spouses.
Source: Own calculations based on CPS-ASEC (2018)

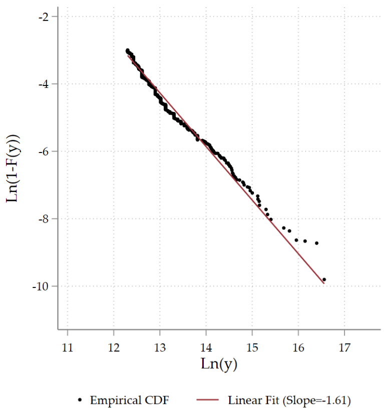
Table B10: Optimal top tax rates, CPS (2018)

Elasticity	P95		P99	
	Prim. earner	Sec. earner	Prim. earner	Sec. earner
.25	71%	55%	74%	55%
.5	55%	38%	59%	38%
.75	45%	29%	49%	29%

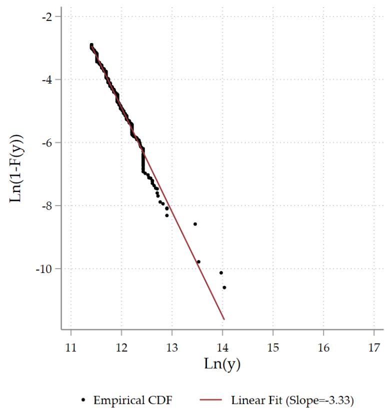
Note: This table shows optimal top tax rates for different values of the Frisch elasticity based on the Pareto coefficients estimated from CPS data under the baseline sample restriction where we focus on all married couples where both spouses are between 25 and 55 years old. Pareto coefficient have been estimated using earnings of primary and secondary earner that include wage, business, and farm income.
Source: Own calculations based on CPS-ASEC (2018)

Figure B41: Pareto coefficient, CPS data

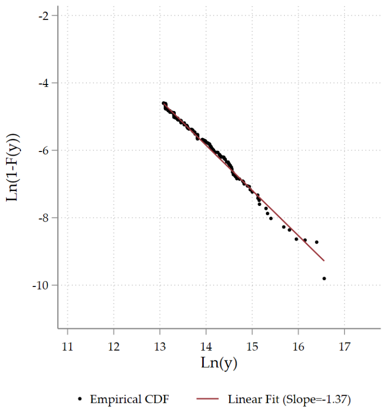
(a) Top 5%, primary



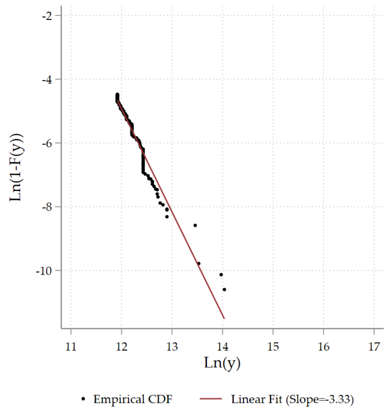
(b) Top 5%, secondary



(c) Top 1%, primary



(d) Top 1%, secondary



Note: This figure displays the top earnings inequality among primary and secondary earners as of 2018. The sample is restricted to all married couples where both spouses are between 25 and 55 years old. The figure shows the shape of the empirical CDF separately for primary and secondary earners in married couples, i.e. $ln(1 - F(y))$ on the vertical and $ln(y)$ on the horizontal axis. The figure also shows the linear fit as a red line, the slope of which is the estimated Pareto coefficient. Earnings of primary and secondary earner include wage, business, and farm income.
Source: Own calculations based on CPS-ASEC (2018)

References

Armour, P., R. V. Burkhauser, and J. Larrimore. 2016. "Using the Pareto distribution to improve estimates of topcoded earnings." *Economic Inquiry* 54 (2): 1263–1273.

Bargain, O., M. Dolls, H. Immervoll, D. Neumann, A. Peichl, N. Pestel, and S. Siegloch. 2015. "Tax policy and income inequality in the United States, 1979–2007." *Economic Inquiry* 53 (2): 1061–1085.

Bargain, O., K. Orsini, and A. Peichl. 2014. "Comparing labor supply elasticities in europe and the united states new results." *Journal of Human Resources* 49 (3): 723–838.

Bierbrauer, F.J., P.C. Boyer, and A. Peichl. 2021. "Politically feasible reforms of nonlinear tax systems." *American Economic Review* 111 (1): 153–91.

Bierbrauer, F.J., P.C. Boyer, A. Peichl, and D. Weishaar. 2023. *The taxation of couples.* CEPR Discussion Paper 18138.

———. 2024. *The political economy of joint taxation.* Mimeo.

Blanchet, T., J. Fournier, and T. Piketty. 2022. "Generalized Pareto Curves: Theory and Applications." *Review of Income and Wealth* 68 (1): 263–288.

Blau, F. D., and L. M. Kahn. 2007. "Changes in the labor supply behavior of married women: 1980–2000." *Journal of Labor Economics* 25 (3): 393–438.

Blundell, R., and T. Macurdy. 1999. "Chapter 27 - Labor Supply: A Review of Alternative Approaches," edited by Orley C. Ashenfelter and David Card, 3:1559–1695. Handbook of Labor Economics. Elsevier.

Eissa, N., and H. W. Hoynes. 2004. "Taxes and the labor market participation of married couples: the Earned Income Tax Credit." *Journal of Public Economics* 88 (9-10): 1931–1958.

Feenberg, D. R., and E. Coutts. 1993. "An Introduction to the TAXSIM Model." *Journal of Policy Analysis and Management* 12 (1): 189–194.

Flood, S., M. King, R. Rodgers, S. Ruggles, J. R. Warren, and M. Westberry. 2021. "Integrated Public Use Microdata Series, Current Population Survey: Version 9.0 [dataset]."

LaLumia, S. 2008. "The effects of joint taxation of married couples on labor supply and non-wage income." *Journal of Public Economics* 92 (7): 1698–1719.

Larrimore, J., R. V. Burkhauser, S. Feng, and L. Zayatz. 2008. "Consistent cell means for topcoded incomes in the public use March CPS (1976–2007)." *Journal of Economic and Social Measurement* 33 (2-3): 89–128.

Neisser, C. 2021. "The elasticity of taxable income: A meta-regression analysis." *The Economic Journal* 131 (640): 3365–3391.

Piketty, T. 2003. "Income Inequality in France, 1901-1998." *Journal of Political Economy* 111 (5): 1004–1042.

Piketty, T., and E. Saez. 2003. "Income Inequality in the United States, 1913-1998." *Quarterly Journal of Economics* 118 (1): 1–41.

Saez, E., J. Slemrod, and S.H. Giertz. 2012. "The elasticity of taxable income with respect to marginal tax rates: A critical review." *Journal of Economic Literature* 50 (1): 3–50.