Taxation and Household Decisions: an Intertemporal Analysis*

Mary Ann Bronson[†] Daniel Haanwinckel[‡] Maurizio Mazzocco[§]

April 2025

Abstract

How do different income taxation systems affect household decisions and welfare? We answer this question by first documenting the strong labor supply disincentives for secondary earners of the U.S. tax system and by using variations from the Bush Tax Cuts to assess their effects on intra-household specialization. We then develop a lifecycle model incorporating labor supply, marriage and divorce decisions with limited commitment, household production, human capital accumulation, and assortative mating. After estimating and validating the model with various datasets, we evaluate four tax systems: a U.S.-like income-splitting system, an individual taxation system, a flexible general joint system, and an income-splitting system with secondary-earner deductions. We find that the individual taxation system provides higher welfare than income splitting but increases inequality. The general joint system offers the highest welfare but is complex to implement. The income-splitting system with a secondary-earner deduction improves welfare and reduces inequality while maintaining simplicity.

*We are grateful to participants at various seminars and conferences for helpful comments.
[†]Georgetown University, Dept. of Economics, Washington, D.C. Email: mary.ann.bronson@gmail.com.
[‡]UCLA, Department of Economics, Bunche Hall, Los Angeles, CA and NBER. Email: haanwinckel@econ.ucla.edu.
[§]UCLA, Department of Economics, Bunche Hall, Los Angeles, CA. Email: mmazzocc@econ.ucla.edu.

1 Introduction

Income taxes constitute the primary source of government revenue in the U.S. and most developed countries. Two important elements of income tax design are the tax system's progressivity and whether marital status and spousal earnings affect how an individual is taxed. Both can have profound effects on output, welfare, and income inequality. While most modern income tax systems are progressive, countries differ significantly in their treatment of married couples. Under individual systems such as the ones used in Canada and Sweden, marital status has little or no effect on an individual's tax schedule. By contrast, the U.S. and Germany use a joint system in which tax rates of married couples are determined by pooled rather than individual income. Other countries, like Spain and France, have hybrid systems that borrow elements from individual and joint schemes.

In this paper, we answer two questions about income taxation. How should a government tax single and married households if the objective is to maximize total welfare? How does this choice affect welfare inequality? We answer them in three steps.

We begin by providing descriptive evidence on how different taxation systems impact individual decision-making. Our analysis reveals that the U.S. joint taxation system creates incentives for specialization within the household compared to individual taxation. These incentives manifest in two primary ways. First, joint taxation imposes significantly higher marginal and average tax rates on secondary earners. Second, it provides marriage bonuses – defined as the difference in a couple's after-tax income between filing jointly or individually. These bonuses are substantial for single-earner couples, decrease as the secondary earner's income rises, and can even turn into marriage penalties for dual high-income couples. Our evidence also documents that individuals respond to these incentives. In the U.S., the higher tax rates for secondary earners, as predicted by the primary earner's income, are associated with lower employment rates for secondary earners. In contrast, Canada, with its individual taxation, exhibits no systematic relationship between the employment of secondary earners and the income of primary earners.

To investigate whether this pattern reflects a causal link, we study the employment effects of a U.S. reform that changed tax rates for some secondary earners: the Jobs and Growth Tax Relief Reconciliation Act of 2003 (the "Bush Tax Cuts"). Using a difference-in-differences specification and data from the Current Population Survey (CPS), we estimate that a six percentage point reduction in the average tax rate of secondary earners increased their labor force participation by four percentage points. The response is about twice as large in families with young children. We obtain similar results using the Panel Study of Income Dynamics (PSID). These findings suggest that moving away from joint taxation, by reducing marriage bonuses and lowering secondary earners' tax rates, could significantly impact how married couples divide their time between work and household production, and the resulting accumulation of human capital. They also indicate that the effects should vary based on the presence of children.

In the second part of the paper, we develop, estimate, and validate a dynamic model of household decisions, which we later use for policy evaluation. In the model, individuals can be single or married, with endogenous marriage and divorce decisions. If married, the couple makes joint and efficient decisions, but cannot commit to future allocations of resources. Whether married or single, individuals choose labor supply on the extensive and intensive margins, consumption, savings, and the time and market inputs they invest in producing a home-produced good. The productivity of time spent in household production increases with the presence of children. An individual's earnings potential depends on education, unobserved ability, and human capital, which accumulates when someone participates in the labor market and depreciates when they do not work.

The distinguishing feature of our model is its ability to capture the dynamic effects of tax reforms on intra-household specialization patterns. By incorporating the concept of limited commitment, a tax system can affect the allocation of time and resources of married couples by influencing the outside options of household members – defined as their lifetime utilities in the event of household dissolution. As these outside options evolve with earnings, accumulated human capital, and the presence of children, our model has welfare implications that, to our knowledge, have not been explored in previous research on optimal income taxation. Having purchased goods as an input in household production is also an essential feature of the model. In the data, secondary earners with a college degree work more than those with a high school diploma. Given assortative matching, our model can match this observation only if higher-educated individuals are able to substitute time with market inputs in household production. Without this adjustment, the model would overestimate intra-household specialization among high-education couples, especially in economies with joint taxation.

Since the model is estimated using U.S. data, we model the U.S. taxation system in detail by coding the year-on-year changes in tax rates. We also account for the institutional features that affect the optimal choice of tax rates, such as state taxes and the main components of the U.S. safety net, including the Earned Income Tax Credit (EITC), the Supplemental Nutrition Assistance Program (SNAP), and a variety of child-related transfers and credits.

Before using the estimated model for policy analysis, we validate it in several ways. We first document that it matches key patterns in the data, both qualitatively and quantitatively. Next, we show that the labor supply elasticities implied by the model compare well to those reported in the literature. Finally, we use simulated data to estimate the difference-in-differences effect of the "Bush Tax Cuts" and obtain results similar to those derived from CPS and PSID data. Importantly, the labor supply elasticities and reduced-form estimates are not directly targeted in the estimation process, lending credibility to our counterfactual simulations.

In the paper's final part, we rely on the estimated model to determine the tax system that maximizes total welfare and the corresponding tax rates. We assess four alternatives: (i) joint taxation in the form of income splitting (similar to the U.S. system), where couples are taxed based on their pooled income and their brackets are set at twice the ones for singles;[1] (ii) individual taxation, where tax rates are based solely on each individual's income; (iii) a general joint system with a flexible relationship between a spouse's tax rates and the income of the other; and (iv) an income splitting system augmented with a flat deduction for secondary earners.

We find that the tax rates currently used in the U.S. are close to the ones that maximize welfare under income-splitting. The optimal rates are moderately more progressive and people on average are willing to pay 0.9% of their income to implement them. The system's main weakness is its limited capacity to redistribute resources from the top to the bottom of the earnings distribution, as increasing tax rates at the top substantially reduces secondary earners' labor supply. By contrast, individual taxation allows for greater progressivity because secondary earners, who have the highest labor supply elasticities, are less common in the top brackets. It also leads to the highest accumulation of human capital among women and a larger increase in aggregate welfare, with individuals willing to pay on average 1.25% of their income. Nonetheless, these aggregate gains mask significant distributional disparities. While the benefits of transitioning to the optimal income-splitting system are relatively uniform, individual taxation results in substantial welfare gains for men with a college degree, while low-ability women experience welfare losses regardless of their education level.

The general joint system attains the highest level of welfare, with individuals willing to pay on average 2.59% of their income for its implementation. Its flexibility allows regulators to effectively redistribute resources by adjusting tax rates and providing substantial marriage bonuses to couples with lower earnings potential, thereby mitigating inequality. Its main shortcoming is its complexity. The previous literature has argued that general systems often lead to optimal tax rates that are challenging to implement (Piketty and Saez (2013)) and policy makers have advocated for a simplification of the tax schedule.

[1] Income splitting approximates well the U.S. system except for the top bracket.

In light of this, we searched for an alternative tax system that could achieve similar levels of redistribution as individual taxation, while avoiding its negative impacts on low-ability women and the complexity of the general joint system. We found that by simply adding a flat deduction for secondary earners to the U.S. income-splitting system, we can attain all three objectives.[2] Accordingly, individuals are willing to pay 1.3% of their income to transition to this system – less than for the general joint system but more than for individual taxation.

By examining how each system affects household behavior along different margins, we confirm the insight that incorporating a home-produced good, which contributes to both spouses' utility in married households, is essential. Under individual taxation, the production of this good declines, negatively affecting welfare. Ignoring household production would therefore overstate the welfare benefits of individual taxation. Additionally, it is important to account for the substitutability between time and market inputs in household production. Our policy evaluations reveal that transitioning to either the general joint system or an income-splitting system with secondary-earner deductions leads to significant substitution of market inputs for time, enabling families to achieve a similar or even higher level of public consumption. Without considering market inputs, we would have underestimated the welfare impacts of these tax systems.

Human capital accumulation is equally important. Our findings indicate that adding a deduction to the income-splitting system can significantly bolster women's labor force participation and subsequent human capital accumulation, without affecting men's human capital growth. This, in turn, leads to increased earnings that directly benefit welfare and also indirectly benefit low-income families through increased tax revenues. The impact of individual taxation on women's human capital is even more pronounced, though it has a contrasting effect on men, thereby tempering the positive welfare implications of such reform.

We directly assess the importance of intra-household specialization and dynamics with a last counterfactual. We focus on the tax system that maximizes welfare among those that are easily implementable: income splitting with a secondary-earner deduction. For this system, we recompute the optimal tax rates and deduction after adjusting a model parameter to reduce divorce probabilities. As expected, in the new environment, married couples choose to increase intra-household specialization because secondary earners are less likely to suffer welfare losses due to diminished human capital in the event of a divorce. Accordingly, the welfare-maximizing tax schedule becomes more progressive. With fewer couples having dual high incomes, it becomes optimal for regulators to set higher tax rates

[2] A policy broadly of this type was proposed by Kearney and Turner (2013) in the discussion paper 'Giving Secondary Earners a Tax Break: A Proposal to Help Low- and Middle-Income Families'.

for medium and high incomes and to slightly increase the deduction for secondary earners.

The paper proceeds as follows. Section 2 discusses our contributions to the existing literature. Section 3 describes the main differences between the most common taxation systems and provides evidence on responses to past U.S. tax reforms. In Section 4, we develop the intertemporal model of household decisions. Section 5 provides heuristic arguments for model identification and describes the moments used in the model estimation. In Section 6, we describe the estimation results and model fit. Section 7 assesses different forms of taxation and presents the key results of the paper. Section 8 concludes.

2 Related Literature and Contribution

We primarily contribute to the literature on optimal income taxation of families. One strand of this literature uses parsimonious static models to study how family considerations affect the optimal design of tax policy. Apps and Rees (1988) argue that the optimal taxation problem is complicated by the possibility that the allocation of resources within families may differ from what a utilitarian central planner would choose (see Alves et al. (2023) for more recent work on this issue). Apps and Rees (1999a, 1999b) also point out the importance of accounting for household production and within-household specialization patterns when discussing the optimality of different family taxation systems. Kleven, Kreiner, and Saez (2009) study optimal nonlinear family taxation as a multidimensional screening extension of the static, classic Mirrleesian framework (Mirrlees (1971)). They argue that *negative jointness* – marginal tax rates that are decreasing in the spouse's earnings – is optimal when a secondary earner's decision to work provides a signal to the central planner that the corresponding household has higher utility than families where the secondary earner does not work. Frankel (2014) shows that the optimality of negative jointness is sensitive to assumptions about assortative mating. Golosov and Krasikov (2023) generalize the static framework by relaxing assumptions on income distributions, sorting, and also include singles in the optimal taxation problem.[3]

To derive closed-form analytical results, each of these insightful papers focuses on particular aspects of the optimal taxation problem, abstracting from other aspects and dynamic channels such as human capital accumulation. Our core contribution relative to them is to develop and estimate a quantitative model that incorporates all of the theoretical considerations listed above while at the same time providing a close fit to the data in terms of the distribution of income (both cross-sectionally and

[3] Alesina, Ichino, and Karabarbounis (2011) study the optimality of gender-based taxation. In our paper, taxes may depend only on income levels of the two spouses and not on their identity.

over the lifecycle), labor supply patterns of single and married individuals, specialization patterns within and between households, and the degree of assortative mating. As a secondary contribution, we evaluate the income-splitting tax system with a flat deduction for secondary earners, showing it has desirable theoretical properties and is operationally simple (Aghion et al. (2017)).

The three most closely related papers to ours also use rich quantitative models to evaluate alternative family income taxation systems. Guner, Kaygusuz, and Ventura (2012) and Borella, De Nardi, and Yang (2022) build dynamic models where individuals derive utility from consumption and leisure, and marriage and divorce are exogenous processes. Gayle and Shephard (2019) study the optimality of the income splitting, individual, and general joint systems using a static collective model that includes household production and endogenous marriage responses to the tax system.

Our framework combines elements of these excellent papers—dynamics on one side and endogenous marriage and household production on the other. It also incorporates new elements that the data suggest are critical for optimal taxation. We combine a dynamic model with endogenous human capital accumulation as in Guner, Kaygusuz, and Ventura (2012) and Borella, De Nardi, and Yang (2022) with a limited commitment framework that allows the tax reforms to affect household decisions through changes in individual outside options and probability of divorce. Our paper is similar to Gayle and Shephard (2019) in that we allow marriage decisions to respond to tax incentives, an important mechanism given the empirical results of Dickert-Conlin (1999) and Moffitt, Phelan, and Winkler (2020). The main difference relative to Gayle and Shephard (2019) is the presence of dynamics with limited commitment, which allows us to consider novel implications of family taxation for intra-household welfare inequality. We also allow for assortative matching on unobserved ability – a feature that Frankel (2014)'s results suggest is important to evaluate the general joint system – and include market inputs in the household production function.

Finally, our descriptive results contribute to the large literature documenting labor supply responses to tax incentives (see Keane (2011) for a review). Within that literature, it is closest to Bick and Fuchs-Schündeln (2017) and Bick et al. (2019), who document that the large differences across countries and over time in married women's labor supply are systematically related to family taxation systems. We provide evidence for the U.S. using the 2003 Bush Tax Cuts.

3 The Labor Supply Incentives of Different Taxation Systems

In this section, we illustrate the distinguishing features of the two most prevalent tax systems: individual taxation and income-splitting. We demonstrate that the labor supply incentives for the secondary earner are very different under alternative taxation systems, show that those incentives correlate with observed employment behavior, and document that secondary earners are more likely to work following a reform that reduces their tax rate. The discussion in this section is important for understanding the effects of the tax reforms we consider in the last part of the paper, while the reduced-form estimates will be used to validate the magnitude of the employment responses in our structural model.

3.1 Taxing Married Households: A Simple Example

Under joint taxation, the tax liability of married individuals is a function of both their own as well as their spouse's earnings. The most common form of joint taxation is income splitting. Under this system, couples face a tax schedule on their pooled earnings that is identical to that of single individuals, but with brackets that are double the size.[4] By contrast, under individual taxation, all workers are taxed as individuals, regardless of their marital status.

To better understand the differences between these two systems, consider the following hypothetical progressive tax schedule under income splitting. Single individuals pay no taxes on the first $10,000, 15% on subsequent income up to $40,000, and 40% on remaining income. For married households these brackets double: they pay no taxes on their first $20,000 of joint income, 15% on subsequent income up to $80,000, and 40% on the remaining joint earnings. This system has a fairly similar (albeit simplified) structure as the one in the U.S. (see Appendix A).

Table 1 reports tax rates and take-home pay for a hypothetical couple with $80,000 and $30,000 in pre-tax income. Comparing the first three and last three columns illustrates differences in how the married couple would be taxed under an individual vs. joint system and how the couple's taxation under a joint system changes before vs. after marriage.

Marginal tax rates. In this example, the primary earner has a marginal tax rate (MTR) of 40% under both systems. By contrast, the secondary earner's MTR is 15% under the individual system and 40% under income splitting. High MTRs for secondary earners are a well-understood consequence of

[4] Married couples are thus taxed as if they were two single individuals, each earning half of the total household income. Germany uses this system. The U.S. uses separate tax schedules for married couples and individuals, but these schedules roughly approximate an income splitting system, except in the highest brackets. Top brackets for married U.S. couples generally start at income levels lower than double the single brackets, thus generating even higher marginal and average tax rates for married secondary earners. For additional details, see Appendix A.

Table 1: Effects on Outcomes: Joint vs. Individual Taxation

	Taxed As Individuals			Taxed Jointly As Married Couple		
	Pooled	Primary Earner	Secondary Earner	Pooled	Primary Earner	Secondary Earner
Pre-Tax Income	\$110k	\$80k	\$30k	\$110k	\$80k	\$30k
Marginal Tax Rate	-	0.40	0.15	0.40	0.40	0.40
After-Tax Income	\$86.5k	\$59.5k	\$27k	\$89k	\$71k†	\$18k†
Marriage Bonus	\$0			\$2.5k		
Average Tax Rate	0.21	0.26	0.10	0.19	0.11†	0.40†

† Calculation supposes that primary earner always works, while secondary earner supplements primary earner income.

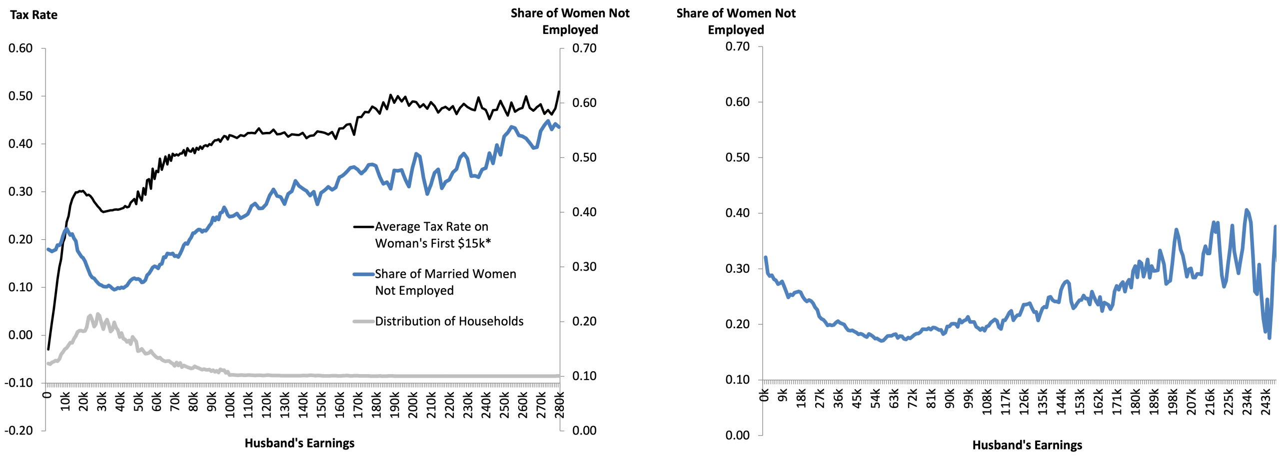
income splitting and a possible reason for low married women's labor supply in countries with highly progressive schedules and joint taxation (Bick and Fuchs-Schündeln (2017)). The higher MTRs also translate into significantly higher average tax rates on secondary-earner's income, which influence discrete choices such as labor force participation decisions.

Marriage bonuses. Table 1 illustrates a second difference between the two systems: married couples under income splitting typically experience an increase in take-home income after marriage. The married couple in the example brings home a total of \$86,500 if taxed as two single individuals but has \$89,000 in after-tax income if taxed jointly. This \$2,500 "marriage bonus" arises from a greater share of the couple's pre-tax earnings being taxed at the lower 15% rate under income splitting. Single-earner couples receive even larger bonuses. If the secondary earner in this example had no earnings, the couple's take-home income would increase from \$59,500 under individual (the primary earner's after-tax income) to \$71,000 under joint – equivalent to a bonus of \$10,500. The example illustrates that under income splitting, marriage bonuses are largest for couples with relatively unequal earnings, creating another potential disincentive for secondary earners' employment in the form of an income effect. This feature of the income-splitting system has also distributional consequences, as marriage bonuses imply transfers from single to married households.

3.2 U.S. Taxation: Descriptive Evidence

Taxation in the U.S. is more complex than in our example, but many of its features are well-captured by the previous discussion. First, U.S. secondary earners face high tax rates. To illustrate this, Figure 1a uses gender as a proxy for primary and secondary earner status and plots two variables as a function of the husband's earnings. The first is the predicted average tax rate (ATR) on a married woman's first \$15,000 of income, which exceeds 40% for a considerable fraction of households. This variable is

Figure 1: Women's Non-employment and Tax Rates, Comparison Between US and Canada



(a) Women's Non-Employment, US

(b) Women's Non-Employment, Canada

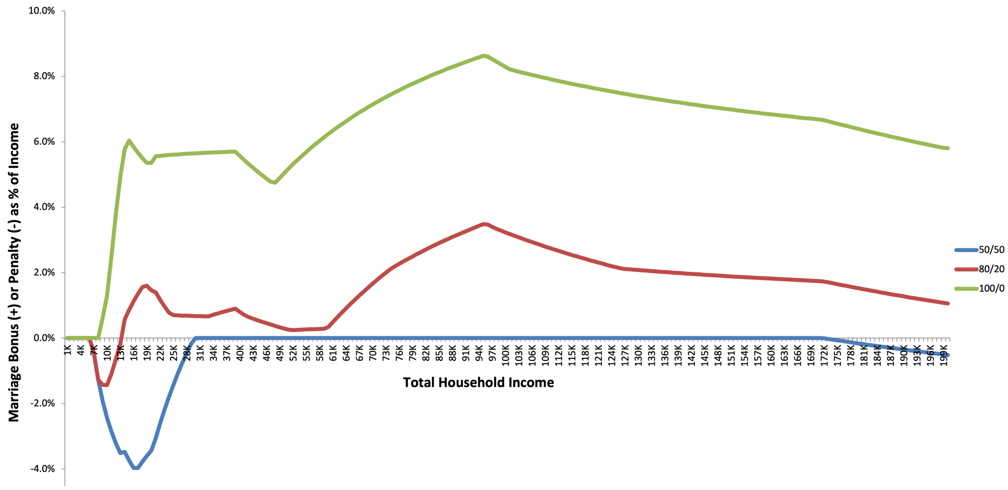
Notes: In the computation of the average tax rate, we include federal income taxes, FICA, state taxes weighted by population, the EITC and the Child Tax Credit. The EITC helps explain the steep increase in the average tax rate at very low levels of income and the subsequent decline for households with income between $\$20,000$ and $\$40,000$.

relevant for the labor force participation decisions of secondary earners in families where the primary earner is strongly attached to the labor force, and secondary earners adjust labor supply based on wages and shocks to the household. The second variable is the share of married women who are not employed. There is a striking positive relationship between these two variables. As the tax rate increases, due to placement along the income tax and EITC schedules, the fraction of married women who are not employed increases. As the ATR decreases, non-employment falls.

For comparison, Figure 1b plots the same relationship in Canada, which employs an individual tax system. We omit graphing the ATR on the first $\$15,000$ since it is similar for all secondary earners under individual taxation. Two features of married women's employment stand out. First, their non-employment varies much less with husband's earnings, ranging from about 20% to 30% (compared to 20% to 55% in the U.S.). Second, there is no distinctive relationship between Canadian women's employment and the husband's earnings. This evidence supports the notion that the U.S. tax system creates substantial disincentives for secondary earners to work, due to the significantly higher tax rates.

Marriage bonuses are also significant in the U.S., accounting for up to 9% of a household's take-home income. Figure 2 plots them as a percent of after-tax income for three types of couples: single-earner, equal-earner, and couples in which one spouse earns 80% of income. Nearly all fully specialized married couples and most partially specialized ones benefit from a marriage bonus, providing further

Figure 2: U.S. Tax Schedule: Marriage Bonuses and Penalties, 2015



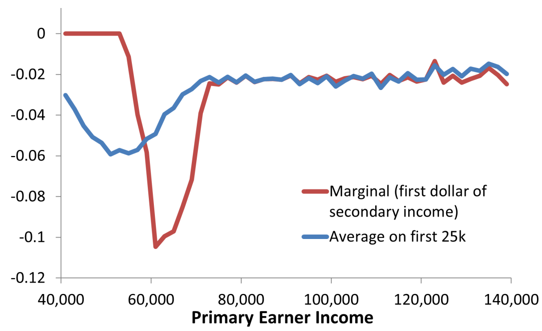
incentive for intra-household specialization. Such bonuses also make marriages between individuals of different earnings potentials more attractive. In the additional figures displayed in Appendix B, we proxy secondary-earner status with gender and document that many married women have low earnings potential, particularly in couples where the husband has low earnings or where only the husband works. Given the significant impact of low-income households on overall welfare calculations, this evidence highlights the importance of accounting for both earning potential and marriage bonuses when assessing the aggregate welfare impact of tax reforms.

3.3 Secondary Earner Responses to Tax Incentives: the 2003 "Bush Tax Cuts"

One limitation of cross-country or cross-sectional comparisons of secondary earner's employment choices is that they may be confounded by unobserved factors. In this section, we provide new evidence that a secondary earner's decision to work responds to incentives embedded in a taxation system by exploiting variation induced by a tax reform. We study the Jobs and Growth Tax Relief Reconciliation Act of 2003, commonly known as the 2003 "Bush Tax Cuts". The reform resulted in large reductions in tax rates for a fairly narrow group of households, which we can exploit in a difference-in-differences analysis. Later in the paper, we estimate similar regression designs using simulated data to validate our structural model.

We start by describing the changes in the secondary earner's tax rates introduced by the 2003 Bush Tax Cuts as a function of the primary earner's income, using the CPS for the tax years 2001-2005. We define the primary earner as the spouse in a married couple who has the highest earned income,

Figure 3: 2003 Bush Tax Cuts, average change in tax rates



adding up wages, business and farm income. To focus on couples in which the role of primary earner is more likely to be stable around the reform, we restrict the sample to households with primary earners that have a strong attachment to the workforce: they work full-time and earn at least $40,000 in 2000 dollars. We also restrict primary earner's income to be lower than $140,000 so that the control group is not too different from the treatment group.

Using this sample, we compute the secondary earners' tax rates using the TAXSIM software based on primary earner's incomes and number of children (Feenberg and Coutts (1993)).[5] We then estimate the following regression:

$$y_{it} = \sum_{k=1}^{50} [\alpha_k inc_k + \beta_k inc_k \cdot Post_t] + \varepsilon_{it}, \tag{1}$$

where y_{it} is the relevant tax rate, inc_k indicates whether the primary earner's income is in one of 50 equally-spaced bins between $40,000 and $140,000, and $Post_t$ is an indicator for tax year 2003, when the Bush Tax Cuts went into effect, or later.

The red curve in Figure 3 shows the estimated β_k coefficients when y_{it} is the secondary earner's marginal tax rate assuming they earn zero income. The general effect of the reform was to reduce the marginal tax rates of all secondary earners in households with medium and high income. But the decline was more dramatic for a subset of couples with medium income, due to an increase in the lower bound of one of the tax brackets that had the effects of moving them to a lower tax bracket. As the extensive margin of labor supply decisions is more likely to respond to average tax rates (Piketty and Saez (2013)), in Figure 3 we also report the estimated β_k coefficients when y_{it} is the secondary earner's average tax rate on the first $25,000. The reform had similar, albeit smoother, effects on the average tax rates.

[5] TAXSIM accounts for standard deductions and child tax benefits.

Table 2: Effects of 2003 Bush Tax Cuts on secondary earner's participation

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS with CPS Data		OLS with PSID Data		2SLS with PSID Data	
$\tilde{\tau}_{i,t}$	-0.651***	-0.314**	-0.450	-0.304	-1.020	-0.741
	(0.138)	(0.144)	(0.405)	(0.428)	(1.610)	(1.608)
$\tilde{\tau}_{i,t} \times YoungChild_{i,t}$		-0.534***		-0.706		-1.634*
		(0.0762)		(0.626)		(0.916)
N	93115	93115	4402	4402	4402	4402
Kleibergen-Paap F-stat					25.89	13.63

CPS regressions are based on repeated cross-sections corresponding to tax years 2001 through 2005. Standard errors are robust. The dependent variable is a dummy for whether the secondary earner (that is, the spouse with the lowest reported earnings) worked full or part-time in the past year. The main regressor $\tilde{\tau}_{i,t}$ is the tax rate on the first \$25k of secondary earner income, taking as given the primary earner's income. **PSID regressions** are based on a panel of households with observations in years 2000, 2002, 2004, and 2006. Standard errors are clustered at the household level. The dependent variable is a dummy for whether the wife was working or searching for work as of the interview. The main regressor $\tilde{\tau}_{i,t}$ is an interaction of an indicator for the years 2004 or 2006 and the change in the tax rate on the first \$25k of wife income, taking as given the husband's income as of 2004. The 2SLS strategy uses as instruments a quartic polynomial of husband's predicted income in 2004, using 2002 characteristics as predictors. See Appendix C for details. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We now analyze the effect of these tax changes on secondary earners' labor force participation, by estimation the following regression:

$$SecondaryEarnerWorks_{it} = \delta_t + \sum_{k=1}^{20} \alpha_k inc_k + \beta\tilde{\tau}_{i,t} + \gamma X_{i,t} + \varepsilon_{it}, \tag{2}$$

where δ_t are time fixed effects, inc_k now denotes slightly wider primary earner's income groups (width \$5,000 instead of \$2,000), $\tilde{\tau}_{i,t}$ is the average tax rate on the first \$25,000 of secondary earner's income, and $X_{i,t}$ is an extensive set of controls fully described in Appendix C. It includes indicators for number of children and number of children younger than six, fine education groups for primary and secondary earners, and a time trend interacted with an indicator for secondary earners' college degree.

The results are reported in Column (1) of Table 2. The estimated coefficient on the average tax rate ($\hat{\beta} = -0.651$) implies that a reduction of 6% in the average tax rate increases labor force participation of the secondary earner by about four percentage points. We can rule out responses smaller than two percentage points at the usual 95% confidence level. In Column (2), we explore whether the responses differ depending on the presence of a child younger than six in the household. We estimate a similar specification, except that it includes an interaction term. The results indicate that the labor supply responses of secondary earners are substantially larger in households with young children. In Appendix Table A.3, we document that these results are robust to several changes in specification.[6]

[6]We find similar qualitative results if the dependent variable is the average tax rate on the first \$15,000 of secondary earner's income instead of \$25,000. In that case, the magnitude of the main effect is significantly lower, consistent with the hypothesis that the measure based on \$25,000 is a better approximation of the tax incentives that matter for the

One may worry that unobserved household-level factors that determine secondary earner participation correlate with the changes in tax rates induced by the reform. One indication that this may not be a significant issue in our context is the observation that our estimates are essentially unchanged regardless of whether one uses a minimal set of controls or the full array (see columns (1) and (2) in Appendix Table A.3). But to further inspect this possibility, we develop an alternative design and estimate it using the PSID. Here, we provide an overview of the strategy and summarize the results. A full description is provided in Appendix C.

As the PSID follows the same families over time, we can evaluate the effect of the Bush Tax Cuts on secondary earners' labor force participation using changes instead of levels. Specifically, we construct our treatment variable, $\tilde{\tau}_{i,t}$, as follow. We first use the primary earner's income in 2004 to calculate the tax rate on the first \$25,000 earned by the secondary earner in that year. We then calculate a second tax rate based on the same numbers, but using tax brackets and rates for the tax year 2002. The treatment variable $\tilde{\tau}_{i,t}$ is computed as the difference between those two tax rates interacted with an indicator for a tax year after the reform. It corresponds to the change in tax incentives induced by the reform, using as the reference point the husband's income in 2004. We then estimate the following OLS regression:

$$SecondaryEarnerWorks_{it} = \alpha_i + \delta_t + \beta\tilde{\tau}_{i,t} + \gamma X_{i,t} + \varepsilon_{it}, \tag{3}$$

where α_i is a household fixed effect, δ_t is a time fixed effect, and $X_{i,t}$ is a set of control variables described in Appendix C. To account for the possibility of that the husband's income responds to the tax reform, which may introduce misspecification biases, we also estimate a two-stage least squares (2SLS) version of equation (3), where the change in the tax rate is instrumented by a polynomial in the primary earner's predicted income in 2004. The prediction is constructed based on pre-reform data.

Columns (3) through (6) in Table 2 report the results. They are qualitatively similar to the ones obtained using the CPS, and close enough quantitatively that it is not possible to reject the hypothesis that the CPS and PSID results are identical. As with the CPS, we also verify that the PSID estimates are robust to specification choices.[7] We conclude that the data supports the hypothesis that the 2003 Bush Tax Cuts led to more significant increases in labor force participation among secondary earners who were more affected by the reform, with these effects being particularly pronounced in households with young children.[8]

extensive margin choices in our context.

[7]Because of the smaller sample size in the PSID, they vary more between specifications than with CPS data.

[8]The variation exploited in our reduced-form specifications can only be measured for existing married couples, pre-

The descriptive results presented in this section provide a solid starting point for understanding the implications of choosing different taxation systems. They suggest that shifting from income splitting to individual taxation should boost labor force participation of secondary earners, especially those with young children. But there are several other policy-relevant questions that are more difficult to address using reduced-form empirical designs. For instance, would this shift to individual taxation enhance welfare after accounting for possible effects on the provision of public goods within the household? Would the benefits be evenly distributed across different levels of earnings potential, or would low earnings-potential workers experience welfare losses under such reform? Are there alternative tax systems that could deliver greater and more equitable welfare benefits? What impacts would such reforms have on primary earners and individuals who are not married? Would the effects vary by age? To answer these questions, we develop and estimate a model of individual decisions and household formation and dissolution that accounts for the main features of the data.

4 Model

To evaluate the long-term effects of tax reforms, a model must accurately reflect the data patterns described in previous sections and possess sufficient flexibility to account for a range of behavioral responses that may arise from transitioning to a new tax system. To achieve this, we develop a model with the following key features.

The first important characteristic is its dynamic nature, which is captured by endogenous variables such as human capital, borrowing and savings, and marriage and divorce status, which are optimally decided under limited commitment. Second, the model includes a household-produced good that constitutes public consumption for couples. As a third distinguishing feature, our model incorporates purchased goods as inputs in the household production function, alongside the usual time inputs. Finally, an individual's earning potential depends on both education and unobserved ability.

4.1 Preferences and Technology

Preferences. Individual i has preferences over private consumption c^i, leisure l^i, and a home-produced good Q. Individuals also receive utility from marriage quality θ, normalized to zero for

venting us from studying marriage decisions using that approach. In principle, one could investigate the impact of the reform on divorce with the PSID. Unfortunately, the sample is not large enough to detect such effects precisely.

singles. Preferences are characterized by the following utility function:

$$u\left(c^i, l^i, Q, \theta | \kappa, s\right) = \frac{\left(c^i\right)^{1-\sigma_c}}{1-\sigma_c} + \gamma_l \frac{\left(l^i\right)^{1-\sigma_l}}{1-\sigma_l} + \gamma_Q^{\kappa,r} \log Q + \theta.$$

The relative taste for the home-produced good, $\gamma_Q^{\kappa,r}$, varies across households depending on the variable κ which takes values 1, 2, or 3 based on whether the couple has, respectively, (i) no children, (ii) children < 6, or (iii) only children 6 or older. It also depends on marriage status r = single (s) or married (ma).[9]

Household Production Function. The good Q – which is a public good for married couples and captures among other things the quality of children, condition of the house, and meal quality – is produced using the time spent by each household member on household production, d^i, and purchased market goods, mg.

The household production function has three main features. First, to account for large differences in household production hours by age of children and parental education (Guryan, Hurst, and Kearney 2008), the productivity of time invested in Q depends on number of children, whether the youngest child is less than 6, and whether the parents have a college degree. Second, the time productivity can differ between women and men, to account for differences in household hours, even conditional on their market hours and wages. Third, the time and market goods invested in the production of Q are possible substitutes to account for the following two empirical observations: (i) women with a college degree have on average spouses with higher earnings due to assortative matching, and (ii) at the same time women with a college degree have higher employment rates (0.79) than women with a high school degree (0.64). The introduction of some degree of substitutability between time and market inputs breaks the link between sorting and the time allocation to the home-produced good and allows our model to match simultaneously both facts.[10]

For couples, these three characteristics are incorporated in the following constant elasticity of substitution (CES) production function:

$$Q = f\left(d^1, d^2, mg, \kappa, g\right) = \left(\eta^{\kappa,g,m} \cdot \left(\alpha^1 \cdot (d^1)^{\upsilon} + \alpha^2 \cdot (d^2)^{\upsilon}\right)^{\frac{\tau}{\upsilon}} + (1-\eta^{\kappa,g,m}) \cdot mg^{\tau}\right)^{\frac{1}{\tau}}.$$

[9] In the estimation of the model and policy evaluations, we replace $\log Q$ with $\log(Q+K)$ to avoid numerical issues when Q is small. K is set equal to 0.5.

[10] For a discussion of the effect of education on the choice of market inputs and time devoted to household and market production, see Flood et al. (2022).

For married individuals, the relative importance of home hours, $\eta^{\kappa,g,m}$, depends on whether the couple has no children, children < 6, or only children 6 or older (κ), and on the education g of the head of the household. Thus, $\eta^{\kappa,g,m}$ takes six possible values depending on the combinations of g and κ. The parameters α^1 and α^2 capture the productivity of women's and men's time inputs, respectively, which are denoted by d^1 and d^2. With υ we measure the substitutability between them. Lastly, τ determines the degree of substitutability between time and market goods. For singles, the production function becomes:

$$Q = f(d, mg, \kappa, g) = \left(\eta^{\kappa,g,s} \cdot (\alpha^j d)^\tau + (1 - \eta^{\kappa,g,s}) \cdot mg^\tau\right)^{\frac{1}{\tau}},$$

where j indexes gender. The α^j parameters are the same ones used in the production function for couples.

The home-produced good is an essential feature of our model. Evidence from previous sections indicates that individual taxation incentivizes secondary earners to increase their labor supply, possibly at the expense of time devoted to household production. Considering the significant amount of time allocated to home production – on average 25 and 12 hours per week by women and men – policy evaluations that ignore this trade-off may overestimate the welfare effects of individual taxation.

Savings and Labor Supply. Households can save using a risk-free asset b with gross return R. They can also borrow at the same rate. Labor supply choices are discrete (no time, part-time, full-time, and over-time), and the sum of the time devoted by an individual to market labor, household production, and leisure adds up to the available total time T.[11]

4.2 Education, Wages, and Taxes

Education. In the model, people differ depending on whether they enter the working stage of their life with or without a college education. Education has three effects. It endows graduates with better wage processes; it potentially affects the productivity of the time invested in the home-good Q, as described earlier; and it increases the probability of matching with other college graduates in the marriage market.

[11] The model includes over-time because in the data a significant fraction of male workers supply more then 40 hours a week. It is therefore important to allow for this labor supply option when assessing tax reforms. Over-time is set equal to 45 hours a week.

Wage Process. Wages depend on gender j, education g, the stock of human capital e_{it}, a permanent unobserved component a_i, and an idiosyncratic shock $\varepsilon_{i,t}$:

$$\ln w_{it} = a_i + \beta_0^{g,j} + \beta_1^{g,j} e_{it} + \beta_2^{g,j} e_{it}^2 + \varepsilon_{it}, \tag{4}$$

where $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^{g,j})$, g is binary (less than college or college), and $a_i \sim N(0, \sigma_a^{g,j})$.

Incorporating a permanent component a_i in the wage process, which we interpret as unobserved ability, is important for three reasons. First, it helps us capture the extensive heterogeneity in earnings potential that remains even after controlling for education. This allows us to more accurately fit the full distribution of earnings in the economy, which is crucial for evaluating tax reforms that lead to significant redistribution of resources across different income levels. Second, it improves our ability to match sorting patterns in the economy. Even after controlling for education, there is considerable assortative matching based on income, which can be influenced by tax reforms. Third, recent work by Heathcote and Tsujiyama (2025) argues that using a continuous distribution of worker types (or a fine grid of types) is important in practical optimal taxation applications, because using coarse grids implies unrealistic incentive compatibility constraints for the planner.

The stock of human capital, e_{it}, evolves endogenously over time. It increases by one unit following a period of full or over-time work; part-time work increases human capital by λ units; e_{it} declines by δ for each year out of the labor force. Thus, the stock of human capital in a given year in equation (4) is given by:

$$e^i = e^i_{FT} + \lambda e^i_{PT} - \delta e^i_{NT}, \tag{5}$$

where e^i_{FT} denotes cumulative years of experience in full and over-time jobs, e^i_{PT} is the corresponding variable for part-time jobs, and e^i_{NT} are the number of years spent outside the labor force.

Human capital accumulation and depreciation and their heterogeneous effects on wages based on gender and education are key features of our model. The labor literature has estimated that wages increase significantly with experience (e.g. Dustmann and Meghir (2005)), and the corresponding rate of return differs considerably by gender and education. Studies have found that women have flatter wage profiles (Goldin (2014)), face smaller penalties for part-time work (Goldin (2014) and Bronson (2015)), and individuals with a high school diploma experience lower rates of returns in part time jobs compared to those with a college degree (Blundell et al. (2016)). Additionally, workers who are absent from the labor market for extended periods, such as during mass layoffs, face significant wage losses (Bertheau et al. (2023)). Given that tax reforms impact labor supply decisions differently based on gender and education, it is essential for our model to account for this heterogeneity in the evolution of wages.

Taxes. The tax schedule depends on the tax system and the marital status of an individual. An individual not married with wage w_t and labor supply h_t pays taxes that depend on own earnings, $\tau^s(w_t h_t)$. For a married couple, the tax liability is a function of both spouses' earnings, $\tau^m(w_t^1 h_t^1, w_t^2 h_t^2)$.

In our baseline model, tax schedules are set to the actual U.S. statutory tax and welfare schedules. We account for both major and minor tax reforms over time, and for year-on-year changes in tax brackets. We assume that individuals have perfect foresight regarding future tax reforms.

To account for interactions between the taxation and social-welfare systems, we allow the tax schedules $\tau^s(w_t h_t)$ and $\tau^m(w_t^1 h_t^1, w_t^2 h_t^2)$ to depend on (i) Social Security Income taxes (SSI), (ii) Medicare taxes, (iii) the Earned Income Tax Credit, (iv) the Child Tax Credit, (v) the Child and Dependent Care Credit, and (vi) the Supplemental Nutrition Assistance Program (SNAP) (food stamps). This is important for welfare analysis. Additionally, economists have argued that non-convexities created by government aid programs have substantial effects on individual decisions and, hence, on the impact of tax reforms (Keane (2011)). We also include state income taxes. To avoid having the state of residence as a state variable, which would make the estimation infeasible, we assign each household the average state tax rate calculated across all states.

4.3 Marriage, Fertility, and Child Care

The Marriage Market. We model the marriage market using a matching framework with search frictions. Each period, single individuals meet a potential spouse with a set of characteristics: education g_i, ability a_i, human capital e_i, savings s_i, and shared match quality θ.

An individual draws a partner with the same education with probability π. The permanent ability of the potential spouse is drawn from a Normal distribution with standard deviation determined by their gender and education. To account for assortative mating in earnings potential conditional on education, we allow for correlation between the individual's own ability and the ability of the spouse, with value $\rho_{g=g_s}$ if the spouse has the same education and $\rho_{g\neq g_s}$ if they have different educational levels. We assume an age gap of two years between male and female partners. Human capital and savings of potential partners are drawn from the observed empirical distributions for those variables for single individuals, conditional on their age.

For potential partners, marriage quality θ is initially drawn from a normal distribution with standard deviation σ_θ. The mean declines linearly with age, to account for the fact that marriage probabilities decrease. Once married, match quality follows a random walk: $\theta_t = \theta_{t-1} + z_t$, where z_t is drawn from a normal distribution with mean 0 and standard deviation σ_z.

Fertility and Child Care. Married women can give birth to a child between the ages of 20 and 35. The probability of a birth depends on education and current number of children. Children affect choices by increasing the value of time spent in household production. Additionally, children under 6 require childcare. A married couple has to purchase childcare hours for the minimum number of hours worked in the labor market by the mother or father. Similarly, a single mother must pay for childcare for the number of hours that she works. We denote by p^c the price per hour of child care. To reduce the size of the state space, children age probabilistically: if younger than 6, they become older with probability 1/6; when older, they leave the households with probability 1/12. If a couple with a child under the age of 6 decides to divorce, the mother retains custody of the child and the child becomes older than 6.

4.4 Timing and Decisions

Timing. Individuals enter adult life with or without a college education. Adult life has a working and a retirement stage. During the working stage, each person makes decisions about labor supply, time and goods used on household production, consumption, savings, marriage, and divorce. At the beginning of every period in the working stage, single individuals meet a potential spouse, learn their characteristics and realization of all shocks, and choose whether to marry. Married individuals learn the realization of the shocks and choose whether to stay married or divorce. At the end of the working life, people enter the retirement stage in which they only choose time and market inputs in household production, consumption, and savings.

Single Individual's Decisions. If individual i enters period t as single and decides to remain single, this person chooses labor supply, the time and goods used in household production, private consumption, and savings that solve a standard single-agent problem. Let V_t^{es} be the value function of an individual who enters period-t as single, V_t^{ds} the value function of a person who decides to stay single in the period, and $S_t^s = \{b_t, e_t, n_t, \kappa_t, a, \tau^s, \tau^m\}$ the period t vector of state variables for a single individual. The single-agent problem can then be written as follows:

$$V_t^{ds}\left(S_t^s\right) = \max_{h_t, d_t, mg_t, c_t, b_{t+1}} \quad u\left(c_t, l_t, Q_t\right) + \beta E_t\left[V_{t+1}^{es}\left(S_{t+1}^s\right)\right]$$
$$s.t. \qquad c_t + mg_t = w_t h_t - \tau^s(w_t h_t) + Rb_t - b_{t+1} - p_t^c h_t \mathbb{1}_{\{\kappa_t=2\}}$$
$$Q_t = f\left(d_t, mg_t, \kappa_t, g\right), \quad h_t + d_t + l_t = T, \quad \text{and} \quad w_t = w\left(e_t, a, \varepsilon_t\right),$$

where $\mathbb{1}_{\{\kappa_t=2\}}$ is an indicator equal to 1 if the household has children younger than 6 and the evolution of κ_t depends on the fertility probability and probabilistic aging discussed above.

Married Couple's Decisions. If two individuals enter period t married, they make efficient decisions with limited commitment. Efficient decisions means that they solve a Pareto problem. However, because they cannot commit to future allocations, the Pareto weights used to make decisions in period t may differ from the Pareto weights with which the two spouses entered the period.[12]

Denote with M_t^1 and M_t^2 the Pareto weights used to make efficient decisions in period t. We will discuss later in this section how they are computed. Also, let $V_t^{em,i}$ be individual i's value function if this person enters period t as married and $S_t^m = \left\{b_t, e_t^1, e_t^2, n_t, \kappa_t, \theta_t, a^1, a^2, \tau^s, \tau^m, M_t^1, M_t^2\right\}$ the period-t vector of state variables for a married individual. Then, the couple chooses labor supply, household production time and goods, private consumption, and savings as the solution to the following problem:

$$\max_{\{h_t^i, d_t^i, mg_t^i, c_t^i, b_{t+1}\}_{i=1}^2} \sum_{i=1}^{2} M_t^i \left\{u^i(c_t^i, l_t^i, Q_t, \theta_t) + \beta E_t\left[V_{t+1}^{em,i}\left(S_{t+1}^m\right)\right]\right\}$$
$$s.t. \qquad c_t^1 + c_t^2 + mg_t = w_t^1 h_t^1 + w_t^2 h_t^2 - \tau^m(w_t^1 h_t^1, w_t^2 h_t^2) + Rb_t - b_{t+1} - p_t^c \min\left(h_t^1, h_t^2\right) \mathbb{1}_{\{\kappa_t=2\}}$$
$$Q_t = f\left(d_t^1, d_t^2, mg_t, \kappa_t, g\right), \quad \theta_{t+1} = \theta_t + z_{t+1}, \quad h_t^i + d_t^i + l_t^i = T, \quad \text{and} \quad w_t^i = w^i\left(e_t^i, a^i, \varepsilon_t\right), \quad i = 1, 2.$$

Denote by $h_t^{i*}, d_t^{i*}, mg_t^*, c_t^{i*}, b_{t+1}^*$, for $i = 1, 2$, the solution of the couple's problem, and let $Q_t^* = f(d_t^{1*}, d_t^{2*}, mg_t^*, \kappa_t, g)$. If person i chooses to stay married, i's value function takes the following form:

$$V_t^{dm,i}\left(S_t^m\right) = u^i(c_t^{i*}, l_t^{i*}, Q_t^*, f_t^{i*}, \theta_t) + \beta E_t\left[V_t^{em,i}\left(S_{t+1}^m\right)\right].$$

Marriage, Divorce, and Renegotiation Decisions. Given the optimal choices and set of value functions above, we can determine the optimal marriage decision. Two individuals who enter period t

[12] The limited-commitment model has been tested and recently used to study policies affecting household choices and outcomes. See for examples Mazzocco (2007), Adams et al. (2014), Mazzocco, Ruiz, and Yamaguchi (2014), Voena (2015), Bronson (2015), Lise and Yamada (2018), Low et al. (2022), Theloudis et al. (2023), Kim (2023), and Reynoso (2024). For a survey of the literature see Chiappori and Mazzocco (2017).

as single and meet will choose to marry if the value of being married is larger than the value of staying single for both of them, i.e.

$$V_t^{dm,i}\left(S_t^m\right) \geq V_t^{ds,i}\left(S_t^s\right), \quad i=1,2.$$

The initial Pareto weights used in the computation of $V_t^{dm,i}\left(S_t^m\right)$ are determined by symmetric Nash bargaining, e.g. by equalizing across spouses the gains in expected lifetime utility relative to the outside option of remaining single. These inequalities are known as participation constraints (Kocherlakota (1996), Mazzocco (2007), Marcet and Marimon (2011)).

The participation constraints determine also whether it is optimal for a married couple to (i) remain married maintaining the current Pareto weights and allocation of resources, (ii) renegotiate the current allocation by changing the Pareto weights, or (iii) divorce. The model allows for an individual cost of divorce K_{div}^g, that varies with education. Thus, for married individuals, the participation constraints become,

$$V_t^{dm,i}\left(S_t^m\right) - K_{div}^g \geq V_t^{ds,i}\left(S_t^s\right), \quad i=1,2.$$

Consider two individuals who enter period t as married with Pareto weights M_{t-1}^1 and M_{t-1}^2. Denote by $z_t^{**} = \{h_t^{i**}, d_t^{i**}, mg_t^{i**}, c_t^{i**}, b_{t+1}^{**}\}$ the solution of the couple's problem computed using these Pareto weights. If at the solution z_t^{**} the participation constraints are satisfied for both spouses, it is optimal for them to remain married at the current optimal allocation of resources. If the participation constraints of both spouses are violated, the marriage produces no surplus that can be shared. It is therefore optimal to divorce. The most interesting case is represented by a situation in which the participation constraint of one spouse is satisfied, but the participation constraint of the other is violated. In this case, there may exist a different allocation of household resources at which both participation constraints are satisfied and, thus, at which both spouses are better off staying married. If such an allocation exists, it can be achieved by increasing the Pareto weight of the constrained individual and, consequently, the share of resources allocated to this spouse. From an ex-ante perspective, the most efficient new allocation is the one that corresponds to new Pareto weights M_t^1 and M_t^2 that make the constrained individual indifferent between staying married or being single Kocherlakota (1996). If such new allocation does not exist, the household does not generate marriage surplus and it is optimal for the spouses to divorce.

Limited commitment is the last key feature of our model. Changes in tax regimes affect the outside options of individuals, which in turn influence their decisions, including the choice to divorce. To capture this relationship and match the divorce rate in the data, our model must incorporate limited

Table 3: Parameters that are Calibrated or Estimated Using Regressions

Parameters	Description	Source or Method
σ_c	Consumption CRRA parameter	Calibrated to 1.5 (Blundell, Browning, and Meghir (1994))
β	Annual discount factor	Calibrated to 0.98 (Attanasio, Low, and Sánchez-Marcos (2008).)
$\beta_0^{g,i}$, β_1^g, β_2^g, σ_ϵ^g, σ_a^g, λ^g, δ^g	Wage process parameters, by gender and education	Estimated in the PSID, following Wooldridge (1995) and Semykina and Wooldridge (2010)
$\pi_0^g - \pi_6^g$	Probability of a birth in a given period, by education	Estimated in the PSID, using a linear probability model saturated with interactions between all discrete variables

commitment, because with full commitment changes to the outside options are inconsequential and only mutual consent divorces are possible.

5 Identification and Moment Selection

The model parameters are described in Tables 3 and 4. Table 3 summarizes the parameters that are either calibrated (two) or estimated directly using a regression model (wage and fertility processes). Table 4 provides heuristic arguments for the identification of all remaining parameters and reports the key moments used for their estimation. Table 5 describes those moments in detail.

We calibrate the discount factor β to 0.98, following Attanasio, Low, and Sánchez-Marcos (2008), and the curvature of the sub-utility for consumption σ_c to 1.5, based on Blundell, Browning, and Meghir (1994) and Attanasio and Weber (1995). Given the model assumptions, we can obtain consistent estimates of all parameters of the wage process using the estimator for panel data regressions with individual fixed effects and selection proposed in Wooldridge (1995) and Semykina and Wooldridge (2010). To estimate different returns to part-time work and possible penalties to non-participation, we incorporate equation (5), which characterizes the accumulation of human capital, into the wage process (4) and re-write it as follows:

$$\ln w^{g,j} = a^i + \beta_0^{g,j} + \beta_1^{g,j}\left(e^i_{FT} + \lambda e^i_{PT} + \delta e^i_{NT}\right) + \beta_2^{g,j}\left(e^i_{FT} + \lambda e^i_{PT} + \delta e^i_{NT}\right)^2 + \epsilon^{g,i},$$

or equivalently,

$$\ln w^{g,j} = a^i + \beta_0^{g,j} + \beta_1^{g,j} e^i_{FT} + \psi_1^{g,j} e^i_{PT} + \psi_2^{g,j} e^i_{NT} + \beta_2^{g,j}\left(e^i_{FT}\right)^2 + \psi_3^{g,j}\left(e^i_{PT}\right)^2 + \psi_4^{g,j}\left(e^i_{NT}\right)^2$$

Table 4: Parameters Estimated Using SMM

Parameter	Interpretation	Heuristic Identification Arguments	Key Moments (see Table 5)
I. Household Production Function			
τ	Substitutablity between time in home production (d_i) and purchased market inputs (m)	It is identified by the elasticity of the ratio of the two inputs with respect to changes in individual wages, for single individuals, which measures the willingness to substitute between them.	V-VII
υ	Substitutability between the primary and secondary earners' time (d_1 and d_2)	It is identified by the elasticity of the ratio of the two time inputs with respect to the relative wages of primary and secondary earners.	V–VI
$\eta^{g,\kappa,r}$	Time inputs' share in the production function, for households with education g, child status κ, and marriage status r	Given τ, a higher η increases the fraction of resources allocated to time inputs in home production, relative to market inputs. Thus, η can be identified for different groups by the amount of market inputs used in household production divided by the sum of market inputs and time inputs valued at the individuals' wage.	V-VII
α_1, α_2	Productivity of time spent in home production for women and men	We normalize one parameter ($\alpha_2 = 2 - \alpha_1$) and identify the other by the difference between single men and women in the ratio of time inputs to the sum of the value of market and time inputs.	V, VI
II. Preferences			
γ_l	Taste for leisure relative to the taste for private consumption.	It can be identified by the average ratio of private consumption to leisure.	IV, VIII, IX
σ_l	Power parameter on leisure. The Frisch leisure elasticity of single individuals is equal to $-\frac{1}{\gamma_l \sigma_l}$.	Given γ_l, it can be identified by changes in leisure of singles in response to expected or marginal changes in their wages, i.e. changes that keep the marginal utility of wealth constant.	I-III, VIII, IX
$\gamma_Q^{\kappa,r}$	Taste for public consumption relative to private consumption, by child and marriage status.	It can be identified by the average ratio of private consumption (or leisure, given γ_l) to the value of inputs used in household production, by child and marriage status.	V-X
III. Marriage, Divorce, and Meeting Parameters			
μ_θ, σ_θ	Mean and standard deviation of initial match quality draw.	We normalize σ_θ to 1. Then, as a higher μ_θ increases the probability that a couple marries, we identify μ_θ by marriage probabilities by education and age.	XI
σ_z, K^g_{div}	Standard deviation of match quality shocks and cost of divorce.	The divorce probability is an increasing (decreasing) function of σ_z (K^g_{div}). We normalize σ_z to 1, and identify K^g_{div} using divorce probabilities for different education-age groups.	XI
π	Probability that an individual with a given education draws a potential spouse with the same education.	Higher π values increase the probabilities of meeting and, thus, marrying. We therefore identify π using the shares of individuals with high school or college education that marry a person with the same education.	XII
ρ_{g,g_s}	Correlation between abilities of potential spouses who have education g and g_s.	Higher values increase the probability that an individual with high ability meets and, thus, marries a partner who also has high ability. The parameters can therefore be identified from the correlation in spousal wages, conditional on education and human capital.	XII

Table 5: Moments Used in Estimation

	Moment Description	Number
I.	Labor force participation by gender, education, marital status, and family status (no children, youngest child age 6+, youngest child under 6)	24
II.	Labor force participation by gender, education, and age group (20-25, 26-54, 56-64)	12
III.	Hours worked by gender, education, marital status, and family status (no children, youngest child age 6+, youngest child under 6)	24
IV.	Married women's labor force participation by income quantile of husband, education, and family status	30
V.	Hours in home production by gender, education, marital status, and family status	20
VI.	Value of hours in home production by gender, education, marital status, and family status	20
VII.	Expenditures on market good inputs on home production by gender, education, marital status, and family status	20
VIII.	Time spent on leisure, by gender, education, and marital and family status	20
IX.	Ratio between leisure and private consumption of married women without children, by husband's income quartile	8
X.	Ratio between private consumption and expenditure on market good inputs, by marital and family status	6
XI.	Share of women married and divorced by education and age group (22-26, 27-36, 37-50, above 50); share of women ever married by 36, by education; share of women ever divorced by 50, by education	12
XII.	Share of women and men with a high school or college education married to individual with the same education; correlation between permanent wage components, conditional on educations of husband and wife	8

$$+ \psi_5^{g,j} e_{FT}^i e_{PT}^i + \psi_6^{g,j} e_{FT}^i e_{NT}^i + \psi_7^{g,j} e_{PT}^i e_{NT}^i + \epsilon^{g,i}, \tag{6}$$

where $\psi_1^{g,j} = \beta_1^{g,j}\lambda$, $\psi_2^{g,j} = \beta_1^{g,j}\delta$, $\psi_3^{g,j} = \beta_2^{g,j}\lambda^2$, $\psi_4^{g,j} = \beta_2^{g,j}\delta^2$, $\psi_5^{g,j} = 2\beta_2^{g,j}\lambda$, $\psi_6^{g,j} = 2\beta_2^{g,j}\delta$, and $\psi_7^{g,j} = 2\beta_2^{g,j}\lambda\delta$. We observe time spent in full and over-time jobs, years in part-time work, as well as absences from the labor force. All parameters in equation (6) are therefore identified. The parameters governing the accumulation of human capital can then be identified by noting that $\lambda = \frac{\psi_2}{\psi_1}$ and $\delta = \frac{\psi_3}{\psi_1}$.

We employ a two-stage procedure that controls for selection into working. The selection equation includes non-labor income, indicators for whether someone is married, never married, widowed, separated, or divorced, an indicator for whether the household includes a child younger than 6, indicators for number of children, year-fixed effects, and the expected average tax rate on own income, calculated based on the mean income of each gender-education group. For married women, we also include the marginal tax rate based on the husband's income. The wage process parameters are estimated separately for each gender-education group using PSID data.

In our model, birth probabilities are functions of discrete variables. Thus, as births and the discrete variables are observed, only one conditional probability function for births is consistent with the data, and the fertility process parameters are identified. We estimate the probability of a birth for a married woman conditional on education and current number of children non-parametrically using PSID data.

The remaining 27 parameters are estimated using the Simulated Method of Moments (SMM). They are the parameters that characterize preferences, the household production function, and the marriage market. To keep our exposition clear and concise, we present the heuristic arguments describing the variation in the data that enables us to identify each of these 27 parameters in Table 4. The last column of the table references the moments most closely tied to each parameter, which are listed in full in the subsequent Table 5.[13]

6 Results

We rely on four main datasets for estimation. The Panel Studies of Income Dynamic (PSID) (1980-2010) and the Current Population Survey (CPS) (1980-2016) provide information about labor force participation, hours of work, earnings, income, education, marital status, and partner characteristics. We use the American Time Use Survey (ATUS) (2000-2011) to compute the time invested in the production of the home-produced good and private leisure, and the Consumer Expenditure Survey (CEX) (1980-2010) to calculate private consumption and the financial resources invested in the production of the home-produced good.

6.1 Parameter Estimates

The parameter estimates are reported in Tables 6, 8, and 9. We will briefly discuss the estimates most relevant for the policy evaluations performed in later sections.

We estimate the preference parameters for leisure σ_l and γ_l to be 2.85 and 1.97, respectively. These parameters are important determinants of labor supply elasticities in the model. In the validation section, we will document that the elasticities produced by our model align with those reported in the literature. Married couples with young children value the home-produced good Q more ($\gamma_Q^{3,m} = 13.28$) than those with older ($\gamma_Q^{2,m} = 5.90$) or without children ($\gamma_Q^{1,m} = 4.43$). Single individuals with older children also value Q more ($\gamma_Q^{2,s} = 10.18$) than those without children ($\gamma_Q^{1,s} = 6.34$). The heterogeneity

[13] Our model parameters affect all targeted moments. The table associates parameters to the moments that are most strongly affected by them.

Table 6: Estimates I: Preference and Production Function Parameters

Parameter	Description:	Value	Std. Error
Utility function: power parameters			
σ_c	Private consumption power parameter (*calibrated*)	1.50	–
σ_l	Leisure power parameter	2.85	(0.04)
Utility function: importance of other factors relative to private consumption			
γ_l	Leisure	1.97	(0.04)
$\gamma_Q^{1,m}$	Public consumption: married, without kids	4.43	(0.08)
$\gamma_Q^{2,m}$	Public consumption: married, kids aged 6+ present	5.90	(0.18)
$\gamma_Q^{3,m}$	Public consumption: married, kids under age 6 present	13.28	(0.24)
$\gamma_Q^{1,s}$	Public consumption: single, without kids	6.34	(0.23)
$\gamma_Q^{2,s}$	Public consumption: single, kids aged 6+ present	10.18	(0.33)
Household public good: importance of time relative to market inputs			
$\eta^{1,1,m}$	Married, high school, no kids	0.33	(0.01)
$\eta^{2,1,m}$	Married, high school, kids age 6+	0.45	(0.01)
$\eta^{3,1,m}$	Married, high school, kids under age 6	0.24	(0.01)
$\eta^{1,2,m}$	Married, college, no kids	0.47	(0.01)
$\eta^{2,2,m}$	Married, college, kids age 6+	0.57	(0.01)
$\eta^{3,2,m}$	Married, college, kids under age 6	0.42	(0.01)
$\eta^{1,1,s}$	Single, high school, no kids	0.62	(0.01)
$\eta^{2,1,s}$	Single, high school, kids age 6+	0.75	(0.01)
$\eta^{1,2,s}$	Single, college, no kids	0.63	(0.01)
$\eta^{2,2,s}$	Single, college, kids age 6+	0.77	(0.01)
Household public good: other production function parameters			
τ	Substitution parameter, home hours vs. market Good	0.38	(0.01)
υ	Substitution parameter, wife's vs husband's home hours	0.40	(0.01)
α^1	Productivity of home hours, women	1.18	(0.02)
α^2	Productivity of home hours, men (equal to $2 - \alpha_1$)	0.82	–

in γ_Q is crucial to explain the large differences in time spent on household production and leisure by child status. In the data and our model simulations, households with young children allocate substantially more resources to household production. As they shift both time and financial resources toward Q, their leisure time unequivocally decreases. Both men and women with children under age six allocate a striking 11 fewer hours weekly to leisure – in both data and simulations – as compared to childless individuals (Table 7); those with older children devote about six fewer hours weekly to leisure. As Table 7 shows, these differences in leisure time by child status are an order of magnitude greater than differences in leisure time by gender ($0.5 - 1.8$ hours) and substantially greater than differences by education ($2.5 - 4.7$ hours). For married women, they are also significantly larger than differences by quartile of husband's income ($1.5 - 1.9$ hours).

Next, we analyze the estimated production function parameters. The CES substitution parameters, $\tau = 0.38$ and $\upsilon = 0.40$, are positive and sizable, indicating that there is considerable substitutability between time and market inputs as well as between the wife's and husband's time inputs. However,

Table 7: Weekly Time Spent on Leisure: Data and Model Simulation

	No Children		Children 6+		Children <6	
	Data	Model	Data	Model	Data	Model
All	66.1	66.4	59.3	59.9	55.2	56.6
Men	65.9	65.9	58.9	58.6	54.3	55.6
Women	66.4	66.9	59.8	61.2	56.1	57.7
High School	66.9	67.3	60.2	60.9	56.6	58.2
College	64.4	64.4	57.3	57.4	51.9	53.0
Married Women,						
By Quartile of Husband's Income						
First Quartile	64.5	65.3	59.9	60.5	56.9	57.6
Second Quartile	63.8	64.6	60.7	62.1	56.0	56.3
Third Quartile	65.0	66.3	59.0	60.4	54.5	55.7
Fourth Quartile	66.0	67.2	61.8	62.3	55.2	56.2

Table 8: Estimates II: Match Quality and Marriage Market Parameters

Description:	Value	Std. Error
Cost of divorce, high school	5.49	(0.11)
Cost of divorce, college	11.52	(0.20)
Mean initial match quality	-0.40	(0.01)
Variance match quality, normalized	1.0	-
Decline in mean of initial match quality with age:	0.14	(0.01)
Probability of drawing a potential match, Normalized	1.0	-
Probability of drawing a partner with the same education	0.75	(0.02)
Corr. between the individual FEs of potential partners, if same education	0.46	(0.01)
Corr. between the individual FEs of potential partners, if different education	0.44	(0.01)

these inputs are far from perfect substitutes – consistent with the findings in Calvo, Lindenlaub, and Reynoso (2021) – which would require τ and υ to equal one. Lastly, we estimate that the productivity of women's hours in household production ($\alpha^1 = 1.18$) is slightly higher than men's ($\alpha^2 = 0.92$).[14]

The estimates of the marriage market parameters indicate substantial sorting on income that is not captured by sorting on education alone. While single individuals are more likely to draw a potential spouse with the same education (with probability $\pi = 0.75$), we estimate additionally a strong positive correlation between individual fixed effects, even conditioning on education. The correlation is similar for potential partners with the same completed education ($\rho_{g=g_s} = 0.44$) and for partners with different degrees ($\rho_{g\neq g_s} = 0.46$).

Finally, the wage parameters have the expected sign and size. Returns to human capital are highest for college-educated men (linear term of 0.079), followed by educated women (0.069), and high school-

[14] We allow for gender-based productivity differences in household production to match the fact that, in the data, wives spend on average more hours in home production than husbands, even in couples where the woman is more educated or earns more.

Table 9: Estimates III: Wage Process Parameters

Description:	Women	Men
Constant:		
High School	2.085	2.229
College	2.340	2.539
Experience:		
High School	0.047	0.050
College	0.069	0.079
Experience Squared:		
High School	-0.0009	-0.0009
College	-0.0014	-0.0016
Part-time Experience Accumulation:		
High School	0.317	-0.098
College	0.630	0.433
No-time Experience Depreciation:		
High School	-0.937	-1.131
College	-1.585	-1.020
Standard Deviation of Wage Shocks:		
High School	0.370	0.373
College	0.387	0.433
Standard Deviation of Fixed Effects:		
High School	0.415	0.414
College	0.475	0.459

educated men and women (0.050 and 0.047). The gender differences are consistent with evidence that women select into occupations with flatter wage profiles (Goldin 2014). Human capital depreciates rapidly for all groups during non-employment, with each year-long absence further reducing the stock of human capital by approximately one year.[15] Women experience fewer penalties for part-time work, consistent with Goldin (2014) and Bronson (2015): college women accumulate 0.63 years of full-time equivalent experience for each year of part-time work, while college men accumulate only 0.43 years. Part-time accumulation rates are significantly lower for high school women and men (0.317 and -0.098), in line with evidence from Blundell et al. (2016) on low returns to part-time work among those without a college degree.

6.2 Model Fit and Validation

Overview. Table 10 gives a high-level snapshot of model performance. We provide evidence about how well we capture overall marriage, divorce, employment, and home production patterns in the economy, after aggregating across ages and individuals. The shares married and divorced in the model are close to the data averages of 0.65 and 0.11. The model combines marriage and fertility probabilities

[15]This is comparable to estimates of long-run wage loss penalties of around 5 to 10% due to firm separations from mass lay-offs. See a recent overview by Bertheau et al. (2022).

Table 10: Summary Statistics, Data and Model Simulation

Moment:	Data	Model
Share Married, Ages 22 to 60	0.61	0.65
Share Divorced, Ages 22 to 60	0.12	0.11
Share w/Child < 6, Ages 22 to 38	0.34	0.38
Share Employed:		
Men, HS	0.74	0.79
Men, College	0.91	0.94
Women, HS	0.64	0.66
Women, College	0.79	0.83
Weekly Hours Spent in Home Production:		
Men	12.3	14.1
Women	25.2	21.2
Share of Households in which Woman is Higher Earner:		
When Husband is High School Graduate	29.3	33.7
When husband is College Graduate	26.2	31.2

to generate a share of households with young children (0.34) only slightly lower than what is observed in the data (0.38). Our simulated data match well the differences in employment rates by gender and education, which range from 0.94 for college men (0.91 in the data) to 0.66 for high school women (0.64 in the data). The model is able to generate substantial specialization, although the gender difference in household production hours is slightly lower in the model (14.1 and 21.2 hours for men and women, respectively) than in the data (12.3 and 25.2 hours).

Income Distributions and Matching. Our primary objective is to assess the impact of changes in the taxation system. These reforms alter the distribution of the tax burden, which in turn affects overall welfare. Since joint taxation systems use the income of both spouses to set tax rates, it is crucial to accurately fit the income distributions of both women and men and how they match in the marriage market, to effectively evaluate these reforms.

Figure 4 depicts the actual and simulated income distribution for men and women, by education. We collapse the right tails for easier visual inspection. The model fits well all four distributions. In the estimation, we target the shares of men and women marrying someone with the same education directly. Thus, it is not surprising that we match these moments almost perfectly. To provide further evidence on the ability of our model to fit the matching patterns in the marriage market, Figure 5 reports the relationship between a husband's earnings and the wife's education over the whole distribution. The model accurately reflects the share of women who have a college degree as a function of the husband's earnings.

Figure 4: Income Distribution By Education (Ages 25-54)

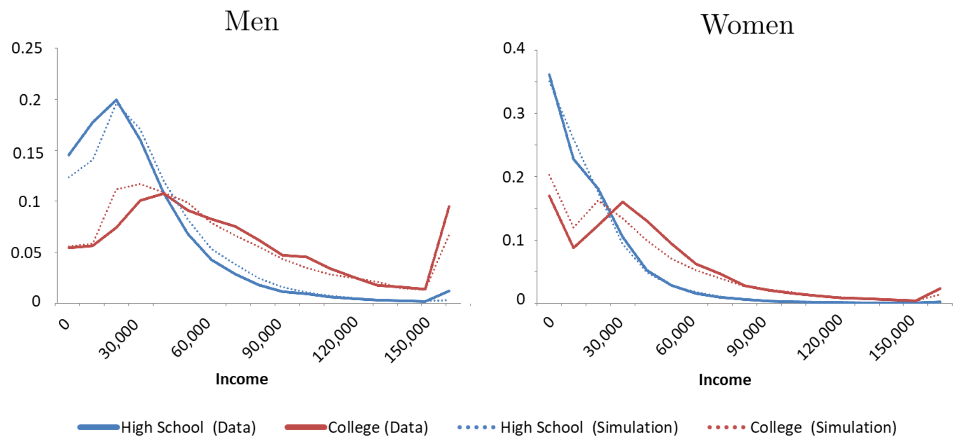


Figure 5: Share of Married Women With College Education, By Husband's Earnings

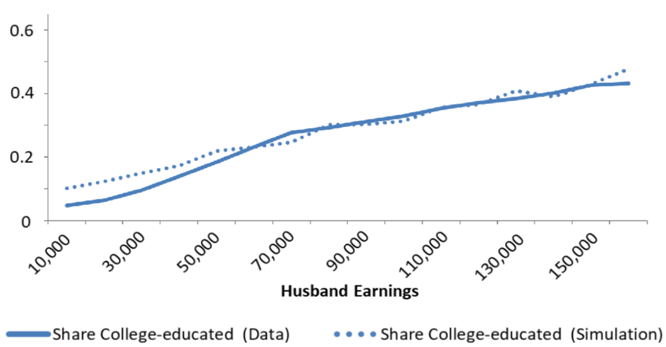


Figure 6: Share of Married Women Employed, By Husband's Earnings and Child Status

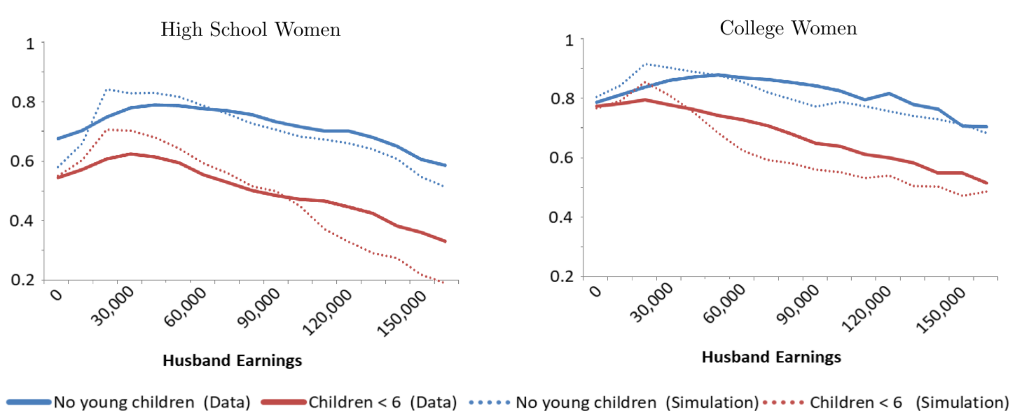


Table 11: Extensive Margin Responses to Temporary Wage Shocks, Simulated Data

			Married	
	All	Single	No yng. ch.	With yng. ch.
Panel A: All women				
Log wage temporary shock	0.580	0.585	0.562	0.633
Constant	0.712	0.718	0.736	0.590
Panel B: Women, less than college				
Log wage temporary shock	0.671	0.693	0.630	0.723
Constant	0.683	0.673	0.727	0.553
Panel C: Women, college				
Log wage temporary shock	0.421	0.321	0.468	0.507
Constant	0.780	0.864	0.753	0.658
Panel D: All men				
Log wage temporary shock	0.301	0.416	0.191	0.212
Constant	0.879	0.821	0.938	0.909
Panel E: Men, less than college				
Log wage temporary shock	0.351	0.484	0.194	0.282
Constant	0.856	0.792	0.934	0.882
Panel F: Men, college				
Log wage temporary shock	0.211	0.267	0.186	0.0969
Constant	0.930	0.904	0.945	0.968

Employment Levels and Heterogeneity Patterns. Figure 6 describes married women's employment rates, separately by education and whether there are young children in the household, as functions of husband's earnings. It shows that we achieve overall a good fit: we match well the patterns by education, child status, and husband's earnings, which are three of the most strongly predictive variables for married women's labor force participation.

Intensive and Extensive Margin Frisch Elasticities. For our policy evaluations to be credible, the labor supply elasticities generated by the model should be sensible. Here we report the labor supply responses to temporary, idiosyncratic shocks to wages (ε_{it} in Equation (4)). We do this by regressing the individual labor supply choices on the idiosyncratic shocks using simulated data. Since the shocks are temporary and uncorrelated with other structural parameters, the coefficients from such regressions approximate Frisch labor supply elasticities.[16]

Table 11 describes the extensive margin responses by marital status and presence of young children. They are larger for women (0.580) than for men (0.301) and vary depending on education, marital status, and presence of young children. Women with less than college education have the largest

[16] As each period in the model corresponds to two years, the estimates are not pure Frisch elasticities as they generate a small, but greater than zero, change in lifetime wealth.

Table 12: Intensive Margin Responses to Temporary Wage Shocks, Simulated Data

	All	Single	Married: No yng. ch.	Married: With yng. ch.
Panel A: All women				
Log wage temporary shock	0.124	0.0489	0.178	0.245
Constant	3.640	3.677	3.616	3.566
Panel B: Women, less than college				
Log wage temporary shock	0.141	0.0505	0.214	0.281
Constant	3.639	3.679	3.608	3.565
Panel C: Women, college				
Log wage temporary shock	0.0994	0.0433	0.133	0.205
Constant	3.640	3.671	3.629	3.564
Panel D: All men				
Log wage temporary shock	0.105	0.0486	0.169	0.0655
Constant	3.696	3.672	3.706	3.760
Panel E: Men, less than college				
Log wage temporary shock	0.109	0.0546	0.178	0.0721
Constant	3.698	3.675	3.711	3.761
Panel F: Men, college				
Log wage temporary shock	0.0964	0.0361	0.156	0.0557
Constant	3.690	3.663	3.696	3.759

responses. The regression coefficient implies that a 10 percent temporary rise in wage increases the employment probability by 6.7 percentage points. Given that the employment rate for this group is 66 percent, this corresponds to a 10.2 percent increase in the probability of being employed following the 10 percent rise in wage—that is, an extensive margin elasticity of one. At the opposite end, men with college education have an elasticity of 0.23 (= 0.211/0.94), which declines to 0.09 if they are married and have young children.

Table 12 reports the intensive margin responses. The heterogeneity in these elasticities across gender, education, and marital status is much less pronounced. They are all small and positive, averaging 0.1 for men and 0.12 for women. Notably, among married individuals, women with young children have the largest elasticity and their spouses have the smallest one. This difference reflects the ability of the model to generate attachment to the workforce that varies depending on the degree of intra-household specialization. When young children are present, the spouse with higher-earning potential responds less to wage fluctuations because they have incentives to work long hours, whereas their partner can respond to shocks by taking advantage of temporary job opportunities.

Are these elasticities consistent with estimates in the labor literature? Many papers have estimated intensive margin Frish elasticities for men, and the majority report small numbers (see Table 6 in Keane

Table 13: Secondary Earners' Employment Responses to the Bush Tax Cuts

	(1)	(2)	(3)	(4)
	CPS Data		Simulated Data	
$\tilde{\tau}_{i,t}$	-1.281***	-0.865***	-2.037***	-1.190***
	(0.259)	(0.274)	(0.119)	(0.122)
$\tilde{\tau}_{i,t} \times YoungChild_{i,t}$		-0.617***		-1.282***
		(0.142)		(0.0550)
N	25937	25937	123594	123594

Regressions shown in columns (1) and (2) are similar to the ones shown in Table 2, but restrict the CPS sample to cohorts born between 1976 and 1980 to make the results compared to the simulated data. Columns (3) and (4) show results of analogous repeated cross-section regressions based on simulated data. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

(2011)). Our small intensive margin elasticities for men are therefore consistent with the literature. There are fewer papers estimating the corresponding elasticities for women and they range from 0.03 to 3 (Table 7 in Keane (2011)). Our elasticites are closer to the lower bound. Not as many papers have estimated extensive margin Frish elasticities. Chetty et al. (2013) report that the estimates obtained using only men or the entire population and quasi-experimental variation range from 0.18 to 0.43 (Table 1). Our estimates for men are within this range. Chetty et al. (2013) do not report estimated extensive margin elasticities for all women or our subgroups (only for single mothers). But it is reassuring that the only elasticity they report that uses both women and men for the U.S. is higher at 0.43 than those obtained only for men. Since we do not target these elasticities directly in our estimation (only employment levels), these results are noteworthy and serve as a validation test of our model.

Responses to the Bush Tax Cuts. As a final validation exercise, we assess the ability of our model to replicate the secondary earners' employment responses to the Bush Tax Cuts. Here, we focus on the CPS results, as they are more precisely estimated and, thus, serve as a stronger test of the model. In Appendix C, we report the model performance against the PSID data and find similar results.

We proceed in two steps. First, we take advantage of the large sample size of the CPS and restrict the sample to the cohorts used in the estimation of the model: those born between 1966 and 1970. We then re-estimate Equation (2) and report the results in columns (1) and (2) of Table 13. They are qualitatively similar to the results obtained using all cohorts but larger in magnitude, indicating that secondary earners who were between 33 and 37 years old at the time of the Bush Tax Cuts were particularly responsive to the changes induced by the reform.

Next, we estimate Equation (2) using the simulated data and describe the results in columns (3) and (4) of the same Table. They are qualitatively similar to the CPS findings: there is a strong association between reduced tax rates and increased participation of secondary earners, and the association is stronger in households with young children. The magnitudes are smaller in the CPS data, with a ratio of real to simulated magnitudes of 60% for households with young children and 73% for families without.

Since we do not match the responses to the Bush Tax Cuts in the estimation, this a noteworthy result. In addition, smaller magnitudes in the CPS data can be expected due to measurement errors. There are two main sources of measurement errors in the tax rates: (i) idiosyncratic family shocks, such as high medical expenditures or donations, that are not accounted for in the model; and (ii) state income taxes that in our model are approximated by taking averages across states.

The outcomes of this and previous validation tests give us confidence that the model can properly evaluate the behavioral responses and welfare effects of relevant tax reforms.

7 Tax Reforms

We now use the estimated model to evaluate potential tax reforms. We consider four alternatives to the income splitting system that approximates the U.S. tax schedule: (i) income splitting with optimal tax rates; (ii) optimal individual taxation; (iii) a general optimal joint system that allows for an arbitrary dependence between a spouse's tax rates and the partner's income; and (iv) optimal income splitting with a flat deduction for secondary earners. In all counterfactuals, we keep the existing welfare system and state taxes constant. Thus, all simulations account for the EITC, SNAP, other transfers and credits, and state taxes.

7.1 Optimal Tax Rates

The Utilitarian Objective Function. For each alternative tax system, we derive the optimal tax rates by maximizing a measure of welfare in the economy, conditional on the constraint that the tax revenues must be greater than or equal to those generated under the baseline model. The baseline model corresponds to the estimated model with tax rates, welfare transfers, and state taxes for 2012. As a measure of welfare, we rely on the widely used utilitarian objective function, constructed as the equally-weighted sum of individual life-time utilities.[17]

[17] This choice imposes planner's preferences towards redistribution that are mostly controlled by the CRRA specification for private consumption.

Given the dynamic nature of the model, we must also choose how to weight the individual utility at different points of the life-cycle. We decided to assign equal weight to the expected lifetime utility of individuals at five different ages: 20, 30, 40, 50, and 60. This choice is analogous to giving equal weights to different generations and is consistent with a democratic political system where individuals vote for alternative policies according to the effects of these policies on their present discounted value of current and future utilities.[18]

Given our choices, the social planner's problem takes the following form:

$$\max_{\tau} \sum_{a=20,30,40,50,60} \int_{S} V^{a}\left(S|\tau\right) df\left(S|a,\tau\right) \quad s.t. \quad TR(\tau) \geq \overline{TR}, \tag{7}$$

where $V^{a}\left(S,\tau\right)$ is the lifetime expected value function of an individual of age a = 20, 30, 40, 50, 60, with state variables S, given the tax policy τ; $f\left(S|a,\tau\right)$ is the distribution of state variables for generation a if the tax system is τ; $TR(\tau)$ is total tax revenues given τ; and $\overline{TR}$ denotes total tax revenues collected by the U.S. government under the baseline model.

The Tax Systems. To be consistent with current practices, the government can condition the tax rates only on individual incomes, which are observed. Since one of our goals is to compare the performance of four alternative tax schemes to the current system, for all of them the optimal tax rates are computed for the seven tax brackets adopted by the U.S. government in 2012. The brackets for singles are approximately half of those for married households, except for the top bracket. We therefore use the brackets for married households and divide them by two to obtain the brackets for singles.[19]

Solution Methods for Alternative Systems. The maximization problem can be solved using standard techniques for the U.S. income splitting system without and with deduction and for the individual system. In these cases, we follow current practices in the U.S. and other countries by allowing for only one optimal marginal tax rate τ_j within each bracket. We can then use standard search methods.

[18] This choice induces an asymmetry, as young individuals value their discounted future utilities while older individuals have zero value for past consumption and leisure. An alternative approach would be to compute the lifetime utility of each generation starting from the period they are born and to assign equal weight to them. Eden (2023) recommends against this alternative approach as, under reasonable preference parameters, at the optimum young individuals consume much more than older ones. In Appendix D, we further discuss this issue and relate our formulation of the social welfare function to the framework in Eden (2023).

[19] After constructing a zero tax bracket equivalent to the standard deduction, the brackets are therefore $\$0k-\$4.5k$, $\$4.5k-\$10.75k$, $\$10.75k-\$31k$, $\$31k-\$57.25k$, $\$57.25k-\$85.5k$, $\$85.5k-\$149k$, $>\$149k$ for singles. For couples, the endpoints of each bracket are multiplied by 2.

Table 14: Optimal Tax Rates in Different Taxation Systems

Brackets	Curr. Splitting	Splitting*	Individual*	Deduction*	Deduction* $0.5\sigma_\theta$
0-4.5k	0.0	0.0	0.0	0.0	0.0
4.5-10.75k	0.11	0.0	0.0	0.0	0.0
10.75-31k	0.19	0.23	0.17	0.25	0.24
31-57.25k	0.30	0.29	0.20	0.32	0.34
57.25-85.5k	0.33	0.35	0.41	0.37	0.39
85.5-149k	0.38	0.40	0.48	0.38	0.40
>149k	0.40	0.50	0.52	0.50	0.52
Optimal Deduction				$9,134	$9,382
Ann. Income Equiv.	–	0.90%	1.25%	1.30%	–

It is more complicated to find the optimal tax rates for a general joint system that allows for flexible dependence between a spouse's tax rates and the partner's income. We approximate this schedule by adapting to our setting the approach used in Gayle and Shephard (2019). Specifically, we allow for different optimal rates for single and married households. For singles, analogously to the other taxation systems, each bracket has one marginal tax rate. For married couples, both spouses have the same seven tax brackets described above. We denote by $n_1, \ldots, n_N$ their endpoints, which form a grid of $(N+1) \times (N+1)$ points. We assume that the grid is symmetric, i.e. the marginal tax rate if the primary earner is at endpoint i and the secondary earner at endpoint j, $\tau_{i,j}$, is equal to $\tau_{j,i}$. To be consistent with the other tax systems, within each bracket the marginal tax rate is constant.

Given this structure, we can solve for the optimal general joint tax schedule as follows. For singles, we select a marginal tax rate for each bracket. For married households, we first assign to each point of the grid the total taxes paid by a couple with total income corresponding to that point, $T_{i,j}$. Next, we approximate in a flexible way the entire tax schedule for couples by (i) dividing each square of the tax grid in upper and lower triangles and (ii) computing the total taxes paid by couples with incomes that correspond to different points in a triangle using the endpoints of the triangle and barycentric coordinates. We then iterate over the marginal tax rates for singles τ_j and total taxes for couples $T_{i,j}$ until total welfare in the planner's problem (7) is maximize.

To apply the approximation method that uses the barycentric coordinates, we need an upper bound for each spouse's income. We choose it to be twice the lower bound of the highest bracket ($\$298k = 2 \times \$149k$). With the seven brackets used in the U.S., we therefore maximize over $8 \times (8+1)/2$ variables for couples and 7 variables for singles. We assume that individuals and couples with zero income pay zero taxes, which reduces the number of variables by 2. We therefore solve a maximization problem with $8 \times (8+1)/2 - 1 + 7 - 1 = 41$ variables. To account for the constraint on total revenues,

Table 15: Marginal tax rates, general joint system

	Income change (in thousands of dollars)						
	0-4.5	4.5-10.75	10.75-31	31-57.25	57.25-85.5	85.5-148.25	148.25-296.5
Marginal tax rates for single individuals							
	0.0%	33.1%	33.4%	46.9%	55.9%	59.6%	65.2%
Marginal tax rates for married individuals, by income of the spouse							
0	**0.0%**	**0.0%**	**5.6%**	**8.5%**	**27.4%**	**38.1%**	**36.0%**
4.5	*11.0%*	**0.0%**	**7.7%**	**8.8%**	**28.4%**	**38.1%**	**36.2%**
10.75	*11.0%*	*9.4%*	**8.3%**	**11.0%**	**30.8%**	**37.8%**	**36.1%**
31	*20.6%*	*11.3%*	*31.0%*	**7.4%**	**36.4%**	**33.8%**	**37.7%**
57.25	*22.4%*	*20.2%*	*26.4%*	*17.2%*	**36.4%**	**38.9%**	**39.4%**
85.5	*28.6%*	*31.2%*	*34.2%*	*17.2%*	*31.6%*	**40.5%**	**38.4%**
148.25	*28.8%*	*28.5%*	*21.6%*	*29.4%*	*35.1%*	*33.1%*	**36.0%**
296.5	*34.3%*	*27.3%*	*33.4%*	*38.5%*	*30.0%*	*27.5%*	*23.6%*
Ann. Income Equiv.							2.59%

we rely on the augmented Lagrangian method.

7.2 Results

We evaluate the transition to each optimal taxation system using two measures: (i) the overall change in total welfare and (ii) the change in welfare distribution. For each tax reform, we compute the changes in the welfare distribution by calculating the average willingness to pay for the policy for each ability-education-gender cell. In our model, these variables are fixed at the individual level, which ensures they remain constant across different taxation systems.[20]

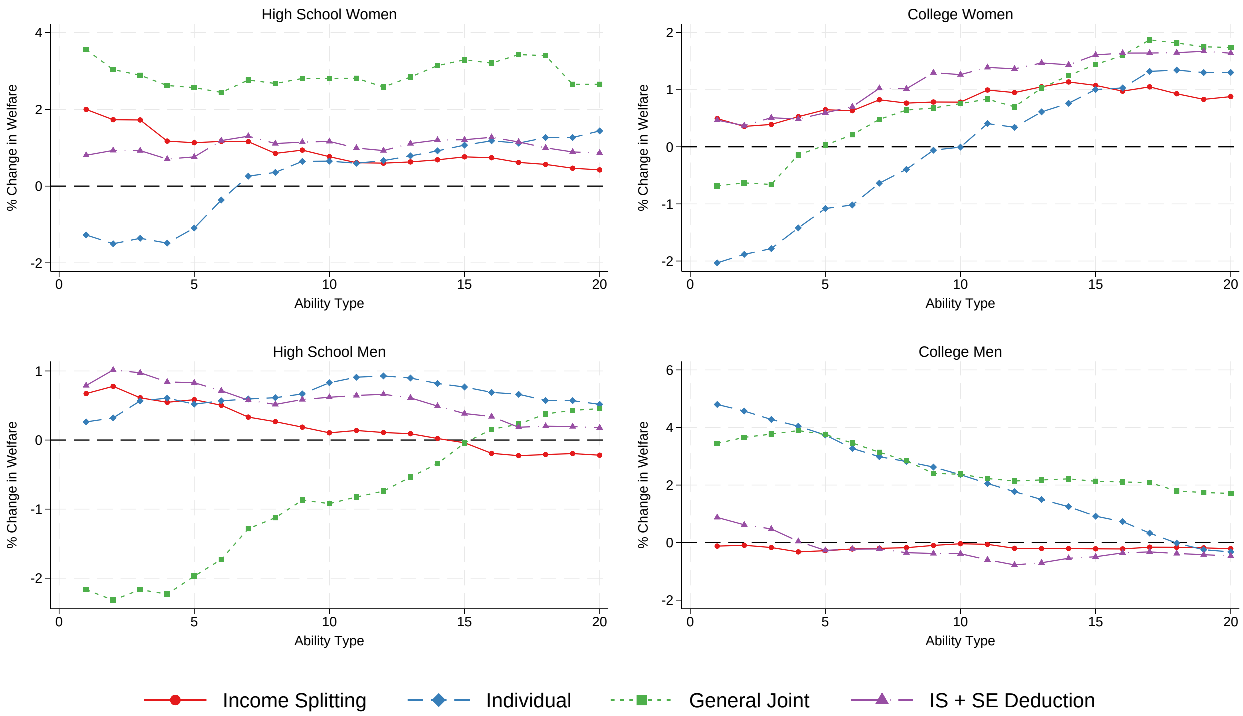
Columns 1-4 of Table 14 contain the current U.S. tax rates, and the optimal tax rates for the income splitting, individual, and income splitting with deduction systems. It also reports the willingness to pay for each reform averaged over ability-education-gender cells. In Table 15, we report the optimal tax rates for the general joint system, separately for primary earners (bold upper triangular part) and secondary earners (script lower part), and the corresponding average willingness to pay. Figure 7 describes the changes in welfare distribution relative to the baseline model.

Optimal vs Current Income Splitting

Rates and Welfare. The first two columns of Table 14 compare the rates currently adopted by the U.S. (first column) with the rates that maximize total welfare (second column) in the same income

[20]We calculate the willingness to pay for each gender-education-ability group by simulating counterfactuals for the baseline model where individuals in the group receive a specific income transfer in every period, with each counterfactual corresponding to different transfer amounts. Then, for each group, we identify the transfer amount that results in an increase in average welfare equivalent to the increase produced by the policy reform. We compute the percent change in willingness to pay by taking the ratio of the transfer to the group's average income.

Figure 7: Percent Changes in Welfare by Ability, Education, Gender, and Tax System



splitting system. The optimal rates are similar to the current ones except at the very bottom and top, as our welfare function prescribes more redistribution with significantly lower rates for incomes below \$10,750 and higher rates for incomes above \$149,000. Individuals are willing to pay on average 0.9% of their income for the more progressive tax system.

Changes in Welfare Distribution. The red lines in Figure 7 describe the changes in welfare distribution as a function of ability for different education and gender groups. Due to more redistribution, women and men with low ability and education benefit the most from the tax rate changes. The lower rates at the bottom increase their take-home income with direct effects on their welfare. They also have an indirect effect: with higher after-tax incomes these families can increase their degree of intra-household specialization and the production of the home-good. These welfare gains are financed by higher revenues from college-educated men and high-ability high school men, whose elevated incomes place them in the highest tax bracket.

General Insights. Under the current system, the U.S. economy would benefit from increasing the tax rate for the top bracket, as advocated by several economists (e.g. Diamond and Saez (2011)). However, our model reveals that there are limits to the effectiveness of such increase. Higher tax rates

Figure 8: Percent Changes in Labor Force Participation over Ability by Education, Gender, and Tax System

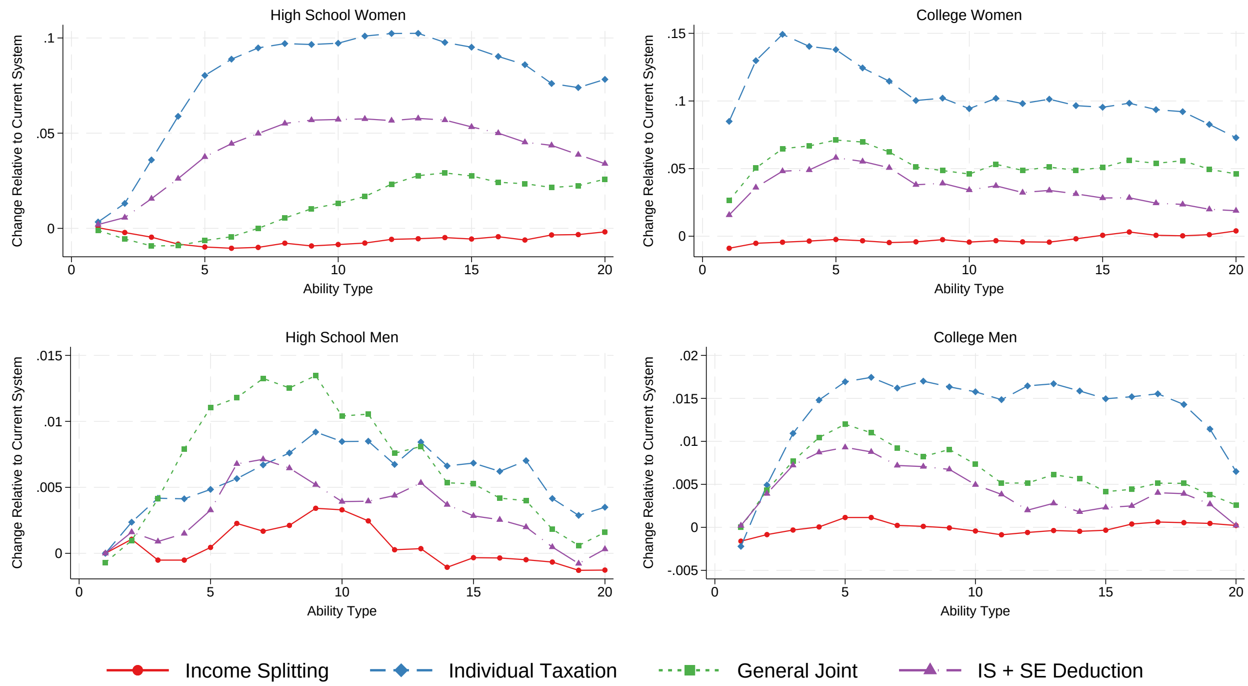
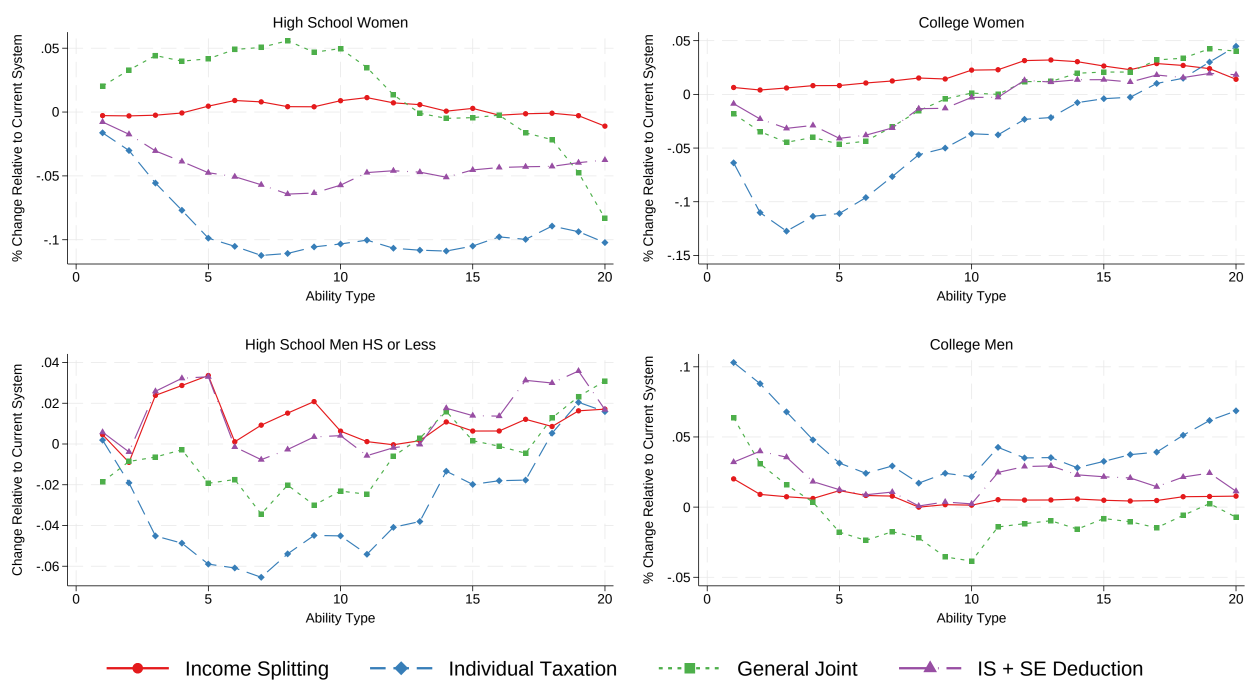


Figure 9: Percent Changes in Household Production Time over Ability by Education, Gender, and Tax System



on top earners affect both primary and secondary earners. The latter, who have higher labor supply elasticities, tend to reduce their labor supply in response to increased tax rates, thereby limiting the potential gains from such policy changes.[21] The red lines in Figures 8, and 9 illustrate these responses (see also Figure A.6 in Appendix G, which accounts for the intensive margin of labor supply). Using women as a proxy for secondary earners, they document a reduction in female labor force participation and labor supply, and an increase in the time invested in household production in response to the more progressive tax schedule. These behavioral changes limit the ability of a government to raise tax revenues and, thus, to redistribute resources under income splitting.

Optimal Individual vs Current Income Splitting

Rates and Welfare. The optimal tax rates under individual taxation differ significantly from those in the current income splitting system, featuring greater redistribution from higher to lower income levels. Marginal tax rates decrease for all income brackets below $57,250, with the most significant reductions occurring in the $4,500-$10,750 and $31,000-$57,250 brackets, where rates fall by 11 and 10 percentage points, respectively. These reductions are financed by substantial increases in tax rates for the top three brackets, which rise by 8, 10, and 12 percentage points.

Individual taxation can achieve higher levels of redistribution because it decouples the tax rates of primary and secondary earners. The government can then increase the tax rates of high-income primary earners with limited effect on the labor supply decisions of secondary earners.

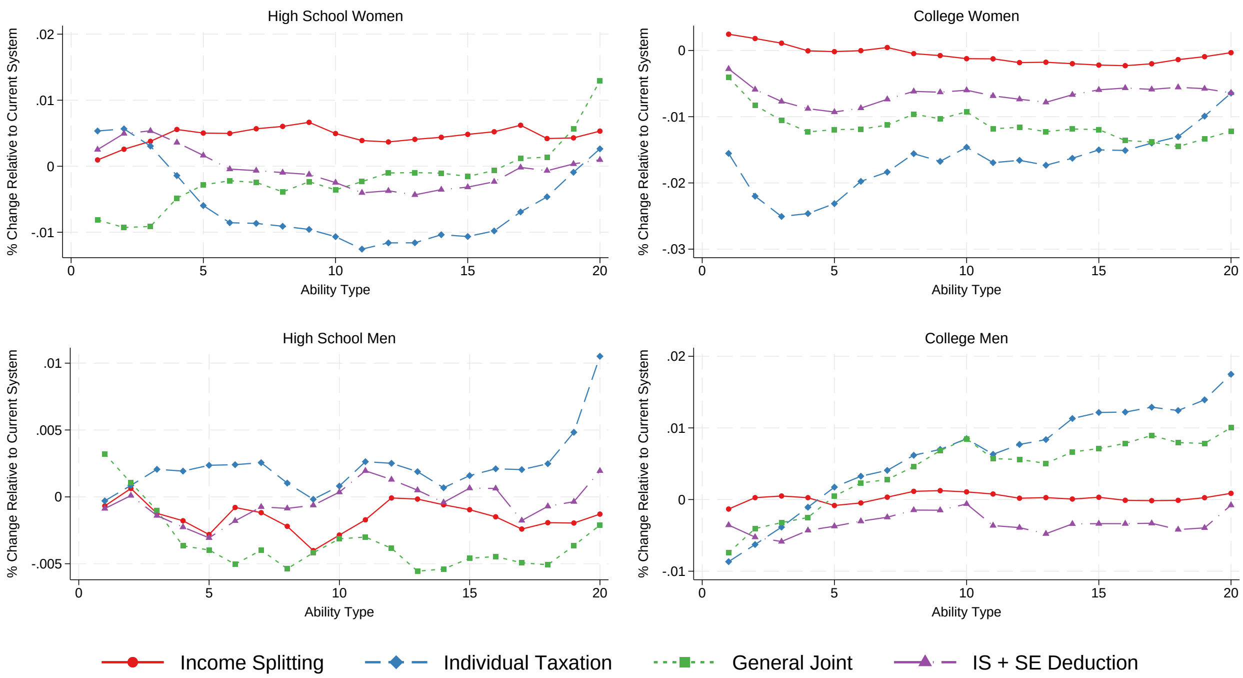
This reform produces a larger increase in aggregate welfare, as individuals are willing to pay on average 1.25% of their income for a switch to individual taxation.

Changes in Welfare Distribution. The substantial increase in aggregate welfare hides an important shortcoming of this reform: it increases welfare inequality. As shown by the blue lines in Figure 7, the individuals benefiting the most are those that have the highest welfare before the reform – college men – with increases that are as high as 5%. By contrast, the most vulnerable group in the population – women with low ability – experience significant welfare loses.

General Insights. The negative welfare effects for low-ability women can be explained by two changes introduced by the reform: lower tax rates for low and medium-income earners and the removal of marriage bonuses. The majority of low-ability women do not benefit from the first change, as they have the lowest participation rates even under individual taxation. Yet, if married, they lose the

[21] For a thorough analysis of the effect of wage changes on female labor force participation see Fernández and Wong (2014).

Figure 10: Percent Changes in Leisure over Ability by Education, Gender, and Tax System



marriage bonuses provided by income splitting. These two changes combine to generate the above-mentioned welfare losses.

For college women, the welfare losses are exacerbated by assortative matching. These women are mostly married to college men, who have on average higher earnings potential than their spouse and lower productivity in household production. Under income splitting, it is optimal for these couples to select a high degree of intra-household specialization. With the switch to individual taxation, the high degree of specialization is no longer optimal and college women, who are typically secondary earners, increase their labor force participation (Figures 8) and reduce the time invested in household production (Figure 9). These adjustments increase the couple's resources – explaining the higher welfare for college men. But they have two adverse effects for women. They reduce their leisure (Figure 10), as the increase in labor supply outweighs the decline in household production hours. They lower the amount of public consumption produced by the couple (Figure 11). For many college women, the welfare losses due to declining leisure and public consumption exceed the positive impact of more resources.

The welfare losses for women occur despite the fact that, under individual taxation, they accumulate the highest levels of human capital. Figure 12 illustrates that, with the exception of college-educated women with very high ability, human capital increases for all of them, with some increases

Figure 11: Percent Changes in Public Consumption over Ability by Education, Gender, and Tax System

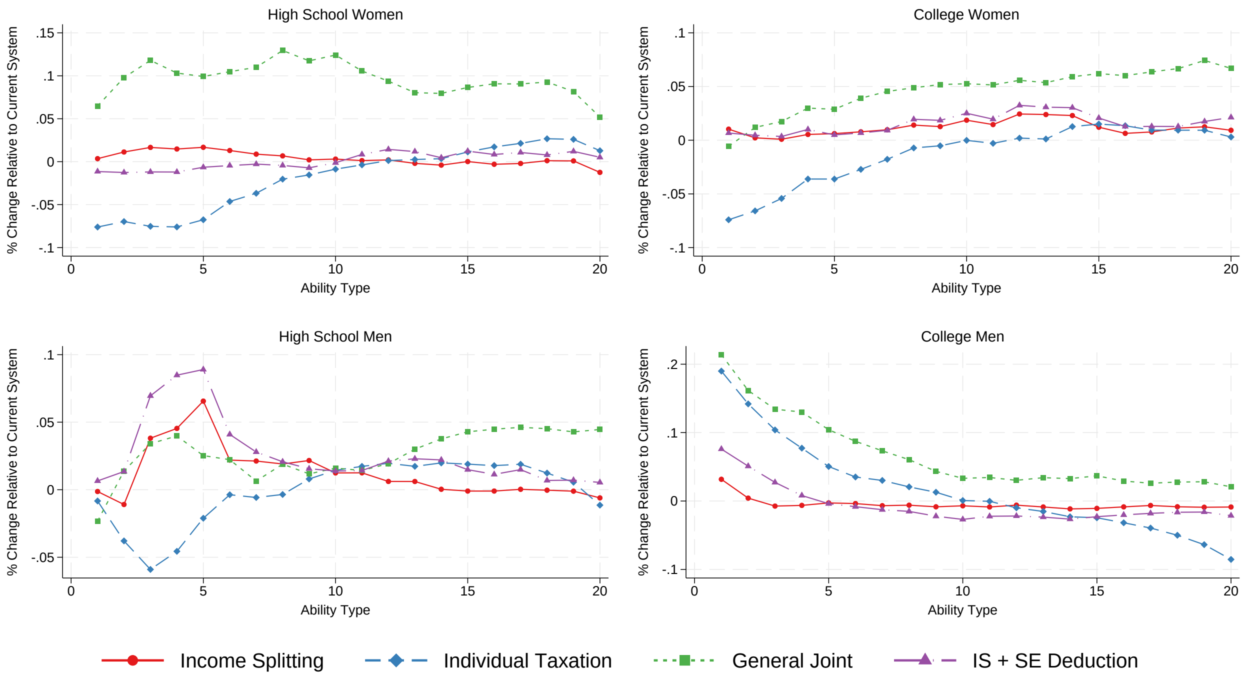
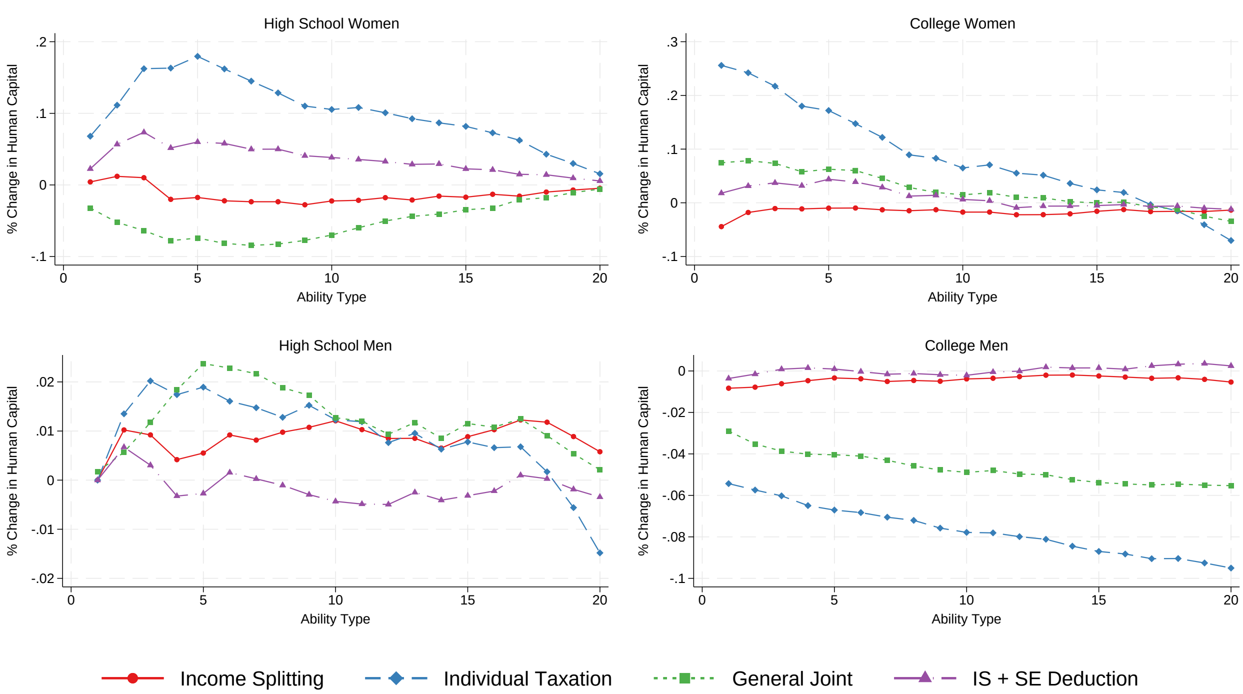


Figure 12: Percent Changes in Human Capital by Tax System



reaching up to 25%. The transition to individual taxation has mixed effects on the human capital of men: college-educated men experience significant declines, primarily due to a shift to part-time work towards the end of their careers, while high school-educated men see increases in their human capital. The decline in human capital of college-educated men – who are among the highest earners – limits the amount of tax revenues the government can collect under this taxation system.

Individuals with a high school degree benefit also from lower divorce rates following the reform (see Panel B in Appendix Table A.6). However, this benefit is offset by a decrease in the share of their lifetimes spent married, driven by reductions in marriage length and total number of marriages among the ever-married. In contrast, college-educated individuals experience an increase in the share of their lifetimes spent married, driven by longer marriages—even though the share of them who ever marry falls (Panel C in the same table).[22]

The previous discussion highlights the value of dividing the population in groups based on ability, education, and gender. For the population as a whole, individual taxation would appear to have positive and similar effects at all ability levels. However, this hides the finding that low-ability women, who have lower welfare even before the reform, in fact suffer welfare loses. Our results also underscore the importance of accounting for the time spent on household production by family members. Individual taxation has the indisputable effect of increasing labor supply of individuals with sufficiently high ability and, thus, the amount of resources available to them. At the same time, it affects the allocation of time to different activities and the production of the home-good in complicated ways. Without properly accounting for them, one can overestimate the positive effects of this system.

Optimal General Joint vs Current Income Splitting

Rates and Welfare. Table 15 reports the optimal marginal tax rates for married and single individuals in the general joint system. For married individuals, the Table describes the rates for primary earners (bold faced upper triangle) and secondary earners (lower triangle). The rates can be read by fixing one of the rows, which determines the level of income of a spouse, and by reading the rates along the row. For example, if a spouse has income $85,500 (sixth row), the marginal tax rate of a primary earner making between $85,500 and $148,250 is 40.5%. The rate for a secondary earner with income between $57,250 and $85,500 is 31.6%.

The key advantage of the general joint tax system is its ability to adjust one spouse's tax rates according to the income of the other spouse. This feature results in a distinct pattern of optimal tax

[22]Appendix E provides additional discussion on marriage-related outcomes.

rates compared to the individual system. Under the general joint system, primary earners benefit from substantially lower marginal tax rates across the income distribution. In contrast, secondary earners face significantly higher tax rates at lower and medium income levels. For instance, under individual taxation, secondary earners with an income of $8,000 pay no taxes, whereas under the general joint system, their marginal tax rate ranges from 9.4% to 31.2%, depending on the primary earner's income. However, secondary earners' tax rates are lower at higher income levels. For example, with an income of $100,000, their tax rate is 48% under individual taxation but between 27.5% and 33.1% under the general joint system. There is a clear rationale behind these rates. The general joint system provides incentives for intra-household specialization in couples with different earnings potentials, by increasing marriage bonuses and secondary earners' tax rates. At the same time it gives strong work incentives for 'power couples' compared to individual taxation.

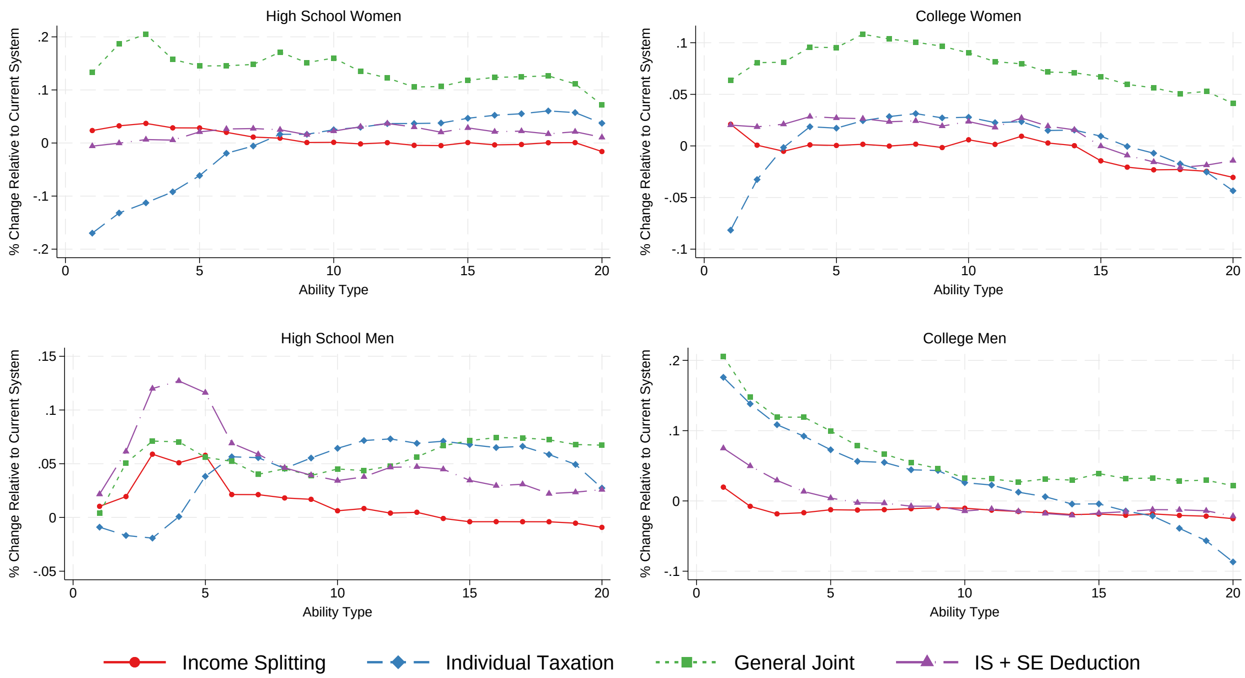
The tax rates of single households are also noteworthy. They are higher than under individual taxation for every income level. Higher rates are optimal from a social planner's perspective because single households have generally lower labor supply elasticities.

Kleven, Kreiner, and Saez (2009) introduced the concept of jointness for the general joint tax system—the relationship between the primary earner's marginal tax rates and the secondary earner's income. Their theoretical results, based on a static model with unitary households, suggest that the optimal tax rates for primary earners can increase (positive jointness) or decrease (negative jointness) with the secondary earner's income depending on whether earnings of the secondary earner signal to the policymaker that those households are worse or better off, respectively. Our results indicate that the optimal general joint system is characterized by a small degree of positive jointness, as the primary earner's tax rates tend to increase slightly with the secondary earner's income. Our findings are only in slight contrast with the results in Gayle and Shephard (2019), which display a small degree of negative jointness. But they are consistent with the current U.S. tax code, which is characterized by positive jointness. One difference between our and Gayle and Shephard (2019)'s model is that we allow for sorting based on unobserved ability. Frankel (2014) finds that the optimality of negative jointness may be overturned if the marriage market is sufficiently assortative.

Under the general joint system, the economy achieves the highest level of welfare, as people are willing to pay on average 2.59% of their income for its implementation.

Change in Welfare Distribution. The green lines in Figure 7 indicate that the general joint system outperforms individual taxation for all groups, with the exception of high-school-educated men. Women, in particular, benefit significantly at all levels of ability and education. The reduced welfare

Figure 13: Percent Changes in Market Inputs by Tax System



for high-school-educated men is explained by shifts in marriage rates, changes in sorting patterns, and higher tax rates for single individuals. As shown in Appendix Table A.6, the general joint system leads to a significant increase in the share of adult life spent married. But not everyone benefits from this change, as detailed in Appendix Figure A.4. The general joint system alters the sorting patterns of married couples, with low earnings-potential men being disproportionately affected. These men face a 0.7% decrease in marriage rates as women of all types increasingly marry higher-ability men. Consequently, these low earnings-potential men miss out on the advantages of marriage and face higher tax rates on their earnings, resulting in the welfare losses outlined earlier.

General Insights. Individuals in the economy exhibit the highest willingness to pay for the general joint tax system because its flexibility allows regulators to boost women's labor supply without compromising household production. As shown in Figures 8, A.6, and 9, with the exception of low-ability, high-school educated women, all other women increase their labor supply and a fraction of them reduce the time spent on the home-produced good. At the same time, the changes in tax rates for married couples raise their disposable income, enabling greater expenditure on market inputs (Figure 13) and, consequently, increased public consumption (Figure 11).

The welfare increase is also partly explained by a higher share of an individual's life spent married and a decline in assortative matching. These findings are relevant to the ongoing discussion about the decline in marriage rates in the United States and its effects on child outcomes (Kearney 2023).

The main limitation of the general joint system is its complexity, requiring regulators to determine 41 tax rates. Over the past four decades, there has been a trend toward simplifying the tax system, as evidenced by the reduction in U.S. tax brackets from 15 in 1980 to 7 in 2024. Thus, while the general joint system is theoretically appealing due to its flexibility, its implementation is highly challenging. It is therefore crucial to identify an alternative system that can offer increased welfare, by introducing some separation between the tax rates of a spouse and the income of the other, while being straightforward to implement and avoiding the adverse effects on low-ability women associated with individual taxation.

Secondary Earners Deduction vs Current Income Splitting

Rates and Welfare. The final reform we evaluate is a transition to an income splitting system with a flat secondary earner's deduction. The optimal rates and deduction for this system, determined by jointly maximizing the social planner's problem (7), are reported in the fourth column of Table 14. With the introduction of a simple deduction, regulators achieve greater redistribution compared to the traditional income splitting. As secondary earners are exempt from taxes on their first \$9,134 regardless of the primary earner's income, the government can raise tax rates for incomes greater than \$10,750, without reducing the labor force participation of secondary earners and diminishing tax revenues. Although the aggregate welfare gains are smaller compared to the general joint system (average willingness to pay 1.30% vs. 2.59%), it still surpasses the individual system and does not adversely affect low-ability women, as shown in Figure 7.

These reform also positively affects women's human capital, with increases of up to 7.5%, though the effect is less pronounced than under individual taxation. But unlike individual taxation, it has minimal impact on men's human capital, enabling higher tax revenues and more effective redistribution. This taxation system has the added benefit of reducing the divorce rate more significantly than any other tax system we have evaluated, leading to more stable marriages (see the discussion in Appendix E for details).

Change in Welfare Distribution and General Insights. With the introduction of a flat deduction for secondary earners, regulators can achieve widespread welfare gains across the ability distribution, with women experiencing the most significant improvements. Only a small fraction of college-educated men experience minor welfare losses. The broad welfare gains result from the system's ability to increase female labor supply with minimal reductions in leisure and virtually no decrease in public consumption, effectively addressing one of the main limitations of the individual taxation system.[23]

Dynamics and Commitment

In the previous discussion, we have highlighted the significance of intra-household specialization, limited commitment, and dynamics. We now take a more direct approach to assess their effects on tax reform evaluation. We focus on the tax system that achieves the highest welfare among those that can be easily implemented – income splitting with a deduction. We then simulate a counterfactual economy where we reduce by half the variance of the match quality shocks that couples receive while married, with the objective of reducing the endogenous probability of divorce. We then recalculate the optimal tax rates and deduction.

When we analyze the simulated data after changing the match quality variance, but before recalculating the optimal rates and deduction, we observe a 24% decrease in the divorce rate. Accordingly, married couples choose to increase their degree of specialization: the fraction of secondary earners working full-time declines by 2%, while the fraction working part-time rises by 2.8%. We then recompute the optimal tax rates and deduction in this new environment and report them in the last column of Table 14. We find a more progressive tax schedule, with lower tax rates for incomes below \$31,000, higher rates for higher incomes, and a slight increase in the deduction. These changes reflect the increased specialization within households. With fewer couples where both spouses work full-time, the social planner can raise tax rates for medium and high earners with minimal impact on secondary earners' labor supply decisions. The extra tax revenues can then be allocated to increase marriage bonuses for low-ability, low-education couples. As a result of these changes, married couples further specialize: full-time work decreases by 3.4% and part-time work increases by 4.6% compared to the scenario before the reduction in the variance. These findings underscore the importance of incorporating limited commitment and dynamics into our counterfactual analyses.

Given the presence of dynamics and uncertainty in our model, we can also evaluate the ability of different taxation systems to insure individuals against consumption shocks. The most relevant finding is that a transition to the general joint system would reduce consumption risk for men. For more details, see Appendix F.

[23] Bierbrauer et al. (2023) discuss the political economy of changes in family taxation, documenting tensions between different welfare measures in determining which reform is optimal or would have political support. The system proposed in this section may provide an alternative to those tensions; investigating this possibility is left to future work.

8 Conclusion

The objective of the paper is to evaluate the effect of different taxation systems on total welfare and welfare distribution. We do this by providing descriptive evidence that distinct taxation systems produce different incentives for primary and secondary earners and that these incentives influence individual choices. We then develop and estimate, using U.S. data, an intertemporal model in which single and married individuals make decisions on labor supply, household production, human capital accumulation, consumption, savings, marriage, and divorce. The model accounts for the main features of the U.S. taxation and welfare systems. Using the estimated model we evaluate four tax reforms: the adoption of tax rates that maximize welfare in an income splitting system similar to the one currently used in the U.S.; a shift to an optimal individual taxation system; a switch to a joint system that allows for flexible dependence between the tax rates of an individual and their spouse's income; and the addition of a flat secondary earner deduction to the income splitting system. Our findings indicate that the general joint system generates the highest level of welfare, as it can achieve a significant redistribution of resources from the top to the bottom of the earnings distribution without introducing significant disincentives to work for secondary earners. But its complexity may prevent its adoption in practice, as it features a total of 41 different tax rates. A simple income-splitting system augmented with a flat deduction for secondary earners achieves significant welfare gains without increasing welfare inequality.

References

Adams, Abi, Laurens Cherchye, Bram De Rock, and Ewout Verriest. 2014. "Consume Now or Later? Time Inconsistency, Collective Choice, and Revealed Preference." *American Economic Review* 104 (12): 4147–83 (December).

Aghion, Philippe, Ufuk Akcigit, Matthieu Lequien, and Stefanie Stantcheva. 2017, November. "Tax Simplicity and Heterogeneous Learning." Working paper 24049.

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 (May).

Alves, Cassiano Breno Machado, Carlos Eugênio da Costa, Felipe Lobel, and Humberto Ataíde Moreira. 2023, May. "Intrahousehold inequality and the joint taxation of household earnings." Unpublished manuscript.

Apps, Patricia, and Ray Rees. 1999a. "On the taxation of trade within and between households." *Journal of Public Economics* 73 (2): 241–263 (August).

Apps, Patricia F., and Ray Rees. 1988. "Taxation and the household." *Journal of Public Economics* 35 (3): 355–369 (April).

———. 1999b. "Individual Versus Joint Taxation in Models With Household Production." *Journal of Political Economy* 107 (2): 393–403.

Attanasio, Orazio, Hamish Low, and Virginia Sánchez-Marcos. 2008. "Explaining Changes in Female Labor Supply in a Life-Cycle Model." *The American Economic Review* 98 (4): 1517–1552.

Attanasio, Orazio P., and Guglielmo Weber. 1995. "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey." *Journal of Political Economy* 103 (6): 1121–1157.

Bertheau, Antoine, Edoardo Maria Acabbi, Cristina Barceló, Andreas Gulyas, Stefano Lombardi, and Raffaele Saggio. 2023. "The Unequal Consequences of Job Loss across Countries." *American Economic Review: Insights* 5 (3): 393–408 (September).

Bick, Alexander, Bettina Brüggemann, Nicola Fuchs-Schündeln, and Hannah Paule-Paludkiewicz. 2019. "Long-term changes in married couples' labor supply and taxes: Evidence from the US and Europe since the 1980s." *Journal of International Economics* 118:44–62.

Bick, Alexander, and Nicola Fuchs-Schündeln. 2017. "Taxation and Labour Supply of Married Couples across Countries: A Macroeconomic Analysis." *The Review of Economic Studies.*

Bierbrauer, Felix, Pierre Boyer, Andreas Peichl, and Daniel Weishaar. 2023. "The taxation of couples." Technical Report 18138.

Blundell, Richard, Martin Browning, and Costas Meghir. 1994. "Consumer Demand and the Life-Cycle Allocation of Household Expenditures." *The Review of Economic Studies* 61 (1): 57–80.

Blundell, Richard, Monica Costa Dias, Costas Meghir, and Jonathan Shaw. 2016. "Female labor supply, human capital, and welfare reform." *Econometrica* 84 (5): 1705–1753.

Borella, Margherita, Mariacristina De Nardi, and Fang Yang. 2022. "Are Marriage-Related Taxes and Social Security Benefits Holding Back Female Labour Supply?" *The Review of Economic Studies* 90 (1): 102–131 (03).

Bronson, Mary Ann. 2015. "Degrees are Forever: Marriage, Educational Investment, and Lifecycle Labor Decisions of Men and Women." *Working Paper, UCLA.*

Calvo, Paula A., Ilse Lindenlaub, and Ana Reynoso. 2021. "Marriage Market and Labor Market Sorting." no. 28883 (June).

Chetty, Raj, Adam Guren, Day Manoli, and Andrea Weber. 2013. "Does Indivisible Labor Explain the Difference between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities." *NBER Macroeconomics Annual* 27:1–56.

Chiappori, Pierre-Andre, and Maurizio Mazzocco. 2017. "Static and Intertemporal Household Decisions." *Journal of Economic Literature* 55 (3): 985–1045 (September).

Diamond, Peter, and Emmanuel Saez. 2011. "The Case for a Progressive Tax: From Basic Research to Policy Recommendations." *Journal of Economic Perspectives* 25 (4): 165–90 (December).

Dickert-Conlin, Stacy. 1999. "Taxes and transfers: their effects on the decision to end a marriage." *Journal of Public Economics* 73 (2): 217–240.

Dustmann, Christian, and Costas Meghir. 2005. "Wages, Experience and Seniority." *The Review of Economic Studies* 72 (1): 77–108 (01).

Eden, Maya. 2023. "The Cross-Sectional Implications of the Social Discount Rate." *Econometrica* 91 (6): 2065–2088.

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

Fernández, Raquel, and Joyce Cheng Wong. 2014. "Divorce Risk, Wages and Working Wives: A Quantitative Life-Cycle Analysis of Female Labour Force Participation." *The Economic Journal* 124 (576): 319–358 (05).

Flood, Sarah, Joel McMurry, Aaron Sojourner, and Matthew Wiswall. 2022. "Inequality in Early Care Experienced by US Children." *Journal of Economic Perspectives* 36 (2): 199–222 (May).

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

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

Goldin, Claudia. 2014. "A grand gender convergence: Its last chapter." *American Economic Review* 104 (4): 1091–1119.

Golosov, Mikhail, and Ilia Krasikov. 2023. "The Optimal Taxation of Couples." Technical Report 31140.

Guner, Nezih, Remzi Kaygusuz, and Gustavo Ventura. 2012. “Taxation and Household Labour Supply.” *The Review of Economic Studies* 79 (3): 1113–1149.

Guryan, Jonathan, Erik Hurst, and Melissa Kearney. 2008. “Parental Education and Parental Time with Children.” *Journal of Economic Perspectives* 22 (3): 23–46 (September).

Heathcote, Jonathan, and Hitoshi Tsujiyama. 2025. “Practical Optimal Income Taxation.”

Keane, Michael P. 2011. “Labor Supply and Taxes: A Survey.” *Journal of Economic Literature* 49 (4): 961–1075.

Kearney, Melissa S. 2023. *The Two-Parent Privilege*. University of Chicago Press Economics Books no. 9780226817781. University of Chicago Press.

Kearney, Melissa S., and Lesley J. Turner. 2013. “Giving Secondary Earners a Tax Break: A Proposal to Help Low- and Middle-Income Families.” *Brookings Discussion Paper 2013-07*.

Kim, Olivia. 2023. “Credit and the Family: The Economic Consequences of Closing the Credit Gap of US Couples.” Technical Report.

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

Kocherlakota, Narayana R. 1996. “Implications of Efficient Risk Sharing without Commitment.” *Review of Economic Studies* 63:595–609.

Lise, Jeremy, and Ken Yamada. 2018. “Household Sharing and Commitment: Evidence from Panel Data on Individual Expenditures and Time Use.” *The Review of Economic Studies* 86 (5): 2184–2219 (10).

Low, Hamish, Costas Meghir, Luigi Pistaferri, and Alessandra Voena. 2022, February. “Marriage, Labor Supply and the Dynamics of the Social Safety Net.” Nber working papers 24356, National Bureau of Economic Research, Inc.

Marcet, Albert, and Ramon Marimon. 2011. “Recursive Contracts.” *Economics Working Paper, European University Institute*, no. eco2011/15.

Mazzocco, Maurizio. 2007. “Household Intertemporal Behaviour: A Collective Characterization and a Test of Commitment.” *Review of Economic Studies* 74:857–895.

Mazzocco, Maurizio, Claudia Ruiz, and Shintaro Yamaguchi. 2014. “Labor Supply and Household Dynamics.” *American Economic Review: Papers & Proceedings* 104 (5): 354–59.

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

Moffitt, Robert A., Brian J. Phelan, and Anne E. Winkler. 2020. "Welfare Rules, Incentives, and Family Structure." *Journal of Human Resources* 55 (1): 1–42.

Piketty, Thomas, and Emmanuel Saez. 2013. "Chapter 7 - Optimal Labor Income Taxation." In *handbook of public economics, vol. 5*, edited by Alan J. Auerbach, Raj Chetty, Martin Feldstein, and Emmanuel Saez, Volume 5 of *Handbook of Public Economics*, 391–474. Elsevier.

Reynoso, Ana. 2024. "The impact of divorce laws on the equilibrium in the marriage market." *Unpublished manuscript.*

Theloudis, Alexandros, Jorge Velilla, Pierre-André Chiappori, J. Ignacio Giménez-Nadal, and José Alberto Molina. 2023. Commitment and the Dynamics of Household Labor Supply.

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

Online Appendix

Taxation and Household Decisions: an Intertemporal Analysis

Mary Ann Bronson, Daniel Haanwinckel, and Maurizio Mazzocco

August 2024

A Additional information on the U.S. Tax System

Below we provide additional background information about the U.S. tax system. In the U.S., married couples are taxed on their combined income. Married couples' tax schedules are always based on their combined income, even though U.S. federal tax forms offer an option for married couples to file separately. However, this is not equivalent to filing as two single individuals, and the option to file separately is rarely exercised.[1]

The U.S. tax system uses separate tax schedules for married couples and individuals, which roughly resemble an income splitting system. Recall that under income splitting, the brackets for married couples are exactly double the brackets for single individuals, with the same marginal tax rate in each bracket. Thus, couples are effectively taxed like two single individuals, each earning half of the total household income. As we discuss in Section 3, such a system generates higher marginal and average tax rates for secondary earners, and lower overall tax liability for two individuals after they marry.

Table A.1 shows tax rates and brackets for single individuals and married couples in 2021. The only deviation from income splitting in the main tax schedule is in the top bracket, which starts at $628,300 for married couples, and is much less than double the starting income for single individuals ($523,600). Thus, spouses with similar earnings and significantly more than $628,300 in combined income will experience an overall marriage penalty. Other brackets for married couples are exactly double those of single individuals; similarly, the 2021 standard deduction (i.e., the zero-tax bracket)

[1] This filing option is often motivated by fairly specific and typically uncommon circumstances (Whalen (2023)). For example, one reason to file separately would be to avoid liability for a spouse's incorrect handling of tax payments, which may be favorable for couples in a second or third marriages with particularly complicated financial histories. Another example is that it may be favorable to take certain deductions separately, such as individual medical expenses, which can be taken only if they exceed a percentage of either own or household income.

Table A.1: 2021 Tax Brackets and Marginal Tax Rates for Single Individuals and Married Couples

Tax Rate	Single Individual	Married Couple
10%	$0 to $9,950	$0 to $19,900
12%	$9,950 to $40,525	$19,900 to $81,050
22%	$40,525 to $86,375	$81,050 to $172,750
24%	$86,375 to $164,925	$172,750 to $329,850
32%	$164,925 to $209,425	$329,850 to $418,850
35%	$209,425 to $523,600	$418,850 to $628,300
37%	$523,600 and up	$628,300 and up

Table A.2: 2000 Tax Brackets and Marginal Tax Rates for Single Individuals and Married Couples

Tax Rate	Single Individual	Married Couple
15%	$0 to $26,250	$0 to $43,850
28%	$26,250 to $63,550	$43,850 to $105,950
31%	$63,550 to $132,600	$105,950 to $161,450
36%	$132,600 to $288,350	$161,450 to $288,350
39.6%	$288,350 and up	$288,350 and up

for married individuals was double that of singles ($25,100 and $12,550, respectively). Consequently, most couples currently experience marriage bonuses, as in a standard income splitting system.[2]

The U.S. deviated more from a traditional income splitting system in the past, with tax schedules that often generated marriage penalties for relatively equal earners (but still generated marriage bonuses for single earner and highly specialized couples). This can be seen in Table A.2, which shows that tax brackets for married couples were significantly less than double those of singles, most dramatically in the higher brackets. The Bush Tax Cuts of 2001 and 2003 changed the tax schedule to eliminate many of these marriage penalties. The 2003 tax cut generated an especially large 12 percentage point reduction in the marginal tax rate for couples with between approximately $47,000 and $57,000 in taxable income (i.e., adjusted income after deductions). It did this by expanding the 15% tax bracket to $56,800 in taxable income for couples, targeting an especially large marriage penalty on income in that range facing many two-earner families (see Figure 3).

Two major reforms before the Bush Tax Cuts include the Tax Reform Act of 1986 during the Reagan administration and the 1993 earned income tax credit (EITC) expansion as part of Omnimus

[2]However, deviations from income splitting can be induced by the EITC and other transfer programs, which offer different size transfers to single and married couples, and may in some cases lead to marriage penalties for low-income households. Similarly, each state sets its own state tax schedule for single and married households, which may also generate deviations from income splitting.

Budget Reconciliation Act (OBRA) under the Clinton administration. The 1986 tax reform effectively reduced the number of brackets from fifteen to just four and, broadly speaking, increased bottom tax rates while reducing top rates. However, it also increased the EITC, the standard deduction, and the personal exemption, thus eliminating taxes altogether for a larger portion of the lowest-earning population. The 1993 reform had an especially large effect on the effective tax schedule in bottom of the income distribution, as it it effectively tripled the EITC (although other welfare programs were reduced around the same period). Notably, it also increased the top federal income tax rate by nearly nine percentage points.

B Additional Descriptive Evidence on Assortative Mating

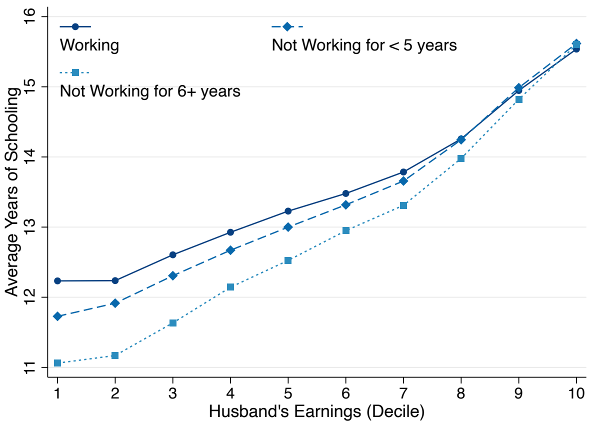


Figure A.1: Married Women's Education

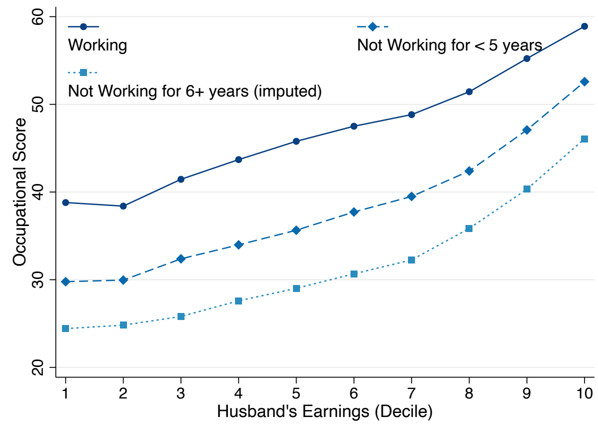


Figure A.2: Married Women's Occupational Score

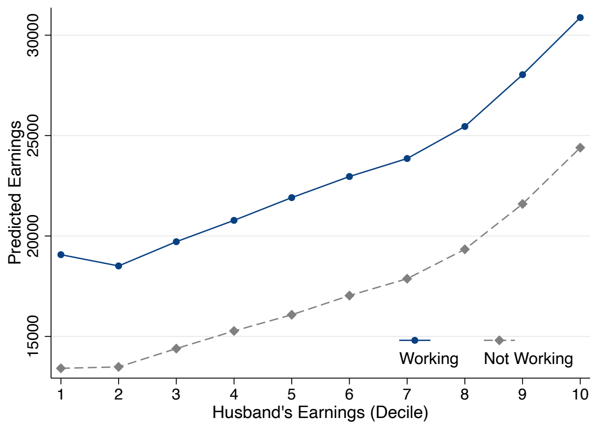


Figure A.3: Married Women's Potential Earnings

Figures A.1, A.2, and A.3 graph three different measures of women's earnings potential, by their employment status and decile of their husband's earnings. The sample includes all couples in which the husband is prime age (25 to 54) and has positive income. Occupational score (in Panel (b)) is available in the IPUMS Census for all individuals who are currently in the labor force or who worked in the last 5 years. For women who have been out of the labor force for more than five years, we impute their occupational score by assigning to them the median score of married women with the same education. The score captures a person's percentile rank in the earnings distribution, based on the median earned income in their occupation. Finally, in Panel (c), we predict potential earnings for all women, working or not, using coefficients from a regression of log earned income on occupational score, and on a set of indicator variables for age and years of education. As before, occupation score is imputed for women who are out of the labor force for more than five years. Source: IPUMS USA, 2000-2007.

C Appendix to the analysis of the 2003 Bush Tax Cuts

C.1 CPS

As mentioned in the main text, we use data from the March CPS from 2002 through 2006, such that the questions on last year's work and earnings refer to tax years 2001 through 2005. The sample includes households where the primary earner earns between \$40,000 and \$140,000 in 2000 dollars, and the secondary earner is between 26 and 45 years of age.[3] The new tax brackets and rates approved in the 2003 Bush Tax Cuts started applying for the 2003 tax year, so that is the first treated year in our analysis. The main regression specification is the one presented in the text, repeated here for convenience:

$$SecondaryEarnerWorks_{it} = \delta_t + \sum_{k=1}^{20} \alpha_k inc_k + \beta \tilde{\tau}_{i,t} + \gamma X_{i,t} + \varepsilon_{it}, \tag{8}$$

where δ_t are time fixed effects, inc_k now denotes primary earner's income groups (width \$5,000), and $\tilde{\tau}_{i,t}$ is the average tax rate on the first \$25,000 of secondary earner's income calculated using TAXSIM based on primary earner's income and number of children in the household.

The controls $X_{i,t}$ are:

- Sex of the respondent in the CPS
- Sex of the secondary worker
- Dummies for education and age of the primary and secondary earner, using the detailed categories in the CPS for both variables
- Dummies for number of children, up to five
- Dummies for the number of children with at most five years of age
- Interactions of the dummies for number of children and number of children with at most five years of age, as above, with a dummy for the sex of the secondary earner
- Quadratic polynomials in primary and secondary age
- Interations of the age polynomials with dummy variables for college completion of the corresponding spouse
- Interactions of time fixed effects with:
  - A dummy for the sex of the secondary earner
  - Dummies for college completion of each spouse

[3] Results are qualitatively similar using secondary earners between 18 and 65, but somewhat weaker in magnitude. As shown in Table 13, when the sample is restricted to cohorts born between 1966 and 1970, the effects are significantly stronger.

– A dummy for having at least one young child in the household

In Table A.3, we exploit the robustness of the results reported in Table 2 to specification choices. Column (1) reports the baseline results. Column (2) shows results using a very sparse set of controls: only year fixed effects, dummies for the number of children, and a dummy for having at least one child.[4] Columns (3) and (4) change the sample inclusion criterion based on husband's income. Column (3) increases the minimum income requirement to $50,000, while Column (4) reduces the maximum income requirement to $100,000. Column (5) excludes tax year 2003 from the analysis.[5] Column (6) uses income bins that are $2,000 wide instead of $5,000 wide. Column (7) reports average marginal effects using a logit model instead of a linear probability model. Finally, Column (8) reports results from regressions where the treatment variable is constructed using the average tax rate on the first $15,000 of secondary earner's income, instead of the first $25,000 of income.

Qualitative results are identical across all specifications. Quantitatively, the biggest change is in the logit specification, where the average marginal effects are smaller (especially so for households with young children). The reason why we are not concerned about this difference in magnitude is that our main purpose with those regressions is to validate the structural model. Specifically, we want to verify that the magnitude of the responses in the data are similar to those of simulated households the model, using the same empirical design on simulated data. To the extent that there are misspecification biases in the linear probability model, they should apply equally to the real data and the simulated data (since the model matches average rates of participation of the secondary spouse very well).

C.2 PSID

The goal of the PSID design is to sidestep concerns about the results from the CPS being spuriously generated by changes in the composition of families in each income bin. The sample includes married couples observed at least in tax years 2000, 2002, and 2004, and we also include data for 2006 if available to increase statistical power. Different from the CPS design, we only use households where the husband is the primary earner (based on 2002 income). This choice is intended to minimize a possible "regression to the mean" concern in a panel setting, if we define the primary earner as

[4] This specific set of sparse controls is the minimal such that the design with the interaction is conceptually sound. The year fixed effects are necessary for the difference-in-differences interpretation of the results. Dummies for number of children are necessary since they are direct inputs used in the construction of the treatment variable. And the dummy for at least one young child is needed for the correct interpretation of the interacted design.

[5] The motivation for dropping this year is that it is the year when the reform was approved, and thus some of the employment choices were made before the actual tax brackets and rates for the year were known.

Table A.3: Effects of 2003 Bush Tax Cuts, CPS data: Robustness Analysis

	(1) Baseline	(2) Few X_{it}	(3) Min. 50k	(4) Max. 100k	(5) Ex. 2003	(6) Fine bins	(7) Logit	(8) 15k
Panel A: Average effects								
$\tilde{\tau}_{i,t}$	-0.651***	-0.657***	-0.504***	-0.496***	-0.650***	-0.573***	-0.376***	-0.252***
	(0.138)	(0.141)	(0.165)	(0.144)	(0.149)	(0.153)	(0.140)	(0.0907)
Panel B: Design with heterogeneity								
$\tilde{\tau}_{i,t}$	-0.314**	-0.395***	-0.221	-0.179	-0.267*	-0.207	-0.202	-0.120
	(0.144)	(0.145)	(0.181)	(0.150)	(0.157)	(0.159)	(0.148)	(0.0916)
$\tilde{\tau}_{i,t} \times YoungChild_{i,t}$	-0.534***	-0.439***	-0.458***	-0.543***	-0.600***	-0.537***	-0.249***	-0.570***
	(0.0762)	(0.0711)	(0.141)	(0.0855)	(0.0855)	(0.0764)	(0.0711)	(0.0737)
N	93115	93115	62894	84872	75112	93115	93115	93115

whichever spouse had the highest income in one specific year.[6] To make the analysis comparable to the CPS one, we also restrict the sample to households where the husband earns between \$40,000 and \$140,000 in 2002, and works full time in both 2002 and 2004. To keep the sample size large enough such that estimates are somewhat precise, we include households where the wife is between the ages of 18 and 65, instead of between 26 and 45 as in the baseline CPS estimates.

The structural model we consider is:

$$SecondaryEarnerWorks_{it} = \alpha_i + \delta_t + \beta\tilde{\tau}_{i,t} + \gamma X_{i,t} + \varepsilon_{it}, \tag{9}$$

where α_i is a household fixed effect, δ_t is the time fixed effect, and $\tilde{\tau}_{i,t}$ is now constructed in the following way. First, we use primary earner's income as of 2004 to calculate the tax rate on the first \$25,000 of secondary earner income in that year. Then, we calculate another tax rate based on the same level of income, but using tax brackets and rates for tax year 2002. The treatment variable $\tilde{\tau}_{i,t}$ is the difference between those two tax rates interacted with a dummy indicating a tax year after the reform. It corresponds to the change in tax incentives induced by the reform, using as the reference point the income of the husband as of 2004.

The controls $X_{i,t}$ are:

- Dummies for number of children up to twelve years of age (with the highest category being three or more)
- A dummy for having at least one child up to six years of age
- A quadratic in wife age interacted with a dummy for college completion
- A dummy for college completion interacted with a post-2003 dummy

Column (3) in Table 2 reports the results of the ordinary least squares (OLS) estimation of equation (9). Column (4) shows the design with the inclusion of an interaction term to measure heterogeneity depending on the presence of young children in the household, similar to the CPS analysis. The magnitudes of the estimated effects are similar what we found using the CPS, but as expected, they are less precise.

As explained in the text, we not only estimate Equation (9) via OLS but also consider an instrumental variables specification. First, we construct a prediction of husband's 2004 income based

[6]Specifically, if we define as secondary earner whichever spouse had the lowest earnings in 2002, we should expect those secondary earners to report higher earnings in 2004 even if there are no behavioral responses to the reform. We choose husbands as the primary earners because, on most households in the data, the husband earns more than the wife. Choosing wives as primary earners would result in even smaller samples.

Table A.4: Effects of 2003 Bush Tax Cuts, PSID data: Robustness Analysis

	(1) Baseline	(2) Nonlinear IV	(3) Min. inc. 50k	(4) Max. inc. 100k	(5) Inc. sec. man
Panel A: Average effects					
$\tilde{\tau}_{i,t}$	-1.020	-1.225	-0.714	-1.429	-0.792
	(1.610)	(2.216)	(1.648)	(2.074)	(1.482)
Panel B: Design with heterogeneity					
$\tilde{\tau}_{i,t}$	-0.741	-1.012	-0.337	-1.200	-0.528
	(1.608)	(2.236)	(1.660)	(2.048)	(1.473)
$\tilde{\tau}_{i,t} \times YoungChild_{i,t}$	-1.634*	-1.681	-1.485	-1.681*	-1.365
	(0.916)	(1.461)	(1.103)	(0.915)	(0.859)
N	4402	4402	3230	3793	4836

on 2002 data using an OLS model with the following variables: a cubic polynomial in income fully interacted with an indicator for college completion, an interaction of age with income, an interaction of highest grade achieved with income, interactions of occupation dummies with income, and interactions of state dummies with income. Note that this income model is estimated in levels, not logs (hence the interaction terms with income being more important than the variables in isolation). The reason is that it is desirable that the averages of predicted income match observed income in 2004. If we use a log model and exponentiate the results to obtain predicted income as of 2004, we would not match the average correctly as the predicted data has less variance than the real one. The specific variables using in the model were chosen because they have high predictive power, that is, they increase the adjusted R2 measure. We restrain from adding many more variables to avoid overfit, given that the PSID sample is not large.

Once predicted husband income in 2004 is known, we estimate 2SLS regressions where we instrument $\tilde{\tau}_{i,t}$ with a quartic polynominal in husbands' predicted 2004 earnings interacted with a post-reform indicator. In the heterogeneity design, we add interactions of the instruments described above with the $YoungChild_{i,t}$ variable as additional excluded instruments.

In Table A.4, we examine the robustness of the 2SLS design to several specification choices. Column (1) repeats the baseline results. In Column (2), we examine the possibility that families where the husband had higher predicted income were on different trajectories compared to other households, even after netting out controls such as the interaction between college status and the post variable. To do so, we change the design such that the linear term of the quadratic polynomial in predicted income, interacted with the post dummy, is now included in the structural regression instead of excluded. Thus, the only excluded instruments are the interactions of higher order terms with the post

Table A.5: Model validation: secondary earner's labor supply responses to the 2003 Bush Tax Cuts, using PSID data

	(1)	(2)	(3)	(4)
	PSID Data		Simulated Data	
$\tilde{\tau}_{i,t}$	-4.041	-3.085	-2.518**	-0.929
	(2.476)	(2.461)	(1.144)	(1.073)
$\tilde{\tau}_{i,t} \times YoungChild_{i,t}$		-2.436**		-6.680***
		(1.104)		(0.163)
N	2251	2251	167926	167926
Kleibergen-Paap F-stat	10.89	5.504	56.41	32.06

dummy. As expected, there is a loss in the precision of the estimates, but the magnitudes are very similar.

The remaining robustness exercises consider changes in the estimating sample. Columns (3) and (4) change the husband income boundaries. Column (5) includes in the sample households where the wife earned more than the husband in 2002. It eliminates any residual concerns regarding regression to the mean, but also introduces some misspecification in the sense that in those households, women are not the secondary earner. In all of those exercises, the estimates are simmilar to the baseline results.

Now we compare the PSID results to similar panel regressions estimated using simulated data. The model-based regression results are reported in Columns (3) and (4) of Table A.5. In the non-interacted design, the estimate is similar to the cross-sectional regression results shown in Table 13. In the interacted design, the estimates are similar for families without young children, but larger in magnitude for families with young children. They are also larger in magnitude compared to the PSID estimates previously discussed.

The CPS results showed that the magnitude of effects can be sensitive to the age range of the sample, with stronger results for the cohort simulated in the model (in their mid-thirties during the reform). The smaller sample size of the PSID makes it difficult to work with narrower age ranges. To verify whether age restrictions can account for smaller estimates using PSID data instead of estimated data, Columns (1) and (2) of Table A.5 report 2SLS estimates from the PSID using a sample between 26 and 45 years of age. The point estimates do increase in magnitude, but the regression becomes much less precise and the instruments are borderline weak. We conclude that the PSID regressions do not provide evidence against the model, but this test is not strong since the estimates become very imprecise once we account for endogeneity of income in 2004 and restrict the age range in the sample to more closely match the simulated cohort.

D Welfare function

In this appendix, we discuss how our social welfare function relates to the work of Eden (2023) concerning tradeoffs between different generations. This analysis focuses on the role of age and cohort; issues such as household production, leisure, marriage, and children are ignored for simplicity. Individuals are assumed to be identical within every cohort.

To start, consider an overlapping generations, stationary economy and define the following social welfare function:

$$W_0 = \sum_{g=-\infty}^{\infty} \omega_g^{Gen} \sum_{a=0}^{T} \omega_a^{Age} \beta^a u(c_{g,a})$$

where g indexes cohorts, a denotes age, $c_{g,a}$ is the consumption of a given cohort at a given age, and u is a concave instantaneous utility function. The innermost summation over a is each cohort's lifetime discounted lifetime utility over T periods, according to the discount factor β. This parameter captures household preferences and can, in principle, be identified from observed savings and borrowing decisions. It does not represent the planner's normative preferences toward specific age groups; that is represented instead by the ω_a^{Age} weights. The planner may also, in principle, assign higher or lower weight to different cohorts through the ω_g^{Gen} terms.

Now, suppose that, for normative reasons, we assume that the social welfare function should not favor one cohort over another. One way to do that is to assign equal weight to every cohort from, say, $-G$ to G for some large G.[7] Given the steady-state assumption, there would be no reason to assign different consumption of individuals of the same age but from different generations. We could thus write the welfare function as:

$$\begin{aligned} W_0 = &\ \textstyle\sum_{g=-G}^{G} \frac{1}{2G+1} \sum_{a=0}^{T} \omega_a^{Age} \beta^a u(c_{g,a}) \\ = &\ \textstyle\sum_{a=0}^{T} \omega_a^{Age} \beta^a u(c_a) \end{aligned} \tag{10}$$

We can use this simple welfare function to illustrate a key result from Eden (2023). If we assume that the planner is indifferent between age groups in addition to being indifferent between generations (that is, $\omega_a^{Age} = 1/(T+1)\ \forall a$), then a utilitarian central planner would assign much higher consumption to young individuals. For instance, if $\beta = 0.98$ (with time defined yearly) and $u'(c) = c^{-1.5}$, as in our structural model, then a 70 year old individual should consume only 30% as much as a 30 year old one.

[7] In general, one would then need to take a limit as G grows to infinity to match the previous definition. However, given the steady-state assumption, this is unnecessary, as cohorts become irrelevant in the welfare function.

Thus, contrary to the intuition that a utilitarian planner wants to equalize consumption across individuals given a concave utility function, this welfare function implies that smooth consumption over the lifecyle is sub-optimal. Starting from a benchmark allocation with equal consumption over time, a tax reform that shifts resources toward younger generations could lead to welfare gains even if it generates sizable aggregate output declines.

Now, we show how the welfare function we use in the paper relates to the results in Eden (2023). Our approach has three components. First, we only include one cross-section of cohorts alive at a given time period. Second, we give them equal weight. Third, and most importantly, we measure the utility of each generation by the discounted sum of utility values *starting from their current age at the same point in time*, instead of their total lifetime utility from the time they are born.

$$W_1 = \textstyle\sum_{g=0}^{T} \frac{1}{T+1} \sum_{a=g}^{T} \beta^{a-g} u(c_{g,a})$$

The cohort index g represents age at time zero. For each generation, the discounted sum of payoffs starts at age g, with no discounting, and continues until the end of life.

When we simulate alternative tax systems, we assume the economy is in a steady-state and the new system has been in place for the whole life of each generation. Thus, abstracting from the many dimensions of heterogeneity in the model, there is no reason why consumption at a particular age should differ between generations. Using this insight, we can rewrite:

$$\begin{aligned} W_1 &= \textstyle\sum_{g=0}^{T} \frac{1}{T+1} \sum_{a=g}^{T} \beta^{a-g} u(c_a) \\ &= \textstyle\sum_{g=0}^{T} \frac{1}{T+1} \sum_{a=0}^{T} \mathbf{1}\{a \geq g\} \beta^{a-g} u(c_a) \\ &= \textstyle\sum_{a=0}^{T} \beta^{a} u(c_a) \underbrace{\sum_{g=0}^{a} \frac{\beta^{-g}}{T+1}}_{\omega_a^{Age}} \end{aligned}$$

The last line clarifies that our welfare function corresponds to the one in Equation (10) when the planner assigns age-specific Pareto weights that are increasing. Thus, examined under the lens of the framework in Eden (2023), our welfare function is paternalistic in that it assigns a higher weight to consumption at older ages compared to what would be inferred by those individual's discount factors. We view our formulation as closer to what could be inferred from a political economy model of voter preferences. It is also more closely related to welfare functions in static models.

Note that the characterization of the welfare function W_1 is slightly different from that used in our numerical exercises. Specifically, in the paper, we use five generations instead of one generation for every period.

E Tax Reforms and Marriage Outcomes

This appendix provides additional information regarding how each tax reform would affect marriage outcomes. We begin with an analysis of Table A.6, which shows how a series of outcomes vary across taxation systems. Starting with the first outcome in Panel A, we note that, in the baseline simulation, individuals spend about 59 percent of their adult life married to someone. That figure accounts for those who never marry and those who marry multiple times. Two reforms have non-negligible impacts on this statistic. The first is the optimal individual taxation system, which reduces that share by about half a percent. The second is the general joint system, which increases it by four percent. Marriage bonuses explain those effects; they are eliminated in the individual system and amplified in the general joint system.

The other outcomes in Panel A provide a more nuanced description of how each reform affects the marriage market. For the sake of brevity, we highlight only a couple of observations. First, the increase in marriage time associated with the general joint system does not reflect more individuals becoming married. That effect arises instead from an increase in the average number of marriages among those ever married, along with an increase in marriage length. Results only appear on those intensive margins due to effects on sorting patterns, which we discuss towards the end of this appendix.

The second observation is that the income splitting system with secondary earner deduction increases marriage stability. The share of individuals who ever divorce falls by more than a percentage point; in relative terms, it falls by about three percent. There is a slight increase in marriage duration. The average marriage time remains constant because there is also a reduction in the number of marriages among those who ever marry.

Before discussing assortative mating, we note that Table A.6 also lists outcomes by education level. An analysis of the first column across panels shows that, compared to high school graduates, college graduates are more likely to marry, less likely to divorce and stay married for longer. All of those observations hold for the data as well. Notably, the model can replicate those detailed qualitative facts despite being relatively parsimonious.[8]

Now, we discuss how tax reforms affect sorting patterns. The last row in Panel A provides a

[8]Those outcomes are targeted for women but not for men. This table shows outcomes for all individuals regardless of sex.

Table A.6: Detailed marriage outcomes by taxation system

	Curr. Splitting	Splitting*	Individual*	General*	Deduction*
Panel A: All individuals					
Share of adult life married	0.591	0.592	0.588	0.615	0.592
Share ever married	0.820	0.824	0.816	0.819	0.821
Share ever divorced	0.435	0.428	0.429	0.433	0.422
Duration of first marriage	10.2	10.2	10.2	10.5	10.3
# Marriages among ever married	1.41	1.40	1.40	1.44	1.40
Corr. in earnings potentials	0.50	0.48	0.52	0.47	0.48
Panel B: High school					
Share of adult life married	0.565	0.564	0.554	0.583	0.559
Share ever married	0.815	0.821	0.814	0.815	0.820
Share ever divorced	0.454	0.444	0.445	0.453	0.437
Duration of first marriage	9.6	9.6	9.4	9.7	9.6
# Marriages among ever married	1.44	1.43	1.42	1.47	1.43
Panel C: College					
Share of adult life married	0.653	0.659	0.668	0.690	0.667
Share ever married	0.832	0.831	0.820	0.828	0.823
Share ever divorced	0.392	0.390	0.393	0.388	0.388
Duration of first marriage	11.7	11.8	11.9	12.2	11.9
# Marriages among ever married	1.34	1.34	1.34	1.36	1.34

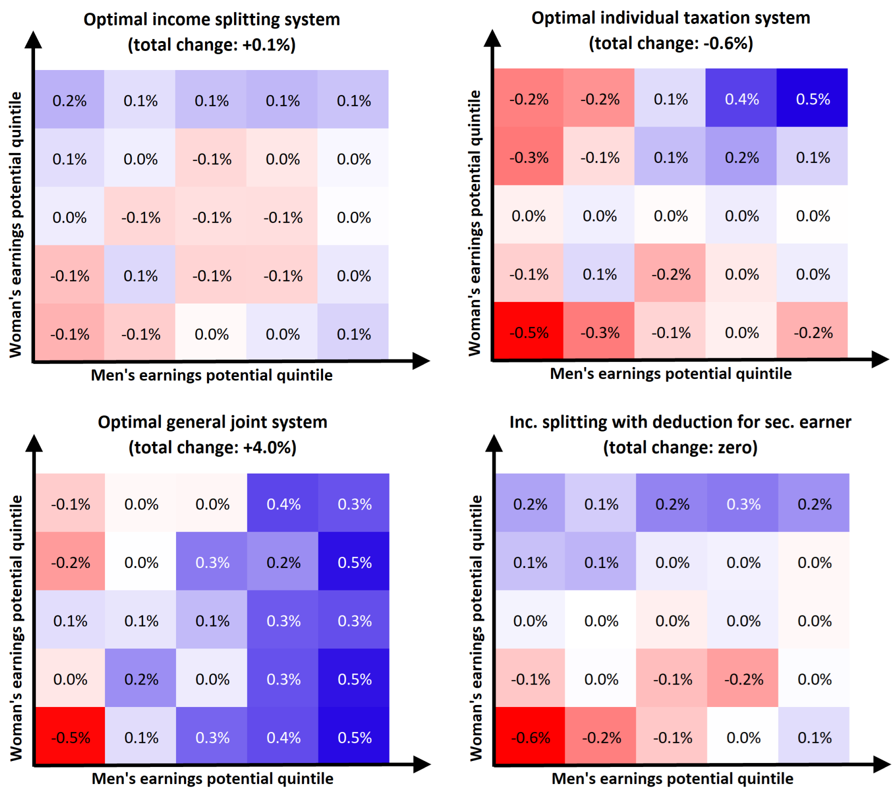
Notes: Never married individuals are included in the calculation of the three "share" outcomes. We restrict the sample to individuals who married at least once when calculating the other outcomes. The last outcome in Panel A corresponds to the correlation in the earnings potentials of husband and wife for all marriages, weighted by marriage length. Earnings potential is defined as expected log earnings of individuals based on sex, education, and ability a, conditioning on a fixed level of experience (eight years if college-educated or 12 years if high school). Because this measure is based on exogenous characteristics, its distribution in the population is stable across tax systems.

preliminary answer to that question. A shift to the individual tax system would increase assortative mating. That is because marriage bonuses are largest in the current U.S. tax system for couples with very different income levels, incentivizing marriages between individuals of different earnings potentials. The individual system eliminates those incentives, which in turn increases sorting. All of the other reforms we analyze go in the opposite direction and tend to increase bonuses for such couples, such that assortative mating falls.

The correlation in earnings potential is a valuable summary statistic regarding sorting, but we can go deeper in our analysis of who marries whom under each tax system. To do so, we split men and women into within-sex quintiles of earnings potential and then analyze how the tax reforms affect marriage rates for couples in each of the 25 husband and wife combinations along those categories. We report the results of that analysis in Figure A.4 (see the footnote of that figure for a description of how we define earnings potential in this exercise).

For brevity, we focus our discussion on the optimal general joint system, which generates the most significant effect on the share of the adult life spent married. Figure A.4 shows an interesting pattern: marriage rates increase in most cells but tend to fall on the left-hand side of the graph. That means

Figure A.4: Changes in Sorting Patterns by Tax System



Notes: This figure shows how each tax reform affects sorting patterns in the marriage market. Earnings potential is defined in the same way as in Table A.6: expected log earnings of individuals based on sex, education, and ability a, conditioning on a fixed level of experience (eight years if college-educated or 12 years if high school). Because this measure is based on exogenous characteristics, its distribution is stable across tax systems. Each number corresponds to the change in total marriage time for individuals in one of 25 groups defined by husband and wife earnings potentials, divided by total marriage time under the current tax system. The sum of all 25 numbers in each graph equals the total effect shown in the header, which corresponds to the numbers from the first row in Table A.6.

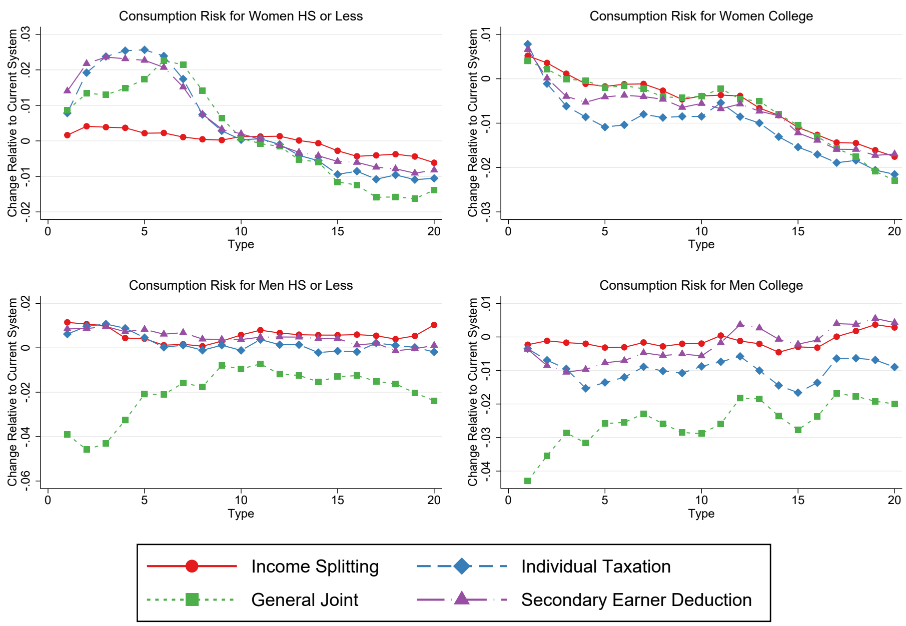
there is a significant reduction in marriage rates for men with low earnings potentials, coupled with increases for men with medium and high earnings potential. There is no such heterogeneity for women, for whom marriage rates increase at all levels of earnings potentials (and, if anything, they increase less at the top). That pattern helps explain why the increase in marriage rates occurs mainly on the intensive margin for that tax system, as mentioned above. The planner chooses to incentivize marriages of high-ability men and de-incentivize marriages for low-ability men. As a result, high-ability men become more likely to marry (sometimes more than once) and to have longer marriages. The extensive margin effects are canceled out by the fall in marriage rates for low-ability men, but the net intensive margin effects remain positive.

F Insurance Against Consumption Risk

The presence of dynamics in our model enables us to evaluate the ability of different taxation systems to insure individuals against consumption shocks. To measure consumption risk under each tax system, we estimate regressions of log individual consumption on age, age squared, and individual fixed effects for different worker groups. Specifically, we run a separate regression for each combination of sex, education, and vintiles of permanent ability (corresponding to 80 regressions for each tax system). Then, we calculate the difference between actual log consumption and predicted consumption using those models. We use the standard deviation of those residuals as a metric of consumption risk at the sex-education-ability level. It measures the variability of consumption over the individuals' lifetime relative to the average trajectory for individuals who start with the same level of education and ability due to different realizations of wage shocks, marriage quality shocks, and fertility outcomes.

In Figure A.5, we report the changes in that measure relative to the current taxation system. A change to the optimal income splitting system has minimal effects on consumption risk, consistent with that system being reasonably similar to the current US tax system. The optimal individual system and the income splitting with deductions lead to small increases in consumption risk for high school women, along with similarly small reductions in consumption risk for high-ability college women. Finally, a shift to the general joint system is very similar to the previous two systems in terms of its effects on women. However, it would lead to non-negligible reductions in consumption risk for men of all types—especially low-ability ones.

Figure A.5: Changes in Consumption Risk by Gender, Education, Age, and Tax System



G Additional Simulation Results

Figure A.6: Percent Changes in Labor Supply over Ability by Education, Gender, and Tax System