

Household chores, taxes, and the labor-supply elasticities of women and men

Dorothee Bahn, Christian Bredemeier*, and Falko Juessen

University of Wuppertal

This version: June 2026

Abstract

We study how the division of household chores and gender-specific preferences shape labor-supply elasticities and optimal income taxation. In a model of dual-earner households, labor-supply responses and optimal marginal tax rates depend jointly on gender and on the within-household division of market work and home production. Using PSID data, we find that chore division is a quantitatively important source of gender differences in market-hours elasticities, whereas gender per se plays a smaller role. We then evaluate how closely simple, feasible tax rules approximate the model-implied optimal within-household tax-rate ratios across households. Gender-based taxation has substantial targeting performance, but gender-neutral rules with realistic progressivity perform better.

Keywords: Elasticity of labor supply, optimal taxation, household specialization, gender-based taxation, separate taxation of spouses

JEL classification: H21, J22, D13, J16

*Corresponding author. Contact: University of Wuppertal, Schumpeter School of Business and Economics, Gaußstraße 20, 42119 Wuppertal, Germany. Mail: bredemeier@uni-wuppertal.de.

1 Introduction

Women are often found to supply market labor more wage-elastically than men (e.g., [Keane 2011](#)). [Alesina, Ichino, and Karabarbounis \(2011\)](#) examine the implications of this fact for optimal taxation and argue for lower marginal tax rates on women, a policy they label “gender-based taxation.” They also offer a mechanism for the elasticity gap: because market work and housework are substitutes in preferences, individuals with more household production have a larger marginal disutility of market work and a steeper disutility schedule, tantamount to a higher labor-supply elasticity. Gender differences in labor-supply elasticities can thus be understood as arising from women’s disproportionate share of housework. They may also reflect gender-specific preferences, which would generate elasticity differences beyond those due to intra-household specialization.

For optimal taxation, these two sources of gender differences in elasticities have distinct implications. If elasticity gaps are generated mainly by the division of household chores, tax policy should target household roles rather than gender itself. Gender is then only a proxy, implying targeting errors: men who perform substantial housework would still face higher marginal tax rates under a purely gender-based rule. Other observables, such as relative income, may be more closely related to intra-household specialization and thus serve as more accurate proxies. By contrast, if systematic differences in preferences between genders are the main driver, gender itself becomes a relevant determinant of optimal taxation rather than a proxy for roles that are difficult for the tax authority to observe.

We examine these issues in a model of joint decision-making in dual-earner households that allocates time between paid market work and unpaid housework. The model captures gender differences in labor-supply elasticities that arise from the within-household allocation of chores, as emphasized by [Alesina, Ichino, and Karabarbounis \(2011\)](#).¹ We use the model

¹In contrast to their framework, however, our model also incorporates gender-specific preferences. This extension enables us to disentangle endogenous specialization from exogenous preference heterogeneity as sources of gendered labor-supply responses and optimal taxation. As a second extension to the model, we introduce concave utility from consumption—a feature supported by empirical evidence on labor supply behavior ([Altonji 1986](#); [Domeij and Flodén 2006](#); [Bredemeier, Gravert, and Juessen 2019](#); [Bredemeier,](#)

to derive estimable labor-supply equations and conditions for intra-household tax efficiency. The model plays two conceptually distinct roles. In the positive analysis, it characterizes and estimates heterogeneous labor-supply responses within households. In the normative analysis, these estimates enter as inputs into sufficient statistics for evaluating alternative tax rules conditional on observed intra-household allocation patterns.

We estimate the labor-supply conditions implied by the model using data from the Panel Study of Income Dynamics (PSID). The estimates both test the model’s predictions and provide inputs for the optimal-tax analysis. The model predicts that market-hours elasticities reflect both preferences and within-household specialization, whereas total-hours elasticities are driven by preferences alone. We find that women exhibit substantially more elastic market labor supply than men. By contrast, total labor supply, defined as the sum of paid market work and unpaid housework, is less responsive to wage changes and displays much smaller gender differences. This pattern supports the housework mechanism proposed by [Alesina, Ichino, and Karabarbounis \(2011\)](#), while also suggesting residual gender-related preference heterogeneity beyond the allocation of chores.

We then combine the theoretical and empirical results to assess how closely implementable tax rules can approximate optimal marginal tax rates within households. Both our model and [Alesina, Ichino, and Karabarbounis \(2011\)](#) imply that optimal tax rates depend on the division of household chores, with lower rates assigned to household members who spend more time in home production and thus supply market labor more elastically. Because housework hours are difficult for the government to observe and verify, however, directly conditioning taxes on them is impractical. We therefore quantify how well gender proxies for the relevant tax determinants and compare gender-based taxation with rules based on other observables correlated with chore division, such as relative income and tags like body height. Specifically, we use estimated labor-supply elasticities and observed mar-

Gravert, and Juessen 2023). This extension enables a quantitative analysis of optimal tax rates based on best-practice estimates of labor-supply elasticities.

ket and housework hours to compute model-implied optimal relative marginal tax rates. We then regress these rates on feasible tax determinants to evaluate how closely simple rules approximate the optimal structure. Importantly, any such rule departs from joint taxation, under which spouses face identical marginal tax rates by construction and which serves as a useful benchmark in our analysis.

We find that feasible tax rules can approximate the optimal intra-household marginal tax structure to a considerable degree. Gender-based taxation has substantial targeting performance, capturing about 44% of the cross-sectional variation in optimal relative tax rates. Achieving this level of targeting performance requires substantial gender-specific tax differentials, with men’s marginal tax rates exceeding women’s by about 14 to 21 percentage points. However, gender-based taxation is outperformed by income-based rules under which married spouses are taxed individually rather than jointly, so that relative marginal tax rates depend on relative incomes. At current U.S. levels of tax progressivity, replacing joint taxation with individual taxation while holding the underlying marginal tax schedule fixed would capture more than 50% of the optimal intra-household tax differentiation. The closest approximation requires somewhat greater progressivity; increasing it by about one-third raises targeting performance to about 59%. Tag-based taxation using BMI and body height also has some substantial targeting performance, but it is dominated by both gender-based taxation and progressive separate taxation. Overall, eliminating joint taxation delivers the largest targeting gains, while conditioning on gender and increasing progressivity provide additional improvements. A partial-equilibrium behavioral exercise, in which labor supply responds to the net wages implied by the tax rule, leaves this ranking unchanged.

The remainder of this paper is organized as follows. Section 2 discusses related literature. Section 3 presents the theoretical model and derives the estimation framework as well as optimal relative tax rates within households. Section 4 presents the empirical analysis and quantifies the accuracy of different implementable tax rules in mimicking optimal within-household tax rates. Section 5 concludes.

2 Related literature

The literature on gender as a determinant of income tax rates has two main strands. The first, initiated by [Alesina, Ichino, and Karabarbounis \(2011\)](#), studies gender-based taxation as a tool for improving intra-household efficiency. Rather than focusing on redistribution across households, this literature asks how a given tax burden can be raised more efficiently within a household by allocating it across spouses. Our paper contributes to this line of research.

Several arguments qualify the [Alesina, Ichino, and Karabarbounis \(2011\)](#) channel. [Hundsdoerfer and Matthaei \(2020\)](#) emphasize disadvantages of gender-based taxation that arise from perceived unfairness and potentially adverse labor supply effects. [Meier and Rainer \(2015\)](#) document a second opposing force: uninternalized externalities in non-cooperative couples. When one spouse raises labor earnings, the other benefits, creating a rationale for Pigouvian subsidies. This externality is stronger for the primary earner, implying lower taxes, or larger subsidies, for that spouse and thereby counteracting the elasticity-based argument for lower taxes on secondary earners. We complement these critiques by showing that implementable income-based tax rules can outperform gender-based taxation in terms of intra-household efficiency.

The second strand treats gender as a tag in redistributive tax policy. Because gender is correlated with income, gender-differentiated taxation can redistribute from high- to low-earning individuals. Gender is also difficult to change, limiting avoidance and evasion responses. Quantitative assessments of redistributive gender-based taxation include [Cremer, Gahvari, and Lozachmeur \(2010\)](#) for the U.S., who find gains for low-wage workers; [Berg \(2023\)](#) for Norway; and [Bastani \(2013\)](#) for Sweden. By focusing on within-household efficiency and abstracting from between-household redistribution, our paper is complementary to this strand of research.

Closely related to the intra-household efficiency motive for gender-based taxation is the

literature on joint versus separate taxation. [Boskin and Sheshinski \(1983\)](#) show that equal marginal tax rates under joint taxation are generally inefficient because they violate Ramsey’s inverse-elasticity rule. With home production, [Piggott and Whalley \(1996\)](#) highlight a potential downside of separate taxation: it can distort spouses’ specialization between market work and home production. In response, [Apps and Rees \(1999\)](#) and [Gottfried and Richter \(1999\)](#) show that Boskin and Sheshinski’s argument outweighs this concern in deterministic settings. [Kleven and Kreiner \(2007\)](#) and [Kleven, Kreiner, and Saez \(2009\)](#) formalize this result under more general conditions, emphasizing the labor-supply distortions for secondary earners created by joint taxation.

More recent work provides quantitative assessments. [Guner, Kaygusuz, and Ventura \(2012\)](#) find that moving from joint to separate taxation in the U.S. substantially raises women’s labor supply and welfare. [Holter, Krüger, and Stepancuk \(2023\)](#) identify significant welfare gains from separate taxation because it induces greater accumulation of labor-market experience among married women. Overall, this literature highlights the importance of differentiating marginal tax rates within households, a conclusion that aligns with our results.

Our analysis also connects to the normative public-finance tradition that derives optimal-tax conditions from micro-founded models and combines them with empirical inputs, including elasticity estimates, to evaluate tax systems and construct efficiency-enhancing tax rules. This approach is exemplified by the seminal contribution of [Saez \(2001\)](#). We share its structure by combining model-based optimality conditions with empirical estimates to evaluate tax systems against efficiency benchmarks. Our paper differs in two respects. First, while much of this literature builds on Mirrleesian models with asymmetric information and incentive constraints, our framework follows the Ramsey tradition of optimal tax theory. Second, rather than studying inter-household efficiency and redistribution, we focus on the efficiency of tax design within households.

Finally, the paper is related to the empirical literature on labor-supply elasticities. We

draw on this literature to disentangle the determinants of gender differences in labor-supply elasticities and to obtain inputs for the optimal-tax analysis. Keane (2011) and Elminejad et al. (2023) provide a survey and a meta-analysis, respectively. One central issue is the discrepancy between labor-supply elasticity estimates from micro and macro data; see Keane and Rogerson (2015) for a review. Numerous studies seek to reconcile this discrepancy by identifying and correcting for downward biases in microeconomic estimates, including Blomquist (1985, 1988), Alogoskoufis (1987), Heckman (1993), Rupert, Rogerson, and Wright (2000), Domeij and Flodén (2006), Faberman (2015), Bredemeier, Gravert, and Juessen (2019), and Bredemeier, Gravert, and Juessen (2023). The estimation strategy we use incorporates these insights and applies the corresponding adjustments.

3 Model

This section presents the theoretical model, derives the labor-supply conditions used in the empirical analysis, and characterizes the model’s predictions for optimal relative tax rates between spouses. The model has two complementary uses. The positive component characterizes and estimates heterogeneous labor-supply elasticities within households. The normative component uses the estimated elasticities to evaluate alternative tax rules.

3.1 Model set-up

The model is populated by households, each consisting of a husband and a wife. Household j , with members indexed by i , maximizes the sum of its members’ weighted expected lifetime utility

$$U_j = E_0 \sum_{t=0}^{\infty} \beta^t v_{j,t} = E_0 \sum_{t=0}^{\infty} \beta^t \sum_i \mu_{ij,t} u_{ij,t}, \quad (1)$$

where E is the expectations operator. Per-period utility is

$$u_{ij,t} = v_{g(i),c}(\{c_{k,j,t}\}_{k=1}^K) + v_{g(i),d}(d_{j,t}) - \psi_{ij} \cdot \frac{(n_{ij,t}^{\text{total}})^{1+1/\eta_{g(i)}}}{1 + 1/\eta_{g(i)}}$$

where $g(i) = m, f$ denotes the gender of individual i , $c_{k,j,t}$ is household j 's consumption of variety k of market-produced goods and services, the function $v_{g(i),c}$ is increasing and concave in all arguments, d refers to household consumption of home-produced goods and services, $v'_d > 0$, $v''_d \leq 0$, and ψ_{ij} and η_g are potentially idiosyncratic and gender-specific preference parameters, respectively. The weight $\mu_{ij,t}$ represents individual i 's bargaining power in period t . It may vary across households and over time, allowing the model to capture shifts in intra-household allocation without imposing a specific structural bargaining mechanism.

Importantly, $n_{ij,t}^{\text{total}}$ measures the *total* working time of household member i and is defined as the sum of hours spent on *market* work, $n_{ij,t}^{\text{market}}$, and *housework*, $n_{ij,t}^{\text{home}}$:

$$n_{ij,t}^{\text{total}} = n_{ij,t}^{\text{market}} + n_{ij,t}^{\text{home}}. \quad (2)$$

These preferences nest the original specification in Alesina, Ichino, and Karabarbounis (2011), where both v_c and v_d are assumed to be linear. For empirical applications, it is crucial to account for variation in marginal utility (Altonji 1986, Domeij and Flodén 2006, Bredemeier, Gravert, and Juessen 2019, Bredemeier, Gravert, and Juessen 2023).

Households act subject to a budget constraint

$$\sum_{k=1}^K p_{k,t} c_{k,j,t} + a_{j,t+1} \leq \sum_i w_{ij,t}^{\text{net}} n_{ij,t}^{\text{market}} + (1 + r_t) a_{j,t}, \quad (3)$$

where a is a risk-free asset, r its interest rate, and $w_{ij,t}^{\text{net}} = (1 - \tau_{ij,t})w_{ij,t}$, with $\tau_{ij,t}$ the individual's linear tax rate and $w_{ij,t}$ the gross wage rate, are net (after-tax) wage rates, a home production function

$$d_{j,t} = f(\{n_{ij,t}^{\text{home}}\}_i), \quad (4)$$

which is increasing and concave in all household members' housework hours, and a borrowing constraint

$$a_{j,t+1} \geq a_t^{\min}, \quad (5)$$

where $a_{\min} \leq 0$ denotes (the negative of) a potentially age-dependent cap on borrowing.

Wage rates are exogenous and evolve stochastically according to a process described by the probability density function $\omega(\{w_{ij,t+1}^{\text{net}}\}_i | \{w_{ij,t}^{\text{net}}\}_i)$.

Discussion of modeling choices. Before turning to the model’s implications, three modeling choices merit discussion. First, we do not impose specific functional forms on either the utility derived from home-produced goods, v , or the home production function f . This choice allows the analysis to remain agnostic about the structural determinants of household specialization, including potential gender differences in preferences or productivity, while still delivering clear implications for labor-supply elasticities and optimal tax rates. The analysis uses observed time allocations as a sufficient statistic for the tax-design problem, rather than analyzing the underlying determinants of these allocations. The results should therefore be read as characterizing how observed time allocation maps into labor-supply elasticities and optimal tax rates, not as causal statements about why those allocation differences arise. In a robustness exercise, we retain this agnostic stance on the underlying determinants of specialization, but allow time allocations to adjust in response to tax-induced changes in net wages as predicted by estimated labor-supply elasticities consistent with the model.

Second, following [Alesina, Ichino, and Karabarbounis \(2011\)](#), market work and housework enter disutility as perfect substitutes and aggregate into total labor, n^{total} . This standard assumption yields a transparent mapping between home production and market labor-supply elasticities. The key mechanism does not require perfect substitutability itself, but rather the weaker property that the marginal disutility of market work increases with the amount of housework performed. Intuitively, an additional hour of market work is more costly when the individual already devotes substantial time to home production. In addition, greater overall workload makes fatigue accumulate more rapidly, causing marginal disutility to increase more strongly with additional work. More generally, the same qualitative mechanism arises with smooth aggregators of market and home work, such as CES specifications, provided the elasticity of substitution between the two is sufficiently high.

In that case, deviations from perfect substitutability have only second-order effects on the relationship between observed time allocation and Frisch elasticities. In regressions with total labor on the left-hand side, we therefore interpret the results as relying on a high-substitutability approximation rather than on the knife-edge case of perfect substitutes.

Third, the model assumes linear taxation of individual labor income. This assumption is used for expositional clarity and tractability because it maps marginal tax rates directly into net wage rates and thus into behavioral incentives. The analysis of labor-supply responses depends only on marginal tax rates, so nonlinear tax systems can be interpreted through their local marginal wedges. The linear-tax formulation should therefore be viewed as a local approximation to more general nonlinear schedules rather than as a restriction on the empirical or normative analysis. In Section 4.3, relative marginal tax rates vary with income and earnings, making this interpretation directly relevant for the evaluation of alternative tax rules.

3.2 Behavior

First-order conditions. With $\lambda_{j,t}$, $\lambda_{j,t}^d$, and $\xi_{j,t}$ denoting the Lagrange multipliers on the budget, housework, and borrowing constraints, (3), (4), and (5), the first-order conditions for the household problem are

$$\partial v_{j,t} / \partial c_{k,j,t} = \lambda_{j,t} p_{k,t}, \quad (6)$$

$$\partial v_{j,t} / \partial d_{j,t} = \lambda_{j,t}^d, \quad (7)$$

$$-\partial v_{j,t} / \partial n_{ij,t}^{\text{market}} = \lambda_{j,t} w_{ij,t}^{\text{net}} \quad \forall i, \quad (8)$$

$$-\partial v_{j,t} / \partial n_{ij,t}^{\text{home}} = \lambda_{j,t}^d \partial f / \partial n_{ij,t}^{\text{home}} \quad \forall i, \quad (9)$$

and

$$\lambda_{j,t} - \xi_{j,t} = \beta E_t \left[\lambda_{j,t+1} (1 + r_{t+1}) + \sum_i \frac{\partial \mu_{ij,t+1}}{\partial a_{j,t+1}} u_{ij,t+1} \right].$$

Relative to the standard Euler equation, the final condition includes the multiplier on the borrowing constraint (5) and a term allowing future intra-household bargaining positions to vary with accumulated state variables.

Frisch elasticities and labor-supply regressions. The first-order condition for market labor supply $n_{ij,t}^{\text{market}}$, (8), is the starting point for deriving Frisch elasticities and estimable labor-supply conditions. Because market work and housework are perfect substitutes in disutility, the marginal utility cost of market work depends only on total work. Applying the functional form of labor disutility to (8) yields $\mu_{ij,t}\psi_{ij} \cdot (n_{ij,t}^{\text{total}})^{1/\eta_{g(i)}} = \lambda_{j,t}w_{ij,t}^{\text{net}}$, which in logs reads

$$\log n_{ij,t}^{\text{total}} = \eta_{g(i)} \log w_{ij,t}^{\text{net}} + \eta_{g(i)} \log \lambda_{j,t} - \eta_{g(i)} \log \mu_{ij,t} - \eta_{g(i)} \log \psi_{ij}. \quad (10)$$

This condition shows that the *Frisch elasticity of total work* n_i^{total} is simply given by the preference parameter $\eta_{g(i)}$.

To derive the elasticity of *market labor supply*, we take logs of the definition (2) and derive the differential $\partial \log n_{ij,t}^{\text{total}} / \partial \log w_{ij,t} = 1/n_{ij,t}^{\text{total}} \cdot (\partial \log n_{ij,t}^{\text{market}} / \partial \log w_{ij,t} \cdot n_{ij,t}^{\text{market}} + \partial \log n_{ij,t}^{\text{home}} / \partial \log w_{ij,t} \cdot n_{ij,t}^{\text{home}})$. Solving for $\partial \log n_{ij,t}^{\text{market}} / \partial \log w_{ij,t}$ gives the Frisch elasticity of market work n_i^{market} as a function of the Frisch elasticity of total hours and the wage responsiveness of home production. Because our focus is on the consequences rather than the origins of household chore division, we treat the latter response as quantitatively small, consistent with the empirical findings in [Bredemeier, Gravert, and Juessen \(2023\)](#). Specifically,

$$e_{ij,t}^{\text{Frisch}} = \frac{\partial \log n_{ij,t}^{\text{market}}}{\partial \log w_{ij,t}^{\text{net}}} \big|_{\lambda,\mu} = \frac{\eta_{g(i)}}{s_{ij,t}^{\text{market}}} - \frac{1 - s_{ij,t}^{\text{market}}}{s_{ij,t}^{\text{market}}} \cdot \frac{\partial \log n_{ij,t}^{\text{home}}}{\partial \log w_{ij,t}^{\text{net}}} \big|_{\lambda,\mu} \approx \frac{\eta_{g(i)}}{s_{ij,t}^{\text{market}}}, \quad (11)$$

where $s_{ij,t}^{\text{market}} = n_{ij,t}^{\text{market}} / n_{ij,t}^{\text{total}}$ is the share of total hours devoted to market work. This approximation should be interpreted as a sufficient-statistic mapping from total-hours elasticities to market-hours elasticities rather than as an exact structural identity.

Equation (11) shows that gender differences in the Frisch elasticity of market work can arise through two channels. First, they can reflect differences in the allocation of total working time between market work and housework. Because women have lower values of $s_{ij,t}^{\text{market}}$ on average, this channel tends to raise women's market-hours elasticities. Second, gender differences can reflect heterogeneity in the preference parameter η_g .

To derive a linear condition for log market work to be used in estimation, we approximate the definition (2) around a gender-specific point of approximation described by $n_{g(i)}^{\text{total}}$, $n_{g(i)}^{\text{market}}$, and $n_{g(i)}^{\text{home}}$ to obtain $\log n_{ij,t}^{\text{total}} \approx s_{g(i)}^{\text{market}} \cdot \log n_{ij,t}^{\text{market}} + (1 - s_{g(i)}^{\text{market}}) \cdot \log n_{ij,t}^{\text{home}}$, where $s_{g(i)}^{\text{market}} = n_{g(i)}^{\text{market}} / n_{g(i)}^{\text{total}}$. Using the approximation in (10) gives

$$\log n_{ij,t}^{\text{market}} = \frac{\eta_{g(i)}}{s_{g(i)}^{\text{market}}} (\log w_{ij,t}^{\text{net}} + \log \lambda_{j,t} - \log \mu_{ij,t} - \log \psi_{ij}) - \frac{1 - s_{g(i)}^{\text{market}}}{s_{g(i)}^{\text{market}}} \log n_{ij,t}^{\text{home}}. \quad (12)$$

Empirical applications must address two unobserved terms in the labor-supply conditions (10) and (12): marginal utility $\log \lambda_{j,t}$ and the bargaining weight $\log \mu_{ij,t}$. We follow Bredemeier, Gravert, and Juessen (2023), who show that these variables can be expressed as log-linear functions of household total consumption expenditures, $\tilde{c}_{j,t} = \sum_k p_{k,t} c_{k,j,t}$, and the expenditure share of a specific consumption category k , $\tilde{c}_{k,j,t} / \tilde{c}_{j,t}$, where $\tilde{c}_{k,j,t} = p_{k,t} c_{k,j,t}$, over which men and women have systematically different preferences. The labor-supply conditions (10) and (12) can then be rewritten as the following regression equations:

$$\begin{aligned} \log n_{ij,t}^{\text{market}} = & \kappa_{ij}^{\text{market}} + \delta_t^{\text{market}} + \frac{\eta_{g(i)}}{s_{g(i)}^{\text{market}}} \log w_{ij,t}^{\text{net}} - \frac{1 - s_{g(i)}^{\text{market}}}{s_{g(i)}^{\text{market}}} \log n_{ij,t}^{\text{home}} \\ & + \alpha_{g(i)}^{\text{market}} \log \tilde{c}_{j,t} + \gamma_{g(i)}^{\text{market}} \log \left(\frac{\tilde{c}_{k,j,t}}{\tilde{c}_{j,t}} \right) + \varepsilon_{ij,t}^{\text{market}} \end{aligned} \quad (13)$$

and

$$\log n_{ij,t}^{\text{total}} = \kappa_{ij}^{\text{total}} + \delta_t^{\text{total}} + \eta_{g(i)} \log w_{ij,t}^{\text{net}} + \alpha_{g(i)}^{\text{total}} \log \tilde{c}_{j,t} + \gamma_{g(i)}^{\text{total}} \log \left(\frac{\tilde{c}_{k,j,t}}{\tilde{c}_{j,t}} \right) + \varepsilon_{ij,t}^{\text{total}}, \quad (14)$$

where the individual fixed effects $\kappa_{ij}^{\text{market}}$ and $\kappa_{ij}^{\text{total}}$ collect $-\eta_{g(i)} \log \psi_{ij}$ and mean approximation and measurement errors, the time fixed effects δ^m and δ^{total} capture the effects of time

variation in goods prices, ε^m and $\varepsilon^{\text{total}}$ are residuals reflecting variation in approximation and measurement errors, and the parameters α^m , γ^m , α^{total} , and γ^{total} combine the proxy relations discussed above with the slope coefficients from (10) and (12). In our empirical analysis, we estimate (13) and (14) separately for men and women.

Because s^{market} is, on average, smaller for women than for men, the model implies two empirical predictions unless a strong countervailing difference in preference parameters, i.e., $\eta_m \gg \eta_f$, offsets this force. First, the coefficient on the wage rate in the market-hours regression (13) should be larger for women. Second, the wage coefficients in the total-hours regression (14) should be more similar between men and women than the corresponding coefficients in the market-hours regression (13). The total-hours regression is particularly important because it identifies the preference parameters η_g , which enter the empirical analysis of optimal tax rates.

3.3 Optimal taxation

We now turn to the model's normative implications for income tax rates. The goal is to characterize how the government should tax the members of a household relative to one another and to express efficient intra-household taxation in terms of objects that can be observed or estimated empirically.

We focus on the structural question of how spouses should be taxed conditional on a given intra-household allocation, abstracting from occasionally binding borrowing constraints and endogenous changes in bargaining positions. Put differently, we isolate intratemporal substitution between market consumption, home-produced consumption, and leisure, while abstracting from intertemporal adjustment and tax-induced changes in household bargaining. This implies fixed Pareto weights μ and stable household preferences in the normative analysis. Because the positive framework intentionally remains agnostic about the structural determinants of bargaining weights, endogenizing their response to tax reform would require imposing a specific bargaining model. Our approach therefore characterizes efficient

tax wedges conditional on a given intra-household allocation and does not capture welfare effects operating through changes in bargaining power. Since the reforms we consider affect married-couple taxation rather than spouses' outside options directly, we view this omitted channel as plausibly attenuated, though not absent.

Together with homotheticity, these simplifications allow us to concentrate on a consumption bundle $c_{j,t}$ that provides one unit of household utility v at minimal cost and has a price index p_t satisfying $p_t c_{j,t} = \sum_{k=1}^K p_{k,t} c_{k,j,t}$. We normalize $p_{k,t}$ to one. With these simplifications, we can combine the first-order conditions (6)–(9) to obtain:²

$$\frac{\partial U_j}{\partial n_{ij,t}^{\text{market}}} = -\frac{\partial U_j}{\partial c_{j,t}} w_{ij,t}^{\text{net}} \quad \forall i, \quad (15)$$

$$\frac{\partial U_j}{\partial d_{j,t}} = \frac{\partial U_j}{\partial c_{j,t}} \frac{w_{ij,t}^{\text{net}}}{\partial f / \partial n_{ij,t}^{\text{home}}} \quad \forall i. \quad (16)$$

We express these relations using upper-tier U as defined in (1), which is convenient for the subsequent tax analysis.

We abstract from redistributive aspects of taxation and take as given the amount of taxes $T_{j,t}$ that the government aims to collect from household j through labor income taxation. We then examine how to raise this amount in the most efficient way.

When deciding on tax rates for the two members of household j at a given period of time, the government maximizes household utility U_j as defined in (1) subject to

$$\sum_i \tau_{ij,t} w_{ij,t} n_{ij,t}^{\text{market}} = T_{j,t}, \quad (17)$$

where w denotes gross (before-government) wage rates and τ denotes tax rates to be set optimally.

²Combining (6) with (8) and using $\partial v_{j,t} / \partial x = \beta^t \partial U_j / \partial x$ for any choice variable x gives (15). Combining (7) with (9) and using $\partial v_{j,t} / \partial x = \beta^t \partial U_j / \partial x$ as well as $\partial U_j / \partial n_{ij,t}^{\text{market}} = \partial U_j / \partial n_{ij,t}^{\text{home}}$ gives (16).

The first-order condition for $\tau_{ij,t}$ is

$$\begin{aligned} & \frac{\partial U_j}{\partial c_{j,t}} \cdot \frac{\partial c_{j,t}}{\partial \tau_{ij,t}} + \frac{\partial U_j}{\partial d_{j,t}} \cdot \frac{\partial d_{j,t}}{\partial \tau_{ij,t}} + \frac{\partial U_j}{\partial n_{ij,t}^{\text{home}}} \cdot \frac{\partial n_{ij,t}^{\text{home}}}{\partial \tau_{ij,t}} + \frac{\partial U_j}{\partial n_{-ij,t}^{\text{home}}} \cdot \frac{\partial n_{-ij,t}^{\text{home}}}{\partial \tau_{ij,t}} \\ & + \frac{\partial U_j}{\partial n_{ij,t}^{\text{market}}} \cdot \frac{\partial n_{ij,t}^{\text{market}}}{\partial \tau_{ij,t}} + \frac{\partial U_j}{\partial n_{-ij,t}^{\text{market}}} \cdot \frac{\partial n_{-ij,t}^{\text{market}}}{\partial \tau_{ij,t}} \\ & + \lambda_{j,t}^G \cdot \left(w_{ij,t} n_{ij,t}^{\text{market}} + \tau_{ij,t} w_{ij,t} \frac{\partial n_{ij,t}^{\text{market}}}{\partial \tau_{ij,t}} + \tau_{-ij,t} w_{-ij,t} \frac{\partial n_{-ij,t}^{\text{market}}}{\partial \tau_{ij,t}} \right) = 0, \end{aligned} \quad (18)$$

where $\lambda_{j,t}^G$ is the Lagrange multiplier on (17) and $-i$ indicates the partner of individual i . This condition simplifies considerably when first-order conditions and constraints of the household problem are substituted. Specifically, using the optimality conditions (15) and (16) as well as the constraints (3) with $a_{j,t+1} = (1+r)a_{j,t}$ holding in a steady state and (4) gives, after collecting terms,

$$-\frac{\partial U_j}{\partial c_{j,t}} + \lambda_{j,t}^G \cdot \left(1 + \frac{\partial n_{ij,t}^{\text{market}}}{\partial \tau_{ij,t}} \frac{\tau_{ij,t}}{n_{ij,t}^{\text{market}}} + \frac{\partial n_{-ij,t}^{\text{market}}}{\partial \tau_{ij,t}} \frac{\tau_{-ij,t} w_{-ij,t}}{w_{ij,t} n_{ij,t}^{\text{market}}} \right) = 0, \quad (19)$$

see Appendix A.1 for a detailed derivation.

For small perturbations around the observed tax system, changes induced by tax rates can be expressed equivalently as changes induced by gross wages because both operate through the net wage. Formally, since $w_{ij,t}^{\text{net}} = (1 - \tau_{ij,t})w_{ij,t}$, a first-order log-linear approximation implies that $\partial x / \partial \tau_{ij,t} \approx -(\partial x / \partial w_{ij,t})w_{ij,t}$ for any variable x that depends on wages and tax rates only through the net wage. Locally, an increase in the tax rate and a proportional decrease in the gross wage therefore induce the same change in decision-relevant net returns to work. Defining

$$\begin{aligned} e_{ij,t}^{\text{own}} &= \frac{\partial n_{ij,t}^{\text{market}}}{\partial w_{ij,t}} \cdot \frac{w_{ij,t}}{n_{ij,t}^{\text{market}}}, \\ e_{ij,t}^{\text{cross}} &= \frac{\partial n_{ij,t}^{\text{market}}}{\partial w_{-ij,t}} \cdot \frac{w_{ij,t}}{n_{-ij,t}^{\text{market}}}, \end{aligned}$$

we can simplify (18) to

$$-\frac{\partial U_j}{\partial c_{j,t}} + \lambda_{j,t}^G \cdot (1 - \tau_{ij,t} e_{ij,t}^{\text{own}} - \tau_{-ij,t} e_{-ij,t}^{\text{cross}}) = 0, \quad (20)$$

which needs to hold for both i and $-i$ and hence describes a system of two equations in two unknowns, $\tau_{ij,t}$ and $\tau_{-ij,t}$. This system can be solved for the ratio of optimal tax rates within the household,

$$\frac{\tau_{ij,t}^*}{\tau_{-ij,t}^*} = \frac{e_{-ij,t}^{\text{own}} - e_{-ij,t}^{\text{cross}}}{e_{ij,t}^{\text{own}} - e_{ij,t}^{\text{cross}}}. \quad (21)$$

This condition is an application of Ramsey's inverse-elasticity rule: tax rates should be higher where behavioral responses are smaller. In this setting, the relevant response includes not only one's own reaction to a tax change but also the effect on the partner's labor supply, which usually has the opposite sign.

Note that $e_{ij,t}^{\text{cross}}$ is not exactly the cross-wage elasticity in the usual sense but multiplies the cross-wage derivative by the ratio of one's own wage to the partner's hours. The term thus measures by how much one's hours change (as a percentage of one's partner's hours) in response to a change in the partner's wage rate (as a percentage of one's own wage). This is important because it implies that income effects embedded in the two derivatives cancel out when they are subtracted from each other.

In Appendix A.2, we show that, in our setting with additively separable preferences, the log ratio of optimal marginal tax rates in a household satisfies

$$\theta_{ij,t}^* \equiv \log(\tau_{ij,t}^*/\tau_{-ij,t}^*) \approx (1 - 2\gamma) (\log e_{-ij,t}^{\text{Frisch}} - \log e_{ij,t}^{\text{Frisch}}). \quad (22)$$

Because wealth effects cancel out, optimal relative tax rates depend only on substitution effects, captured by Frisch elasticities. The slope $1 - 2\gamma$ reflects the household's marginal propensity to earn out of unearned income, 2γ . Since γ is negative and empirical estimates are typically modest, the slope exceeds one but is unlikely to exceed 1.5. We review the empirical literature on the marginal propensity to earn out of unearned income in Appendix B.

Independent of the specific value of 2γ , the optimal tax rate ratio depends on the difference between spouses' Frisch elasticities of market work. The government should tax

more heavily the spouse whose labor supply is less elastic. In our model, such differences in elasticities can arise from gender-specific preferences and from differences in time allocation, with individuals who spend more time on housework exhibiting higher Frisch elasticities. Substituting the Frisch elasticities of market work from (11), optimal relative marginal tax rates within a household are given by:

$$\theta_{ij,t}^* = \log(\tau_{ij,t}^*/\tau_{-ij,t}^*) = (1 - 2\gamma) \left(-\log \eta_{g(i)} + \log \eta_{g(-i)} + \log s_{ij,t}^{\text{market}} - \log s_{-ij,t}^{\text{market}} \right). \quad (23)$$

The model therefore allows for both gender-based taxation and taxation based on the division of household chores. Which approach performs better, and how tax codes can be designed to mimic optimal intra-household marginal tax rates, is an empirical question that we address in the next section.

Efficiency criterion. As a final step before the empirical analysis, we develop a normative criterion for evaluating existing and counterfactual tax rules. Let Θ^z denote a system that generates relative intra-household marginal tax rates $\{\theta_{ij,t}^z\}$. The efficiency of Θ^z can be assessed by its deviations from the optimal relative tax rates $\{\theta_{ij,t}^*\}$. A standard approach is to approximate the government's value function with a second-order Taylor expansion around the optimum. We define the government welfare function, aggregated over all households, as $V_t^G = \sum_j \omega_j \{U_{j,t} + \lambda_{j,t}^G (\sum_i \tau_{ij,t} w_{ij,t} n_{ij,t}^{\text{market}} - G_{j,t})\}$, where ω_j are welfare weights. Expanding this function around the optimal set $\{\theta_{ij,t}^*\}$ gives

$$V_t^G(\Theta^z) \approx V_t^G(\Theta^*) + \sum_j \omega_j \left\{ \frac{\partial V_t^G}{\partial \theta_{ij,t}} \Big|_{\theta_{ij,t}^*} \cdot (\theta_{ij,t}^z - \theta_{ij,t}^*) - \frac{1}{2} \cdot \frac{\partial^2 V_t^G}{\partial \theta_{ij,t}^2} \Big|_{\theta_{ij,t}^*} \cdot (\theta_{ij,t}^z - \theta_{ij,t}^*)^2 \right\}. \quad (24)$$

The optimality of $\theta_{ij,t}^*$ implies, by definition, that $\frac{\partial V_t^G}{\partial \theta_{ij,t}} \Big|_{\theta_{ij,t}^*} = 0$ and $\frac{\partial^2 V_t^G}{\partial \theta_{ij,t}^2} \Big|_{\theta_{ij,t}^*} < 0$. Hence, we can define a loss function

$$\mathcal{L}(\Theta^z) \equiv V_t^G(\Theta^*) - V_t^G(\Theta^z) = \frac{1}{2} \cdot \sum_j \omega_j \left| \frac{\partial^2 V_t^G}{\partial \theta_{ij,t}^2} \Big|_{\theta_{ij,t}^*} \right| \cdot (\theta_{ij,t}^z - \theta_{ij,t}^*)^2, \quad (25)$$

which is a weighted sum of squared deviations between actual and optimal relative marginal tax rates in a household. Weights are determined by social welfare weights ω_j and the local curvature of welfare, i.e., $\frac{\partial^2 V_t^G}{\partial \theta_{ij,t}^2} | \theta_{ij,t}^*$, which captures how sensitive household j 's well-being is to policy deviations.

To maintain our focus on intra-household efficiency and abstract from redistributive concerns, we assume that for each household j , the social welfare weight ω_j equals the absolute inverse of the local welfare curvature $\frac{\partial^2 V_t^G}{\partial \theta_{ij,t}^2} | \theta_{ij,t}^*$. This technical assumption implies a local, second-order form of redistributive neutrality: the planner treats all deviations from the optimal relative tax rate as equally costly, regardless of which household is affected.³ Formally, the loss function simplifies to

$$\mathcal{L}(\Theta^z) = \frac{1}{2} \sum_j (\theta_{ij,t}^z - \theta_{ij,t}^*)^2. \quad (26)$$

In the empirical analysis, we construct a distribution of optimal relative intra-household marginal tax rates $\theta_{ij,t}^*$ from gender-specific elasticity estimates and observable information on $s_{ij,t}^{\text{market}}$. We also construct a distribution of relative tax rates $\theta_{ij,t}^z$ implied by a linear tax rule Θ^z , using fitted values from a regression of optimal $\theta_{ij,t}^*$ on the rule's tax determinants. The R^2 of this regression, given by

$$1 - \frac{\sum_j (\theta_{ij,t}^z - \theta_{ij,t}^*)^2}{\sum_j (\theta_{ij,t}^*)^2}$$

is a linear function of the loss $\mathcal{L}(\Theta^z)$ associated with the tax rule. The R^2 has a direct interpretation. It equals one for the optimal tax rule Θ^* itself and zero for rules with no intra-household differences in marginal tax rates, i.e., $\theta_{ij,t}^z = 0 \ \forall ij, t$. This zero case corresponds to the status quo of joint taxation of married couples, under which spouses face identical marginal tax rates by construction. Between these two extremes, the R^2 is linear in the loss $\mathcal{L}(\Theta^z)$ under the assumption that household allocation patterns remain fixed. It therefore measures how closely a tax rule approximates the cross-sectional variation in

³In the empirical analysis, we also consider alternative welfare weights to assess how they affect results.

optimal relative tax rates and can be interpreted as a diagnostic of targeting performance with respect to intra-household efficiency.

4 Empirical Analysis

This section quantifies how closely different tax rules, including gender-based taxation, approximate the model-implied optimal relative marginal tax rates within couples. We first estimate regressions for market and total labor supply. These regressions serve two purposes: they provide reduced-form evidence on the relative importance of household specialization versus preference heterogeneity in explaining gender differences in market labor supply elasticities, and they deliver estimates of the preference parameter η_g , which is identified as the Frisch elasticity of total hours worked. We then combine these estimates with observed housework and market hours to predict optimal relative marginal tax rates within couples. Finally, we regress these predicted optimal relative tax rates on observable characteristics that could plausibly enter tax rules, such as gender or income. These regressions show how strongly marginal tax rates should depend on each observable and the targeting accuracy achieved by conditioning on them.

4.1 Variable definitions

Data and sample selection. We use biennial waves 1999–2023 of the Panel Study of Income Dynamics (PSID), which contain information on earnings, time use, and other relevant variables for calendar years 1998–2022. Because the sample includes the Covid period, robustness checks below account for potentially unusual conditions during this period. Our sample selection closely follows [Bredemeier, Gravert, and Juessen \(2023\)](#).

We consider household-year observations in which the household contains a married heterosexual couple whose spouses are both between 25 and 60 years old and both report positive market labor supply. We exclude the Survey of Economic Opportunity and immigrant samples, drop observations with wages below half the minimum wage, and exclude

observations with extreme asset values or transfer income, as well as those with extreme inter-wave jumps (see [Blundell, Pistaferri, and Saporta-Eksten \(2016\)](#) for details). Each retained household-year observation yields two individual-year observations, one for each spouse, with partner and household variables attached. If an individual i is observed in the PSID with different spouses over time, we treat each marriage as a separate household j . PSID sampling weights are used throughout.

Market hours, housework, and wages. Market hours are measured by annual hours worked, covering all jobs and including overtime. The PSID housework variable covers cooking, cleaning, and other work around the house.⁴ We treat missing housework values as zeros and add one before taking logs when including housework in the regressions.

Hourly wages are annual earnings divided by annual market hours. Below, we explain how we address the division bias created by this procedure. Annual earnings are measured in real year-2000 dollars and include labor earnings and the labor component of business income. As emphasized by [Blomquist \(1985, 1988\)](#), using gross wages in labor-supply regressions biases estimated elasticities. To construct decision-relevant net wages, we compute household net income after federal income taxes, the EITC, and SNAP using year-specific program rules and household characteristics, i.e., family size and the number of children. We define the local net-of-tax rate as the change in net income, after taxes and transfers, generated by a \$500 increase in gross annual earnings, divided by \$500. Equivalently, this is one minus the local marginal tax rate. Net wages are then obtained by multiplying gross wages by this local net-of-tax rate.

Additional variables. The approach for estimating labor-supply elasticities developed by [Bredemeier, Gravert, and Juessen \(2023\)](#) uses expenditure variables to characterize the composition of household consumption, which in turn can be used to account for relative bargaining power within the household in regressions. We follow their preferred specification

⁴Shopping and caring for children or adult family members needing assistance are addressed in separate questions, but only since 2017.

and include the food expenditure share together with total household consumption in the labor-supply regressions. Total household consumption is defined as the sum of expenditures on individual consumption items.

The labor-supply regressions also include individual and time fixed effects. Individual fixed effects capture heterogeneity in tastes for work and other time-invariant unobserved characteristics. Time effects absorb price variation and macroeconomic fluctuations, including economy-wide factors that shift labor demand. To control for time-varying taste shifters, we include a third-order polynomial in age and the numbers of young children, below age 7, and older children, ages 7–17, in the household. Other determinants of work preferences that are mostly time-invariant, such as education, are effectively absorbed by individual fixed effects.

Wage regression. Labor-supply regressions are subject to division bias when wages are computed as earnings divided by hours worked; see, e.g., [Altonji \(1986\)](#), [Borjas \(1980\)](#), [Pencavel \(1986\)](#), and [Keane \(2011\)](#). Measurement error in hours then appears on both sides of the regression equation, generating a spurious negative correlation between constructed wages and hours worked. Following [Bredemeier, Gravert, and Juessen \(2023\)](#), we address this issue by estimating an initial wage regression separately for men and women and using it to construct predicted wage rates that are uncorrelated with measurement error in hours. The labor-supply regressions then use predicted log net wages, denoted $\log \tilde{w}_{i,j,t}$.

In addition to being uncorrelated with measurement error in hours worked, the right-hand-side variables in the wage regression should be uncorrelated with idiosyncratic shocks to tastes for work. This condition ensures that predicted wage variation reflects shifts in labor *demand*, driven by changes in productivity or business conditions, and therefore identifies the slope of the labor *supply* curve, that is, the labor-supply elasticity.

The specification of the wage regression closely follows [Bredemeier, Gravert, and Juessen \(2023\)](#) and exploits education-specific life-cycle wage profiles. Specifically, it includes a third-

Table 1: Market hours and total hours regressions.

	(1)	(2)	(3)	(4)
	log market hours, $\log n_{ij,t}^{\text{market}}$		log total hours, $\log n_{ij,t}^{\text{total}}$	
	men	women	men	women
log wage rate, $\log w_{ij,t}$	0.579*** (0.043)	1.135*** (0.079)	0.375*** (0.033)	0.461*** (0.042)
log housework hours, $\log n_{ij,t}^{\text{home}}$	-0.036*** (0.004)	-0.126*** (0.009)		
Observations	18801	18801	18801	18801

Notes: Dependent variables are log hours worked in the market, $\log n_{ij,t}^{\text{market}}$ (columns (1) and (2)), and log total hours worked, $\log n_{ij,t}^{\text{total}}$ (columns (3) and (4)). All regressions include individual and time fixed effects, taste shifters (number of young children, number of older children, cubic in age), log household consumption, log share of food expenditures, and a constant. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

order polynomial in age and interactions of these terms with education, firm tenure, firm tenure squared, state dummies, year dummies, and, following [Altonji \(1986\)](#), the controls from the labor-supply equation, including the child, consumption, and food-share controls, as well as individual fixed effects.

4.2 Labor-supply elasticities

We begin by estimating the market-hours regression [\(13\)](#), regressing log market hours on predicted log net wages, log housework time, individual and time fixed effects, a cubic in age, the numbers of young and older children, as well as log household consumption and the food expenditure share. We then estimate the corresponding regression for total hours, [\(14\)](#).

The first two columns of [Table 1](#) report the market-hours results for men and women. The male and female samples have the same number of observations by construction, as the sample consists of couples in which both spouses work for pay. To improve readability, the main-text tables omit control variables; [Appendix Table 7](#) reports the full estimates. For men, the estimated wage elasticity is 0.579, close to the values reported in [Bredemeier](#),

Gravert, and Juessen (2023). For women, the estimate is 1.135, about twice as large. Thus, consistent with the literature, the estimated Frisch elasticity of market work is substantially larger for women than for men.

Columns (3) and (4) report the results of the total-hours regressions, in which log total hours is the dependent variable and the log wage rate is the main explanatory variable; see equation (14). As discussed above, the model predicts that the wage coefficients in this regression should be more similar between men and women than in the market-hours regressions. The results support this prediction: the estimated total-hours elasticity is 0.375 for men and 0.461 for women. Relative to the pronounced gender gap in market-hours elasticities, this difference is small.

Taken together, Table 1 is consistent with the mechanism proposed by Alesina, Ichino, and Karabarbounis (2011). In their model, there are no deep gender differences in preferences, so the wage elasticities of total time are identical and gender differences in market-hours elasticities arise endogenously from household decisions. Our estimates point in the same direction: women have much higher market-hours elasticities, while the corresponding total-hours elasticities are much closer.

Covid-period interactions. The baseline labor-supply regressions include time fixed effects and therefore absorb shocks common to all households in a given survey year. As an additional check, we assess whether potential Covid-period changes in the slope coefficients of the wage equations affect our results. We define the Covid period as income years 2020 and 2022 (the 2021 income year is not observed due to the PSID’s biennial survey design). In constructing the predicted log net wages for this exercise, we also let the wage regressions include interactions between education categories and the Covid-period indicator. This allows the predicted wage measure to reflect education-specific wage changes during the pandemic, rather than only common survey-year shifts. We then re-estimate the four labor-supply equations after interacting the predicted log net wage with an indicator for Covid-period

income years. This specification separates the non-Covid wage slope from the incremental Covid-period slope and therefore addresses the pandemic episode without changing the estimation support. Appendix Table 9 reports the results. The non-Covid slopes are 0.571 for men’s market hours, 1.107 for women’s market hours, 0.368 for men’s total hours, and 0.444 for women’s total hours.⁵ These estimates are very close to the baseline estimates, suggesting that the pandemic years do not drive the estimated Frisch elasticities.

Couples without young children. We next re-estimate the baseline specification for couples without young children. In these households, the specialization channel emphasized by Alesina, Ichino, and Karabarbounis (2011) is likely to be less important because the scope for specialization is smaller. We therefore expect gender differences in market-hours elasticities to be smaller than in columns (1) and (2) of Table 1. In the restricted sample, the estimated market-hours elasticity is 0.590 for men and 0.865 for women, while the corresponding total-hours elasticities are 0.373 and 0.379, respectively; see Table 8 in Appendix C. The estimated market-hours elasticities are therefore more similar between men and women than in the full sample, and the estimated total-hours elasticities are nearly identical. The reduction in the market-hours gender gap comes mainly from the lower female elasticity, consistent with the specialization mechanism. These results suggest that the pronounced gender differences in market labor-supply elasticities are partly tied to couples with young children, for whom specialization in home production is likely more pronounced.

Selection. In our sample, 94.7% of married men aged 25–60 are observed in paid work with a usable wage measure, compared with 81.6% of married women. To assess whether selection into observed work affects the results, we estimate a discrete-choice model of female labor-force participation and compute an inverse Mills ratio, which we include as an additional regressor in the female wage equation. We follow Blundell, Pistaferri, and Saporta-Eksten (2016) and Del Boca and Lusardi (2003) in the choice of instruments. The latter study shows

⁵The corresponding Covid-period slopes are 0.580, 1.133, 0.375, and 0.462.

that female participation rises when households move into home ownership and therefore proposes the presence of first and second mortgages, interacted with year dummies, as instruments. The selection equation has predictive power: the mortgage-by-year instruments are jointly significant, with $\chi^2(26) = 163.67$ and $p < 0.001$. However, the coefficient on the resulting inverse Mills ratio is not statistically distinguishable from zero in the female wage equation ($p = 0.250$) and the selection-corrected labor-supply elasticities are nearly identical to the baseline estimates: 0.577 for men’s market hours, 1.150 for women’s market hours, 0.374 for men’s total hours, and 0.462 for women’s total hours. Thus, in line with the existing literature, see, e.g., [Blundell, Pistaferri, and Saporta-Eksten \(2016\)](#), correcting for selection has little impact on the results.

4.3 Intra-household efficiency of alternative tax rules

We now evaluate alternative tax rules empirically by comparing them to the model-implied optimal tax structure. We use the labor-supply elasticities, together with observed market and housework hours, to construct sufficient statistics for evaluating alternative tax rules. This approach deliberately abstracts from mechanisms that determine intra-household allocations. Instead, we take observed household allocation patterns as given and evaluate their consequences using a sufficient-statistics approach. The analysis therefore captures the tax-design implications of existing intra-household specialization, while abstracting from general-equilibrium effects of tax reforms on household behavior. While this may affect the level of implied optimal tax differentials, comparisons across tax systems are more robust unless reforms induce large changes in household specialization. In a partial-equilibrium robustness exercise, we introduce behavioral adjustments in household time allocations based on estimated labor-supply responses.

Using equation (23), we compute an implied optimal relative tax rate, $\theta_{ij,t}^*$, for each individual in the sample. We then regress $\theta_{ij,t}^*$ on potential elements of a tax rule, including income, gender, and other tags. The estimated coefficients indicate how strongly marginal

tax rates should be conditioned on each observable from the perspective of intra-household efficiency. The R^2 statistics summarize how well each rule captures the observed variation in implied optimal relative tax rates across households. They therefore measure the rule’s ability to target heterogeneous intra-household labor-supply distortions; see Section 3.3. All counterfactual comparisons are conducted in terms of relative marginal tax rates within households, conceptually holding fixed the underlying schedule of household tax liabilities that determines aggregate revenue.

We take gender-specific Frisch elasticities of total working time from the estimates in the previous section. We use the estimates of η_m and η_f from columns (3) and (4) of Table 1. Market hours n^{market} and total hours $n^{\text{total}} = n^{\text{market}} + n^{\text{home}}$ are taken directly from the PSID. For the multiplicative constant $1 - 2\gamma$, the literature suggests values between 1 and 1.5; see Section 3. Its exact value is secondary for our analysis because it rescales the estimated coefficients but leaves the R^2 , our main object of interest, unchanged. For simplicity, we set $1 - 2\gamma = 1$, while noting that coefficients could be up to 50% larger for other plausible values of γ .

Across all individuals, the log relative optimal tax rate has mean zero by construction and its standard deviation is 0.73. The distribution is symmetric between genders, with a mean of -0.48 for women and 0.48 for men, and a common standard deviation of 0.55 for both groups. The implied average gender gap in log relative tax rates is 0.96 for $1 - 2\gamma = 1$; the higher scaling of $1 - 2\gamma = 1.5$ implies a gap of 1.44 . Note that the reported gender gap differs from the log male-to-female tax-rate ratio. Each couple contributes two relative tax-rate observations with opposite signs, one for each spouse, so the gender gap is twice the log ratio of men’s to women’s marginal tax rates. The implied log tax-rate wedges are therefore 0.48 and 0.72 under the baseline and upper scaling, respectively. If the average marginal tax rate across both genders is 30% , these wedges imply average marginal tax rates of 37.1% for men and 22.9% for women in the baseline scaling, and 40.4% for men and 19.6% for women under the upper scaling.

This average difference can be implemented through gender-based taxation. However, optimal relative tax rates also vary substantially *within* gender, and this variation cannot be captured by gender alone. The relevant quantitative question is therefore how closely gender-based taxation can approximate efficient intra-household taxation.

Results. We now use regressions to quantify how closely simple tax rules approximate optimal intra-household relative tax rates. We begin with three regimes. The first is joint taxation of married couples, which corresponds to the U.S. status quo. In this system, both spouses in a couple face identical marginal tax rates, i.e., $\theta_{ij,t} = \theta \forall ij, t$. We represent this regime by regressing optimal relative within-household tax rates on a constant only. The second regime is gender-based taxation, implemented by regressing optimal rates on the individual’s gender and a constant. Third, we consider a progressive income-tax system in which married spouses file individually rather than jointly, so that relative marginal tax rates within couples depend on relative earnings. Specifically, we use log relative earnings as a determinant of relative marginal tax rates. We first restrict the relative-income coefficient to 0.311, the value implied by the estimates of Wu and Krueger (2021) for the current U.S. tax-transfer system.⁶ This evaluation illustrates the consequences of abolishing joint taxation of couples while maintaining the overall progressivity of the system. Thereafter, we estimate in an unrestricted regression the income-sensitivity of marginal tax rates that best approximates intra-household optimal relative tax rates. We also consider combinations of gender-based and income-based rules.⁷

Table 2 compares these tax regimes. By construction, joint taxation of married couples

⁶Wu and Krueger (2021) use the Bénabou (2000) tax function $T = Y - (1 - \chi)Y^{1-\mu}$, where T is taxes, Y is pre-tax income, and χ and μ are parameters determining average taxes and tax progressivity, respectively. The marginal tax rate is $T' = 1 - (1 - \chi)(1 - \mu)Y^{-\mu}$ and its elasticity to pre-tax income is $\mu(1 - T')/T'$. The (income-weighted) average marginal tax rate is $\mu = (E T' - E(T/Y))/(1 - E(T/Y))$. Wu and Krueger (2021) estimate an elasticity of after-tax income to pre-tax income of 0.1327. With a ratio of taxes to GDP of roughly 0.192 (Heathcote, Storesletten, and Violante 2020), this implies an elasticity of marginal tax rates to pre-tax income of 0.311. This elasticity refers to the local slope of the marginal tax rate schedule in log-log space and is used as a summary statistic for the curvature of the tax system.

⁷In the tax analysis, there is no need to account for year effects. The dependent variable is the log optimal tax-rate ratio between spouses. Because each household-year contributes one observation for each spouse and the wife’s value is the negative of the husband’s by construction, the weighted mean of the dependent variable is zero in every year.

Table 2: Comparison of intra-household tax efficiency under joint taxation of married couples, gender-based taxation, progressive separate taxation, and combinations.

	(1)	(2)	(3)	(4)	(5)	(6)
	joint taxation	gender- based	progressive separate taxation			
			restricted		unrestricted	
			base	+ gender	base	+ gender
Constant	0.000 (0.004)	-0.481 (0.004)	0.000 (0.003)	-0.284 (0.003)	0.000 (0.002)	-0.282 (0.003)
Male		0.962 (0.006)		0.568 (0.004)		0.564 (0.005)
log relative earnings			<i>0.311</i> —	<i>0.311</i> —	0.413 (0.002)	0.315 (0.002)
Observations	37602	37602	37602	37602	37602	37602
R^2	0.000	0.438	0.551	0.703	0.586	0.704

Notes: Dependent variable is log relative optimal tax rate, $\theta_{ij,t}^*$. Relative earnings are measured in logs. Standard errors in parentheses. Coefficients without standard errors (in italics) are constrained coefficients.

cannot capture any variation in optimal intra-household relative tax rates; see column (1). Relative to this benchmark, gender-based taxation improves intra-household tax efficiency because optimal tax rates vary systematically by gender. The coefficient on gender is 0.962, equal to the difference between the male and female means of the dependent variable. Because the dependent variable is an individual's log optimal tax rate relative to the spouse, the gender coefficient is twice the log male-to-female tax-rate ratio implied by a gender-based schedule. Thus, the estimated coefficient translates into a log gap between marginal tax rates of 0.481, corresponding to a relative gap of about three-fifths. Under the same illustrative average marginal tax rate of 30% used above, the implied log tax-rate ratio is 0.481, so the corresponding marginal tax rates are 37.1% for men and 22.9% for women, a pronounced difference of about 14 percentage points. This policy captures 44% of the variation in optimal relative marginal tax rates within households; see column (2).

We next consider separate progressive taxation, in which relative tax rates vary with relative income. We begin by holding tax progressivity fixed at current U.S. levels. Column

(3) shows that the R^2 of this restricted regression is 55%.⁸ Thus, abolishing joint filing while preserving the current degree of tax progressivity would approximate optimal intra-household relative tax rates with targeting performance above 50%. Parallel adjustments of all tax rates to achieve revenue neutrality would not affect this relative-marginal-rate result.

Column (4) examines a progressive tax system in which tax rates are also conditioned on gender. Conditioning on gender yields additional efficiency gains even when tax rates already depend on individual income: the R^2 rises to 70%, and the coefficient on being male is significantly positive. The coefficient is about three-fifths of the coefficient in the purely gender-based regime in column (2), however, and corresponds to a log male-to-female tax-rate wedge of 0.284 or a relative gap of about one-third. Under the 30% average-rate illustration, this wedge implies men's marginal tax rates about 8 percentage points above women's, conditional on income.

In columns (5) and (6), we relax the restriction on the relative-earnings coefficient and estimate the degree of progressivity that best approximates optimal intra-household relative tax rates. Without conditioning on gender, column (5), the unrestricted regression yields that relative marginal tax rates depend on relative income with an elasticity of 0.413. Such a system would replicate optimal relative intra-household tax rates with an accuracy of approximately 59%.

The unrestricted coefficient is above the 0.311 value implied by current U.S. progressivity. Thus, if the tax rule uses relative income alone, the best approximation to optimal intra-household relative tax rates calls for a steeper relative-income gradient than the current system.

Finally, column (6) allows both gender and income to determine marginal tax rates without imposing parameter restrictions. Both the estimated income coefficient and the resulting R^2 are almost identical to those in column (4). Conditional on a gender-specific tax

⁸We compute the R^2 for specifications with parameter constraints as $R^2 = 1 - \text{SSR}/\text{SST}$, where the sum of squared residuals (SSR) and the total sum of squares (SST) are calculated using the model's predicted values and the sample-weighted deviations from the mean.

rebate, the implied optimal degree of progressivity therefore remains close to its status quo level. Put differently, conditional on observed household allocation patterns, an appropriate gender-specific tax-rate component approximates the implied optimal tax structure nearly as closely as jointly re-optimizing both the gender and income components of the tax rule.

In summary, Table 2 shows that feasible alternatives to the current joint-taxation benchmark can capture a substantial share of the cross-sectional variation in implied optimal intra-household tax differentiation. Progressive separate taxation already outperforms gender-based taxation, and combining relative earnings with gender delivers the highest targeting performance, about 70%. Adjusting the relative-income gradient matters mainly when gender is not included; once gender is included, fixing the income coefficient at current U.S. progressivity performs almost the same as choosing it optimally.

We also repeat the baseline tax analysis using the non-Covid total-hours elasticities from the Covid-interaction labor-supply specification. Appendix Table 10 reports the full results. In this robustness exercise, gender-based taxation has targeting performance of 0.422. The corresponding values are 0.555 for restricted progressive separate taxation and 0.587 for unrestricted progressive separate taxation. These movements are small relative to the corresponding baseline estimates and leave the ranking of tax rules unchanged.

Ignoring the impact of relative housework time on optimal taxes. It is informative to assess the bias that arises when the endogenous dependence of labor-supply elasticities on the intra-household division of labor is ignored in the analysis and all gender differences in labor-supply elasticities are attributed directly to gender. To quantify this bias, we repeat the previous analysis using counterfactual optimal marginal tax ratios that would arise if the wage coefficients from the *market labor supply regressions*, columns (1) and (2) of Table 1, were incorrectly interpreted as the Frisch elasticities of *all* men and *all* women, respectively. These coefficients are in fact estimates of *average* Frisch elasticities by gender, while the model implies within-gender heterogeneity that follows from differences in relative

Table 3: Comparison of intra-household tax efficiency under gender-based taxation, progressive separate taxation, and combinations, when the endogenous dependence of labor-supply elasticities on the division of household chores is ignored.

	(1) gender- based	(2) progressive separate taxation restricted	(3) unrestricted
Constant	-0.673 (0.000)	0.000 (0.003)	0.000 (0.003)
Male	1.347 (0.000)		
Log rel. earn.		<i>0.311</i> —	0.235 (0.002)
Observations	37602	37602	37602
R^2	1.000	0.198	0.221

Notes: Dependent variable is the counterfactual log relative optimal tax rate, $\theta_{ij,t}^*$, constructed under the assumption that the wage coefficients from the market-hours regressions correspond to gender-specific Frisch elasticities of market labor supply that are constant within gender. Relative earnings are measured in logs. Standard errors in parentheses. Coefficients without standard errors (in italics) are constrained coefficients.

housework time. Table 3 reports results using as dependent variables the counterfactual tax ratios implied by average gender-specific elasticities, thereby ignoring within-gender heterogeneity. The resulting coefficient on being male is 1.35, above the 0.962 coefficient in the baseline gender-based regression, because this exercise assigns the entire market-hours elasticity gap to gender itself. By construction, this misspecified framework implies that gender-based taxation yields perfect intra-household tax efficiency; the R^2 in column (1) is one. In contrast, progressive separate taxation appears much less effective, capturing only 20–22% of the cross-sectional variation in implied optimal relative tax rates, columns (2) and (3). This is far below the 55–59% captured by the corresponding income-based specifications in Table 2. Ignoring the intra-household division of labor as a determinant of labor-supply elasticities therefore leads to substantially overstated efficiency gains from gender-based taxation and understated targeting performance of progressive separate taxation.

Weighted tax analysis. The baseline efficiency criterion treats a given deviation from the optimal relative tax rate as equally costly for all households in the estimation sample. This is the natural benchmark for the intra-household efficiency exercise above, but it abstracts

Table 4: Comparison of intra-household tax efficiency under inverse-income welfare weights.

	(1)	(2)	(3)	(4)	(5)	(6)
	joint taxation	gender- based	progressive separate taxation			
			restricted		unrestricted	
			base	+ gender	base	+ gender
Constant	0.000 (0.004)	-0.517 (0.005)	-0.000 (0.003)	-0.341 (0.003)	-0.000 (0.003)	-0.297 (0.004)
Male		1.033 (0.007)		0.681 (0.005)		0.593 (0.005)
log relative earnings			<i>0.311</i> —	<i>0.311</i> —	0.480 (0.002)	0.389 (0.002)
Observations	37602	37602	37602	37602	37602	37602
R^2	0.000	0.373	0.520	0.682	0.594	0.695

Notes: Dependent variable is log relative optimal tax rate, $\theta_{ij,t}^*$. Relative earnings are measured in logs. Regressions are weighted by the product of the PSID sample weight and normalized inverse household labor income. Standard errors are in parentheses. Coefficients without standard errors (in italics) are constrained coefficients.

from the fact that the marginal social value of resources may be higher in lower-income households. To assess the importance of this consideration, we repeat the tax analysis using welfare weights that depend on the inverse of household earnings. Specifically, the regression weight is the PSID sample weight multiplied by the inverse of household earnings and normalized to have a sample mean of one. This exercise does not recompute the optimal tax ratios themselves; it only changes the loss function used to evaluate how closely simple tax rules approximate those ratios.

Table 4 shows that reweighting the criterion toward lower-income households leaves the main ranking of tax rules unchanged. Gender-based taxation captures 37% of the weighted variation, while progressive separate taxation captures 52% when progressivity is fixed at the current U.S. value and 59% when the coefficient on relative income is chosen freely. Combining gender and income performs best, with R^2 values of 68–70%. The gender-only coefficient is 1.03, tantamount to a log male-to-female tax-rate wedge of 0.52, or a relative gap of about two-thirds. At the same time, the weighted criterion favors a steeper income gradient than current U.S. progressivity: the unrestricted coefficient on log relative

earnings is 0.480 in the income-only specification and 0.389 when gender is included. Thus, giving more weight to lower-income households gives relative income a larger role in the approximating tax rule, but it does not overturn the conclusion that gender and income together approximate the model-implied optimal tax structure most closely.

Are there better tags than gender? Although taxation based on relative income aligns more closely with optimal relative tax rates than pure gender-based taxation, it relies on an endogenous, tax-dependent variable. This feature can induce inefficient behavioral responses to taxation. By contrast, gender-based taxation can be interpreted as a form of tagging (Cremer, Gahvari, and Lozachmeur 2010). We therefore ask whether other simple observable characteristics, or tags, outperform gender in achieving intra-household tax efficiency.

Mankiw, Weinzierl, and Yagan (2009) discuss several potential tags for optimal tax systems, including the presence and number of children, gender, height, skin color, physical attractiveness, health, and parental education. For example, Mankiw and Weinzierl (2010) provide a quantitative analysis of height-based taxation in a redistributive tax framework. We evaluate how well selected tags from Mankiw, Weinzierl, and Yagan (2009) approximate predicted optimal intra-household marginal tax-rate ratios. Because the number of children in the household, skin color, and parental education are strongly correlated within couples, through joint choices or assortative mating, we focus on body height and the body mass index (BMI), which we use as a proxy for physical attractiveness.

Table 5 compares tagging strategies in income taxation, using the baseline estimates of optimal marginal tax ratios as the dependent variable. For completeness, we first re-estimate gender-based taxation in the slightly smaller sample for which height and BMI are available. As before, gender-based taxation captures 44% of the variation in intra-household relative marginal tax rates. Conditioning tax rates only on BMI, column (2), has low targeting performance, capturing 9% of the variation, relative to gender-based taxation. Moreover, the positive coefficient on BMI conflicts with the redistributive logic of taxing physical

Table 5: Comparison of intra-household tax efficiency under different forms of tagging in income taxation.

	(1) gender-based	(2) BMI	(3) body height	(4) BMI, height	(5) gender, height
Constant	-0.483 (0.004)	0.000 (0.004)	-0.000 (0.003)	-0.000 (0.003)	-0.484 (0.006)
Male	0.966 (0.006)				0.968 (0.011)
Log rel. BMI		0.905 (0.016)		0.368 (0.014)	
Log rel. body height			4.082 (0.032)	3.776 (0.034)	-0.015 (0.055)
Observations	35692	35692	35692	35692	35692
R^2	0.437	0.090	0.314	0.327	0.437

Notes: Dependent variable is the log relative optimal tax rate, $\theta_{ij,t}^*$. All right-hand-side variables except gender are measured in intra-household differences. Standard errors in parentheses. Body mass index (BMI) is $weight/height^2$.

attractiveness, because a higher BMI is typically associated with lower earnings. Tagging based on relative body height performs better than BMI, with an R^2 of 31%, and yields a coefficient consistent with redistributive motives, but still underperforms gender-based taxation. A combination of BMI and body height without gender, column (4), is also inferior to purely gender-based taxation, column (1), raising the R^2 only to 33%. Column (5) adds body height to the gender-based specification. The targeting performance is unchanged, and the coefficient on relative body height is close to zero and statistically insignificant.

Thus, among the tags considered here, gender provides the most effective basis for approximating optimal intra-household relative tax rates. The exercise does not overturn the baseline role of relative earnings; rather, it clarifies the relevant trade-off. While relative earnings are more closely related to intra-household specialization, they are, unlike gender, themselves influenced by the tax system.

Behavioral responses and tax-rule comparison. The baseline tax evaluation computes optimal intra-household relative tax rates using the observed allocation of working time between market work and household production. A valid concern is that tax reforms

may alter labor supply and thereby change the allocation of working time entering the efficiency benchmark. To assess the importance of this feedback, we conduct a partial-equilibrium behavioral robustness exercise.

Here, we consider two tax rules: gender-based taxation and progressive separate taxation with fixed progressivity. For each tax rule, we compute a behavioral equilibrium in which relative time allocations respond, according to the estimated labor-supply conditions and holding estimated coefficients constant, to the net wages implied by the rule and households' empirical marginal tax rates.⁹ In the case of gender-based taxation, determining the equilibrium includes re-optimizing the coefficient on the male indicator to match most closely the implied distribution of efficient intra-household tax rates. For progressive separate taxation, relative marginal tax rates adjust to changing relative earnings with a fixed elasticity of 0.311. With gross hourly wages constant, relative earnings move through the counterfactual market-hours response. In either case, we find the equilibrium through a fixed-point iteration of time allocations (and relative earnings) and marginal tax rates.

The analysis remains partial equilibrium in the sense that gross wages are held fixed because the model has no endogenous labor-demand responses, and we determine the level of marginal tax rates by holding household-specific average marginal tax rates constant rather than total tax liabilities. The resulting equilibrium allocation of working time between market work and housework determines the model-implied efficient relative tax rates, and these efficient tax rates are then compared to those implied by the tax rule itself.

Table 6 reports the results.¹⁰ They show that behavioral responses to the two rules are

⁹This step calculates, for every couple-year observation in our empirical sample, a pair of individual marginal tax rates that is consistent with the rule and holds the couple's mean marginal tax rate constant at its empirical level. We then determine counterfactual time allocations based on conditions (13) and (14) and the counterfactual net wage rates implied by the resulting marginal tax rates. The market-work share s^{market} is updated using a bounded logit transformation, so the simulated shares remain between zero and one. We can rely on Frisch labor-supply conditions due to our focus on intra-household log differences, where income effects cancel out.

¹⁰For gender-based taxation, the algorithm converges after four iterations and retains all observations. For progressive separate taxation, it converges after 70 iterations and excludes four household-year observations for which preserving the household's average marginal net-of-tax rate would imply a nonpositive or near-zero net-of-tax rate for one spouse.

Table 6: Behavioral response to gender-based taxation and progressive separate taxation

	Gender-based taxation		Progressive separate taxation	
	Emp. alloc.	Behav. equ.	Emp. alloc.	Behav. equ.
Avg. market work share s^{market}				
men	0.8553	0.8535	0.8553	0.8546
women	0.6896	0.6986	0.6896	0.6935
Coefficient on male/rel. income (standard error)	0.962 (0.006)	0.926 (0.006)	<i>0.311</i> —	<i>0.311</i> —
Avg. marginal tax rates				
men	0.311	0.309	0.276	0.275
women	0.192	0.194	0.227	0.229
R^2	0.438	0.426	0.551	0.530

Notes: The GBT coefficient is the estimated male dummy coefficient from the gender-based tax regression (twice the log male-to-female tax-rate wedge). The income coefficient in the progressive separate taxation rule is fixed at 0.311 (italicized). s^{market} denotes the share of total work time allocated to market production. “Emp. alloc.” corresponds to results under the observed time allocations. “Behav. equ.” corresponds to the converged behavioral equilibrium under each tax rule.

similar and modest, and that they leave the ranking of the tax rules unchanged. Under both rules, men’s average marginal tax rates are above the status quo and women’s are below it. In response, their time allocations converge moderately, with men on average reducing their share of working time devoted to market labor and women on average increasing theirs, as shown in the upper part of Table 6.

The model implies that as spouses’ allocations of working time become more similar, differences in labor-supply elasticities shrink and the case for differentiated taxation weakens. This mechanism is present under both reforms. Under gender-based taxation, gender becomes a slightly weaker predictor of efficient relative tax rates. Consequently, the coefficient on the male indicator falls, average marginal tax rates converge, and the R^2 declines, as shown in the lower-left part of Table 6.

Under progressive separate taxation, secondary earners, independent of gender, face lower marginal tax rates and therefore increase market work. Consequently, the behavioral response tends to reduce specialization. The model thus implies converging efficient relative

tax rates, and such convergence is generated by progressive separate taxation automatically. The adjustment is incomplete, however, because rule-implied relative tax rates move less than proportionally with relative earnings, which are rather closely linked to relative market-work shares and efficient relative tax rates. The lower-right part of Table 6 shows that the R^2 of progressive separate taxation declines slightly from 0.551 to 0.530 but that it remains above gender-based taxation in targeting performance.

The results should be interpreted with appropriate caution. Unlike the baseline tax evaluation, the behavioral exercise does not compare alternative tax rules against a common efficiency benchmark. Because labor supply adjusts differently under different tax rules, the model-implied efficient relative tax rates also differ across rules. Consequently, changes in R^2 should not be interpreted as welfare effects. Rather, the exercise asks how closely a given tax rule approximates the efficient relative tax structure that emerges under the behavioral responses induced by that same rule.

With this caveat in mind, the exercise provides useful insights on the relative importance of behavioral adjustments across tax regimes. Behavioral responses are modest and leave the gap between the targeting performance of gender-based taxation and progressive separate taxation nearly unchanged, thereby reinforcing the main conclusion of the tax evaluation. Progressive separate taxation remains a closer approximation to the model-implied efficient tax structure than gender-based taxation, and this advantage persists once behavioral responses are taken into account.

5 Conclusion

This paper studies how household specialization and gender-specific preferences shape labor-supply elasticities and the optimal taxation of dual-earner couples. In the model, optimal intra-household relative marginal tax rates depend on spouses' relative housework time and, when preferences differ by gender, on gender itself. The empirical results show that household specialization is central for understanding gender differences in market-hours elastic-

ities, while also leaving room for gender-related differences beyond the division of chores. In the policy analysis, we construct model-implied optimal intra-household tax ratios and evaluate how closely alternative implementable tax rules approximate them. Gender-based taxation delivers substantial targeting performance relative to joint taxation of married couples, but it is dominated by gender-neutral progressive tax systems with separate taxation. This ranking remains unchanged in the partial-equilibrium behavioral exercise that lets labor supply respond to the net wages implied by each tax rule.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process. During the preparation of this work, the authors used ChatGPT Pro 5.5 to improve the clarity and quality of the language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

- Akee, Randall K. Q., William E. Copeland, Gordon Keeler, Adrian Angold, and Elizabeth J. Costello. 2010. “Parents’ Incomes and Children’s Outcomes: A Quasi-Experiment.” *American Economic Journal: Applied Economics* 2 (1): 86–115.
- Alesina, Alberto, Andrea Ichino, and Loukas Karabarbounis. 2011. “Gender-Based Taxation and the Division of Family Chores.” *American Economic Journal: Economic Policy* 3 (2): 1–40.
- Alogoskoufis, George S. 1987. “On Intertemporal Substitution and Aggregate Labor Supply.” *Journal of Political Economy* 95 (5): 938–960.
- Altonji, Joseph G. 1986. “Intertemporal Substitution in Labor Supply: Evidence from Micro Data.” *Journal of Political Economy* 94 (3, Part 2): S176–S215.
- Apps, Patricia F., and Ray Rees. 1999. “Individual Versus Joint Taxation in Models With Household Production.” *Journal of Political Economy* 107 (2): 393–403.
- Auclert, Adrien, Bence Bardóczy, and Matthew Rognlie. 2023. “MPCs, MPEs, and Multipliers: A Trilemma for New Keynesian Models.” *Review of Economics and Statistics* 105 (3): 700–712.

- Banerjee, Abhijit V., Rema Hanna, Gabriel E. Kreindler, and Benjamin A. Olken. 2017. “Debunking the Stereotype of the Lazy Welfare Recipient: Evidence from Cash Transfer Programs.” *The World Bank Research Observer* 32 (2): 155–184.
- Bastagli, Francesca, Jessica Hagen-Zanker, Luke Harman, Valentina Barca, Georgina Sturge, Tanja Schmidt, and Luca Pellerano. 2016. *Cash transfers: What does the evidence say? A rigorous review of programme impact and the role of design and implementation features*. Technical report. London: Overseas Development Institute.
- Bastani, Spencer. 2013. “Gender-based and couple-based taxation.” *International Tax and Public Finance* 20 (4): 653–686.
- Bengtsson, Niklas. 2012. “The Marginal Propensity to Earn and Consume out of Unearned Income: Evidence Using an Unusually Large Cash Grant Reform.” *The Scandinavian Journal of Economics* 114 (4): 1393–1413.
- Berg, Kristoffer. 2023. *Revealing Inequality Aversion from Tax Policy and the Role of Non-Discrimination*.
- Blomquist, N. Sören. 1985. “Labour Supply in a Two-Period Model: The Effect of a Non-linear Progressive Income Tax.” *The Review of Economic Studies* 52 (3): 515–524.
- Blomquist, N. Sören. 1988. “Nonlinear taxes and labor supply.” *European Economic Review* 32 (6): 1213–1226.
- Blundell, Richard, Luigi Pistaferri, and Itay Saporta-Eksten. 2016. “Consumption Inequality and Family Labor Supply.” *American Economic Review* 106 (2): 387–435.
- Borjas, George J. 1980. “The Relationship between Wages and Weekly Hours of Work: The Role of Division Bias.” *The Journal of Human Resources* 15 (3): 409.
- Boskin, Michael J., and Eytan Sheshinski. 1983. “Optimal tax treatment of the family: Married couples.” *Journal of Public Economics* 20 (3): 281–297.
- Bredemeier, Christian, Jan Gravert, and Falko Juessen. 2019. “Estimating labor supply elasticities with joint borrowing constraints of couples.” *Journal of Labor Economics* 37 (4): 1215–1265.
- Bredemeier, Christian, Jan Gravert, and Falko Juessen. 2023. “Accounting for limited commitment between spouses when estimating labor-supply elasticities.” *Review of Economic Dynamics* 51: 547–578.
- Bredemeier, Christian, Falko Juessen, and Roland Winkler. 2023. “Bringing Back the Jobs Lost to Covid-19: The Role of Fiscal Policy.” *Journal of Money, Credit and Banking* 55 (7): 1703–1747.

- Bénabou, Roland. 2000. "Unequal Societies: Income Distribution and the Social Contract." *American Economic Review* 90 (1): 96–129.
- Cesarini, David, Erik Lindqvist, Matthew J. Notowidigdo, and Robert Östling. 2017. "The effect of wealth on individual and household labor supply: Evidence from Swedish lotteries." *American Economic Review* 107 (12): 3917–3946.
- Chaudhuri, Anita. 1995. "On the relationship between the Frisch and Slutsky decompositions." *Economics Letters* 47 (3-4): 283–290.
- Cremer, Helmuth, Firouz Gahvari, and Jean-Marie Lozachmeur. 2010. "Tagging and Income Taxation: Theory and an Application." *American Economic Journal: Economic Policy* 2 (1): 31–50.
- Del Boca, Daniela, and Annamaria Lusardi. 2003. "Credit market constraints and labor market decisions." *Labour Economics* 10 (6): 681–703.
- Domeij, David, and Martin Flodén. 2006. "The labor-supply elasticity and borrowing constraints: Why estimates are biased." *Review of Economic Dynamics* 9 (2): 242–262.
- Dyrda, Sebastian, and Marcelo Pedroni. 2023. "Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Income Risk." *The Review of Economic Studies* 90 (2): 744–780.
- Elminejad, Ali, Tomas Havranek, Roman Horvath, and Zuzana Irsova. 2023. "Intertemporal substitution in labor supply: A meta-analysis." *Review of Economic Dynamics*.
- Faberman, Jason. 2015. "Revisiting the role of home production in life-cycle labor supply."
- Gelber, Alexander, Timothy J. Moore, and Alexander Strand. 2017. "The Effect of Disability Insurance Payments on Beneficiaries' Earnings." *American Economic Journal: Economic Policy* 9 (3): 229–261.
- Golosov, Mikhail, Michael Graber, Magne Mogstad, and David Novgorodsky. 2021. *How Americans Respond to Idiosyncratic and Exogenous Changes in Household Wealth and Unearned Income*. Working Paper, Working Paper Series 29000. Cambridge, MA: National Bureau of Economic Research.
- Gottfried, Peter, and Wolfram F. Richter. 1999. "The Tax Unit and Household Production Comment." *Journal of Political Economy* 107 (2): 404–409.
- Greenwood, Jeremy, Zvi Hercowitz, and Gregory W. Huffman. 1988. "Investment, capacity utilization, and the real business cycle." *The American Economic Review* 78 (3): 402–417.
- Gromadzki, Jan. 2023. *Labor supply effects of a universal cash transfer*. Technical report.

- Guner, N., R. Kaygusuz, and G. Ventura. 2012. "Taxation and Household Labour Supply." *The Review of Economic Studies* 79 (3): 1113–1149.
- Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. 2020. "How Should Tax Progressivity Respond to Rising Income Inequality?" *Journal of the European Economic Association* 18 (6): 2715–2754.
- Heckman, James J. 1993. "What Has Been Learned About Labor Supply in the Past Twenty Years?" *American Economic Review* 83 (2): 116–121.
- Holter, Hans A., Dirk Krüger, and Serhiy Stepanchuk. 2023. *Until the IRS do us part: (Optimal) taxation of households*. Mimeo, University of Pennsylvania.
- Houthakker, H. S. 1960. "Additive preferences." *Econometrica* 28 (2): 244–257.
- Hundsdoerfer, Jochen, and Eva Matthaei. 2020. *Gender discriminatory taxes, fairness perception, and labor supply*.
- Imbens, Guido W., Donald B. Rubin, and Bruce I. Sacerdote. 2001. "Estimating the effect of unearned income on labor earnings, savings, and consumption: Evidence from a survey of lottery players." *American Economic Review* 91 (4): 778–794.
- Jacob, Brian A., and Jens Ludwig. 2012. "The Effects of Housing Assistance on Labor Supply: Evidence from a Voucher Lottery." *American Economic Review* 102 (1): 272–304.
- Jaimovich, Nir, and Sergio Rebelo. 2009. "Can News about the Future Drive the Business Cycle?" *American Economic Review* 99 (4): 1097–1118.
- Jones, Damon, and Ioana Marinescu. 2022. "The Labor Market Impacts of Universal and Permanent Cash Transfers: Evidence from the Alaska Permanent Fund." *American Economic Journal: Economic Policy* 14 (2): 315–340.
- Keane, Michael P. 2011. "Labor Supply and Taxes: A Survey." *Journal of Economic Literature* 49 (4): 961–1075.
- Keane, Michael P., and Richard Rogerson. 2015. "Reconciling Micro and Macro Labor Supply Elasticities: A Structural Perspective." *Annual Review of Economics* 7 (1): 89–117.
- Kimball, Miles S., and Matthew D. Shapiro. 2008. *Labor Supply: Are the Income and Substitution Effects Both Large or Both Small?* Working Paper, Working Paper Series 14208. National Bureau of Economic Research.
- Kleven, Henrik Jacobsen, and Claus Thustrup Kreiner. 2007. "Optimal Taxation of Married Couples with Household Production." *American Economic Journal: Economic Policy* 1 (1): 200–219.

- Kleven, Henrik Jacobsen, Claus Thustrup Kreiner, and Emmanuel Saez. 2009. "The Optimal Income Taxation of Couples." *Econometrica* 77 (2): 537–560.
- Mankiw, N. Gregory, and Matthew Weinzierl. 2010. "The Optimal Taxation of Height: A Case Study of Utilitarian Income Redistribution." *American Economic Journal: Economic Policy* 2 (1): 155–176.
- Mankiw, N. Gregory, Matthew Weinzierl, and Danny Yagan. 2009. "Optimal Taxation in Theory and Practice." *Journal of Economic Perspectives* 23 (4): 147–174.
- Marinescu, Ioana. 2018. *No Strings Attached: The Behavioral Effects of U.S. Unconditional Cash Transfer Programs*. Technical report. Cambridge, MA: National Bureau of Economic Research.
- Meier, Volker, and Helmut Rainer. 2015. "Pigou meets Ramsey: Gender-based taxation with non-cooperative couples." *European Economic Review* 77: 28–46.
- Pencavel, John. 1986. "Chapter 1 Labor supply of men: A survey." In *Handbook of Labor Economics*, edited by Orley Ashenfelter, R. Layard, and David Edward Card, 1: 3–102. North Holland.
- Picchio, Matteo, Sigrid Suetens, and Jan C. van Ours. 2018. "Labour Supply Effects of Winning a Lottery." *The Economic Journal* 128 (611): 1700–1729.
- Piggott, John, and John Whalley. 1996. "The Tax Unit and Household Production." *Journal of Political Economy* 104 (2): 398–418.
- Robins, Philip K. 1985. "A Comparison of the Labor Supply Findings from the Four Negative Income Tax Experiments." *The Journal of Human Resources* 20 (4): 567–582.
- Rupert, Peter, Richard Rogerson, and Randall Wright. 2000. "Homework in labor economics: Household production and intertemporal substitution." *Journal of Monetary Economics* 46 (3): 557–579.
- Saez, Emmanuel. 2001. "Using Elasticities to Derive Optimal Income Tax Rates." *The Review of Economic Studies* 68 (1): 205–229.
- Schmitt-Grohé, Stephanie, and Martín Uribe. 2012. "What's News in Business Cycles." *Econometrica* 80 (6): 2733–2764.
- Wolf, Christian K. 2023. "The Missing Intercept: A Demand Equivalence Approach." *American Economic Review* 113 (8): 2232–2269.
- Wu, Chunzan, and Dirk Krueger. 2021. "Consumption Insurance against Wage Risk: Family Labor Supply and Optimal Progressive Income Taxation." *American Economic Journal: Macroeconomics* 13 (1): 79–113.

Appendix

A Theoretical model

A.1 Government optimization and optimal taxation

The government's maximization problem is

$$\max_{\tau_{ij,t}, \tau_{-ij,t}} U_j(c_{j,t}, d_{j,t}, n_{ij,t}^{\text{home}}, n_{-ij,t}^{\text{home}}, n_{ij,t}^{\text{market}}, n_{-ij,t}^{\text{market}}),$$

subject to (17). The first-order condition is (18).

Substituting the household optimality conditions (15) and (16) into the government's first-order condition gives

$$\begin{aligned} & \frac{\partial U_j}{\partial c_{j,t}} \frac{\partial c_{j,t}}{\partial \tau_{ij,t}} + \frac{\partial U_j}{\partial c_{j,t}} \frac{(1 - \tau_{ij,t})w_{ij,t}}{\partial f / \partial n_{ij,t}^{\text{home}}} \frac{\partial d_{j,t}}{\partial \tau_{ij,t}} \\ & - \frac{\partial U_j}{\partial c_{j,t}} (1 - \tau_{ij,t})w_{ij,t} \frac{\partial n_{ij,t}^{\text{home}}}{\partial \tau_{ij,t}} - \frac{\partial U_j}{\partial c_{j,t}} (1 - \tau_{-ij,t})w_{-ij,t} \frac{\partial n_{-ij,t}^{\text{home}}}{\partial \tau_{ij,t}} \\ & - \frac{\partial U_j}{\partial c_{j,t}} (1 - \tau_{ij,t})w_{ij,t} \frac{\partial n_{ij,t}^{\text{market}}}{\partial \tau_{ij,t}} - \frac{\partial U_j}{\partial c_{j,t}} (1 - \tau_{-ij,t})w_{-ij,t} \frac{\partial n_{-ij,t}^{\text{market}}}{\partial \tau_{ij,t}} \\ & + \lambda_{j,t}^G \cdot \left(w_{ij,t} n_{ij,t}^{\text{market}} + \tau_{ij,t} w_{ij,t} \frac{\partial n_{ij,t}^{\text{market}}}{\partial \tau_{ij,t}} + \tau_{-ij,t} w_{-ij,t} \frac{\partial n_{-ij,t}^{\text{market}}}{\partial \tau_{ij,t}} \right) = 0 \quad \forall i. \end{aligned}$$

Rearranging yields

$$\begin{aligned} & \frac{\partial U_j}{\partial c_{j,t}} \left[\frac{\partial c_{j,t}}{\partial \tau_{ij,t}} + \frac{(1 - \tau_{ij,t})w_{ij,t}}{\partial f / \partial n_{ij,t}^{\text{home}}} \frac{\partial d_{j,t}}{\partial \tau_{ij,t}} \right. \\ & \quad - (1 - \tau_{ij,t})w_{ij,t} \frac{\partial n_{ij,t}^{\text{home}}}{\partial \tau_{ij,t}} - (1 - \tau_{-ij,t})w_{-ij,t} \frac{\partial n_{-ij,t}^{\text{home}}}{\partial \tau_{ij,t}} \\ & \quad \left. - (1 - \tau_{ij,t})w_{ij,t} \frac{\partial n_{ij,t}^{\text{market}}}{\partial \tau_{ij,t}} - (1 - \tau_{-ij,t})w_{-ij,t} \frac{\partial n_{-ij,t}^{\text{market}}}{\partial \tau_{ij,t}} \right] \\ & + \lambda_{j,t}^G \cdot \left(w_{ij,t} n_{ij,t}^{\text{market}} + \tau_{ij,t} w_{ij,t} \frac{\partial n_{ij,t}^{\text{market}}}{\partial \tau_{ij,t}} + \tau_{-ij,t} w_{-ij,t} \frac{\partial n_{-ij,t}^{\text{market}}}{\partial \tau_{ij,t}} \right) = 0. \end{aligned}$$

The responses of $c_{j,t}$ and $d_{j,t}$ to tax rates can be determined from the household constraints (3) and (4). Applying $a_{j,t+1} = (1 + r_t)a_{j,t}$ in the household budget constraint (3)

gives

$$c_{j,t} = \sum_i (1 - \tau_{ij,t}) w_{ij,t} n_{ij,t}^{\text{market}},$$

which yields

$$\frac{\partial c_{j,t}}{\partial \tau_{ij,t}} = -w_{ij,t} n_{ij,t}^{\text{market}} + (1 - \tau_{ij,t}) w_{ij,t} \frac{\partial n_{ij,t}^{\text{market}}}{\partial \tau_{ij,t}} + (1 - \tau_{-ij,t}) w_{-ij,t} \frac{\partial n_{-ij,t}^{\text{market}}}{\partial \tau_{ij,t}}.$$

Further, for the home production function (4), it holds that

$$\frac{\partial d_{j,t}}{\partial \tau_{ij,t}} = \frac{\partial f}{\partial n_{ij,t}^{\text{home}}} \frac{\partial n_{ij,t}^{\text{home}}}{\partial \tau_{ij,t}} + \frac{\partial f}{\partial n_{-ij,t}^{\text{home}}} \frac{\partial n_{-ij,t}^{\text{home}}}{\partial \tau_{ij,t}}.$$

Using these results in the optimal-tax condition gives

$$\begin{aligned} \frac{\partial U_j}{\partial c_{j,t}} & \left[-w_{ij,t} n_{ij,t}^{\text{market}} + (1 - \tau_{ij,t}) w_{ij,t} \frac{\partial n_{ij,t}^{\text{market}}}{\partial \tau_{ij,t}} + (1 - \tau_{-ij,t}) w_{-ij,t} \frac{\partial n_{-ij,t}^{\text{market}}}{\partial \tau_{ij,t}} \right. \\ & + \frac{(1 - \tau_{ij,t}) w_{ij,t}}{\partial f / \partial n_{ij,t}^{\text{home}}} \left(\frac{\partial f}{\partial n_{ij,t}^{\text{home}}} \frac{\partial n_{ij,t}^{\text{home}}}{\partial \tau_{ij,t}} + \frac{\partial f}{\partial n_{-ij,t}^{\text{home}}} \frac{\partial n_{-ij,t}^{\text{home}}}{\partial \tau_{ij,t}} \right) \\ & - (1 - \tau_{ij,t}) w_{ij,t} \frac{\partial n_{ij,t}^{\text{home}}}{\partial \tau_{ij,t}} - (1 - \tau_{-ij,t}) w_{-ij,t} \frac{\partial n_{-ij,t}^{\text{home}}}{\partial \tau_{ij,t}} \\ & \left. - (1 - \tau_{ij,t}) w_{ij,t} \frac{\partial n_{ij,t}^{\text{market}}}{\partial \tau_{ij,t}} - (1 - \tau_{-ij,t}) w_{-ij,t} \frac{\partial n_{-ij,t}^{\text{market}}}{\partial \tau_{ij,t}} \right] \\ & + \lambda_{j,t}^G \cdot \left(w_{ij,t} n_{ij,t}^{\text{market}} + \tau_{ij,t} w_{ij,t} \frac{\partial n_{ij,t}^{\text{market}}}{\partial \tau_{ij,t}} + \tau_{-ij,t} w_{-ij,t} \frac{\partial n_{-ij,t}^{\text{market}}}{\partial \tau_{ij,t}} \right) = 0 \end{aligned}$$

which can be simplified to

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \tau_{ij,t}} & = \frac{\partial U_j}{\partial c_{j,t}} \left[-w_{ij,t} n_{ij,t}^{\text{market}} + \frac{(1 - \tau_{ij,t}) w_{ij,t}}{\partial f / \partial n_{ij,t}^{\text{home}}} \left(\frac{\partial f}{\partial n_{ij,t}^{\text{home}}} \frac{\partial n_{ij,t}^{\text{home}}}{\partial \tau_{ij,t}} + \frac{\partial f}{\partial n_{-ij,t}^{\text{home}}} \frac{\partial n_{-ij,t}^{\text{home}}}{\partial \tau_{ij,t}} \right) \right. \\ & \quad \left. - (1 - \tau_{ij,t}) w_{ij,t} \frac{\partial n_{ij,t}^{\text{home}}}{\partial \tau_{ij,t}} - (1 - \tau_{-ij,t}) w_{-ij,t} \frac{\partial n_{-ij,t}^{\text{home}}}{\partial \tau_{ij,t}} \right] \\ & \quad + \lambda_{j,t}^G \cdot \left(w_{ij,t} n_{ij,t}^{\text{market}} + \tau_{ij,t} w_{ij,t} \frac{\partial n_{ij,t}^{\text{market}}}{\partial \tau_{ij,t}} + \tau_{-ij,t} w_{-ij,t} \frac{\partial n_{-ij,t}^{\text{market}}}{\partial \tau_{ij,t}} \right) = 0. \end{aligned}$$

From (16), $(1 - \tau_{ij,t}) w_{ij,t} / (\partial f / \partial n_{ij,t}^{\text{home}})$ is the same for both household members i and $-i$.

Using this in the optimal-tax condition and multiplying out gives

$$\begin{aligned} \frac{\partial U_j}{\partial c_{j,t}} & \left[-w_{ij,t} n_{ij,t}^{\text{market}} + \frac{(1 - \tau_{ij,t}) w_{ij,t}}{\partial f / \partial n_{ij,t}^{\text{home}}} \frac{\partial f}{\partial n_{ij,t}^{\text{home}}} \frac{\partial n_{ij,t}^{\text{home}}}{\partial \tau_{ij,t}} + \frac{(1 - \tau_{-ij,t}) w_{-ij,t}}{\partial f / \partial n_{-ij,t}^{\text{home}}} \frac{\partial f}{\partial n_{-ij,t}^{\text{home}}} \frac{\partial n_{-ij,t}^{\text{home}}}{\partial \tau_{ij,t}} \right. \\ & \quad \left. - (1 - \tau_{ij,t}) w_{ij,t} \frac{\partial n_{ij,t}^{\text{home}}}{\partial \tau_{ij,t}} - (1 - \tau_{-ij,t}) w_{-ij,t} \frac{\partial n_{-ij,t}^{\text{home}}}{\partial \tau_{ij,t}} \right] \\ & + \lambda_{j,t}^G \cdot \left(w_{ij,t} n_{ij,t}^{\text{market}} + \tau_{ij,t} w_{ij,t} \frac{\partial n_{ij,t}^{\text{market}}}{\partial \tau_{ij,t}} + \tau_{-ij,t} w_{-ij,t} \frac{\partial n_{-ij,t}^{\text{market}}}{\partial \tau_{ij,t}} \right) = 0. \end{aligned}$$

In the square brackets, all but the first term cancel, which gives

$$-\frac{\partial U_j}{\partial c_{j,t}} w_{ij,t} n_{ij,t}^{\text{market}} + \lambda_{j,t}^G \cdot \left(w_{ij,t} n_{ij,t}^{\text{market}} + \tau_{ij,t} w_{ij,t} \frac{\partial n_{ij,t}^{\text{market}}}{\partial \tau_{ij,t}} + \tau_{-ij,t} w_{-ij,t} \frac{\partial n_{-ij,t}^{\text{market}}}{\partial \tau_{ij,t}} \right).$$

Dividing by $w_{ij,t} n_{ij,t}^{\text{market}}$ yields (19).

A.2 Simplification of relative optimal marginal tax rates

As shown by [Chaudhuri \(1995\)](#), $\partial n_{ij,t}^{\text{market}} / \partial w_{ij,t}$ and $\partial n_{ij,t}^{\text{market}} / \partial w_{-ij,t}$ can be decomposed as

$$\begin{aligned} \frac{\partial n_{ij,t}^{\text{market}}}{\partial w_{ij,t}} &= \frac{\partial n_{ij,t}^{\text{market}}}{\partial w_{ij,t}} \Big|_{\lambda} + n_{ij,t}^{\text{market}} \frac{\partial n_{ij,t}^{\text{market}}}{\partial y_{j,t}} - \xi_{ij,t}^{\text{own}}, \\ \frac{\partial n_{ij,t}^{\text{market}}}{\partial w_{-ij,t}} &= \frac{\partial n_{-ij,t}^{\text{market}}}{\partial w_{ij,t}} \Big|_{\lambda} + n_{-ij,t}^{\text{market}} \frac{\partial n_{ij,t}^{\text{market}}}{\partial y_{j,t}} - \xi_{ij,t}^{\text{cross}}, \end{aligned}$$

where $y_{j,t}$ is the period- t value of household j 's stream of unearned income. In both equations, the first term is the derivative of the Frisch labor-supply function, the second is the classical income effect from the Slutsky decomposition, and the final term, $\xi_{ij,t}^{\text{own}}$ or $\xi_{ij,t}^{\text{cross}}$, is the general substitution effect ([Houthakker 1960](#)). The general substitution effect captures income changes induced by behavioral responses, specifically changes in all supply and demand decisions according to the relevant Frisch supply or demand functions.

With additively separable preferences, the only Frisch responses are to a decision vari-

able's own price. This implies

$$\begin{aligned}\xi_{ij,t}^{\text{own}} &= \frac{\partial n_{ij,t}^{\text{market}}}{\partial y_{j,t}} w_{ij,t} \frac{\partial n_{ij,t}^{\text{market}}}{\partial w_{ij,t}} \Big|_{\lambda}, \\ \xi_{ij,t}^{\text{cross}} &= \frac{\partial n_{ij,t}^{\text{market}}}{\partial y_{j,t}} w_{-ij,t} \frac{\partial n_{-ij,t}^{\text{market}}}{\partial w_{-ij,t}} \Big|_{\lambda}.\end{aligned}$$

Additive separability also implies that the Frisch cross-wage derivative in $\partial n_{ij,t}^{\text{market}} / \partial w_{-ij,t}$ is zero. Consequently, the difference between the own-wage and cross-wage elasticities in (21) simplifies to

$$\begin{aligned}e_{ij,t}^{\text{own}} - e_{ij,t}^{\text{cross}} &= \frac{\partial n_{ij,t}^{\text{market}}}{\partial w_{ij,t}} \cdot \frac{w_{ij,t}}{n_{ij,t}^{\text{market}}} - \frac{\partial n_{ij,t}^{\text{market}}}{\partial w_{-ij,t}} \cdot \frac{w_{ij,t}}{n_{-ij,t}^{\text{market}}} \\ &= \left(\frac{\partial n_{ij,t}^{\text{market}}}{\partial w_{ij,t}} \Big|_{\lambda} + n_{ij,t}^{\text{market}} \frac{\partial n_{ij,t}^{\text{market}}}{\partial y_{j,t}} \right) \cdot \frac{w_{ij,t}}{n_{ij,t}^{\text{market}}} - n_{-ij,t}^{\text{market}} \frac{\partial n_{ij,t}^{\text{market}}}{\partial y_{j,t}} \cdot \frac{w_{ij,t}}{n_{-ij,t}^{\text{market}}} \\ &\quad - \xi_{ij,t}^{\text{own}} \cdot \frac{w_{ij,t}}{n_{ij,t}^{\text{market}}} + \xi_{ij,t}^{\text{cross}} \cdot \frac{w_{ij,t}}{n_{-ij,t}^{\text{market}}} \\ &= e_{ij,t}^{\text{Frisch}} - \frac{\partial n_{ij,t}^{\text{market}}}{\partial y_{j,t}} \cdot w_{ij,t} \cdot (e_{ij,t}^{\text{Frisch}} - e_{-ij,t}^{\text{Frisch}})\end{aligned}$$

and the analogous steps for $e_{-ij,t}^{\text{own}} - e_{-ij,t}^{\text{cross}}$ give

$$e_{-ij,t}^{\text{own}} - e_{-ij,t}^{\text{cross}} = e_{-ij,t}^{\text{Frisch}} - \frac{\partial n_{-ij,t}^{\text{market}}}{\partial y_{j,t}} \cdot w_{-ij,t} \cdot (e_{-ij,t}^{\text{Frisch}} - e_{ij,t}^{\text{Frisch}})$$

Thus, optimal relative marginal tax rates depend on Frisch substitution effects, while income effects cancel out. The remaining term tends to be small when heterogeneity in Frisch elasticities within the household is limited.

To simplify further, we apply a first-order Taylor approximation of $e_{ij,t}^{\text{own}} - e_{ij,t}^{\text{cross}}$ around a situation in which spouses are identical in all respects and therefore have, for example, equal Frisch elasticities. This gives, in logs,

$$\log(e_{ij,t}^{\text{own}} - e_{ij,t}^{\text{cross}}) \approx (1 - \gamma) \log e_{ij,t}^{\text{Frisch}} + \gamma \log e_{-ij,t}^{\text{Frisch}},$$

with $\gamma = w \cdot \partial n / \partial y$ being the (individual) propensity to earn out of (family) unearned income in the point of approximation where no individual indices are needed due to symmetry. Using

the approximation, we can write

$$\begin{aligned}
& \log(e_{ij,t}^{\text{own}} - e_{ij,t}^{\text{cross}}) - \log(e_{-ij,t}^{\text{own}} - e_{-ij,t}^{\text{cross}}) \\
& \approx (1 - \gamma) \log e_{ij,t}^{\text{Frisch}} + \gamma \log e_{-ij,t}^{\text{Frisch}} - (1 - \gamma) \log e_{-ij,t}^{\text{Frisch}} - \gamma \log e_{ij,t}^{\text{Frisch}} \\
& = (1 - 2\gamma) \log e_{ij,t}^{\text{Frisch}} - (1 - 2\gamma) \log e_{-ij,t}^{\text{Frisch}}.
\end{aligned}$$

Using this result in (21) gives (22).

B Marginal propensity to earn out of unearned income

We briefly review the empirical literature on the marginal propensity to earn out of unearned income, or the wealth effect on labor supply. A large body of research suggests that this wealth effect is small or close to zero. For instance, Schmitt-Grohé and Uribe (2012) estimate an RBC model and find the wealth effect on labor supply to be essentially zero. This result is consistent with quasi-experimental evidence on cash transfers in developed and developing countries. Marinescu (2018) summarizes quasi-experimental designs from high-income countries, while Banerjee et al. (2017) and Bastagli et al. (2016) review field experiments in low-income settings. All three conclude that cash transfers have little to no adverse effect on labor supply.¹¹

Several studies exploit the quasi-random nature of lottery wins to estimate the marginal propensity to earn out of unearned income. Imbens, Rubin, and Sacerdote (2001), Cesarini et al. (2017), and Picchio, Suetens, and van Ours (2018) all find modest labor-supply responses.¹² The extensive empirical evidence pointing to small wealth effects has motivated

¹¹Synthesizing the findings from the negative income tax experiments in the 1970s, Robins (1985) estimates the marginal propensity to earn out of unearned income to lie between -0.06 and -0.10 . Using a difference-in-differences strategy based on casino payments to Native American families, Akee et al. (2010) find no labor-supply response to unearned income. Jones and Marinescu (2022) use a synthetic-control design for the Alaska Permanent Fund Dividend and similarly find no effect on employment. However, macroeconomic spillovers, such as wage adjustments, may offset underlying wealth effects.

¹²Cesarini et al. (2017) and Picchio, Suetens, and van Ours (2018) examine how couples adjust labor earnings after receiving one-time windfall gains. In the year of the shock, household earnings decline by 1.4% (Cesarini et al. 2017) and 1.8% (Picchio, Suetens, and van Ours 2018) of the amount received, respectively. Multiplying this initial response by the average remaining years in the labor market, as in Cesarini et al. (2017), yields upper-bound estimates for the marginal propensity to earn out of unearned income of -0.225 and -0.358 , respectively. Imbens, Rubin, and Sacerdote (2001) study lottery prizes paid out in annual installments and

the development of utility functions consistent with this property (Greenwood, Hercowitz, and Huffman 1988; Jaimovich and Rebelo 2009). As a result, many macroeconomic studies adopt models in which income effects are assumed to be negligible (Auclert, Bardóczy, and Rognlie 2023; Bredemeier, Juessen, and Winkler 2023; Dyrda and Pedroni 2023; Wolf 2023).

Some studies estimate larger wealth effects on labor supply. Gromadzki (2023) exploits the design of a child benefit program in Poland and finds a marginal propensity to earn out of unearned income of -0.14 . Gelber, Moore, and Strand (2017) use a regression discontinuity in eligibility for disability insurance payments in the U.S. and estimate a propensity of -0.2 . Kimball and Shapiro (2008) rely on survey responses to hypothetical lottery wins and report values close to -0.3 , similar to the estimate in Bengtsson (2012) from a reform to unconditional cash transfers in South Africa. The largest microeconomic estimate we are aware of is -0.51 , reported by Golosov et al. (2021), who use an event-study design based on variation in the timing of lottery wins.

Overall, most estimates suggest that the marginal propensity to earn out of unearned income is negligible or moderate at most, typically ranging between -0.5 and 0 .

report a marginal propensity to earn of approximately -11% for individuals. Jacob and Ludwig (2012) analyze the effects of winning a housing voucher lottery in Chicago and find a similarly small response, with a marginal propensity to earn of about -0.09 .

C Additional regression results

For completeness, Table 7 reports the full estimation results for the baseline labor-supply specification underlying Table 1. The main-text table reports the key wage and housework coefficients; the appendix table displays the full set of included controls.

Table 7: Market hours and total hours regressions (complete estimation results)

	(1) log market hours, $\log n_{ij,t}^{\text{market}}$	(2) log market hours, $\log n_{ij,t}^{\text{market}}$	(3) log total hours, $\log n_{ij,t}^{\text{total}}$	(4) log total hours, $\log n_{ij,t}^{\text{total}}$
	men	women	men	women
log wage rate, $\log w_{ij,t}$	0.579*** (0.043)	1.135*** (0.079)	0.375*** (0.033)	0.461*** (0.042)
log housework hours, $\log n_{ij,t}^{\text{home}}$	-0.036*** (0.004)	-0.126*** (0.009)		
age	-0.180*** (0.022)	-0.368*** (0.037)	-0.094*** (0.017)	-0.151*** (0.020)
age ²	0.004*** (0.001)	0.008*** (0.001)	0.002*** (0.000)	0.003*** (0.000)
age ³	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)
# kids age 0–5 = 1	-0.034*** (0.009)	-0.276*** (0.015)	-0.010 (0.007)	-0.080*** (0.008)
# kids age 0–5 = 2	-0.045*** (0.013)	-0.412*** (0.021)	-0.005 (0.010)	-0.109*** (0.011)
# kids age 0–5 = 3	-0.012 (0.026)	-0.555*** (0.044)	0.047** (0.020)	-0.074*** (0.023)
# kids age 6–17 = 1	-0.002 (0.008)	-0.011 (0.014)	0.002 (0.006)	0.013* (0.007)
# kids age 6–17 = 2	0.013 (0.010)	-0.025 (0.019)	0.018** (0.008)	0.022** (0.010)
# kids age 6–17 = 3	-0.025 (0.016)	-0.054* (0.030)	-0.007 (0.012)	0.030* (0.016)
# kids age 6–17 = 4	-0.011 (0.032)	0.056 (0.056)	0.016 (0.024)	0.112*** (0.030)
# kids age 6–17 = 5	-0.093 (0.067)	-0.061 (0.110)	0.019 (0.051)	0.102* (0.059)
log consumption, $\log \tilde{c}_{j,t}$	-0.010 (0.009)	0.077*** (0.015)	0.007 (0.007)	0.042*** (0.008)
log food share, $\log \left(\frac{\tilde{c}_{k,j,t}}{\tilde{c}_{j,t}} \right)$	0.008 (0.006)	-0.012 (0.010)	0.004 (0.005)	-0.001 (0.005)
Constant	9.076*** (0.302)	9.550*** (0.477)	4.287*** (0.228)	4.477*** (0.253)
Year dummies included	Yes	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes	Yes
Observations	18801	18801	18801	18801

Table 8 reports the corresponding restricted-sample estimates for couples without children below age 7, as discussed in the robustness analysis in the main text.

Table 8: Market hours and total hours regressions, sample of couples without young children.

	(1) log market hours, $\log n_{ij,t}^{\text{market}}$	(2) log market hours, $\log n_{ij,t}^{\text{market}}$	(3) log total hours, $\log n_{ij,t}^{\text{total}}$	(4) log total hours, $\log n_{ij,t}^{\text{total}}$
	men	women	men	women
log wage rate, $\log w_{ij,t}$	0.590*** (0.050)	0.865*** (0.081)	0.373*** (0.037)	0.379*** (0.044)
log housework hours, $\log n_{ij,t}^{\text{home}}$	-0.037*** (0.005)	-0.106*** (0.010)		
Observations	13402	13402	13402	13402

Notes: Restricted sample without children below age 7. Dependent variables are log hours worked in the market, $\log n_{ij,t}^{\text{market}}$ (columns (1) and (2)), and log total hours worked, $\log n_{ij,t}^{\text{total}}$ (columns (3) and (4)). All regressions include individual and time fixed effects, taste shifters (a cubic in age and the number of older children), log household consumption, log share of food expenditures, and a constant. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 9 reports the labor-supply robustness specification that allows the wage slope to differ in Covid-period income years, as discussed in the main text.

Table 9: Market hours and total hours regressions with Covid-period wage-slope interactions.

	(1) log market hours, $\log n_{ij,t}^{\text{market}}$	(2) log market hours, $\log n_{ij,t}^{\text{market}}$	(3) log total hours, $\log n_{ij,t}^{\text{total}}$	(4) log total hours, $\log n_{ij,t}^{\text{total}}$
	men	women	men	women
<i>Direct regression coefficients</i>				
Non-Covid log wage slope	0.571*** (0.044)	1.107*** (0.078)	0.368*** (0.033)	0.444*** (0.041)
Covid-period interaction	0.009 (0.007)	0.025** (0.012)	0.007 (0.005)	0.017*** (0.006)
<i>Post-estimation linear combination</i>				
Covid-period log wage slope	0.580*** (0.043)	1.133*** (0.079)	0.375*** (0.033)	0.462*** (0.042)
Observations	18801	18801	18801	18801
Covid-period observations	2453	2453	2453	2453

Notes: Covid-period income years are 2020 and 2022. The non-Covid log wage slope is the coefficient on predicted log net wages outside the Covid period. The Covid-period interaction is the incremental slope in Covid-period income years. The Covid-period log wage slope is reported as a post-estimation linear combination equal to the sum of the non-Covid slope and the Covid-period interaction, not as a separate regression coefficient. Each column is estimated from the specification with the Covid-period wage-slope interaction and includes individual and time fixed effects, taste shifters, consumption controls, and the same Covid-period education interactions in the wage regressions described in the main text. Standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

D Additional tax evaluations

Table 10 reports the tax-rule comparison obtained when the optimal tax-rate ratios are computed using the non-Covid total-hours slopes from the Covid-interaction labor-supply specification.

Table 10: Comparison of intra-household tax efficiency using non-Covid total-hours slopes from the Covid-interaction labor-supply specification.

	(1) joint taxation	(2) gender- based	(3)	(4)	(5)	(6)
			progressive separate taxation			
			restricted		unrestricted	
			base	+ gender	base	+ gender
Constant	0.000 (0.004)	-0.465 (0.004)	0.000 (0.002)	-0.268 (0.003)	0.000 (0.002)	-0.266 (0.003)
Male		0.930 (0.006)		0.536 (0.004)		0.531 (0.005)
log relative earnings			<i>0.311</i> —	<i>0.311</i> —	0.407 (0.002)	0.315 (0.002)
Observations	37602	37602	37602	37602	37602	37602
R^2	0.000	0.422	0.555	0.695	0.587	0.695

Notes: Dependent variable is log relative optimal tax rate, $\theta_{ij,t}^*$, recomputed using the non-Covid total-hours wage slopes from the Covid-interaction labor-supply specification. Relative earnings are measured in logs. Standard errors are in parentheses. Coefficients without standard errors (in italics) are constrained coefficients.