PUEY UNGPHAKORN INSTITUTE
FOR ECONOMIC RESEARCH

Wages, taxes, and labor supply elasticities: The role of social preferences

by

Janjala Chirakijja and Pinchuan Ong

June 2026

Discussion Paper

No. 259

Wages, taxes, and labor supply elasticities: The role of social preferences*

Janjala Chirakijja[1] and Pinchuan Ong[2]

[1]Monash University, Department of Econometrics and Business Statistics
[2]National University of Singapore, NUS Business School

June 12, 2026

ABSTRACT: Economists typically treat labor supply responses to wages and taxes as equivalent. We show that social preferences towards tax-funded government expenditures induce differences between the wage and net-of-tax rate elasticities of labor supply in canonical models. We use a large-scale vignette experiment to show that wage elasticities of labor supply are meaningfully larger than their net-of-tax rate counterparts, consistent with social preferences affecting labor supply. We show relevance for real labor market decisions by leveraging an existing elasticity of taxable income meta-analysis. Hence, models calibrated using net-of-tax rate elasticities when wage elasticities are more suitable understate individuals' labor supply responses.

JEL: J22, H24, H41.

*Chirakijja: janjala.chirakijja@monash.edu; Ong: ong.pinchuan@nus.edu.sg. We thank Sascha Becker, Marianne Bertrand, Adeline Delavande, Jason Faberman, Lorenz Goette, Jie Gong, Naijia Guo, Jon Guryan, Hilary Hoynes, Anders Humlum, Supreet Kaur, Patrick Kline, John Klopfer, Matthew Notowidigdo, Amanda Pallais, Jessica Pan, Simon Quach, Jesse Rothstein, Benjamin Schoefer, Uta Schoenberg, Dmitry Taubinsky, Basit Zafar, Anson Zhou, James Ziliak, and participants of the NUS SIG and Macro Brownbags, NUS Labor Mini-Conference, CUHK Lunch Brownbag, HKU seminar, SOLE 2024 Meeting, Berkeley Labor Lunch, Australasian Public Choice Conference, PIER seminar, EEA-ESEM 2024 Congress, WZB Seminar, EALE 2024, ESAM 2024, and NTU seminar for valuable comments and suggestions. We are grateful to Carina Neisser for making her paper's replication kit publicly available. Justin Goh, Ke Li, and Hongqing Qiu provided outstanding research assistance. Research reported in this paper received financial support from the MOE AcRF Tier 1 Grant A-0003317-00-00. This study received IRB Approval from the NUS Business School under Protocol SP-23-0103 and was preregistered as AEARCTR-0011052.

1 Introduction

The responsiveness of labor supply to wages and taxes — the labor supply elasticity — plays a key role in business cycle models and for assessing the efficiency costs of labor income taxation, and differences in its magnitude can give rise to different conclusions (Keane 2011). When thinking about this elasticity, economists seldom distinguish between responses to wages and taxes. Despite a nominal distinction — the tax response is more accurately called the elasticity with respect to the net-of-tax rate (e.g. Saez et al. 2012) — virtually all labor supply models assume that wages and income taxes matter only through their effects on disposable income. Since wage increases and tax cuts equivalently raise disposable income, most empirical labor supply elasticity estimates are based on responses to after-tax wages (see e.g. Eissa and Hoynes 2004; Ziliak and Kniesner 2005; Blau and Kahn 2007; Keane and Wasi 2016; Blundell et al. 2016).

This paper identifies a reason — beyond salience (see Chetty et al. 2009; Finkelstein 2009; Blumkin et al. 2012; Taubinsky and Rees-Jones 2018; Kroft et al. 2024) — to distinguish between wage and net-of-tax rate elasticities of labor supply. A large literature documents other-regarding or social preferences (Andreoni 1989, 1990; Fehr and Schmidt 1999; Charness and Rabin 2002). Applied to income taxation, people might like or dislike the fact that tax dollars eventually go towards (or are wasted in the course of providing) public goods expenditure. Moreover, while wage changes are often viewed as individual-specific, tax rate changes almost always apply to broad groups (e.g. entire tax brackets). Hence, tax changes potentially affect the aggregate provision of public goods. Such government-related social preferences would lead individuals to derive utility or disutility from income taxation over and above what would be explained by its impact on their disposable income.

We show that incorporating government-related social preferences in a labor supply model drives a wedge between the responses of labor supply to wages and the net-of-tax rate. Intuitively, if people derive satisfaction from their or others' contributions to public good provision, a tax increase hurts less than an equivalent wage decrease — despite reduced disposable income, the public good or warm glow compensates. In contrast, a wage increase benefits both the government budget and personal consumption. Thus, if people derive satisfaction from their or others' tax dollars, the wage elasticity of labor supply should exceed its net-of-tax rate counterpart. Of course, the opposite is also possible: individuals might perceive government spending as wasteful or resent paying taxes, making a tax increase hurt more than an equivalent wage decrease. The wage elasticity would then be smaller than its net-of-tax rate counterpart. The sign and magnitude of the wedge are therefore empirical questions.

To estimate the size of this "wage-tax elasticity wedge" (the wage less the net-of-tax rate elasticities), we field a large online vignette experiment in the US. Our vignette experiment presented respondents with pairs of hypothetical scenarios that randomly varied wages, work hours, and tax rates. For each scenario, respondents indicated whether they would take the job, and for each pair, they selected the scenario they preferred. Our vignette experiment is similar in spirit to Kosar et al. (2019), with an additional tax dimension.[1]

A vignette experiment allows us to obtain causal estimates of the two different elasticities (wage and net-of-tax rate) for the same population, which is difficult to achieve using observational data.[2] We obtain wedge estimates with high internal validity since we control the randomization process. Its main limitation is ecological validity: although design choices such as pivoting scenarios around respondents' actual situations can reduce experimental artificiality, hypothetical bias may still cause respondents' stated choices to differ from decisions they would make in real-world labor markets.[3] A further feature — not inherently a strength or weakness, but useful for our purpose — is that the stated-preference method allows us to abstract from market frictions, as in Ameriks et al. (2020), and focus on the social-preference channel. Additionally, presenting wages, tax rates, and the implied dollar amounts together reduces the scope for salience to drive the difference.

Consistent with social preferences affecting labor supply, we find a meaningful wage-tax elasticity wedge: wage elasticities are larger than net-of-tax rate elasticities across specifications. This has an implication for model calibration. Models that plug in a net-of-tax rate elasticity when a wage elasticity is called for will understate labor supply responses. Crucially, this implication does not depend on the channel generating the wedge — part of it could reflect status utility from higher gross wages or other employer-related social preferences. (Prior work has shown that employer-specific reciprocity (Krueger and Mas 2004; Gneezy and List 2006; Mas 2006; Kube et al. 2012, 2013; DellaVigna et al. 2022) and pay inequality (Cohn et al. 2014; Bracha et al. 2015; Breza et al. 2018; Dube et al. 2019; Cullen

[1] Vignette experiments have been used in labor economics to estimate preferences for nonpecuniary job (dis)amenities (Eriksson and Kristensen 2014; Mas and Pallais 2017; Wiswall and Zafar 2018; Mas and Pallais 2019; Kosar et al. 2019; Folke and Rickne 2022; Maestas et al. 2023; Burbano et al. forthcoming), including distaste for work by including work hours in the list of vignette disamenities (Wiswall and Zafar 2018; Mas and Pallais 2019; Kosar et al. 2019; Maestas et al. 2023). However, only Kosar et al. and Mas and Pallais report labor supply elasticities — the relevant parameter in many applications — after estimating distaste for work.

[2] Suppose we used observational data with one instrument for wages and another for tax rates to estimate the two types of elasticities. Even if both estimates came from the same sample, there is no guarantee that compliers to the two instruments (Angrist et al. 1996) are the same people. This would bias the estimated wedge.

[3] That said, estimates based on stated preferences have been shown to correspond to actual market decisions in several other settings (Hainmueller et al. 2015; Mas and Pallais 2017; Wiswall and Zafar 2018; Parker and Souleles 2019; Maestas et al. 2023; Ong and Png forthcoming).

and Perez-Truglia 2022; Huet-Vaughn 2025) affect worker behavior, but few consider implications for interpreting wage versus net-of-tax rate elasticities, with Huet-Vaughn a notable exception.) That said, we find evidence that government-related social preferences explain at least part of the wedge: preference for higher-tax scenarios is concentrated among respondents with more favorable views of government, and the wedge is larger when taxes are earmarked for specific programs that the respondent likes more.

Lastly, to address ecological validity, we build on a recent meta-analysis of the elasticity of taxable income (ETI) with respect to the net-of-tax rate by Neisser (2021). Because it examines taxable income, the ETI generalizes the standard labor supply elasticity by additionally capturing other behavioral responses (e.g. compensation timing, tax avoidance) that are relevant for income taxation analysis (Feldstein 1995; Saez et al. 2012). Important for our purposes, government-related social preferences should also be reflected in the ETI (including in the additional behavioral margins) — a better opinion of the government should be correlated with a lower ETI.[4] Using Neisser's replication kit, we show that the ETI is negatively correlated with measures of institutional trust in the World Values Survey (WVS) and European Values Survey (EVS), even after controlling for a large number of study-related covariates and institutional factors. This pattern's alignment with our vignette experiment results suggests that government-related social preferences also matter for labor supply responses to taxation in real-world settings, although as with any cross-country correlational design, we cannot fully rule out all institutional confounders.

A large literature studies labor supply elasticities treating wage and tax variation as interchangeable (see Keane 2011, 2022 for reviews). A small but growing literature uses real-effort experiments to challenge this assumption, yet provides mixed evidence on whether framing a reduction in take-home pay as a tax induces a different response than an equivalent wage cut (Djanali and Sheehan-Connor 2012; Kessler and Norton 2016; Rick et al. 2018; Huet-Vaughn et al. 2019; Mori et al. 2022; Esslinger et al. 2024). Closest to us, Huet-Vaughn et al. show in a real-effort experiment that directing political moderates' taxes to favored rather than unfavored federal agencies reduces the net-of-tax rate elasticity from about 0.77 to essentially zero. We advance knowledge on this issue in several important ways — real-effort experiments are incentivized but also have particular limitations that our study is able to address. First, we use a large and nationally representative sample, and examine tax variation closer to policy-relevant magnitudes; our tax rates varied by at most 20 percentage points, whereas previous studies often rely on extreme jumps such as from zero to 50 percent. Second, we frame our tax to apply to the entire tax bracket, reflecting how actual tax changes are implemented. Replicating this feature is difficult in real-effort experiments because participants

[4] To our knowledge, prior work has not discussed social preferences as a source of ETI variation.

must know that the tax rates do not apply to individuals outside the experiment. Third, we frame the tax as the federal income tax paid to the US Government (i.e. the policy-relevant tax). This is an advantage over many experimental settings where "taxes" are paid to the experimenter (the most common case), other participants, or the university. The distinction is nontrivial: if taxes go to the experimenter and subjects perceive the experimenter to be wealthier than average, redistributive preferences or fairness considerations should generate tax aversion. Of the studies we know of, only two direct the tax to the government — Kessler and Norton, who acknowledge that their design is predisposed to finding a large tax-aversion effect, and Huet-Vaughn et al., described above. We provide evidence on responses to both the general government budget (the relevant broad income-tax margin) and usage-specific taxes.

A separate body of work documents that taxes which are perceived to fund worker benefits (e.g. individual pension benefits) induce small labor supply responses (Gruber 1997; French et al. 2022; Bozio et al. 2023). We generalize beyond a tax-benefit linkage that benefits the self to government-related social preferences broadly defined — our results suggest factors beyond self-benefits (in particular a fair redistribution system) also matter. More broadly, we provide a theoretical framework that connects the real-effort experiment and tax-benefit linkage strands of the literature. For the former, our model generalizes the static framework of Djanali and Sheehan-Connor (2012) to an intertemporal context and provides, to our knowledge, the most detailed derivations for how social preferences shape the net-of-tax rate elasticity. (For example, we go further than Djanali and Sheehan-Connor by providing the functional form of the elasticities and wedge with more transparent interpretation.) For the tax-benefit linkage literature, our model provides the microfoundation for the reduced-form labor supply specifications commonly used in that work.

Our paper also contributes to the question of whether microdata estimates of the intertemporal elasticity of labor supply (Frisch elasticity) can be used to calibrate labor supply responses to business cycle fluctuations. Previous work has argued that estimates based on life-cycle models are underestimated due to credit constraints (Domeij and Floden 2006), precautionary savings (Low 2005), endogenous human capital accumulation (Imai and Keane 2004; Keane and Wasi 2016), optimization frictions (Chetty 2012), and failure to model participation margins (Chang and Kim 2006; Rogerson and Wallenius 2009; Keane and Wasi 2016). Recent Frisch elasticity estimates based on tax holiday natural experiments are designed to be robust to many of the above (Martinez et al. 2021; Sigurdsson 2019; Stefánsson 2019), but remain too small to account for business-cycle variation in work hours. We partially explain the discrepancy — labor supply responds to the wage signal of the business cycle, implying larger elasticities than estimates based on tax variation. Similar considera-

tions apply to estimates based on life-cycle models that exploit exogenous variation in taxes for identification of Frisch elasticities (e.g. Blundell et al. 2016). More generally, microdata-based estimates are frequently based on tax variation, likely because quasi-experimental tax variation is more easily available to the econometrician.[5] However, the social preferences-induced wage-tax elasticity wedge means that these tax variation-based estimates should not be directly used to model business cycles.

Our paper draws insights from a large tax compliance literature (see reviews in Luttmer and Singhal 2014; Alm 2019). A long-standing puzzle in this literature is that taxpayers comply far more than the canonical deterrence model of Allingham and Sandmo (1972) would predict at realistic audit probabilities and penalties (Andreoni et al. 1998), motivating a large body of work on the non-pecuniary motives — collectively, "tax morale" — that sustain compliance. We build on an active sub-literature that links tax morale to attitudes about government spending (Cowell and Gordon 1988; Alm et al. 1993; Hall and Preston 2000; Cullen et al. 2021; Giaccobasso et al. 2022; Falsetta et al. 2023). We add a new perspective to the literature on intrinsic motivations for paying taxes (Kirchler et al. 2008; Dwenger et al. 2016) — whereas this work has focused on the evasion and avoidance margins, we show that the same government-related motives also shape labor supply and, consistent with Huet-Vaughn et al. (2019), appear to reduce the labor supply distortion from taxation.[6]

Finally, we contribute to a growing literature on people's attitudes towards taxation. While recent work establishes that social preferences often outweigh efficiency considerations in shaping views on taxation and inequality (Fisman et al. 2020; Almås et al. 2020; Stantcheva 2021), the downstream implications for economic behavior remain underexplored. We show that these preferences translate into tangible labor supply responses.

The paper proceeds as follows. Section 2 formalizes the intuition for the wage-tax elasticity wedge. Section 3 describes our survey, vignette experiment, and estimation strategy. Section 4 presents the vignette experiment results. Section 5 presents correlations between the elasticity of taxable income and proxies for government-related social preferences. Section 6 concludes.

[5]Compared with papers exploiting tax variation (Eissa 1995; Blundell et al. 1998; Bianchi et al. 2001; Eissa and Hoynes 2004; Chetty et al. 2011; Gelber 2014; Blundell et al. 2016; Sigurdsson 2019; Stefánsson 2019; Unrath 2020; Martinez et al. 2021; Elder et al. 2023; Sigaard 2023), papers exploiting wage variation tend to be less quasi-experimental or focus on specific occupations (MaCurdy 1981; Altonji 1986; Camerer et al. 1997; Oettinger 1999; Pistaferri 2003; Ziliak and Kniesner 2005; Farber 2005; Blau and Kahn 2007; Fehr and Goette 2007; Stafford 2015; Giné et al. 2017; Chen et al. 2019).

[6]More broadly, while researchers have studied social preferences in charitable giving (review in Andreoni and Payne 2013), volunteering (Freeman 1997; Bauer et al. 2013; Lilley and Slonim 2014), redistribution (reviews in Alesina and Giuliano 2011; Mengel and Weidenholzer 2023), and optimal taxation (Saez 2004; Diamond 2006), we demonstrate that social preferences also matter for non-volunteering labor supply.

2 Model

We start by incorporating government-related social preferences in the canonical dynamic model of labor supply and derive implications for the wage and net-of-tax rate elasticities of labor supply. While our model below can be interpreted on its own, we show in Appendix B.1 that it can be derived from a more complex model with utility that depends separately on own and others' contributions, as well as total government expenditure.

In each period, an individual i with initial assets a_i chooses consumption c_i, next period assets a_i', and labor supply h_i to solve (prime denotes the next period)

$$V\left(a_i, w_i, \tau, N_i\right) = \max_{c_i, a_i', h_i} U\left(c_i, h_i, g_i\right) + \beta E\left[V\left(a_i', w_i', \tau', N_i'\right)\right] \tag{1}$$

$$\text{s.t. } c_i + \frac{1}{1+r} a_i' = a_i + \left(1-\tau\right) w_i h_i + N_i, \tag{2}$$

$$g_i = \tau w_i h_i, \tag{3}$$

where w_i is the person's wage, τ is the tax rate on earnings from work $w_i h_i$, N_i is non-labor income, β is an impatience parameter, r is the interest rate, $V\left(\cdot\right)$ is the indirect utility function, and $E\left[\cdot\right]$ is the expectations operator. Equation (2) is a usual budget constraint: consumption c_i and savings $\frac{1}{1+r} a_i' - a_i$ equal post-tax earnings and non-labor income.

We make only one change compared to the usual labor supply model: per-period utility in Equation (1) also depends on the amount of government expenditure from the individual's taxes paid g_i (tax-funded government expenditure), equivalent to the tax rate applied to their total earnings from work by Equation (3). In our main analysis, we specialize utility to

$$U\left(c_i, h_i, g_i\right) = u\left(c_i, h_i\right) + v\left(g_i\right) \tag{4}$$

so that we have an easily interpretable social preferences term $v\left(g_i\right)$. Similar to DellaVigna et al. (2012), $v\left(g_i\right)$ can encompass both pure altruism (people value the public goods being provided) or warm glow (people like contributing to nation building, regardless of the public goods' value). We do not take a stand on which is more important, nor whether $v\left(g_i\right)$ is driven by the perceived benefit of taxes to oneself (e.g. French et al. 2022; Bozio et al. 2023) or to society. $v\left(g_i\right)$ allows for both positive and negative social preferences. We expect individuals to value tax-funded public goods to some degree, and there might be diminishing marginal returns; this translates to $v\left(g_i\right)$ having positive first and negative second derivatives. However, in principle, $v\left(g_i\right)$ might have a negative first derivative if the individual sees public goods provision as net-detrimental to society or himself.

We model τ as a common tax rate that does not depend on i to reflect tax rate changes

in practice — taxes often affect broad groups of people in similar circumstances. Thus, individuals may get social preferences utility from the increased (or decreased) contributions of others, in addition to their own, as well as from changes in the aggregate provision of public goods. Everyone is affected when τ goes up and hence $v(g_i)$ should be interpreted as a summary measure that also captures social preferences from others' contributions. Our vignette experiment wording also reflects this feature of taxes, although we do not seek to separately identify social preferences from others' contributions versus one's own contribution.

In Appendix B.2, we derive the form for three commonly-used labor supply elasticity types: the uncompensated Marshallian elasticity, the compensated Hicksian elasticity, and the Frisch elasticity which holds the marginal utility of wealth constant. For all three elasticity types, the wage elasticity differs from the net-of-tax rate elasticity. We start our discussion with the Frisch elasticity which has the simplest form. Denoting partial derivatives using subscripts and dropping utility's dependence on its arguments and i, the Frisch wage elasticity is

$$\epsilon_w^F = \frac{-w_i(1-\tau)U_{cc}\left[U_c + \frac{\tau}{1-\tau}(U_g + U_{gg}g_i)\right]}{h_i(U_{cc}U_{hh} - U_{ch}^2 + \tau^2 w_i^2 U_{cc}U_{gg})} \tag{5}$$

while the Frisch net-of-tax rate elasticity is

$$\epsilon_{1-\tau}^F = \frac{-w_i(1-\tau)U_{cc}\left[U_c - (U_g + U_{gg}g_i)\right]}{h_i(U_{cc}U_{hh} - U_{ch}^2 + \tau^2 w_i^2 U_{cc}U_{gg})}. \tag{6}$$

The key difference between the two elasticities is the sign on the $U_g + U_{gg}g_i$ term. This term captures the intuition for the wage-tax wedge. Typical concavity assumptions on $U(\cdot)$ ensure that $\frac{-w_i(1-\tau)U_{cc}}{h_i\left(U_{cc}U_{hh} - U_{ch}^2 + \tau^2 w_i^2 U_{cc}U_{gg}\right)}$ and U_c are positive, with their product capturing the usual intuition that increases in disposable earnings encourage substitution towards working. For the Frisch wage elasticity, $U_g + U_{gg}g_i$ reinforces this substitution effect, assuming that taxes were not so high that diminishing marginal utility starts to kick in. (We find limited diminishing marginal utility in our data.) In contrast, in the Frisch net-of-tax rate elasticity, $U_g + U_{gg}g_i$ counters the substitution effect, as increases in tax rates fund more public goods even as they reduce disposable income available for consumption. The immediate implication is that social preferences cause the Frisch wage elasticity to be different from the Frisch net-of-tax rate elasticity, with the wage-tax elasticity wedge

$$\Delta\epsilon^F = \epsilon_w^F - \epsilon_{1-\tau}^F = \frac{-w_i U_{cc}(U_g + U_{gg}g_i)}{h_i(U_{cc}U_{hh} - U_{ch}^2 + \tau^2 w_i^2 U_{cc}U_{gg})} \tag{7}$$

widening as social preferences strengthen (i.e. $U_g + U_{gg}g_i$ increases).

Compared to the Frisch elasticities, the Marshallian analogs

$$\epsilon_w^M = \frac{-w_i(1-\tau)\left\{\left[U_c + \frac{\tau}{1-\tau}(U_g + U_{gg}g_i)\right] + h_i\left[(1-\tau)w_iU_{cc} + U_{ch}\right]\right\}}{h_i\left[(1-\tau)^2 w_i^2 U_{cc} + 2(1-\tau)w_iU_{ch} + U_{hh} + \tau^2 w_i^2 U_{gg}\right]} \tag{8}$$

and

$$\epsilon_{1-\tau}^M = \frac{-w_i(1-\tau)\left\{\left[U_c - (U_g + U_{gg}g_i)\right] + h_i\left[(1-\tau)w_iU_{cc} + U_{ch}\right]\right\}}{h_i\left[(1-\tau)^2 w_i^2 U_{cc} + 2(1-\tau)w_iU_{ch} + U_{hh} + \tau^2 w_i^2 U_{gg}\right]} \tag{9}$$

contain the income effect term $h_i\left[(1-\tau)w_iU_{cc} + U_{ch}\right]$ that we are familiar with. (We ignore the denominator in this discussion for brevity.) The wage-tax elasticity wedge

$$\Delta\epsilon^M = \epsilon_w^M - \epsilon_{1-\tau}^M = \frac{-w_i(U_g + U_{gg}g_i)}{h_i\left[(1-\tau)^2 w_i^2 U_{cc} + 2(1-\tau)w_iU_{ch} + U_{hh} + \tau^2 w_i^2 U_{gg}\right]} \tag{10}$$

is again increasing in social preferences. Notably, the usual income-effect term affects the Marshallian elasticities themselves, but cancels out of the Marshallian wedge.

Somewhat surprisingly, the Hicksian analogs

$$\epsilon_w^H = \frac{-w_i(1-\tau)\left\{\left[U_c + \frac{\tau}{1-\tau}(U_g + U_{gg}g_i)\right] - \frac{\tau}{1-\tau}h_i\frac{U_g}{U_c}\left[(1-\tau)w_iU_{cc} + U_{ch}\right]\right\}}{h_i\left[(1-\tau)^2 w_i^2 U_{cc} + 2(1-\tau)w_iU_{ch} + U_{hh} + \tau^2 w_i^2 U_{gg}\right]}, \tag{11}$$

$$\epsilon_{1-\tau}^H = \frac{-w_i(1-\tau)\left\{\left[U_c - (U_g + U_{gg}g_i)\right] + h_i\frac{U_g}{U_c}\left[(1-\tau)w_iU_{cc} + U_{ch}\right]\right\}}{h_i\left[(1-\tau)^2 w_i^2 U_{cc} + 2(1-\tau)w_iU_{ch} + U_{hh} + \tau^2 w_i^2 U_{gg}\right]}, \tag{12}$$

and

$$\Delta\epsilon^H = \epsilon_w^H - \epsilon_{1-\tau}^H = \frac{-w_i\left\{(U_g + U_{gg}g_i) - h_i\frac{U_g}{U_c}\left[(1-\tau)w_iU_{cc} + U_{ch}\right]\right\}}{h_i\left[(1-\tau)^2 w_i^2 U_{cc} + 2(1-\tau)w_iU_{ch} + U_{hh} + \tau^2 w_i^2 U_{gg}\right]} \tag{13}$$

still contain the income effect term involving $h_i\left[(1-\tau)w_iU_{cc} + U_{ch}\right]$, except modified by a $\frac{U_g}{U_c}$ factor. This income-effect component reflects the compensation needed to hold utility fixed, rather than the direct labor-leisure substitution response; it therefore affects the Hicksian elasticity, but not by changing the marginal work incentive directly. In fact, an immediate implication of Equation (12) is that the Hicksian net-of-tax rate elasticity may still be negative. This occurs because tax-funded government expenditure is outside the control of the individual (an agency problem), and hence does not envelope out when holding utility constant. For further intuition, consider the extreme case of $U_g = U_c$ (i.e. when the person derives as much marginal utility from tax-funded government expenditure as from consumption). In this case, the Hicksian net-of-tax rate elasticity in Equation (12) is the same as the Marshallian counterpart in Equation (9) — no "Hicksian compensation" is needed to keep utility constant because government expenditures automatically compensate for the income

effect. (Valuing consumption and taxation equally means that there is no longer an agency problem.) On the other hand, the Hicksian wage elasticity in Equation (11) is larger than the Marshallian counterpart in Equation (8) — on top of compensating for the income effect, we have to add a $-\frac{\tau}{1-\tau}h_i\left[(1-\tau)\,w_iU_{cc}+U_{ch}\right]$ term to compensate for the difference induced by the agency problem.

Lastly, our model does not rule out the possibility that other social preferences channels such as employer-specific reciprocity or pay inequality affect labor supply. These channels can contribute to the wedge (even in our vignette experiment which shuts down job- and environment-specific reasons) and do not detract from our main point that the wage and net-of-tax rate elasticities are different in general. That said, we focus on tax-funded government expenditure for the direct policy relevance, and also because it is of first-order importance: g_i depends on labor supply via Equation (3), so the derivatives U_g and U_{gg} enter the elasticities directly. This captures the intuition that increasing labor supply provides a direct utility benefit (or cost) through higher g_i, holding constant its effect on disposable income. By contrast, Appendix B.3 shows that when utility also depends directly on wages, as in $U\left(c_i,h_i,g_i,w_i\right)$, wage preferences affect labor-leisure substitution only through complementarities or substitutabilities with consumption, hours, or tax-funded government expenditure. This is because the wage is exogenous in our model and vignette experiment: working more changes tax-funded government expenditure, but not the wage. A useful contrast is with human capital accumulation models, where working more can raise future wages and hence future consumption (Imai and Keane 2004; Keane and Wasi 2016); no similar endogenous-wage channel is present in our setting. Formally, the derivative U_w does not appear in the Marshallian or Frisch wage elasticities. In the Hicksian elasticity, it appears only in the compensation term, not in the labor-leisure substitution term that relates to direct work incentives.[7] Reassuringly, we find that government-related social preferences must account for at least part of the wedge — the heterogeneity patterns in Sections 4.1 and 4.4, and the meta-analysis results in Section 5 are difficult to explain with wage-related reasons alone.

We next turn to our vignette survey experiment to estimate the size of the wedge.

3 Survey, vignette experiment, and estimation strategy

Our empirical goal is to estimate how individuals' labor supply decisions react to wage changes compared to tax changes. Since observed labor supply choices are influenced by various other determinants, it is difficult to identify the labor supply elasticities using obser-

[7] Specifically, U_w modifies the factor multiplying the income-effect term (the second term) in Equation (11), from $-\frac{\tau}{1-\tau}h_i\frac{U_g}{U_c}$ to $-\left(\frac{\tau}{1-\tau}h_i\frac{U_g}{U_c}+\frac{U_w}{(1-\tau)U_c}\right)$.

vational data without imposing some additional assumptions. To overcome this challenge, we designed a vignette experiment that allows us to obtain causal estimates for the two types of elasticities (wages and net-of-tax rates) that are internally consistent. Our experimental design was inspired by Kosar et al. (2019), who estimated labor supply elasticities using a vignette method in order to investigate heterogeneity across demographic groups. The key advantages of this approach are that the variations are all exogenous, and we do not need to make assumptions about other factors that affect individuals' work hours. Throughout the survey, we employed several approaches to ensure that respondents understood and paid attention to the questions we asked — Appendix C.1 discusses the details. (Appendix C also provides webpage links to the full questionnaire and a live version of the online survey.)

3.1 Survey structure and sample

We conducted a large-scale vignette experiment embedded in an online survey fielded between March and June 2023. Our survey targeted US residents aged 25–64 who were working, excluding self-employed workers as they might have difficulty relating to our vignettes. We recruited participants via Bilendi (previously called Respondi), a survey company commonly used in social science research (see e.g. Alesina et al. 2021; Stantcheva 2021; Jäger et al. 2023).

Bilendi maintains panels of respondents to whom they distribute survey links via email. Respondents were informed about the expected length of the survey but not its topic, to avoid selection based on subject matter. On clicking the survey link, respondents were directed to a consent page that identified the study as an academic research project conducted by non-partisan researchers and assured that their responses are strictly confidential. We asked participants to read the questions carefully and answer honestly and informed them that responses would be subject to statistical checks and low-quality responses might be dropped from the study. Respondents were compensated by Bilendi upon completion of the survey at rates consistent with the platform's standard practices. We asked Bilendi to target participants such that our sample is representative along age, gender, and income group dimensions, providing target proportions based on the working population in the 2022 Current Population Survey data. Respondents who did not meet our eligibility criteria (US residents aged 25–64, currently employed, excluding the self-employed) were screened out.[8] The median time for our survey completion was 20.1 minutes.

Our survey was structured as follows:

1. Demographic and employment questions: We collected information on demographic

[8]See Stantcheva (2023) for a general discussion of recruitment strategies employed by survey companies including Bilendi.

and work characteristics, including work hours, wage/salary income of self and spouse, assets, and consumption. The survey instrument was programmed to use this information to compute each respondent's hourly wage and simulate their income tax rate, needed for our vignette experiment.[9] Appendix C.2 provides details about our tax simulation procedure.

2. Vignette experiment: Respondents first clicked through an animated slideshow that walked them through the instructions for the vignettes. These ten non-skippable slides explained the hypothetical scenarios, the assumptions the respondents should be making and the choices they would be asked to make. Respondents then proceeded to answer the questions in our vignette experiment, described in Section 3.2.

3. Views on government and tax questions: Lastly, we elicited opinions on the federal government and tax system, and the extent to which they liked or disliked their tax money funding different spending categories (e.g. national defense, health insurance programs, and social security). The survey closed with miscellaneous questions, including an open-ended feedback question on the survey.

7,122 respondents completed our survey. To arrive at our main analysis sample, we drop individuals whose work hours, consumption, or tax amounts are below the 5th or above the 95th percentiles, to avoid the influence of outliers on our simulated elasticities.[10] We also exclude a small number of respondents who skipped our compulsory instructions slide show — this could be achieved if the respondent knew how to manipulate JavaScript in a browser console — and respondents whose consumption information in the vignettes could not be imputed (e.g., if they always reported the maximum or minimum values on the consumption scale). Lastly, we drop one respondent whose utility function is not locally concave in our main specification; this issue does not arise for other respondents, and including this individual makes some results less precise but qualitatively unchanged.

Appendix Table A1 compares the characteristics of our final sample of 5,439 individuals with those of the US working population. By construction, our sample is representative along the targeted dimensions of age, gender, and income. The sample also broadly mirrors the population in terms of non-targeted dimensions such as marital status and race. However, consistent with patterns documented for online survey samples more generally (see e.g.

[9] Respondents were asked to verify the hourly wage and amend their reports of income or work hours if the hourly wage did not make sense. Simulated tax rates were not similarly verified to avoid emphasizing taxes early in the survey. We did not inquire directly about taxes paid for the same reason.

[10] For example, if the utility function in Equation (1) contained $\log c$, the elasticities in Equation (5) and (6) would contain $1/c$ terms. Observations with very small consumption values would then have a big influence on the elasticities that we simulate. Additionally, consumption is likely measured with considerable error given survey time constraints, reinforcing the importance of excluding extreme reported consumption values.

Stantcheva 2022, Table 1), our respondents are somewhat more educated, less likely to be Hispanic, and more likely to identify as Democrat. In Section 4.3, we show that our results are robust to post-stratification reweighting that adjusts for these imbalances. Appendix C.3 provides more details about our sample.

3.2 Vignette experimental design

The vignette experiment showed respondents ten pairs of hypothetical scenarios. In all scenarios, respondents were asked to assume that everything about them was the same as their situation at the time of the survey except that they had to leave their job and find a new one. Thus, our vignette experiment by design abstracts from market frictions, and corresponds to a model in which individuals can freely choose their hours of work (Blundell and MaCurdy 1999, p. 1588). Appendix Figure A1 shows the first two instructional slides. The scenario description read:

> "You have to leave your current job and find a new one. You have received a job offer that will pay you $*[wage]* per hour and require that you work *[hour]* hours per week. You will have to pay the federal income tax on the amount you earn from this job at the rate of *[tax rate]*%."

The scenarios randomized the three bracketed variables in the description above around the respondents' actual values to reduce hypothetical bias.[11] Wages and tax rates were drawn uniformly from the 21 and 11 even numbers closest to the actual wage and tax rate, respectively, and work hours were drawn uniformly from the 9 multiples of five closest to the actual work hours. We constrained the wage and hours grids to be positive, and the tax rate grid to be nonnegative.[12] In each pair of scenarios, respondents were asked if they would take up the job in each scenario and which of the two scenarios they preferred.

We instructed respondents to assume the following as they thought about the hypothetical scenarios:

- Jobs and family members' situations were otherwise identical to actual situations (hence our estimates are partial equilibrium responses);

[11] A potential downside of randomizing around the actual values is that identification of curvature in $v(g_i)$ is local to each respondent's status quo rather than global. That said, our tax rates vary by up to 20 percentage points, so the local curvature we identify remains relevant for policy analysis.

[12] For example, a respondent earning 31 dollars per hour at a 40 hour-per-week job and facing a tax rate of 8 percent would see wages drawn from 12, 14, ..., 52, work hours drawn from 20, 25, ..., 60, and tax rates drawn from 0, 2, ..., 20 (these being the eleven even numbers closest to the actual tax rate while still being nonnegative).

- The income tax rate in the scenarios was the federal income tax, and applied to people "with similar income and employment situation to you" (reflecting both reality and our modeling of the common tax rate in Section 2);
- The federal government balanced its budget; and
- If they chose not to take the job, they would have access to the same financial resources that are available if they were to quit their actual job, and the same chance of finding a new job.

Appendix Figure A2 shows the basic layout of our scenario pairs. To facilitate understanding, we computed all relevant details — specifically monthly pre-tax earnings, take-home income, and tax amounts paid. The pre-vignette instructional slideshow and the information icons on the vignette screens also explained these computations in detail. Participants answered the questions step-by-step, starting with the choice of whether to take up the job in each of the two scenarios before selecting their preferred scenario. Additionally, we repeated the instructions at the top of the screen for respondents' reference.

Pairs 1 to 4 followed the basic layout. Pairs 5 to 8 (shown in Appendix Figure A3) involved an additional step — delayed to minimize cognitive load — by also collecting information on consumption, required for estimating our model. Specifically, we asked respondents how much their household would spend in total each month if they took up the job in each scenario. Pairs 9 and 10 added an earmarked tax for a specific program that allows us to investigate whether the size of the wage-tax elasticity wedge varies with whether the taxes fund specific programs that respondents like. We discuss the design of pairs 9 and 10 together with the analysis in Section 4.4.[13] The vignette part of our survey is necessarily more complex than other sections; to help respondents along and to maintain attention, we showed an interactive "guided website tour" every time we introduced a change in the layout (pairs 1, 5, and 9).

We built in two attention checks in our survey. The first doubles as a comprehension check, and asked participants a simple question following the instructional slideshow mentioned in Section 3.1 point 2 — whether the income tax rate in the scenarios referred to federal, state or local tax rate (the correct answer being federal). Second, pair 7 of our vignette experiment presented a strictly dominating relationship, with both scenarios showing the same tax rate and work hours but different hourly wages. We expect attentive participants to prefer the scenario with the higher wage. About 7 and 8 percents of our sample failed the two checks, respectively, a rate comparable with Mas and Pallais (2017) and the

[13] Appendix C.4 accounts for the compliance with the preregistration.

literature on inattention in online surveys (see e.g. Peer et al. 2022). (Combined, 13% of respondents failed at least one of the two attention checks).

Estimating the model in Section 2 requires consumption information, which we collected in only pairs 5 to 8 by design. To use information in all pairs, we impute consumption for all scenarios. We first assess if respondents are correctly basing their consumption decisions on take-home earnings rather than pre-tax earnings in pairs 5 to 8. Appendix Table A2 reports that, in a horse-race regression of log consumption on log take-home earnings and log pre-tax earnings, the effect of "more earnings" loads on take-home earnings.[14] We therefore predict log consumption for all scenarios using log take-home earnings at the individual-level.[15] That is, the variation in consumption comes from each respondent's take-home earnings (which varies across scenarios) and a respondent-specific slope based on eight data points collected in pairs 5 to 8.

The hypothetical nature of the choices in our vignette deserves emphasis here. Accepting or rejecting a scenario generates no realized change in income, consumption, or wealth. This is especially relevant for income effects, so they may be less reliably captured if respondents focus on how they would adjust work to wage or tax changes without fully accounting for resulting downstream changes in household resources. The design helps them do so — each scenario displays the implied monthly pre-tax earnings, tax payment, and take-home earnings, and the instructions hold financial resources outside the job fixed — but we nevertheless place greater interpretive weight on substitution responses than on the implied income effects.

3.3 Main empirical strategy

Using the vignette data, we estimate the wage and net-of-tax rate elasticities for the three types that are commonly used in the literature (Marshallian, Hicksian, and Frisch). We start by estimating the structural parameters of a choice model using the data. These structural parameters are then used to simulate the elasticities of interest.

To estimate the structural utility parameters using our vignette experiment, we param-

[14] Our estimate implies that a 1 percent increase in take-home income is associated with a 0.37 percentage increase in consumption. This elasticity is of a similar magnitude to the elasticity of consumption with respect to permanent income shocks estimated by Baker (2018).

[15] We control for pair and scenario order (whether the scenario is the first or second) fixed effects, and drop observations in which consumption is at the minimum or maximum of the range slider used to collect consumption information.

eterize the utility function in Equation (4) with

$$\begin{aligned} U\left(c_i, h_i, g_i\right) = & \beta_c \log c_i + \beta_h \log\left(\bar{L} - h_i\right) + \beta_{cc}\left(\log c_i\right)^2 \\ & + \beta_{hh}\left(\log\left(\bar{L} - h_i\right)\right)^2 + \beta_{ch} \log c_i \log\left(\bar{L} - h_i\right) \\ & + \beta_g g_i + \beta_{gg} g_i^2 + \beta_{ggg} g_i^3, \end{aligned} \tag{14}$$

where the β's are coefficients to be estimated and $\bar{L}$ is the maximum time available so that $\bar{L} - h_i$ represents leisure. The first two lines of Equation (14) model preferences over consumption and leisure with a translog specification as in prior work (Hotz et al. 1988; Shaw 1989; Van Soest 1995; Ziliak and Kniesner 2005; Elder et al. 2023). Our results are robust to a quadratic utility specification that models c_i and $\bar{L} - h_i$ in levels instead of logs (Blundell et al. 2000; Bargain et al. 2010), but we prefer the translog specification for our main analysis due to better precision. The third line models social preferences as a flexible function of tax-funded government expenditure. Importantly, the polynomial specification for social preferences imposes no restriction on whether respondents like or dislike (in the manner of Levine 1998) tax-funded government expenditure, or the degree of diminishing marginal utility (Carpenter 2021).

In the main analysis of our vignette experiment, each respondent i saw eight pairs t of scenarios j (pairs 1 to 8 of our vignettes). The scenarios randomly varied three parameters (wage, tax rate, and work hours), which map to consumption c_{itj}, work hours h_{itj}, and tax-funded government expenditure g_{itj} (the last two were presented to the respondent directly). To estimate the model, we assume that respondents made their choices in the vignette experiment — namely to work or not for each scenario Work_{itj} and which scenario she preferred Choice_{itj}(our two outcome variables) — by maximizing $U\left(c_{itj}, h_{itj}, g_{itj}\right) + \boldsymbol{x}_{itj}\boldsymbol{\kappa} + \xi_{itj}$, where $\boldsymbol{x}_{itj}$ is a vector of controls (listed below) with associated coefficient vector $\boldsymbol{\kappa}$, and ξ_{itj} is independent and identically distributed Type I extreme value, conditional on the respondent-pair. In other words, choices are made based on the utility function evaluated at the scenario's parameter values, plus a ξ_{itj} term that captures all other factors that might affect utility when making the choice, including possible model misspecification. This allows us to estimate the coefficients in Equation (14) using standard fixed-effects logit regressions (McFadden 1974):

$$\Pr\left(\text{Work}_{itj} = 1 | c_{itj}, h_{itj}, g_{itj}, \boldsymbol{x}_{itj}\right) = \mathcal{F}\left[\gamma_{it} + U\left(c_{itj}, h_{itj}, g_{itj}\right) + \boldsymbol{x}_{itj}\boldsymbol{\kappa}\right] \tag{15}$$

and

$$\Pr\left(\text{Choice}_{itj} = j | c_{itj}, h_{itj}, g_{itj}, \boldsymbol{x}_{itj}\right) = \mathcal{F}\left[\gamma_{it} + U\left(c_{itj}, h_{itj}, g_{itj}\right) + \boldsymbol{x}_{itj}\boldsymbol{\kappa}\right], \tag{16}$$

where $\mathcal{F}\left(z\right) = \exp\left(z\right) / \left[1 + \exp\left(z\right)\right]$ is the cumulative logistic distribution, and γ_{it} is the individual-pair level fixed effect.

In $\boldsymbol{x}_{itj}$, we control for the order of the scenario (first of the pair shown versus second) and for whether the scenario's work hours was below 35 hours per week, the latter following previous papers which model a part-time utility cost (Van Soest 1995; Euwals and Van Soest 1999; Elder et al. 2023).[16] Equation (1) includes the expected indirect utility of the next period; we account for this by controlling for $\log(assets_i + spouseincome_i + savings_{itj})$ and its square, where $assets_i$, $spouseincome_i$, and $savings_{itj}$ are respectively wealth, spouse's income, and the after-tax disposable income less consumption in the scenario. Although this variable is not consequential for estimating the wage-tax elasticity wedge, it ensures the empirical specification is consistent with the intertemporal model without explicitly modeling expectations about future wages or tax rates. Intuitively, time-separability in the model allows for two-stage budgeting, so that the total wealth carried into the next month acts as a summary statistic for all future events (Blundell and MaCurdy 1999, p. 1596).[17] Standard errors are clustered at the respondent level.

After estimating the model, we simulate the elasticities and their differences using Equations (5) to (13). These elasticities are computed using the estimated model coefficients (the β's) and the respondents' actual (non-vignette) consumption, work hours, wages, and tax rates. Standard errors are computed using the delta method.

To investigate heterogeneity with respect to an opinion of the government K_i, we allow the coefficients to depend on K_i with the parameterization $\beta_l = \alpha_{0l} + \alpha_{1l}K_i$, where $l \in \{c, h, cc, hh, ch, g, gg, ggg\}$ is an index of a coefficient in Equation (14). We estimate the model as before, and then simulate the difference between high and low opinions of the government as the average partial effect of K_i on the wage-tax elasticity wedge.

4 Results

4.1 Reduced form analysis

We start by investigating whether respondents' choices vary with taxes in a way consistent with social preferences. The relationship between taxes and choices is not straightforward because higher tax rates, which we directly randomized, mechanically increase taxes but reduce disposable income. Since respondents are unlikely to value the additional aggregate

[16] These papers usually also include a fixed cost of working and model coefficient heterogeneity in demographics. The former is not relevant in our analysis because we only showed scenarios with positive work hours. We show robustness to modeling demographic heterogeneity in Section 4.3.

[17] For example, if respondents interpret high wages in a scenario as implying that wages will fall in the future, they will want to save more today to smooth consumption over time; this control variable captures the effect of such expectations.

government expenditure enough to offset their loss in private consumption, choices should be negatively related to the tax rate, as is typically found in the literature.

To isolate the role of social preferences, we therefore need to hold disposable income constant in a sufficiently flexible way. We implement this idea using regressions of whether respondents preferred the scenario over the pair's alternative against the tax rate in the vignette experiment. An observation is a respondent-pair-scenario, and we include individual, pair, and order fixed effects. Crucially, we control for cells constructed from the interaction of the 200-quantiles of disposable income with all possible levels of work hours, and allow these cells to have heterogeneous linear slopes in disposable income and work hours. Controlling for work hours in conjunction with disposable income is important because, holding disposable income fixed, a higher tax rate mechanically implies higher work hours, which would confound our estimates of social preferences with disutility from work.[18] Conceptually, this isolates variation from higher wages that are immediately channeled into taxes. As shown in Figure 1 Panel A and Table 1 Column 1, holding constant disposable income and work hours, respondents were likely to prefer scenarios with higher tax rates. This tax-choice relationship appears relatively linear and is not driven by outliers.

The above result indicates that choices depend on more than disposable income and hours and suggests that social preferences broadly defined matter. However, it does not by itself show that the relevant social preferences are government-related. A preference for scenarios with higher wages that come together with higher taxes could also reflect non-government-related channels, such as status utility from higher gross wages. To sharpen the evidence for the government-related channel, we next examine whether the relationship between taxes and choices is stronger among respondents who should plausibly value tax-funded government expenditure more. We focus on two dimensions: political affiliation and opinion of government.

Our survey was fielded when a Democrat president was in power, and our scenarios specified that vignette taxes were federal taxes. Hence, we expect Democrats to have more positive government-related social preferences. We also elicited respondents' satisfaction with the government and with the way it spends tax money. We standardize the two questions, average them, and dichotomize the average to construct an indicator for positive opinion of government.

Figure 1 Panels B and C and Table 1 Columns 2 and 3 show the results. In both cases, the comparison group has a slope close to zero, while the government-aligned group has a

[18] To see this, consider two scenarios with the same disposable income $(1-\tau)\,wh$ and wage w. The scenario with the higher tax rate τ must then also have higher work hours h. If we held disposable income fixed but not work hours, the tax-choice relationship would mix tax-related social preferences with the disutility from work.

positive slope. The fact that social preferences are driven mainly by the government-aligned group suggests that government-related social preferences play a role in our experiment.

In addition to the general opinion of government discussed above, our survey included four other opinion dimensions that might contribute to this general opinion (two slightly different questions each, aggregated into one measure as before): whether the government has an important role to play, trust in the government, whether the respondent thought she benefited from government policies, and whether the existing redistribution system is fair. Appendix Figure A4 repeats our main reduced-form heterogeneity analysis by these four other factors and shows similar results. Holding constant disposable income and hours, government-aligned respondents preferred scenarios with higher tax rates, while non-aligned respondents did not.

We perform two robustness checks. First, although we use tax rates as the independent variable above because we directly varied them, the theory in Section 2 shows that the amount of taxes paid is what matters. Indeed, Appendix B.3 shows that the tax rate has first-order effects on labor supply only through its correlation with taxes paid. Appendix Figure A5 and Appendix Table A3 Panel A show similar, and sometimes stronger, results using taxes paid as the main regressor. Second, our main analysis isolates variation from higher wages that are channeled into taxes based on the intuition that we should hold constant disposable income and work hours. We could instead directly hold constant the net-of-tax wage $(1-\tau)\,w$. Appendix Figure A6 and Appendix Table A3 Panel B show similar results based on such a specification.

4.2 Main vignette experiment analysis

We next turn to our structural estimates of utility parameters in order to investigate the wage-tax elasticity wedge. Table 2 reports estimates of the utility parameters estimated via logit. The first two columns report estimates based on the choice of whether to work in the scenario, and the next two columns report estimates based on the choice of the preferred scenario. Odd columns report estimates ignoring social preferences, and even columns report estimates with a third-degree polynomial in tax-funded government expenditure. The social preferences parameters are quite precisely estimated.

At the bottom of the table, we report that the average marginal utility of tax-funded government expenditure $E\left(U_g\right)$ is positive, consistent with our reduced form results. We also report $E\left(U_{gg}g\right)$, the curvature of social preferences multiplied by the amount of tax-funded government expenditure. Its negative value implies that the social preferences function is concave (i.e. there is diminishing marginal utility). Despite this, $E\left(U_g\right)$ is still posi-

tive at a high tax rate (50 percent) often used in real-effort experiments (e.g. Djanali and Sheehan-Connor 2012; Kessler and Norton 2016; Rick et al. 2018). Note, however, that even $E(U_g|\tau = 50\%)$ is not directly comparable with the results in such literature for at least two reasons. First, previous papers often start participants off in an environment with no taxes before introducing a large tax increase, which would likely lead to more tax aversion than when a pre-existing tax rate changes (conceptually closer to our setting). Second, our tax variation also allows for social preferences to act through other people's contributions and changes in the aggregate public goods provision, a factor that is missing thus far in the literature. Another implication of concavity is that it would reduce the impact of positive social preferences on the wage-tax elasticity wedge, by Equations (5) and (6). Comparing the magnitudes of $E(U_g)$ and $E(U_{gg}g)$, we still obtain a net-positive $E(U_g + U_{gg}g)$.

Table 3 reports estimates of the wage elasticities, net-of-tax rate elasticities, and the differences between the two. The columns correspond to those from Table 2, with each cell estimated using the parameters of the model and the actual (or tax-simulated) values of the respondents (see Section 3.3 for details). The three panels show estimates for, respectively, the Marshallian, Hicksian, and Frisch elasticities of labor supply.

Columns 1 and 3 show the estimates of the elasticities without incorporating social preferences, as assumed by all previous models of labor supply that we are aware of. Without social preferences, the differences between the wage and net-of-tax rate elasticities are necessarily zero. The estimated elasticities are at the high end of what others have estimated in real labor market situations (see e.g. Chetty et al. 2013; Bargain and Peichl 2016; Elminejad et al. 2023), which is not surprising given that the literature thus far conflates wage and tax variation. Furthermore, the labor supply elasticities in this paper are better interpreted as (larger) frictionless elasticities, since many market frictions — including salience and constraints on work hours (e.g. imposed by employers) — should not be in play as our respondents made their choices. The larger Frisch elasticity estimate in column 3 compared to column 1 also points to the role of frictions. Whether one should accept a job offer depends on one's available time and monetary resources and the job-finding probability — these frictions are less likely to play a role when asked "which scenario do you prefer". Our estimates of the Frisch wage elasticity are in line with several market-based estimates that are more likely to be frictionless — 1.2 for bicycle messengers (Fehr and Goette 2007), 1.3 for fishermen (Stafford 2015; Giné et al. 2017), and 1.9 for Uber drivers (Chen et al. 2019).[19]

Our key results are shown in columns 2 and 4 of the table. With social preferences, the

[19]These estimates are based on shorter decision-making time horizons and, as the authors emphasize, are for workers who are more likely to freely choose their labor supply decisions. Additionally, Mui and Schoefer (2025) use representative surveys in the US and Germany to elicit wage changes that would induce a person to work (or stop working), and find implied Frisch elasticities that exceed 3 for small wage changes.

wage elasticity is always larger than the net-of-tax rate elasticity, and the difference — the wage-tax elasticity wedge — is statistically significant for all three types. In our preferred specification in column 4 (which generally has better precision), the Hicksian net-of-tax rate elasticity becomes essentially zero, and the Frisch wage elasticity is 50 percent larger than the corresponding net-of-tax rate elasticity.

This wedge is meaningfully large. In a recent meta-analysis, Elminejad et al. (2023) report 8 quasi-experimental studies in their appendix that look at the intensive margin Frisch elasticity, of which we classify 6 as based on tax variation. (We focus on the quasi-experimental intensive-margin estimates because the wage-or-tax variation is more transparent, the framework fits our model best, and we assess the "quasi-experimental" nature of the intensive-margin estimates to be more believable.) These 6 studies that use tax variation have an average Frisch elasticity of 0.5 (our computation). Our results imply that this average maps to a wage-Frisch of around 0.8, closer to levels often used in macroeconomic real business cycle models.

4.3 Robustness

Table 4 shows robustness of our estimated wage-tax elasticity wedge to a variety of specification changes. Row 0 of the table repeats our main estimates in Table 3, column 4, and all subsequent rows show a deviation in specification from row 0.

The first five rows examine robustness to varying the specification of the utility function in Equation (14). In row 1, we use a more flexible social preferences term $v(g_i)$, increasing the polynomial order of g_i from three to five. Row 2 allows tax-funded government expenditure to be complements or substitutes to consumption and leisure by including interactions of g_{itj} with $\log(c_{itj})$ and $\log\left(\bar{L} - h_{itj}\right)$. Another commonly-used parameterization in discrete choice labor supply models is the quadratic utility function $\tilde{u}(c,h) = \beta_c c + \beta_h\left(\bar{L} - h\right) + \beta_{cc}c^2 + \beta_{hh}\left(\bar{L} - h\right)^2 + \beta_{ch}c\left(\bar{L} - h\right)$, although Löffler et al. (2018) and Elder et al. (2023) report that estimated elasticities are not sensitive to the choice of translog versus quadratic utility. To check this, row 3 replaces the terms involving $\log(c_{itj})$ and $\log\left(\bar{L} - h_{itj}\right)$ in Equation (14) with $\tilde{u}(c_{itj}, h_{itj})$. (Appendix Table A4, formatted similarly to Table 3, shows that the estimated elasticities and wedges are generally less precise under the quadratic model.) Previous papers also frequently model parameter heterogeneity in demographics. Row 4 allows β_c and β_h to vary with sex, marital status, their interaction, education, and quadratic terms in log age and number of children.[20] Row 5 uses an alternative social preferences term

[20] Appendix Table A5 uses this model to simulate elasticities for the four combinations of sex and marital status. Unlike papers using market data, we do not find larger labor supply elasticities for women than for men. This suggests that institutional setting or market frictions, more than workers' preferences, drive the

$v(g_i, \tau)$ implied by the more complex model (discussed in Appendix B.1) that provides a foundation for the theory in Section 2 (elasticities are shown in Appendix B.4). In general, the estimated wage-tax elasticity wedge is qualitatively similar to our main estimates.

Our main specification balances survey burden and statistical power by collecting consumption information only for pairs 5 to 8 and then imputing using the observed income-consumption relationship. Row 6 shows that our estimates are not sensitive to the imputation procedure by directly using the non-imputed consumption information (restricting the sample to pairs 5 to 8). Row 7 shows that our results are qualitatively similar if we alternatively impute using empirical Bayes prediction, treating the individual-specific slope of log consumption on log take-home earnings as random. Besides showing robustness regarding the imputation procedure, this specification also addresses potential noise in respondents' stated consumption in the scenarios. By combining individual-level consumption responses with the aggregate income-consumption relationship, the empirical Bayes prediction reduces the influence of noisy individual consumption estimates.

Our main specification accounts for the expected indirect utility term in Equation 1 by controlling for a quadratic in the log of the sum of the individual's assets, spouse's income, and the scenario's unconsumed income (or money borrowed, if applicable). Row 8 shows that omitting this control results in an even larger wedge. Because our vignettes were designed with an intertemporal framework in mind to better reflect real-world decision making, our preferred specification includes this control.

By preference transitivity, a respondent who would work in one scenario of a pair but not the other should also prefer the former in the "choice between scenarios" question. When designing the survey, we worried that allowing the inconsistent choices might lead respondents (who noticed) to view the survey as poorly designed and lose motivation. To avoid this, we included a non-intrusive prompt about the inconsistency whenever it was made. A natural worry is whether this leads to experimenter effects (e.g. the respondent may want to avoid seeing the text even though it does not interrupt the flow of his choices). To assess this, we tracked when the prompt was observed. In our final sample, 19% of respondents saw the prompt at least once, with the median first-observation of the prompt at pair 4, and 4% of respondents still made the intransitive choice despite the prompt. In row 9, we show that estimates are similar when we restrict the sample to pairs before any prompt was observed, suggesting that neither experimenter effects nor intransitivity meaningfully affect our results.[21]

differences across gender.

[21]The logit functional form used in this paper imposes an independence of irrelevant alternatives (IIA) assumption, and papers that estimate discrete choice labor supply models frequently relax this by including random parameter slopes (Van Soest 1995; Euwals and Van Soest 1999; Löffler et al. 2018). Since IIA is

Alternatively, intransitive choices might indicate waning attention. Assuming that intransitive choices were unintentional, the 19% of respondents who ever made an intransitive choice are less careful respondents whom we might want to exclude. In row 10, we show that estimates are robust to this exclusion.

We tracked various measures of response quality throughout the survey. Row 11 reports that our results are robust to using a quality sample that excludes respondents who were inattentive, impatient, or responded too quickly. Specifically, we exclude respondents who failed either of our two attention check questions described in Section 3.2, clicked "next slide" more than twenty times in our instructions slides (only ten clicks were needed, and the user interface made this obvious), or answered any of the overall, vignette, or government expenditure liking questions in less than half the median time taken by all respondents.[22]

As discussed in footnote 10, actual consumption is likely measured with noise in our survey. Row 12 checks whether the estimated wedge is driven by noisy individual-level inputs, especially consumption, by computing the wedge using median respondent characteristics. The resulting wedge is similar in magnitude to our main estimate. Relatedly, our main sample trims respondents whose work hours, consumption, or tax amounts fall below the 5th or above the 95th percentiles, to limit the influence of outliers. Row 13 relaxes this to trimming only the 2.5th and 97.5th percentiles, retaining 95% of respondents on each variable. The Marshallian and Hicksian wedges are quantitatively similar to our main estimates. The Frisch wedge is somewhat larger, consistent with outliers having a larger effect on this estimate. Next, scenarios with work hours far from the respondent's current hours may be effectively ruled out on the basis of hours alone, making them less informative about wage and tax preferences. In such scenarios, choices may instead be governed by reference-point or minimum-income considerations. Row 14 restricts the sample to scenarios within ± 10 hours per week of the respondent's actual work hours and obtains wedge estimates close to our main estimates.

As noted in Section 3.1, Appendix Table A1 shows minor demographic imbalances in our sample compared to the US population. To address this, we construct post-stratification weights using a raking procedure based on all the variables in the table (DeBell and Krosnick 2009), and apply it in row 15 of Table 4. Further, because the conditional logit uses implied

related to transitivity through the weak axiom of revealed preference (Peters and Wakker 1991, Lemma 3.4), the robustness of our result suggests that the IIA assumption is benign. This aligns with Elder et al. (2023) who report that their results are similar whether random slopes were included or not, despite the added computational cost. Because of this, and because the primary goal of this paper is not in investigating heterogeneity in social preferences (e.g. as in Kosar et al. 2019; Burbano et al. forthcoming), we do not use a random slopes specification.

[22]Responses submitted too quickly are generally associated with lower quality. While defining "too fast" is difficult, using a threshold relative to median time is frequently done in practice.

consumption, leisure, and tax-funded government expenditure rather than the randomized wage, tax-rate, and hours variables directly, and because the randomization is done around respondent-specific baseline values, the realized regressors are not fully orthogonal and do not have uniform support over the population. To assess whether this introduces bias, row 16 reweights observations to approximate a full-factorial design that balances these regressors. The estimated wedge remains similar with larger standard errors in both cases.

4.4 The role of attitudes towards the government and programs

Similar to our discussion at the end of Section 4.1, if government-related social preferences play a role in driving the wage-tax wedge, the wedge should be larger among respondents who have more positive expectations about how taxes are used. We examine this in three ways, with the first two building on the heterogeneity dimensions introduced in Section 4.1.

Table 5 Panel A examines heterogeneity in the wage-tax elasticity wedge by political affiliation. The choice variable in the logit specification is whether the respondent preferred the scenario over the other. Democrats have larger wedges than non-Democrats, although the difference is only marginally significant for the Marshallian and Hicksian elasticity wedges.

Political affiliation is a noisy measure of one's opinion of the government — Democrat-aligned respondents might still disapprove of the policies implemented by a Democrat administration.[23] For this reason, our survey directly elicited the respondents' opinions of the government including satisfaction with the government and the way it spends tax money, whether the government has an important role to play, trust in the government, whether the respondent thought she benefited from government policies, and whether the existing redistribution system is fair. In Table 5 Panel B, with the exception of *Importance of government*, all heterogeneity coefficients are positive, and most of the heterogeneity coefficients for the Frisch and Hicksian wedges are statistically or marginally significant as well.[24]

The wedge should also be larger if taxes fund specific programs that respondents like. (In a similar vein, previous papers have found that the program that taxes fund matters for tax compliance [Alm et al. 1993; Hall and Preston 2000; Giaccobasso et al. 2022; Falsetta et al. 2023].) To investigate this, the last two pairs of scenarios in our survey showed respondents randomly-drawn wages and work hours as before, but the randomly-drawn tax rates were now earmarked for a specific program (e.g. education). Appendix Figure A7 shows the instructions and layout. Each pair of scenarios received the same program, and the programs

[23] In the months that we fielded the survey (March–June 2023), more Americans disapproved (56%) than approved (40%) of President Biden's job performance, suggesting disapproval even among Democrats. In fact, his approval rating achieved its lowest point (37%) in April 2023 (Gallup 2025).

[24] Appendix Table A6 shows the simulated wedges that correspond to each level of the heterogeneity variable.

differed across pairs 9 and 10 for greater within-person variation.[25] The program-specific tax rate was drawn from a more restricted grid (2, 4, ..., 10 percent), and to maintain continuity with previous pairs, respondents were shown a general tax rate set at the center of the tax rate grid from previous scenarios.

Near the end of the survey (six screens later), we asked respondents how much they liked or disliked their tax money being used to fund five different programs, including these two. We placed these items after the vignette experiment to avoid priming respondents with their views on government spending, which could otherwise influence their vignette responses. The five programs were randomly drawn from a list of ten US federal budget functions — we did not elicit opinions on the full list to minimize respondent burden. Liking was measured on a 5-point Likert scale, with the two extremes corresponding to "dislike a lot" and "like a lot" (we randomized the scale direction between respondents). To ensure that respondents understood the programs, we included a short description in both the last two pairs of vignettes and the liking questions. Appendix Figure A8 plots the mean liking of the ten programs that we use, additionally broken down by political affiliation. On average, respondents liked their tax monies being used for retirement-related programs the most (social security and Medicare) and international affairs (which includes international aid) the least, with Democrats liking tax-funded expenditure more than Republicans on all programs except for national defense. This pattern is broadly consistent with what one might expect and suggests that respondents were careful in their responses.

We model the general and specific tax rates separately in the utility function:

$$V\left(a_i, w_i, \tau, \sigma, N_i\right) = \max_{c_i, a_i', h_i} U\left(c_i, h_i, g_i, s_i\right) + \beta E\left[V\left(a_i', w_i', \tau', \sigma', N_i'\right)\right] \tag{17}$$

$$\text{s.t. } c_i + \frac{1}{1+r} a_i' = a_i + \left(1 - \tau - \sigma\right) w_i h_i + N_i, \tag{18}$$

$$g_i = \tau w_i h_i, \tag{19}$$

$$s_i = \sigma w_i h_i, \tag{20}$$

where the new terms s_i and σ stand for the program-specific taxes paid and tax rate respectively. (Appendix B.4 shows the functional forms for the elasticities based on this model.)

[25] We elected not to show two programs on the same screen (i.e. in the same pair) to avoid experimenter effects. Assuming respondents focused on the decision between the two scenarios on the screen, different programs on different screens would avoid this.

Utility is parameterized with

$$U\left(c_{itj}, h_{itj}, g_{itj}, s_{itj}\right) = \beta_c \log c_{itj} + \beta_h \log\left(\bar{L} - h_{itj}\right) + \beta_{cc}\left(\log c_{itj}\right)^2 + \beta_{hh}\left(\log\left(\bar{L} - h_{itj}\right)\right)^2 + \beta_{ch} \log c_{itj} \log\left(\bar{L} - h_{itj}\right) + \beta_g g_{itj} + \beta_s s_{itj}, \tag{21}$$

$$\beta_l = \alpha_{0l} + \alpha_{1l} K_{it}, \quad l \in \{c, h, cc, hh, ch, g, s\}, \tag{22}$$

where K_{it} is respondent i's liking score for the program shown in pair t, modeled as a continuous variable.

Table 5 Panel C reports the relationship between the wage-tax elasticity wedge and liking score for the program. For better precision, we use the quality sample that excluded respondents who were inattentive, impatient, or responded too quickly from Section 4.3. As the table shows, the wage-tax elasticity wedge is larger if taxes go into funding programs that respondents like more, consistent with government-related social preferences contributing to our findings.

Overall, while we do not claim that government-related social preferences are the only type of social preference driving the wedge in our vignette experiment, the results in this section suggest that they account for at least part of the wedge. The wedges are systematically larger for respondents with more favorable views of the government, and larger when taxes are earmarked for programs the respondent likes — a pattern that wage-related social preferences or other channels operating through gross income would not predict.

5 Government-related social preferences and the ETI

The vignette experiment allows us to experimentally manipulate wages and tax rates to obtain wage and net-of-tax rate elasticities for the same individuals. Its main drawback is that the result is based on stated preferences, and we have no way of verifying if the respondents would make the same choices in the real labor market. In this section, we examine correlational evidence between proxies for government-related social preferences and estimates of the ETI with respect to the net-of-tax rate collected by Neisser (2021). Similar to Section 4.4, this ETI meta-analysis provides complementary evidence for the role of government-related social preferences as a channel for labor supply. As with any cross-country correlational design, our specifications cannot fully rule out institutional confounders or reverse causality, though we address several of the most plausible candidates with controls below. The proxies we use are unrelated to status utility or wage-related social preferences so they cannot drive the results in this section.

We use the ETI meta-analysis for two reasons. First, ETI estimates should also be affected by government-related social preferences since they are always estimated using tax variation. Second, this literature has a sizable body of estimates with comparable methodologies. In contrast, the more mature labor supply elasticities literature has much greater variability in methods and modeling decisions, which makes comparability more difficult, and in some cases it is unclear if the variation comes from wage or tax changes.

In total, Neisser (2021) examined 61 studies covering 17 countries. We drop Israel (one study) from the list because it participated in only one wave of the WVS, and its inclusion of WVS questions was only partial in that wave. (Results including Israel are qualitatively similar.) We examine seven government-related social preferences proxies: trust in the government, political parties, parliament, and civil service; whether income should be made more equal (to proxy for a desire for redistribution); whether the government should increase its ownership of businesses (more state ownership implies a more important government), and national pride (a general state-related social preference). Question wordings of the proxies are similar in the WVS and EVS, and all proxies appear in multiple waves. We collapse the proxies to the country-wave level, and match to Neisser's dataset based on the mean study year in the latter.

Empirically, we follow Neisser's main specification in column 6 of her Tables 2 and 3 with three modifications. First, we include the proxy for government-related social preferences. Second, we exclude her country group variable in order to exploit all variation across countries. Third, we combine both after- and before-deductions elasticities for more statistical power.[26] Note that using Neisser's specification means that we already control for a large number of study-related covariates (e.g. the estimation technique used).

Table 6 Panel A shows results based on this specification. In general, the proxies for government-related social preferences are negatively correlated with the ETI. Two proxies, the level of confidence in the civil service and the desire for equality in incomes, are not statistically significant, with the former significant at the 10% level. To ease interpretation, we report the implied change in the ETI when each proxy increases from its 25th to its 75th percentile. Neisser (2021) reports a mean ETI of 0.34 and a standard deviation of 0.975 across all ETI estimates, implying that 95 percent of estimates lie within $1.96 \times 0.975 \approx 1.9$ of the mean. The conceptual 25th-75th percentile shifts in the proxies would account for about 1 to 7 percent of this range, which we interpret as economically meaningful but not transformative.

[26] The fact that the after-deductions ETI is larger than the before-deductions ETI — a key point in Neisser (2021) — is not relevant for our purposes. We therefore analyze both ETI types as one sample, include a dummy for whether the ETI is before- or after-deductions, and interact it with all control variables.

Panel B of the table shows robustness checks, with each row showing a change in specification compared to that in Panel A. The first two rows show that estimates based on only after- or before-deductions elasticities are qualitatively similar to our main estimates. Correlations for the after-deductions elasticities in row 1 remain strong and slightly larger, consistent with after-deductions elasticities having more possible margins of responses — including avoidance responses like deductions-claiming and changes in the timing of income — that might still be affected by social preferences (Slemrod and Kopczuk 2002; Kopczuk 2005; Neisser 2021). That said, these avoidance margins do not drive the correlations entirely — while correlations for the before-deductions elasticities in row 2 are smaller in magnitude, they are all still negative, and several are statistically significant.

Differences in the ETIs can be driven by institutional differences as well, and these might drive the correlations if institutional factors are correlated with social preferences. In row 3, we control for six institutional factors noted in Neisser (2021): the Gini coefficient, the top 10 percent of income shares, whether the tax reform also introduced a top tax bracket, the unemployment rate, the fraction of self-employed, and the share of taxes that are exposed to third-party information reporting (Kleven et al. 2011, 2016). In row 4, we control for the share of public spending on the provision of child care, preschool, and elderly care. This variable is from Kleven (2014), who notes that these public goods are complementary to labor supply, and hence the mechanism could be due to the way that taxes are used rather than government-related social preferences. The correlations are robust to the inclusion of these variables.

In rows 5 and 6, we further address two specific institutional factors that could be correlated with both institutional trust and the ETI: the prevalence of collective bargaining and the breadth of the tax base. In row 5, we control for trade union density from the OECD (2025b), defined as the number of net union members as a proportion of employees (excluding the unemployed, those out of the labor force, and the self-employed). In row 6, we control for a proxy for the breadth of the tax base, constructed from the OECD's Taxing Wages tables (OECD 2026): for each country, we compute taxable income as a percentage of gross earnings across eight household–earnings combinations (single and married households, with no or two children, at varying levels of principal and spouse earnings), and take the first principal component across these as our control.[27] The correlations remain robust to the inclusion of either control.

The correlations are also robust to controlling for audit yield in row 7, a proxy for audit

[27]Both this variable and the audit yield variable in the next row are available only after 2000, so these specifications rely more on cross-country variation in the controls. Since only about one-third of the ETI sample is post-2000, we lack power to estimate the correlations on that subsample alone.

intensity defined as additional assessments from audits and verification actions as a share of net tax revenue (OECD 2025a). This speaks to the concern that high-trust countries may have lower ETIs because of stronger enforcement rather than social preferences. In the "slippery slope" framework of Kirchler et al. (2008), enforcement variables such as audit yield and third-party reporting (row 3) capture the power-based route to compliance, distinct from the trust-based, voluntary route that our proxies are meant to capture. The robustness of our results to the inclusion of these controls is consistent with the latter social-preferences interpretation.

Our main estimates exploit all variation across countries because repeat studies within a country are less common, and furthermore social preferences are unlikely to change in a big way over the time span available in the data. In rows 8 and 9, we probe robustness by including country-group fixed effects that Neisser (2021) includes in some of her specifications, and further including time fixed effects, where time corresponds to the waves of the WVS. The correlations are still robust to the inclusion of these fixed effects.

Finally, row 10 reports a specification that simultaneously includes all of the controls used in rows 3–9 (Neisser's contextual factors, the labor share subsidy, union density, tax breadth, audit yield, and country-group and time fixed effects). The point estimates remain comparable in sign and magnitude to those in Panel A — if anything, larger for several proxies — though standard errors widen as expected.

6 Conclusion

This paper studies the implications of social preferences for the response of labor supply to wages and taxes. Using a vignette experiment, we show that when social preferences are incorporated in the canonical labor supply model, estimates of the labor supply elasticity that exploit exogenous variation in wages are larger than those that exploit variation in tax rates. This wage-tax elasticity wedge is meaningfully large — in our main estimate, the frictionless Frisch wage elasticity is 1.5 times as large as the corresponding net-of-tax rate elasticity. The wedge would be even larger if income tax changes are less salient and induce under-reaction compared to wage changes. Part of it may reflect channels other than government-related social preferences, such as status utility conferred by a higher gross wage, but we find corroborating evidence for the government-related channel: the wedge is larger when the respondent has a better opinion of the government, or when taxes are earmarked for a program the respondent likes more. In a complementary analysis, we find that estimates of the ETI with respect to the net-of-tax rates are correlated with proxies for government-related social preferences, consistent with these preferences influencing real-life

labor market decisions.

Our results have implications for how we use estimates of the labor supply elasticity from the literature in model calibration. In particular, macroeconomic models frequently model the labor supply response to wage changes over the real business cycle. If a net-of-tax rate elasticity is used (e.g. from tax holiday natural experiments), the calibrated parameter would be too small, with implications for the possible recommendations made by such models.

Conversely, the profession is likely already using the correct elasticity when assessing tax changes — most quasi-experimental estimates tend to be net-of-tax rate elasticities. That said, incorporating government-related social preferences can still subtly alter model outcomes. In Appendix B.5, we replicate the main arguments in Piketty and Saez (2013) to draw two conclusions about optimal taxation. First, admitting government-related social preferences into individuals' utility functions naturally changes the social welfare function, and if these preferences are positive as our results suggest, optimal tax rates would be higher relative to a world without government-related social preferences. Intuitively, knowing that people have preferences on taxation beyond its immediate impact on disposable income means that the government can (and should) raise taxes. Second, heterogeneity in social preferences affects the optimal nonlinear tax rate. We leave further detailed explorations along this line for future work.

Lastly, our findings that the wedge varies by (and the ETI is correlated with) opinion of the government and funded programs have tax policy implications. The tax morale literature emphasizes that attitudes towards government and programs affect tax compliance; we show such buy-in also reduces the distortionary effect of taxes on labor supply. This supports policies promoting government transparency and accountability — for example, highlighting how tax monies are used, or reminders that they fund services. To the extent that people have more positive attitudes towards local rather than geographically larger communities, our results also highlight an overlooked benefit of fiscal devolution (e.g. a review by Martinez-Vazquez et al. 2017 barely discusses labor responses). More speculatively, our findings also suggest that tax collection practices that engender negative attitudes (e.g. differential opportunities to reduce or avoid tax liabilities, or distrusting tax systems described in Frey [1997]) are particularly harmful, because of their additional consequences for labor supply.

References

Alesina, Alberto and Paola Giuliano, "Preferences for redistribution," in "Handbook of Social Economics," Vol. 1, Elsevier, 2011, pp. 93–131.

_ , **Matteo F Ferroni, and Stefanie Stantcheva**, "Perceptions of racial gaps, their causes, and ways to reduce them," Working Paper 29245, National Bureau of Economic Research 2021.

Allingham, Michael G. and Agnar Sandmo, "Income tax evasion: a theoretical analysis," *Journal of Public Economics*, 1972, *1* (3), 323–338.

Alm, James, "What motivates tax compliance?," *Journal of Economic Surveys*, 2019, *33* (2), 353–388.

_ , **Betty R Jackson, and Michael McKee**, "Fiscal exchange, collective decision institutions, and tax compliance," *Journal of Economic Behavior and Organization*, 1993, *22* (3), 285–303.

Almås, Ingvild, Alexander W Cappelen, and Bertil Tungodden, "Cutthroat capitalism versus cuddly socialism: Are Americans more meritocratic and efficiency-seeking than Scandinavians?," *Journal of Political Economy*, 2020, *128* (5), 1753–1788.

Altonji, Joseph G, "Intertemporal substitution in labor supply: Evidence from micro data," *Journal of Political Economy*, 1986, *94* (3, Part 2), S176–S215.

Ameriks, John, Joseph Briggs, Andrew Caplin, Minjoon Lee, Matthew D Shapiro, and Christopher Tonetti, "Older Americans would work longer if jobs were flexible," *American Economic Journal: Macroeconomics*, 2020, *12* (1), 174–209.

Andreoni, James, "Giving with impure altruism: Applications to charity and Ricardian equivalence," *Journal of Political Economy*, 1989, *97* (6), 1447–1458.

_ , "Impure altruism and donations to public goods: A theory of warm-glow giving," *The Economic Journal*, 1990, *100* (401), 464–477.

_ **and A Abigail Payne**, "Charitable giving," in "Handbook of Public Economics," Vol. 5, Elsevier, 2013, pp. 1–50.

_ , **Brian Erard, and Jonathan Feinstein**, "Tax Compliance," *Journal of Economic Literature*, 1998, *36* (2), 818–860.

Angrist, Joshua D, Guido W Imbens, and Donald B Rubin, "Identification of causal effects using instrumental variables," *Journal of the American statistical Association*, 1996, *91* (434), 444–455.

Baker, Scott R., "Debt and the response to household income shocks: Validation and application of linked financial account data," *Journal of Political Economy*, 2018, *126* (4), 1504–1557.

Bargain, Olivier and Andreas Peichl, "Own-wage labor supply elasticities: variation across time and estimation methods," *IZA Journal of Labor Economics*, 2016, *5* (1), 1–31.

_ , **Marco Caliendo, Peter Haan, and Kristian Orsini**, ""Making work pay" in a rationed labor market," *Journal of Population Economics*, 2010, *23* (1), 323–351.

Bauer, Thomas K, Julia Bredtmann, and Christoph M Schmidt, "Time vs. money — The supply of voluntary labor and charitable donations across Europe," *European Journal of Political Economy*, 2013, *32*, 80–94.

Bianchi, Marco, Bjorn R Gudmundsson, and Gylfi Zoega, "Iceland's natural experiment in supply-side economics," *American Economic Review*, 2001, *91* (5), 1564–1579.

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

Blumkin, Tomer, Bradley J Ruffle, and Yosef Ganun, "Are income and consumption taxes ever really equivalent? Evidence from a real-effort experiment with real goods," *European Economic Review*, 2012, *56* (6), 1200–1219.

Blundell, Richard, Alan Duncan, and Costas Meghir, "Estimating labor supply responses using tax reforms," *Econometrica*, 1998, pp. 827–861.

_ , _ , Julian McCrae, and Costas Meghir, "The labour market impact of the working families' tax credit," *Fiscal Studies*, 2000, *21* (1), 75–104.

_ and Thomas MaCurdy, "Labor supply: A review of alternative approaches," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 3, Amsterdam, The Netherlands: Elsevier, 1999, chapter 27, pp. 1559–1695.

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

Bozio, Antoine, Thomas Breda, Julien Grenet, and Arthur Guillouzouic, "Does tax-benefit linkage matter for the incidence of payroll taxes?," Working Paper No. 2019-43, Paris School of Economics 2023. Retrieved May 12, 2024, from https://www.parisschoolofeconomics.com/grenet-julien/wp/Bozio_Breda_Grenet_Guillouzouic_2023.pdf.

Bracha, Anat, Uri Gneezy, and George Loewenstein, "Relative Pay and Labor Supply," *Journal of Labor Economics*, 2015, *33* (2), 297–315.

Breza, Emily, Supreet Kaur, and Yogita Shamdasani, "The Morale Effects of Pay Inequality," *The Quarterly Journal of Economics*, 10 2018, *133* (2), 611–663.

Burbano, Vanessa, Nicolas Padilla, and Stephan Meier, "Gender differences in preferences for meaning at work," *American Economic Journal: Economic Policy*, forthcoming.

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler, "Labor supply of New York City cabdrivers: One day at a time," *The Quarterly Journal of Economics*, 1997, *112* (2), 407–441.

Carpenter, Jeffrey, "The shape of warm glow: Field experimental evidence from a fundraiser," *Journal of Economic Behavior and Organization*, 2021, *191*, 555–574.

Chang, Yongsung and Sun-Bin Kim, "From individual to aggregate labor supply: A quantitative analysis based on a heterogeneous agent macroeconomy," *International Economic Review*, 2006, *47* (1), 1–27.

Charness, Gary and Matthew Rabin, "Understanding social preferences with simple tests," *The quarterly journal of economics*, 2002, *117* (3), 817–869.

Chen, M Keith, Peter E Rossi, Judith A Chevalier, and Emily Oehlsen, "The value of flexible work: Evidence from Uber drivers," *Journal of political economy*, 2019, *127* (6), 2735–2794.

Chetty, Raj, "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply," *Econometrica*, 2012, *80* (3), 969–1018.

_ , Adam Guren, Day Manoli, and Andrea Weber, "Does indivisible labor explain the difference between micro and macro elasticities? A meta-analysis of extensive margin elasticities," *NBER macroeconomics Annual*, 2013, *27* (1), 1–56.

_ , Adam Looney, and Kory Kroft, "Salience and taxation: Theory and evidence," *American Economic Review*, 2009, *99* (4), 1145–1177.

_ , John N Friedman, Tore Olsen, and Luigi Pistaferri, "Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records," *The Quarterly Journal of Economics*, 2011, *126* (2), 749–804.

Cohn, Alain, Ernst Fehr, Benedikt Herrmann, and Frédéric Schneider, "Social Comparison and Effort Provision: Evidence from a Field Experiment," *Journal of the European Economic Association*, 08 2014, *12* (4), 877–898.

Cowell, Frank A and James P F Gordon, "Unwillingness to pay: Tax evasion and public good provision," *Journal of Public Economics*, 1988, *36* (3), 305–321.

Cullen, Julie Berry, Nicholas Turner, and Ebonya Washington, "Political alignment, attitudes toward government, and tax evasion," *American Economic Journal: Economic Policy*, 2021, *13* (3), 135–166.

Cullen, Zoë and Ricardo Perez-Truglia, "How much does your boss make? The effects of salary comparisons," *Journal of Political Economy*, 2022, *130* (3), 766–822.

DeBell, Matthew and Jon A Krosnick, "Computing weights for American National Election Study Survey data," ANES Technical Report series, no. nes012427, American National Election Studies, Ann Arbor, MI, and Palo Alto, CA 2009.

DellaVigna, Stefano, John A List, and Ulrike Malmendier, "Testing for altruism and social pressure in charitable giving," *The Quarterly Journal of Economics*, 2012, *127* (1), 1–56.

_ , John A. List, Ulrike Malmendier, and Gautam Rao, "Estimating Social Preferences and Gift Exchange at Work," *American Economic Review*, March 2022, *112* (3), 1038–74.

Diamond, Peter, "Optimal tax treatment of private contributions for public goods with and without warm glow preferences," *Journal of Public Economics*, 2006, *90* (4-5), 897–919.

Djanali, Iwan and Damien Sheehan-Connor, "Tax affinity hypothesis: Do we really hate paying taxes?," *Journal of Economic Psychology*, 2012, *33* (4), 758–775.

Domeij, David and Martin Floden, "The labor-supply elasticity and borrowing constraints: Why estimates are biased," *Review of Economic Dynamics*, 2006, *9* (2), 242–262.

Dube, Arindrajit, Laura Giuliano, and Jonathan Leonard, "Fairness and Frictions: The Impact of Unequal Raises on Quit Behavior," *American Economic Review*, February 2019, *109* (2), 620–63.

Dwenger, Nadja, Henrik Kleven, Imran Rasul, and Johannes Rincke, "Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany," *American Economic Journal: Economic Policy*, August 2016, *8* (3), 203–32.

Eissa, Nada, "Taxation and labor supply of married women: The Tax Reform Act of 1986 as a natural experiment," Working Paper No. 5023, National Bureau of Economic Research 1995.

_ **and Hilary Williamson Hoynes**, "Taxes and the labor market participation of married couples: The earned income tax credit," *Journal of Public Economics*, 2004, *88* (9-10), 1931–1958.

Elder, Todd E, Steven Haider, and Cody Orr, "The evolution of the wage elasticity of labor supply over time," Discussion Paper 16393, IZA 2023. Retrieved Nov 4, 2023, from https://ssrn.com/abstract=4547697.

Elminejad, Ali, Tomas Havranek, Roman Horvath, and Zuzana Irsova, "Intertemporal substitution in labor supply: A meta-analysis," *Review of Economic Dynamics*, 2023, *51*, 1095–1113.

Eriksson, Tor and Nicolai Kristensen, "Wages or fringes? Some evidence on trade-offs and sorting," *Journal of Labor Economics*, 2014, *32* (4), 899–928.

Esslinger, Anna, Katharina Pfeil, and Lars P Feld, "Asymmetric labor supply responses to taxation," Working Paper No. 11317, CESifo 2024.

Euwals, Rob and Arthur Van Soest, "Desired and actual labour supply of unmarried men and women in the Netherlands," *Labour Economics*, 1999, *6* (1), 95–118.

EVS, "EVS Trend File 1981-2017. ZA7503 Data file Version 3.0.0," 2022. GESIS Data Archive, Cologne. https://doi.org/10.4232/1.14021.

Falsetta, Diana, Jennifer K Schafer, and George T Tsakumis, "How government spending impacts tax compliance," *Journal of Business Ethics*, 2023, pp. 1–18.

Farber, Henry S, "Is tomorrow another day? The labor supply of New York City cabdrivers," *Journal of political Economy*, 2005, *113* (1), 46–82.

Fehr, Ernst and Klaus M Schmidt, "A theory of fairness, competition, and cooperation," *The Quarterly Journal of Economics*, 1999, *114* (3), 817–868.

_ **and Lorenz Goette**, "Do workers work more if wages are high? Evidence from a randomized field experiment," *American Economic Review*, 2007, *97* (1), 298–317.

Feldstein, Martin, "The effect of marginal tax rates on taxable income: a panel study of the 1986 Tax Reform Act," *Journal of Political Economy*, 1995, *103* (3), 551–572.

Finkelstein, Amy, "E-ztax: Tax Salience and Tax Rates," *The Quarterly Journal of Economics*, 08 2009, *124* (3), 969–1010.

Fisman, Raymond, Keith Gladstone, Ilyana Kuziemko, and Suresh Naidu, "Do Americans want to tax wealth? Evidence from online surveys," *Journal of Public Economics*, 2020, *188*, 104207.

Folke, Olle and Johanna Rickne, "Sexual harassment and gender inequality in the labor market," *The Quarterly Journal of Economics*, 2022, *137* (4), 2163–2212.

Freeman, Richard B, "Working for nothing: The supply of volunteer labor," *Journal of Labor Economics*, 1997, *15* (1, Part 2), S140–S166.

French, E, A Lindner, C O'Dea, and T Zawisza, "Labor supply and the pension contribution-benefit link," Working Paper No. 30184, National Bureau of Economic Research 2022.

Frey, Bruno S, "A constitution for knaves crowds out civic virtues," *The Economic Journal*, 1997, *107* (443), 1043–1053.

Gallup, "Presidential Job Approval Center," 2025. Retrieved June 25, 2025, from https://news.gallup.com/interactives/507569/presidential-job-approval-center.aspx.

Gelber, Alexander M, "Taxation and the earnings of husbands and wives: evidence from Sweden," *Review of Economics and Statistics*, 2014, *96* (2), 287–305.

Giaccobasso, Matias, Brad C Nathan, Ricardo Perez-Truglia, and Alejandro Zentner, "Where do my tax dollars go? Tax morale effects of perceived government spending," Working Paper 29789, National Bureau of Economic Research 2022.

Giné, Xavier, Monica Martinez-Bravo, and Marian Vidal-Fernández, "Are labor supply decisions consistent with neoclassical preferences? Evidence from Indian boat owners," *Journal of Economic Behavior & Organization*, 2017, *142*, 331–347.

Gneezy, Uri and John A List, "Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments," *Econometrica*, 2006, *74* (5), 1365–1384.

Gruber, Jonathan, "The incidence of payroll taxation: Evidence from Chile," *Journal of Labor Economics*, 1997, *15* (S3), S72–S101.

Hainmueller, Jens, Dominik Hangartner, and Teppei Yamamoto, "Validating vignette and conjoint survey experiments against real-world behavior," *Proceedings of the National Academy of Sciences*, 2015, *112* (8), 2395–2400.

Hall, John and Ian Preston, "Tax price effects on attitudes to hypothecated tax increases," *Journal of Public Economics*, 2000, *75* (3), 417–438.

Hotz, V Joseph, Finn E Kydland, and Guilherme L Sedlacek, "Intertemporal preferences and labor supply," *Econometrica: Journal of the Econometric Society*, 1988, pp. 335–360.

Huet-Vaughn, Emiliano, "Do social comparisons motivate workers? A field experiment on relative earnings, labor supply and the inhibitory effect of pay inequality," *Experimental Economics*, 2025, *28* (1), 155–180.

_ , Andrea Robbett, and Matthew Spitzer, "A taste for taxes: Minimizing distortions using political preferences," *Journal of Public Economics*, 2019, *180*, 104055.

Imai, Susumu and Michael P Keane, "Intertemporal labor supply and human capital accumulation," *International Economic Review*, 2004, *45* (2), 601–641.

Inglehart, R., C. Haerpfer, A. Moreno, C. Welzel, K. Kizilova, J. Diez-Medrano, M. Lagos, P. Norris, E. Ponarin, and B. Puranen, "World Values Survey Trend File (1981-2022) Cross-National Data-Set. Data File Version 3.0.0.," 2022. Madrid, Spain & Vienna, Austria: JD Systems Institute & WVSA Secretariat. https://doi.org/10.14281/18241.23.

Jäger, Simon, Christopher Roth, Nina Roussille, and Benjamin Schoefer, "Worker beliefs about outside options," Working Paper 29623, National Bureau of Economic Research 2023.

Keane, Michael P, "Labor supply and taxes: A survey," *Journal of Economic Literature*, 2011, *49* (4), 961–1075.

_ , "Recent research on labor supply: Implications for tax and transfer policy," *Labour Economics*, 2022, *77*, 102026.

_ **and Nada Wasi**, "Labour supply: The roles of human capital and the extensive margin," *The Economic Journal*, 2016, *126* (592), 578–617.

Kessler, Judd B. and Michael I. Norton, "Tax aversion in labor supply," *Journal of Economic Behavior & Organization*, 2016, *124*, 15–28.

Kirchler, Erich, Erik Hoelzl, and Ingrid Wahl, "Enforced versus voluntary tax compliance: The "slippery slope" framework," *Journal of Economic Psychology*, 2008, *29* (2), 210–225.

Kleven, Henrik Jacobsen, "How can Scandinavians tax so much?," *Journal of Economic Perspectives*, 2014, *28* (4), 77–98.

_ **, Claus Thustrup Kreiner, and Emmanuel Saez**, "Why can modern governments tax so much? An agency model of firms as fiscal intermediaries," *Economica*, 2016, *83* (330), 219–246.

_ **, Martin B Knudsen, Claus Thustrup Kreiner, Søren Pedersen, and Emmanuel Saez**, "Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark," *Econometrica*, 2011, *79* (3), 651–692.

Kopczuk, Wojciech, "Tax bases, tax rates and the elasticity of reported income," *Journal of Public Economics*, 2005, *89* (11-12), 2093–2119.

Kosar, Gizem, Aysegul Sahin, Basit Zafar et al., "The work-leisure tradeoff: Identifying the heterogeneity," in "2019 Meeting Papers" number 410 Society for Economic Dynamics 2019.

Kroft, Kory, Jean-William Laliberté, René Leal-Vizcaíno, and Matthew J Notowidigdo, "Salience and taxation with imperfect competition," *Review of Economic Studies*, 2024, *91* (1), 403–437.

Krueger, Alan B. and Alexandre Mas, "Strikes, Scabs, and Tread Separations: Labor Strife and the Production of Defective Bridgestone/Firestone Tires," *Journal of Political Economy*, 2004, *112* (2), 253–289.

Kube, Sebastian, Michel André Maréchal, and Clemens Puppe, "The Currency of Reciprocity: Gift Exchange in the Workplace," *American Economic Review*, June 2012, *102* (4), 1644–62.

_ **,** _ **, and** _ , "Do Wage Cuts Damage Work Morale? Evidence from a Natural Field Experiment," *Journal of the European Economic Association*, 08 2013, *11* (4), 853–870.

Levine, David K, "Modeling altruism and spitefulness in experiments," *Review of Economic Dynamics*, 1998, *1* (3), 593–622.

Lilley, Andrew and Robert Slonim, "The price of warm glow," *Journal of Public Economics*, 2014, *114*, 58–74.

Löffler, Max, Andreas Peichl, and Sebastian Siegloch, "The sensitivity of structural labor supply estimations to modeling assumptions," Discussion Paper 11425, IZA 2018. Retrieved Nov 18, 2023, from https://www.iza.org/publications/dp/11425/the-sensitivity-of-structural-labor-supply-estimations-to-modeling-assumptions.

Low, Hamish W, "Self-insurance in a life-cycle model of labour supply and savings," *Review of Economic Dynamics*, 2005, *8* (4), 945–975.

Luttmer, Erzo FP and Monica Singhal, "Tax morale," *Journal of Economic Perspectives*, 2014, *28* (4), 149–168.

MaCurdy, Thomas E, "An empirical model of labor supply in a life-cycle setting," *Journal of Political Economy*, 1981, *89* (6), 1059–1085.

Maestas, Nicole, Kathleen J Mullen, David Powell, Till Von Wachter, and Jeffrey B Wenger, "The value of working conditions in the United States and implications for the structure of wages," *American Economic Review*, 2023, *113* (7), 2007–2047.

Martinez, Isabel Z, Emmanuel Saez, and Michael Siegenthaler, "Intertemporal labor supply substitution? Evidence from the Swiss income tax holidays," *American Economic Review*, 2021, *111* (2), 506–546.

Martinez-Vazquez, Jorge, Santiago Lago-Peñas, and Agnese Sacchi, "The impact of fiscal decentralization: A survey," *Journal of Economic Surveys*, 2017, *31* (4), 1095–1129.

Mas, Alexandre, "Pay, Reference Points, and Police Performance," *The Quarterly Journal of Economics*, 08 2006, *121* (3), 783–821.

_ **and Amanda Pallais**, "Valuing alternative work arrangements," *American Economic Review*, December 2017, *107* (12), 3722–3759.

_ **and** _ , "Labor supply and the value of non-work time: Experimental estimates from the field," *American Economic Review: Insights*, 2019, *1* (1), 111–126.

McFadden, Daniel, "Conditional logit analysis of qualitative choice behavior," in Paul Zarembka, ed., *Frontiers in Econometrics*, New York, NY: Academic Press, 1974, chapter 4.

Mengel, Friederike and Elke Weidenholzer, "Preferences for redistribution," *Journal of Economic Surveys*, 2023, *37* (5), 1660–1677.

Mori, Tomoharu, Hirofumi Kurokawa, and Fumio Ohtake, "Labor supply reaction to wage cuts and tax increases: A real-effort experiment," *FinanzArchiv*, 2022, *78* (3), 1–16.

Mui, Preston and Benjamin Schoefer, "Reservation Raises: The Aggregate Labour Supply Curve at the Extensive Margin," *Review of Economic Studies*, 2025, *92* (1), 442–475.

Neisser, Carina, "The elasticity of taxable income: A meta-regression analysis," *The Economic Journal*, 2021, *131* (640), 3365–3391.

OECD, *Tax Administration: Comparative Information on OECD and Other Advanced and Emerging Economies*, Paris: OECD Publishing, 2025. Retrieved May 1, 2026, from https://doi.org/10.1787/cc015ce8-en.

_ , "Trade Unions and Collective Bargaining: OECD/AIAS ICTWSS Database Version 2.0," OECD Data Explorer 2025. Retrieved May 1, 2026, from https://data-viewer.oecd.org?chartId=3caf2eb8-8f5c-4fd0-8ba7-bc9ccaa99c5c.

_ , *Taxing wages 2026: The progressivity of labour taxation in OECD Countries*, Paris: OECD Publishing, 2026. Retrieved May 1, 2026, from https://doi.org/10.1787/3a5169ef-en.

Oettinger, Gerald S, "An empirical analysis of the daily labor supply of stadium vendors," *Journal of Political Economy*, 1999, *107* (2), 360–392.

Ong, Pinchuan and Ivan PL Png, "Deskilling technology affords work amenity, increases labor supply," *Strategic Management Journal*, forthcoming. doi:https://doi.org/10.1002/smj.70017.

Parker, Jonathan A and Nicholas S Souleles, "Reported effects versus revealed-preference estimates: Evidence from the propensity to spend tax rebates," *American Economic Review: Insights*, 2019, *1* (3), 273–290.

Peer, Eyal, David Rothschild, Andrew Gordon, Zak Evernden, and Ekaterina Damer, "Data quality of platforms and panels for online behavioral research," *Behavior Research Methods*, 2022, p. 1.

Peters, Hans and Peter Wakker, "Independence of irrelevant alternatives and revealed group preferences," *Econometrica*, 1991, pp. 1787–1801.

Piketty, Thomas and Emmanuel Saez, "Optimal labor income taxation," in "Handbook of Public Economics," Vol. 5, Elsevier, 2013, pp. 391–474.

Pistaferri, Luigi, "Anticipated and unanticipated wage changes, wage risk, and intertemporal labor supply," *Journal of Labor Economics*, 2003, *21* (3), 729–754.

Rick, Scott, Gabriele Paolacci, and Katherine Burson, "Income tax and the motivation to work," *Journal of Behavioral Decision Making*, 2018, *31* (5), 619–631.

Rogerson, Richard and Johanna Wallenius, "Micro and macro elasticities in a life cycle model with taxes," *Journal of Economic Theory*, 2009, *144* (6), 2277–2292.

Saez, Emmanuel, "The optimal treatment of tax expenditures," *Journal of Public Economics*, 2004, *88* (12), 2657–2684.

_ , Joel Slemrod, and Seth H Giertz, "The elasticity of taxable income with respect to marginal tax rates: A critical review," *Journal of Economic Literature*, 2012, *50* (1), 3–50.

Shaw, Kathryn L, "Life-cycle labor supply with human capital accumulation," *International Economic Review*, 1989, pp. 431–456.

Sigaard, Hans Schytte, "Estimating labor supply responses to Danish tax reforms," *Journal of Public Economics*, 2023, *224*, 104943.

Sigurdsson, Jósef, "Labor Supply Responses and Adjustment Frictions: A Tax-Free Year in Iceland," Working Paper 2019. Available at SSRN: https://ssrn.com/abstract=3278308.

Slemrod, Joel and Wojciech Kopczuk, "The optimal elasticity of taxable income," *Journal of Public Economics*, 2002, *84* (1), 91–112.

Soest, Arthur Van, "Structural models of family labor supply: A discrete choice approach," *Journal of Human Resources*, 1995, pp. 63–88.

Stafford, Tess M, "What do fishermen tell us that taxi drivers do not? An empirical investigation of labor supply," *Journal of Labor Economics*, 2015, *33* (3), 683–710.

Stantcheva, Stefanie, "Understanding tax policy: How do people reason?," *The Quarterly Journal of Economics*, 2021, *136* (4), 2309–2369.

_ , "Understanding of trade," Working Paper 30040, National Bureau of Economic Research 2022.

_ , "How to run surveys: A guide to creating your own identifying variation and revealing the invisible," *Annual Review of Economics*, 2023, *15*, 205–234.

Stefánsson, Arnaldur, "Labor supply response to a tax holiday: The take-home from a large and salient shock," Working Paper, Uppsala University 2019.

Taubinsky, Dmitry and Alex Rees-Jones, "Attention variation and welfare: Theory and evidence from a tax salience experiment," *The Review of Economic Studies*, 2018, *85* (4), 2462–2496.

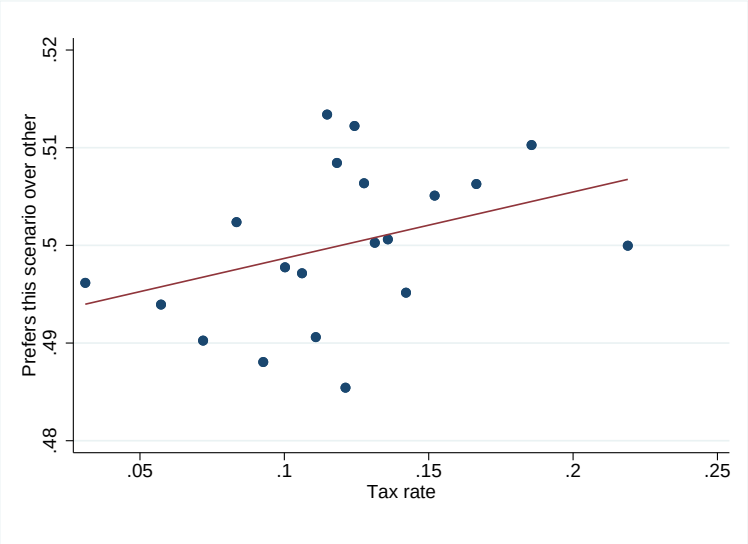
Unrath, Matthew, "Beating the clock: Using year-end changes to identify intensive margin labor supply responses to taxation," Working Paper, UC Berkeley 2020. Retrieved Nov 4, 2023, from https://www.mattunrath.com/research.

Wiswall, Matthew and Basit Zafar, "Preference for the workplace, investment in human capital, and gender," *The Quarterly Journal of Economics*, 2018, *133* (1), 457–507.

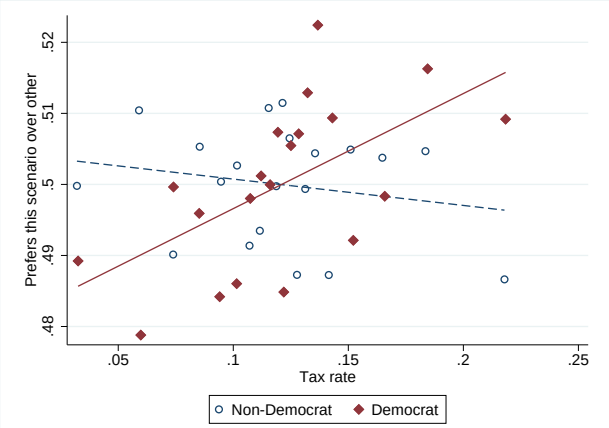
Ziliak, James P and Thomas J Kniesner, "The effect of income taxation on consumption and labor supply," *Journal of Labor Economics*, 2005, *23* (4), 769–796.

Figure 1: Binscatter plots of the choice of preferred scenario against tax rates

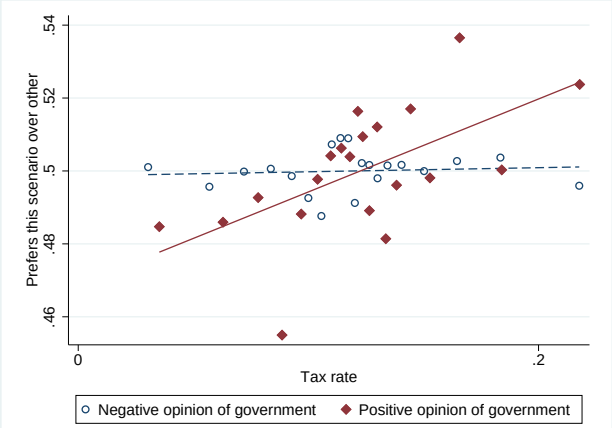
(a) No heterogeneity



(b) By political affiliation



(c) By opinion of the government



Notes: The sample comprises respondent-pair-scenarios with valid in-vignette consumption data, excluding respondents in the top and bottom 5 percent of actual work hours, consumption, and taxes paid. All panels show binscatter plots of whether the respondent preferred the scenario over the other scenario in the pair against the tax rate, further broken down by the specified heterogeneity variable in Panels B and C. All binscatter plots partial out individual, pair, and scenario-order fixed effects, as well as fixed effects for the interaction of the 200 quantiles of disposable income and all possible levels of work hours, further interacted separately with linear slopes in disposable income and work hours. In Panels B and C, this residualization is done separately by heterogeneity group.

Table 1: Reduced form evidence for social preferences

	Dependent variable: Prefers this scenario over other. Heterogeneity variable is:		
	No heterogeneity	Democrat	Positive opinion of government
	(1)	(2)	(3)
Tax rate	0.068*	-0.037	0.0068
	(0.038)	(0.054)	(0.045)
Heterogeneity variable × Tax rate		0.20**	0.24***
		(0.083)	(0.091)
Observations	86,878	80,398	86,685
Respondents	5,439	5,046	5,439
Implied effect when heterogeneity = 1		0.163***	0.244***
		(0.063)	(0.079)

Notes: Standard errors clustered by respondents in parentheses. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample comprises respondent-pair-scenarios with valid in-vignette consumption data, excluding respondents in the top and bottom 5 percent of actual work hours, consumption, and taxes paid. Column 1 controls for individual, pair, and scenario-order fixed effects. Column 1 also controls non-parametrically for disposable income and work hours by including fixed effects for the interaction of the 200 quantiles of disposable income and all possible levels of work hours, further interacted separately with linear slopes in disposable income and work hours. Columns 2 and 3 include the same controls as in column 1, and further interacted with the heterogeneity variable. The heterogeneity variable in column 2 equals 1 for respondents who identified as Democrats, and 0 for those who identified as Republicans or Independents. The heterogeneity variable in column 3 is constructed by standardizing two questions on the respondent's opinion of the government such that the extreme positive score (likes the government) equals 1 and the extreme negative score (dislikes the government) equals -1, taking the average of the two scores, and then dichotomizing such that average scores above zero equals 1 (and zero otherwise).

Table 2: Estimates of utility parameters

	Dependent variable:			
	Will work in this scenario		Prefers this scenario over other	
	(1)	(2)	(3)	(4)
Log(Consumption)	2.28	0.81	1.20	0.12
	(1.63)	(1.65)	(1.37)	(1.37)
Log(672 - Work hours)	219.6***	201.5***	143.9***	131.4***
	(20.7)	(20.9)	(13.5)	(13.6)
Log(Consumption) × Log(Consumption)	0.18***	0.17***	0.19***	0.18***
	(0.018)	(0.018)	(0.032)	(0.032)
Log(672 - Work hours) × Log(672 - Work hours)	-16.9***	-15.5***	-11.0***	-10.1***
	(1.63)	(1.65)	(1.07)	(1.07)
Log(Consumption) × Log(672 - Work hours)	-0.44*	-0.21	-0.30*	-0.13
	(0.25)	(0.25)	(0.18)	(0.18)
Part-time	-0.23***	-0.23***	-0.17***	-0.17***
	(0.056)	(0.057)	(0.036)	(0.036)
Tax paid ($1000)		1.02***		0.69***
		(0.13)		(0.054)
Tax paid ($1000) ^ 2		-0.24***		-0.16***
		(0.071)		(0.025)
Tax paid ($1000) ^ 3		0.015*		0.0092***
		(0.0085)		(0.0027)
Observations	36,032	36,032	80,422	80,422
Respondents	4,670	4,670	5,380	5,380
$E(U_g) \times 1,000$		0.77***		0.52***
$E(U_{gg} g_i) \times 1,000$		-0.23***		-0.15***
$E(U_g \mid \tau = 50\%) \times 1,000$		0.19***		0.14***
Log pseudo-likelihood	-8606.2	-8427.5	-21761.8	-21545.1

Notes: Standard errors clustered by respondents in parentheses. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample comprises respondent-pair-scenarios with valid in-vignette consumption data, excluding respondents in the top and bottom 5 percent of actual work hours, consumption, and taxes paid. All variables are on a per-month basis. Estimates are based on a conditional (on respondent-pair) logit model, controlling for scenario-order fixed effects, an indicator for whether the scenario's weekly work hours is below 35, and a squared polynomial term in $\log(assets_i + spouseincome_i + savings_{itj})$, where $assets_i$ is the wealth in dollars of the individual, $spouseincome_i$ is the individual's spouse's income (or zero if not applicable), and $savings_{itj}$ is the after-tax disposable income less consumption for the scenario. Consumption is the imputed consumption for the scenario, see Section 3.2 for details. Work hours are the work hours for the scenario, so that 672 (the number of hours in 4 weeks) is the amount of leisure time for the scenario. Tax paid is amount of taxes to be paid for the scenario, in thousand dollars. $E(U_g)$ is the marginal utility with respect to tax-funded government expenditure, i.e. the derivative of the utility function with respect to g_i in Equation (14), computed using the estimated parameters in the table and the actual values reported (for non-tax variables) or simulated (for taxes paid and tax rates) by the respondent, averaged over respondents in the sample (one per respondent). $E(U_{gg} g_i)$, which measures the curvature of the social preferences function, is the similarly-computed second-order derivative multiplied by the product of the simulated tax rate, reported hourly wage, and reported work hours of the respondent. $E(U_g \mid \tau = 50\%)$ is computed similar to $E(U_g)$, except that it is computed at a tax rate of 50% instead of the actual tax rate. We multiply utility derivatives by 1,000 to ease interpretation. Standard errors are computed using the delta method.

Table 3: Elasticities of labor supply

	Based on dep. var.: Will work in this scenario		Based on dep. var.: Prefers this scenario over other	
	No social preferences (1)	With social preferences (2)	No social preferences (3)	With social preferences (4)
Panel A: Marshallian elasticity of labor supply with respect to:				
Wage	-0.12***	-0.10***	-0.11***	-0.095***
	(0.012)	(0.012)	(0.016)	(0.016)
Net-of-tax rate	-0.12***	-0.38***	-0.11***	-0.34***
	(0.012)	(0.024)	(0.016)	(0.021)
Wage-tax elasticity wedge	0	0.27***	0	0.24***
	(.)	(0.025)	(.)	(0.019)
Panel B: Hicksian elasticity of labor supply with respect to:				
Wage	0.42***	0.47***	0.50***	0.54***
	(0.019)	(0.021)	(0.021)	(0.022)
Net-of-tax rate	0.42***	-0.26***	0.50***	-0.078*
	(0.019)	(0.055)	(0.021)	(0.041)
Wage-tax elasticity wedge	0	0.73***	0	0.62***
	(.)	(0.062)	(.)	(0.045)
Panel C: Frisch elasticity of labor supply with respect to:				
Wage	1.29***	1.21***	1.84***	1.74***
	(0.11)	(0.10)	(0.15)	(0.14)
Net-of-tax rate	1.29***	0.64***	1.84***	1.15***
	(0.11)	(0.085)	(0.15)	(0.11)
Wage-tax elasticity wedge	0	0.57***	0	0.59***
	(.)	(0.068)	(.)	(0.063)
Respondents	4,670	4,670	5,380	5,380

Notes: Each cell shows the average simulated elasticity or difference in elasticities based on the model specified in the column header. The sample comprises respondents with valid in-vignette consumption data, excluding respondents in the top and bottom 5 percent of actual work hours, consumption, and taxes paid. Elasticities and differences are simulated using the parameters of the model (shown in Table 2), at the reported (for non-tax variables) or tax-simulated values of each respondent (one for each respondent), and then averaged over respondents. In each panel, the wage-tax elasticity wedge is the wage elasticity less the net-of-tax rate elasticity. Standard errors computed by delta method in parentheses. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 4: Robustness of the wage-tax elasticity wedge

		Wage-tax elasticity wedge		
		Marshallian (1)	Hicksian (2)	Frisch (3)
0	Main estimates	0.24***	0.62***	0.59***
		(0.019)	(0.045)	(0.063)
1	Fifth order polynomial in g_i	0.21***	0.63***	0.50***
		(0.019)	(0.045)	(0.071)
2	g_i interacted with $\log(c_i)$ and $\log(\bar{L} - h_i)$	-0.0074	0.38***	0.46***
		(0.050)	(0.059)	(0.097)
3	Quadratic utility	0.43***	0.62***	0.57***
		(0.040)	(0.049)	(0.12)
4	β_c and β_h varies with individual characteristics	0.23***	0.59***	0.56***
		(0.018)	(0.043)	(0.060)
5	Social preferences function is $v(g_i, \tau)$	0.40***	1.65***	1.21***
		(0.025)	(0.090)	(0.14)
6	Using non-imputed consumption	0.26***	0.64***	0.73***
		(0.024)	(0.056)	(0.10)
7	Consumption imputed based on Empirical Bayes	0.083***	0.21***	0.38***
		(0.010)	(0.024)	(0.071)
8	No $\log(assets_i + spouseincome_i + savings_{itj})$ controls	0.42***	1.18***	0.98***
		(0.029)	(0.083)	(0.084)
9	Exclude pairs since observation of the intransitivity prompt	0.23***	0.61***	0.53***
		(0.019)	(0.045)	(0.059)
10	Exclude respondents who ever made an intransitive choice	0.23***	0.61***	0.54***
		(0.020)	(0.046)	(0.062)
11	Exclude inattentive or impatient or speeding respondents	0.19***	0.49***	0.43***
		(0.019)	(0.045)	(0.052)
12	Wedge at median respondent characteristics	0.29***	0.74***	0.73***
		(0.021)	(0.057)	(0.073)
13	Trim top and bottom 2.5%	0.23***	0.50***	0.84***
		(0.019)	(0.043)	(0.081)
14	Restrict to ±10 hours per week	0.19***	0.48***	0.58***
		(0.030)	(0.067)	(0.15)
15	Reweight for demographics	0.23***	0.64***	0.58***
		(0.036)	(0.086)	(0.13)
16	Full-factorial reweighting	0.25***	0.70***	0.55***
		(0.029)	(0.085)	(0.097)

Notes: Standard errors computed by delta method in parentheses unless otherwise stated. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. Each cell shows the average simulated wage-tax elasticity wedge, estimated analogous to Table 3, column 4, repeated in Row 0. Each row shows a specification deviation from Row 0. Row 1: The utility function uses a fifth-order polynomial in the tax-funded government expenditure. Row 2: The utility function interacts tax-funded government expenditure with log consumption and leisure. Row 3: The utility function uses a quadratic utility instead of a translog specification. Row 4: The coefficients on consumption and leisure vary linearly with sex, marital status (married or not), their interaction, a quadratic in log age, a quadratic in the number of children in the family (top-coded at 9), and indicators for high-school-or-less and four-year-college-or-more. Row 5: The utility function is Equation 14 plus an additional interaction between g_i and the scenario tax rate τ. Row 6: We restrict the sample to pairs 5 to 8 and use each scenario's non-imputed consumption information directly. Row 7: We impute consumption using empirical Bayes prediction: the coefficient on log take-home earnings is random, with additional (non-random) interactions of log take-home earnings with sex, marital status, their interaction, a quadratic in log age, a quadratic in the number of children in the family, and indicators for high-school-or-less and four-year-college-or-more. Row 8: We do not control for the quadratic in the log of the sum of assets, spousal income, and the scenario's savings. Row 9: We use only pairs before respondents observed the intransitivity prompt. Row 10: We exclude respondents who ever made an intransitive choice. Row 11: We exclude respondents who were inattentive in our two attention check questions, clicked "next" more than twenty times when viewing the instruction slides, or answered any of the overall, vignette, and government expenditure liking questions in less than half the median time. Row 12: We evaluate the wage-tax elasticity wedge at the median values of consumption, work hours, hourly wage, average tax rate, and taxes paid across respondents. Row 13: We trim the top and bottom 2.5% of actual work hours, consumption, and taxes paid instead of 5%. Row 14: We restrict to vignettes whose work hours are within ±10 hours per week of the respondent's current hours. Row 15: We weight the estimation and simulated differences using post-stratification weights constructed based on the variables in Appendix Table A1. Row 16: We weight to approximate a full-factorial design over consumption, leisure, and tax-funded government expenditure. Specifically, we bin each of the three variables into deciles to form a grid comprising $10^3 = 1,000$ joint cells, count the observations N_c in each cell, and set each respondent-pair's weight to the average of $1/N_c$ across its two scenarios.

Table 5: The effect of attitudes towards government and programs on the wedge

	Average partial effect of the specified heterogeneity variable on the wage-tax elasticity wedge		
	Marshallian	Hicksian	Frisch
	(1)	(2)	(3)
Panel A: Effect of political affiliation on the wedge			
Respondent is Democrat	0.032*	0.070*	0.13**
	(0.017)	(0.041)	(0.062)
Number of Respondents	4,994	4,994	4,994
Panel B: Effect of attitudes towards the government on the wedge			
General opinion of government	0.045**	0.12**	0.26***
	(0.021)	(0.048)	(0.085)
Importance of government	-0.0041	0.0054	0.017
	(0.019)	(0.044)	(0.060)
Trust in government	0.044	0.16**	0.29**
	(0.029)	(0.068)	(0.13)
Programs benefit people like me	0.042*	0.11**	0.17**
	(0.024)	(0.056)	(0.083)
Government revenue allocation is fair	0.047***	0.12***	0.20***
	(0.018)	(0.042)	(0.065)
Number of Respondents	5,380	5,380	5,380
Panel C: Effect of liking the program that taxes fund on the wedge			
Likes program funded by taxes	0.15**	0.41**	0.40*
	(0.071)	(0.19)	(0.21)
Number of Respondents	3,136	3,136	3,136

Notes: Standard errors computed by delta method in parentheses. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. Estimation method follows that of Table 3 column 4; details for the estimation method and specifications are in Section 3.3 (for Panels A and B), Section 4.4 (for Panel C) and Appendix B.4 (for Panel C). Sample for Panels A and B is that from Table 3 column 4, additionally excluding respondents who reported *Non-Affiliated* or *Others* political affiliations for Panel A; sample for Panel C is the same sample excluding respondents who failed either of two inattention check questions, clicked "next slide" more than 20 times in the instructions slides, or answered any of the overall, vignette, and government expenditure liking questions in less than half the median time taken by all respondents. Each cell shows the average partial effect of the heterogeneity variable specified in the row title on the wage-tax elasticity wedge (type specified in the column title). For Panel C, the net-of-tax rate elasticity used for computing the wedge is specific to the program of the vignette. The heterogeneity variable in Panel Aequals 1 for respondents who identified as Democrats, and 0 for those who identified as Republicans or Independents. Each heterogeneity variable in Panel B is constructed by standardizing each of the two questions such that the extreme positive score (e.g. strongly agree) equals 1 and the extreme negative score (e.g. strongly disagree) equals -1, taking the average of the two scores, and then dichotomizing such that average scores above zero equals 1 (and zero otherwise). The liking of a program in Panel C was elicited on a 5-point Likert scale (5 being liking the program the most), and is modeled as a continuous variable.

Table 6: The relationship between government-related social preferences and the elasticity of taxable income

	Dependent variable: ETI. Government-related social preferences proxy is:						
	Confidence in government	Confidence in political parties	Confidence in parliament	Confidence in civil service	Income should be made more equal	Gvt should increase ownership of businesses	Proud to be a citizen
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Main estimates							
Gvt-related social preferences proxy	-0.50***	-0.36***	-0.41***	-0.49*	-0.047	-0.16**	-0.37***
	(0.079)	(0.073)	(0.12)	(0.25)	(0.057)	(0.078)	(0.13)
Observations	1,701	1,701	1,701	1,701	1,701	1,701	1,701
Number of studies	60	60	60	60	60	60	60
Implied ETI effect, p75 − p25 of proxy	-0.19	-0.11	-0.14	-0.13	-0.051	-0.26	-0.13
Panel B: Robustness							
1 After-deduction elasticities only	-0.53***	-0.38***	-0.46***	-0.76**	-0.051	-0.30**	-0.48*
	(0.082)	(0.074)	(0.13)	(0.36)	(0.12)	(0.11)	(0.25)
2 Before-deduction elasticities only	-0.36***	-0.28	-0.23*	-0.19	-0.044	-0.00071	-0.28***
	(0.12)	(0.17)	(0.12)	(0.11)	(0.046)	(0.053)	(0.097)
3 Include Neisser's contextual factors	-0.44***	-0.48***	-0.48***	-0.73***	-0.12	-0.21***	-0.010
	(0.078)	(0.091)	(0.072)	(0.21)	(0.094)	(0.060)	(0.12)
4 Control for labor share subsidy	-0.54***	-0.42***	-0.50***	-0.56**	0.012	-0.14	-0.38**
	(0.092)	(0.073)	(0.12)	(0.25)	(0.038)	(0.088)	(0.14)
5 Control for union density	-0.50***	-0.36***	-0.42***	-0.48*	-0.037	-0.16	-0.36***
	(0.083)	(0.077)	(0.13)	(0.26)	(0.060)	(0.095)	(0.14)
6 Control for tax breadth	-0.53***	-0.42***	-0.45***	-0.54**	-0.022	-0.15	-0.36**
	(0.11)	(0.098)	(0.14)	(0.23)	(0.058)	(0.098)	(0.16)
7 Control for audit yield	-0.52***	-0.43***	-0.42***	-0.50*	-0.050	-0.16*	-0.34**
	(0.085)	(0.067)	(0.13)	(0.26)	(0.065)	(0.081)	(0.13)
8 Include country-group FE	-0.51***	-0.36***	-0.41***	-0.42	-0.027	-0.24***	-0.28*
	(0.089)	(0.082)	(0.15)	(0.29)	(0.061)	(0.083)	(0.15)
9 Include country-group and time FE	-0.54***	-0.42***	-0.43***	-0.68*	-0.042	-0.31***	-0.29
	(0.080)	(0.064)	(0.14)	(0.38)	(0.10)	(0.11)	(0.20)
10 Include all controls (rows 3–9)	-0.64***	-0.70***	-0.39**	-0.57*	-0.18	-0.36*	-0.49
	(0.18)	(0.19)	(0.18)	(0.29)	(0.11)	(0.19)	(0.57)

Notes: Standard errors clustered by study in parentheses. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample comprises all ETI estimates in Neisser (2021), excluding one study from Israel that participated in only one wave of the WVS. The government-related social preferences proxies are country-wave level averages from the WVS and the EVS (Inglehart et al. 2022; EVS 2022); we match observations to the nearest WVS or EVS year using the mean study year. All estimates control for the regression technique used (five possible categories, including the omitted base category), whether and how the study controls for income (five categories), the difference length (e.g. 1 year) for the first-differences technique used in estimation (4 categories), whether the study weights by income, whether the study restricts the sample by age, the income restriction used by the study if any (five categories), the publication decade (3 categories), and the decade that the tax policy changed (three categories). See Neisser's Table 2 and 3 column 6 for further details on the above variables. All estimates additionally control for whether the elasticity is an after-deductions (AD) or before-deductions (BD) elasticity, and the interactions of this dummy variable with all control variables used in the table. Panel A shows the main estimates; Panel B shows robustness estimates that deviate in the specified manner. *Implied ETI effect, p75 − p25 of proxy* is the predicted change in the ETI when the proxy variable increases from its 25th to its 75th percentile across all country-waves. Rows 1 and 2: The sample is restricted to AD and BD elasticities respectively; the AD-or-BD dummy and its interactions are not relevant in these specifications. Row 3: We additionally control for all contextual factors except the top 1% income share from Neisser's Table 5. Row 4: We additionally control for the share of public spending on the provision of child care, preschool, and elderly care, available in Kleven (2014). Row 5: We additionally control for the share of workers who are union members in OECD countries, available from the OECD (2025b). Row 6: We additionally control for tax breadth, the ratio of taxable income to gross earnings; we use the first principal component of all eight measures (for eight household types) available in OECD (2026). Row 7: We additionally control for the total value of additional assessments raised through audits and verification actions divided by net tax revenue collected (gross collections minus refunds), available from OECD (2025a). Row 8: We include the country-group FEs. Row 9: We include the fixed effects from Row 8, and additionally include FEs for the year ranges 1981–1984, 1989–1993, 1994–1998, 1999-2004, 2005–2009, 2010–2014, 2017–2022. These years are WVS survey years, and the five EVS surveys match (uniquely) to the first, second, fourth, fifth, and seventh year ranges. Row 10: We include all controls and fixed effects in rows 3–7, and 9.

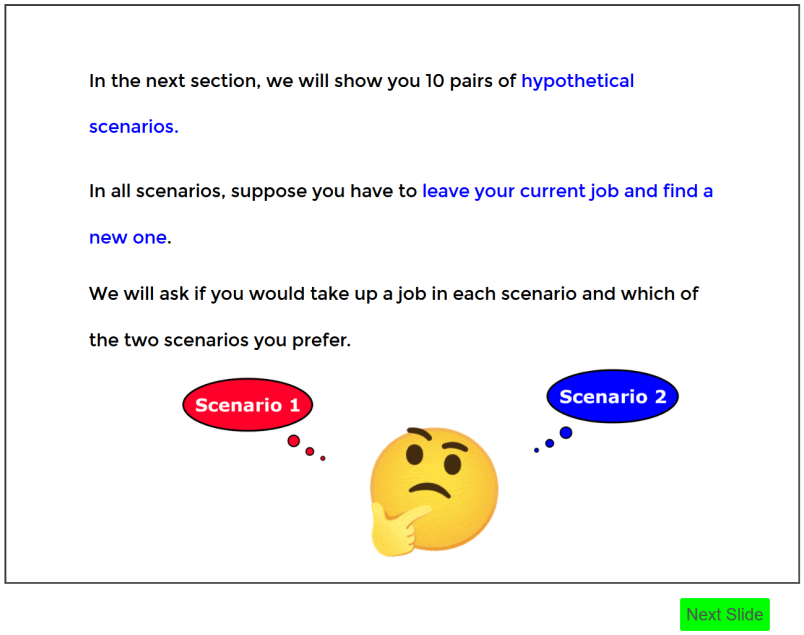
Online Appendix for
Wages, taxes, and labor supply elasticities: The role of social preferences

By Janjala Chirakijja and Pinchuan Ong

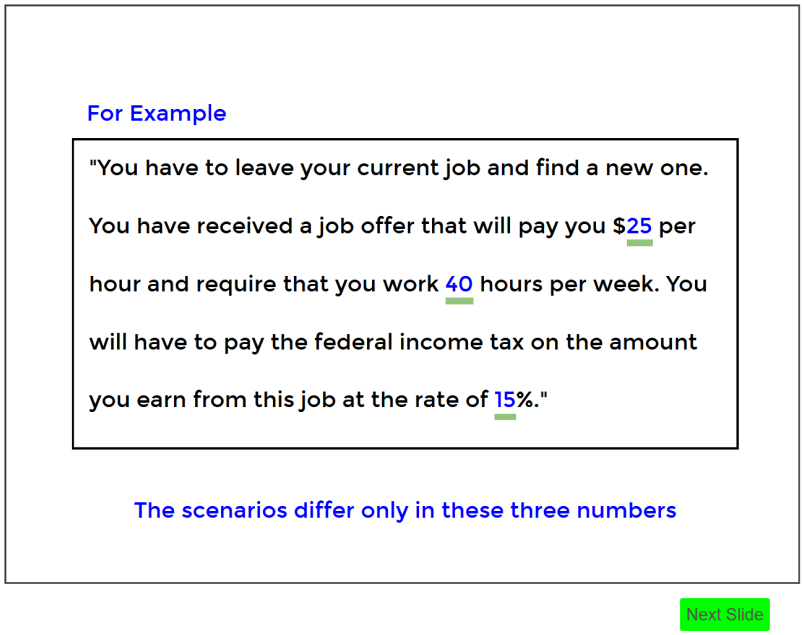
A Appendix tables and figures

Appendix Figure A1: First two instructional slides

(a) Slide 1



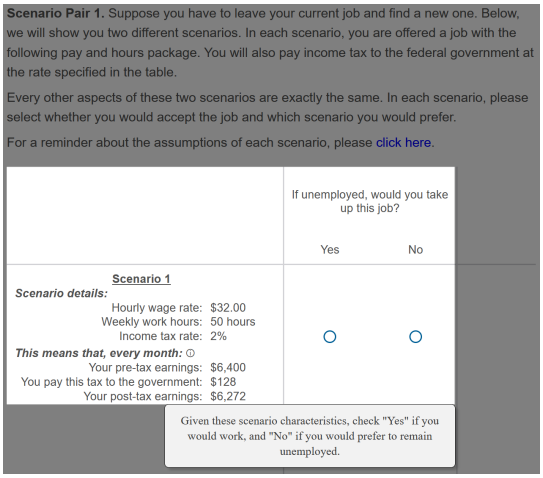
(b) Slide 2



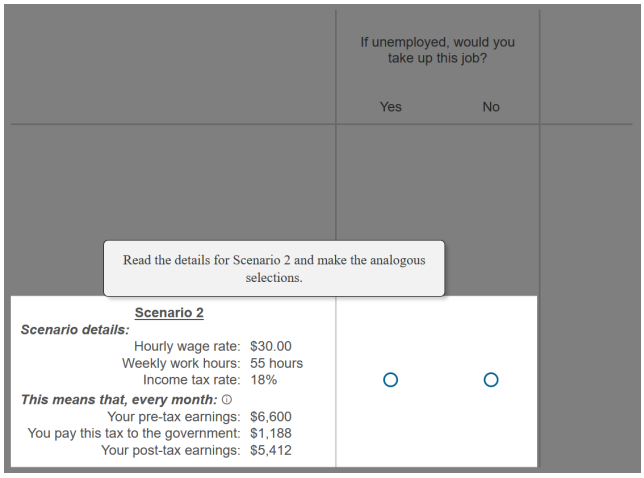
Notes: The figure shows the first two slides presented to respondents. Sentences appeared at 250 words per minute (between the usual reading and speaking speed), and the "Next Slide" button was grayed out (i.e., not responsive) until all the text had appeared.

Appendix Figure A2: Example of scenarios presentation for pairs 1 to 4

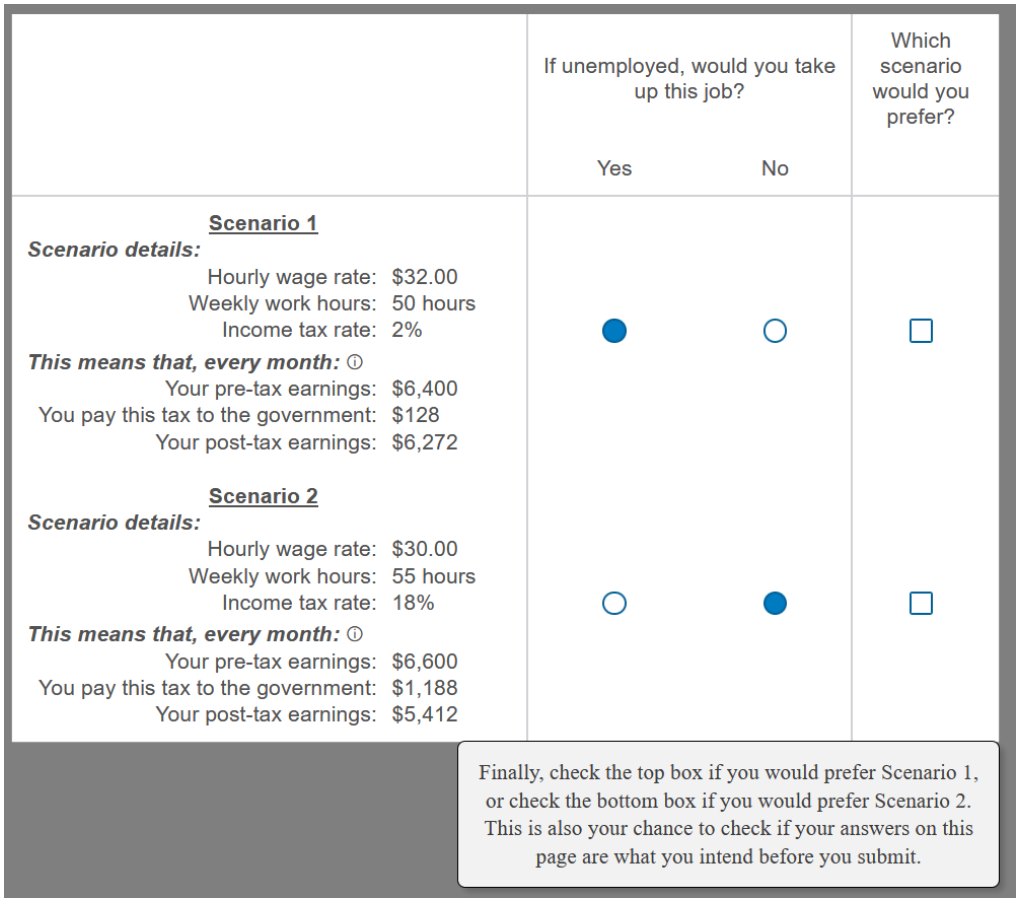
(a) Work or not in first scenario



(b) Work or not in second scenario



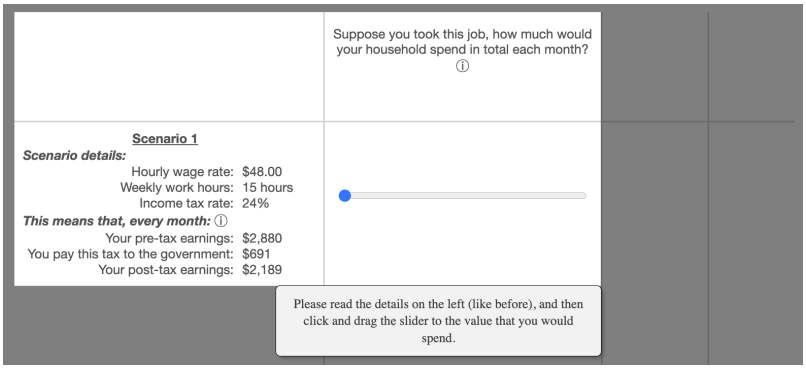
(c) Choice between scenarios



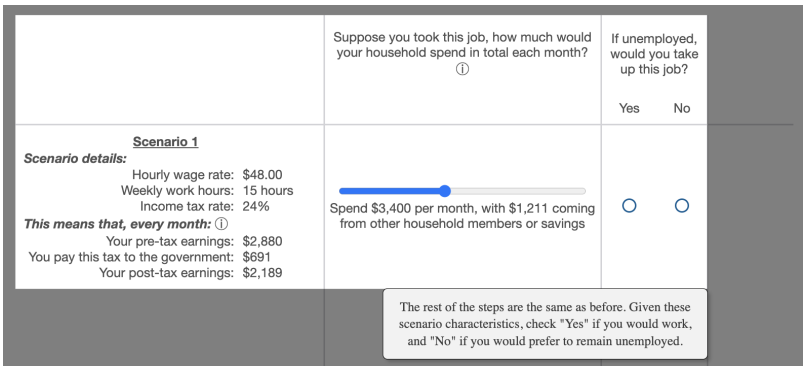
Notes: The figure shows an example of the scenario presentation for pairs 1 to 4. Panels A and B show that the other scenario of the pair was hidden when the respondent was asked to make a choice of whether to work or not. Panel C shows that both scenarios were shown again when the respondent was asked to choose between the two. The guided website tour (white focus on black background, with explanatory text) was shown only in pair 1 of the first four pairs.

Appendix Figure A3: Example of scenarios presentation for pairs 5 to 8

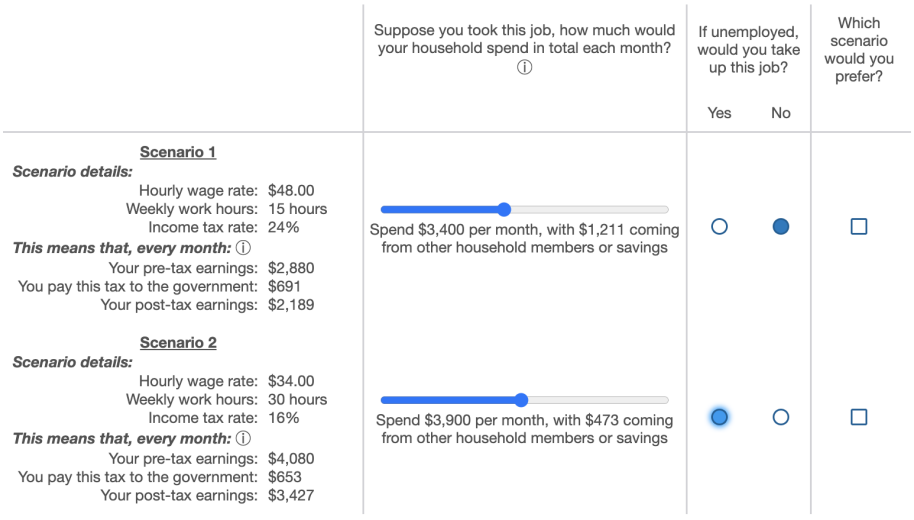
(a) Consumption amount



(b) Work or not



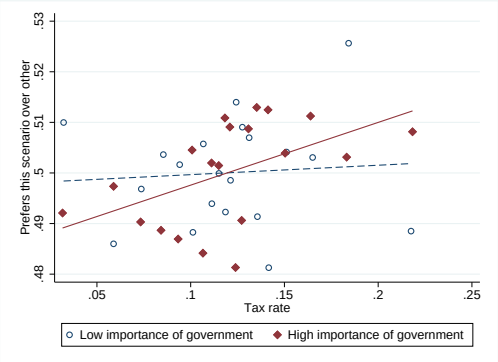
(c) Choice between scenarios



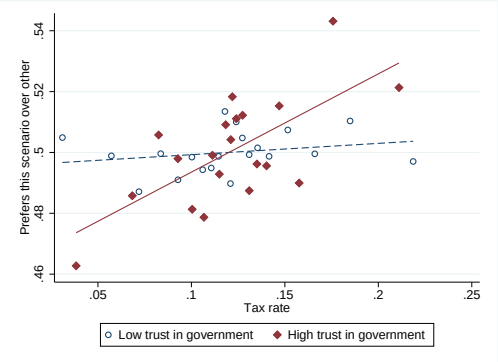
Notes: The figure shows an example of the scenario presentation for pairs 5 to 8, omitting a repeat of the instructions shown above the questionnaire matrix. Panels A and B show that the choice of whether to work was hidden when eliciting the amount of consumption using the slider bar shown, and revealed only after the respondent made the consumption choice. The second scenario was hidden when the respondent made choices for the first, and vice versa, analogous to Appendix Figure A2. Panel C shows the full layout, revealed only when the respondent was asked to choose between the two scenarios. The guided website tour (white focus on black background, with explanatory text) was shown only in pair 5 of pairs 5 to 8. The slider buttons were initialized at 0 (an illogical choice for consumption) and the respondent had to drag the slider to proceed. Both sliders of each pair were set to the same maximum: the respondent's actual monthly spending plus the larger of the two scenarios' post-tax earnings, rounded up to the nearest $1,000.

Appendix Figure A4: Binscatter plots by other government opinion heterogeneity dimensions

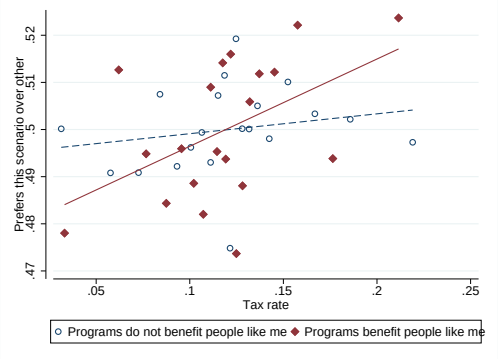
(a) By importance of government



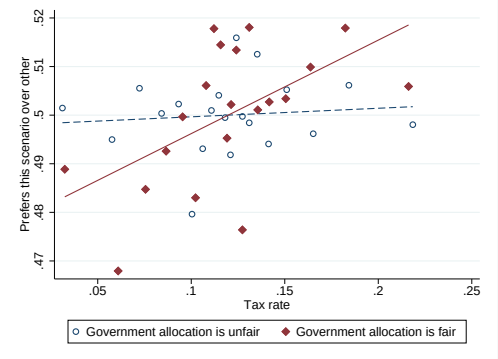
(b) By trust in government



(c) By whether programs benefit me



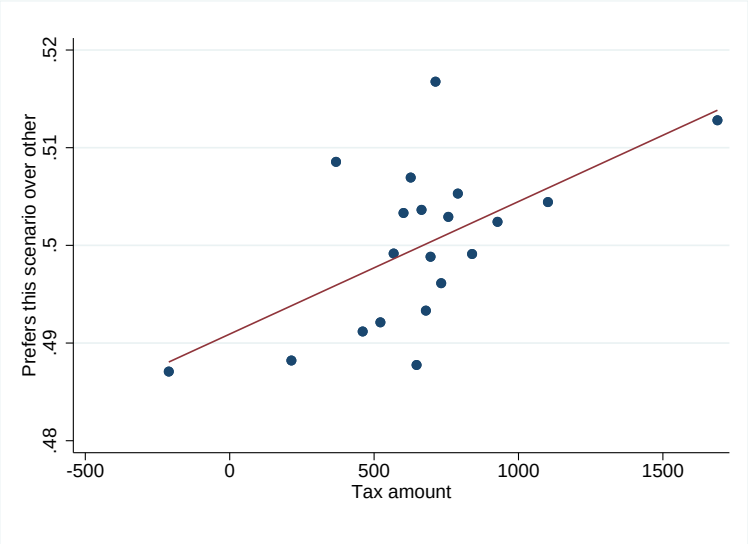
(d) By fairness of revenue allocation



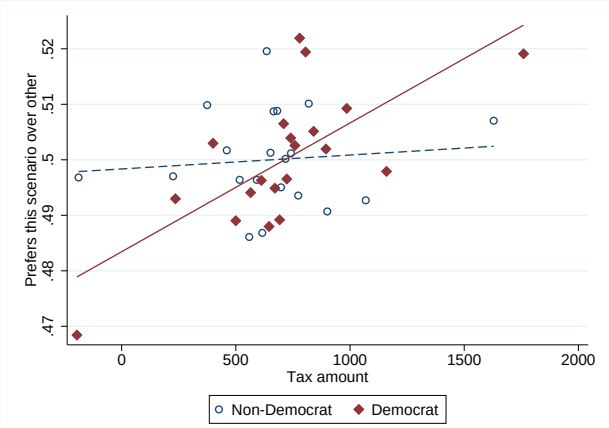
Notes: The sample comprises respondent-pair-scenarios with valid in-vignette consumption data, excluding respondents in the top and bottom 5 percent of actual work hours, consumption, and taxes paid. All panels show binscatter plots of whether the respondent preferred the scenario over the other scenario in the pair against the tax rate, broken down by the specified heterogeneity variable. All binscatter plots partial out individual, pair, and scenario-order fixed effects, as well as fixed effects for the interaction of the 200 quantiles of disposable income and all possible levels of work hours, further interacted separately with linear slopes in disposable income and work hours. Residualization is done separately by heterogeneity group.

Appendix Figure A5: Binscatter plots of the choice of preferred scenario against taxes paid

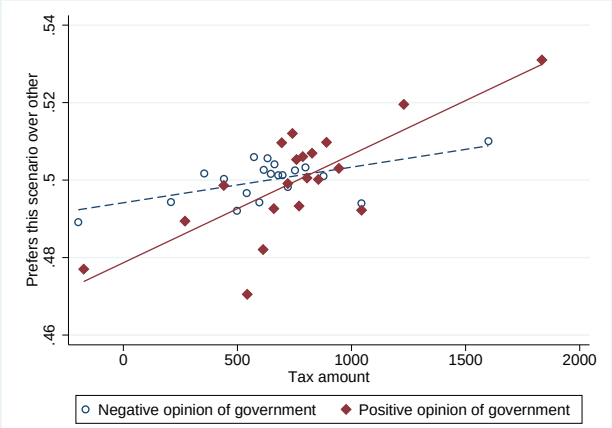
(a) No heterogeneity



(b) By political affiliation



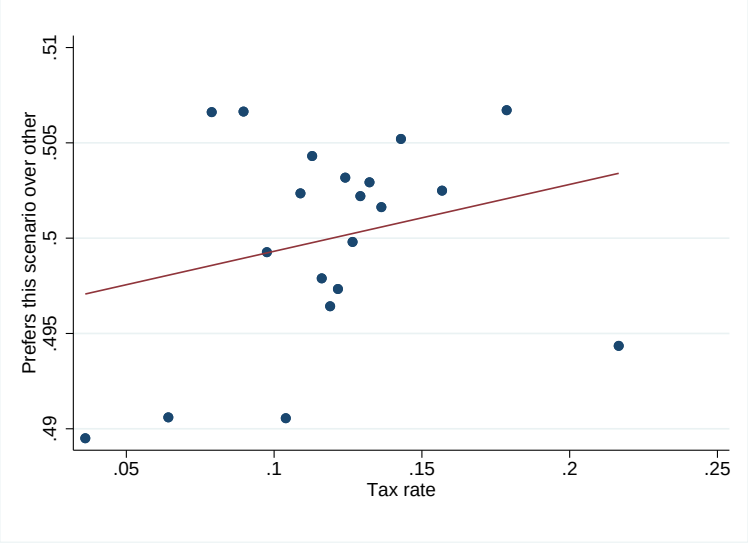
(c) By opinion of the government



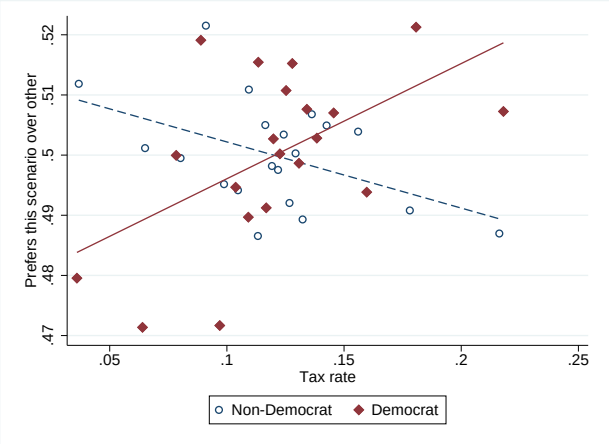
Notes: The sample comprises respondent-pair-scenarios with valid in-vignette consumption data, excluding respondents in the top and bottom 5 percent of actual work hours, consumption, and taxes paid. All panels show binscatter plots of whether the respondent preferred the scenario over the other scenario in the pair against the amount of taxes paid, further broken down by the specified heterogeneity variable in Panels B and C. All binscatter plots partial out individual, pair, and scenario-order fixed effects, as well as fixed effects for the interaction of the 200 quantiles of disposable income and all possible levels of work hours, further interacted separately with linear slopes in disposable income and work hours. In Panels B and C, this residualization is done separately by heterogeneity group.

Appendix Figure A6: Binscatter plots of the choice of preferred scenario against tax rates, controlling instead for the net-of-tax wage

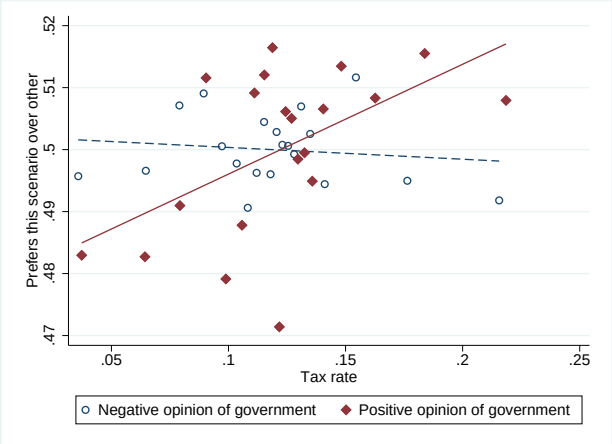
(a) No heterogeneity



(b) By political affiliation



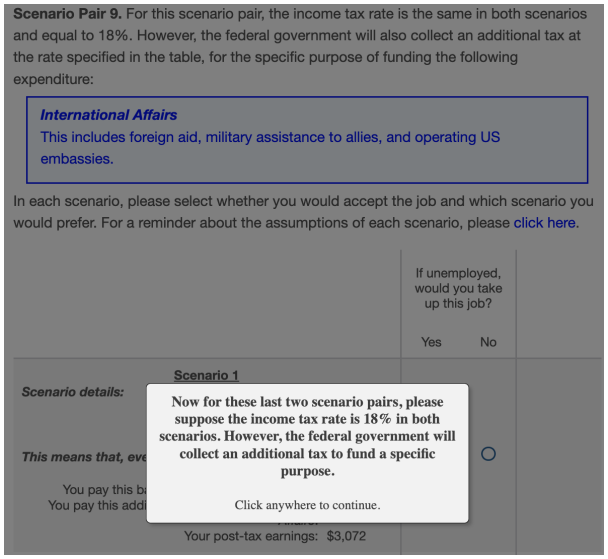
(c) By opinion of the government



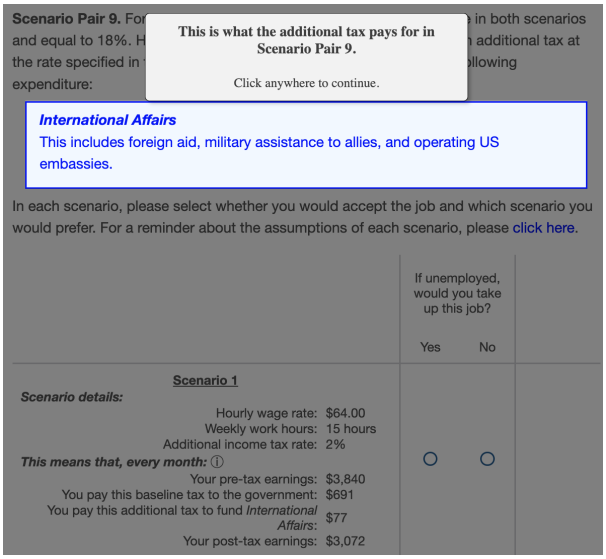
Notes: The sample comprises respondent-pair-scenarios with valid in-vignette consumption data, excluding respondents in the top and bottom 5 percent of actual work hours, consumption, and taxes paid. All panels show binscatter plots of whether the respondent preferred the scenario over the other scenario in the pair against the tax rate, further broken down by the specified heterogeneity variable in Panels B and C. All binscatter plots partial out individual, pair, and scenario-order fixed effects, as well as fixed effects for the 200 quantiles of the net-of-tax wage interacted with linear slopes in the net-of-tax wage. In Panels B and C, this residualization is done separately by heterogeneity group.

Appendix Figure A7: Example of scenarios presentation for pairs 9 and 10

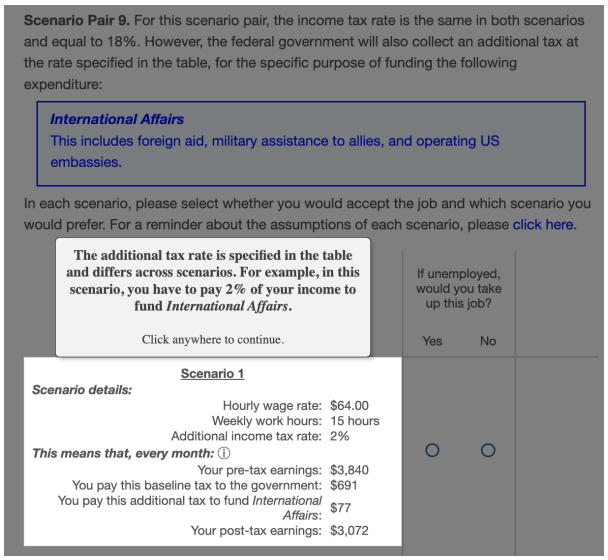
(a) Screen on entering the ninth pair



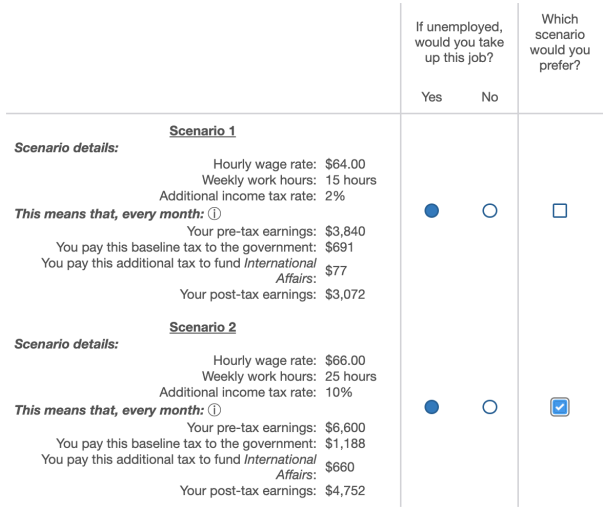
(b) First transition of website tour



(c) Second transition of website tour

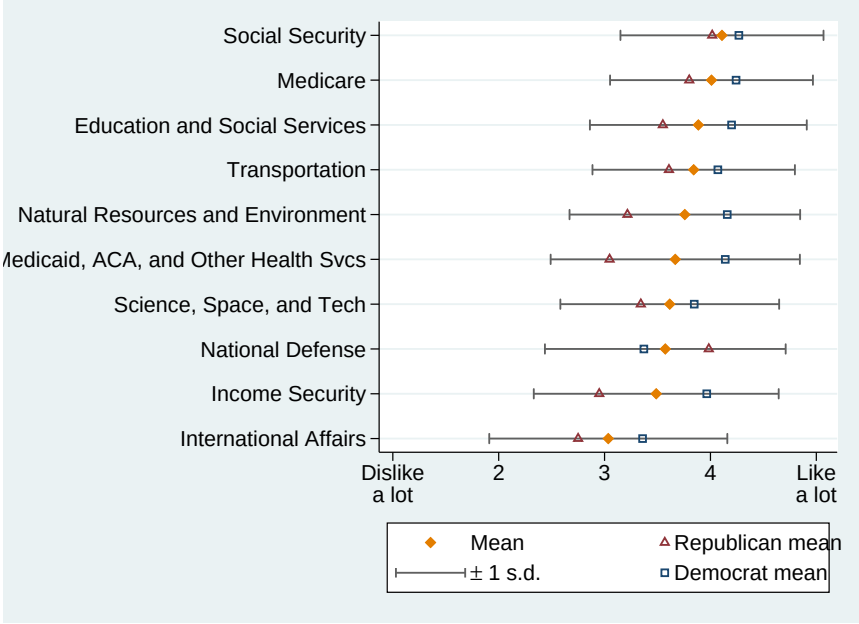


(d) Full matrix



Notes: The figure shows an example of the scenario presentation for pairs 9 and 10. Panel A shows the instructions displayed on entering the ninth pair. Panels B and C similarly show the information that we highlighted to respondents on the first two transitions of the guided website tour. Panel D shows the full layout. The second scenario was hidden when the respondent made choices for the first, and vice versa, analogous to Appendix Figure A2. The transition in Panel B was also displayed in pair 10 to highlight the new program that the change in program being funded.

Appendix Figure A8: How much respondents like government programs



Notes: The figure shows means and standard deviations of responses to the question "To what extent do you like or dislike that your tax money is used to fund each [federal government spending] category?" The sample comprises respondents with valid in-vignette consumption data, excluding respondents in the top and bottom 5 percent of actual work hours, consumption, and taxes paid. Each respondent received five of the ten categories shown. Each category was displayed with a tooltip which, on mouse hover, would provide more details. Republican (Democrat) mean indicates the mean computed among respondents who reported a Republican (Democrat) political affiliation.

Appendix Table A1: Survey sample and U.S. population characteristics

	U.S. Population	Survey Sample
Male	0.52	0.51
25-34 years old	0.29	0.28
35-44 years old	0.27	0.26
45-54 years old	0.24	0.24
55+ years old	0.21	0.22
0-19,999 dollars	0.12	0.08
20,000-39,999 dollars	0.23	0.24
40,000-59,999 dollars	0.23	0.27
60,000-99,999 dollars	0.24	0.28
100,000+ dollars	0.18	0.13
Four-year college degree or more	0.43	0.60
High-school graduate or less	0.31	0.12
Married	0.61	0.52
White	0.76	0.82
Black	0.12	0.08
Asian	0.08	0.07
Others	0.04	0.04
Hispanic	0.20	0.09
Republican	0.30	0.27
Democrat	0.29	0.40
Independent	0.38	0.26
Voted for Biden in the 2020 presidential election	0.51	0.51
Voted for Trump in the 2020 presidential election	0.47	0.32
Sample Size		5,439

Notes: This table presents statistics for the overall U.S. population aged 25-64 who are employed and compares them to the characteristics of our survey respondents. National statistics on gender, age, income brackets, education, marital status, and race are from the IPUMS Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS) data set for 2022 (Flood et al. 2022). National statistics on party affiliation for May 2023 are from Gallup (2023). Note that these statistics are based on surveys of adults 18 years and older and are not restricted to employed individuals. Presidential election results from 2020 are from Leip (2023).

Appendix Table A2: Horse-race regression of consumption and earnings

	Log(consumption) (1)
Log(Take-home earnings)	0.37***
	(0.036)
Log(Pre-tax earnings)	0.024
	(0.036)
Observations	43494
Respondents	5439

Notes: Standard errors clustered by respondents in parentheses. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample comprises respondent-pair-scenarios with valid in-vignette consumption data, excluding respondents in the top and bottom 5 percent of actual work hours, consumption, and taxes paid. The specification controls for individual, pair, and scenario-order fixed effects.

Appendix Table A3: Reduced form evidence for social preferences (robustness)

	Dependent variable: Prefers this scenario over other. Heterogeneity variable is:		
	No heterogeneity	Democrat	Positive opinion of government
	(1)	(2)	(3)
Panel A: Main regressor: Tax amount			
Tax paid ($1000)	0.014***	0.0014	0.0087
	(0.0045)	(0.0063)	(0.0055)
Heterogeneity variable × Tax paid ($1000)		0.022**	0.015
		(0.0095)	(0.0099)
Observations	86,878	80,398	86,685
Respondents	5,439	5,046	5,439
Implied effect when heterogeneity = 1		0.023***	0.024***
		(0.007)	(0.008)
Panel B: Controlling for net-of-tax wage			
Tax rate	0.027	-0.12**	-0.021
	(0.044)	(0.059)	(0.050)
Heterogeneity variable × Tax rate		0.29***	0.17*
		(0.091)	(0.10)
Observations	87,024	80,736	87,024
Respondents	5,439	5,046	5,439
Implied effect when heterogeneity = 1		0.171**	0.151*
		(0.069)	(0.087)

Notes: Standard errors clustered by respondents in parentheses. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$. The sample comprises respondent-pair-scenarios with valid in-vignette consumption data, excluding respondents in the top and bottom 5 percent of actual work hours, consumption, and taxes paid. Column 1 controls for individual, pair, and scenario-order fixed effects. In Panel A, Column 1 controls non-parametrically for disposable income and work hours by including fixed effects for the interaction of the 200 quantiles of disposable income and all possible levels of work hours, further interacted separately with linear slopes in disposable income and work hours. In Panel B, Column 1 controls non-parametrically for the net-of-tax wage by including fixed effects for the 200 quantiles of the net-of-tax wage, further interacted with a linear slope in the net-of-tax wage. Columns 2 and 3 include the same controls as in column 1, and further interacted with the heterogeneity variable. The heterogeneity variable in column 2 equals 1 for respondents who identified as Democrats, and 0 for those who identified as Republicans or Independents. The heterogeneity variable in column 3 is constructed by standardizing two questions on the respondent's opinion of the government such that the extreme positive score (e.g. strongly agree) equals 1 and the extreme negative score (e.g. strongly disagree) equals -1, taking the average of the two scores, and then dichotomizing such that average scores above zero equals 1 (and zero otherwise).

Appendix Table A4: Elasticities of labor supply based on a quadratic utility specification

	Based on dep. var.: Will work in this scenario		Based on dep. var.: Prefers this scenario over other	
	No social preferences (1)	With social preferences (2)	No social preferences (3)	With social preferences (4)
Panel A: Marshallian elasticity of labor supply with respect to:				
Wage	0.41***	0.41***	0.59***	0.57***
	(0.076)	(0.074)	(0.074)	(0.060)
Net-of-tax rate	0.41***	-0.047	0.59***	0.14***
	(0.076)	(0.049)	(0.074)	(0.047)
Wage-tax elasticity wedge	0	0.46***	0	0.43***
	(.)	(0.054)	(.)	(0.040)
Panel B: Hicksian elasticity of labor supply with respect to:				
Wage	0.67***	0.68***	0.83***	0.83***
	(0.044)	(0.044)	(0.054)	(0.051)
Net-of-tax rate	0.67***	-0.031	0.83***	0.21***
	(0.044)	(0.056)	(0.054)	(0.054)
Wage-tax elasticity wedge	0	0.71***	0	0.62***
	(.)	(0.062)	(.)	(0.049)
Panel C: Frisch elasticity of labor supply with respect to:				
Wage	1.06***	0.96***	1.40***	1.35***
	(0.11)	(0.11)	(0.13)	(0.50)
Net-of-tax rate	1.06***	0.40***	1.40***	0.78
	(0.11)	(0.11)	(0.13)	(0.59)
Wage-tax elasticity wedge	0	0.56***	0	0.57***
	(.)	(0.070)	(.)	(0.12)
Respondents	4,670	4,670	5,380	5,380

Notes: Each cell shows the average simulated elasticity or difference in elasticities based on the model specified in the column header. The sample comprises respondents with valid in-vignette consumption data, excluding respondents in the top and bottom 5 percent of actual work hours, consumption, and taxes paid. Elasticities and differences are simulated using the procedure in Table 3, with utility specification $U(c_i, h_i, g_i) = \beta_c c_i + \beta_h\left(\bar{L} - h_i\right) + \beta_{cc} c_i^2 + \beta_{hh}\left(\bar{L} - h_i\right)^2 + \beta_{ch} c_i\left(\bar{L} - h_i\right) + \beta_g g_i + \beta_{gg} g_i^2 + \beta_{ggg} g_i^3$ in place of that in Equation (14). In each panel, the wage-tax elasticity wedge is the wage elasticity less the net-of-tax rate elasticity. Standard errors computed by delta method in parentheses. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$.

Appendix Table A5: Elasticities of labor supply, by sex and marital status

	Subsample:			
	Non-married men (1)	Non-married women (2)	Married men (3)	Married women (4)
Panel A: Marshallian elasticity of labor supply with respect to:				
Wage	-0.20***	-0.17***	-0.10***	0.023
	(0.015)	(0.013)	(0.014)	(0.015)
Net-of-tax rate	-0.43***	-0.39***	-0.29***	-0.24***
	(0.020)	(0.019)	(0.021)	(0.023)
Wage-tax elasticity wedge	0.23***	0.23***	0.19***	0.26***
	(0.020)	(0.019)	(0.021)	(0.024)
Panel B: Hicksian elasticity of labor supply with respect to:				
Wage	0.49***	0.51***	0.50***	0.56***
	(0.018)	(0.018)	(0.021)	(0.025)
Net-of-tax rate	-0.11**	-0.077	-0.059	-0.054
	(0.049)	(0.051)	(0.043)	(0.045)
Wage-tax elasticity wedge	0.60***	0.59***	0.56***	0.62***
	(0.050)	(0.050)	(0.045)	(0.048)
Panel C: Frisch elasticity of labor supply with respect to:				
Wage	2.01***	2.06***	1.60***	1.40***
	(0.20)	(0.19)	(0.16)	(0.12)
Net-of-tax rate	1.30***	1.38***	1.22***	0.91***
	(0.16)	(0.17)	(0.14)	(0.10)
Wage-tax elasticity wedge	0.71***	0.68***	0.38***	0.49***
	(0.072)	(0.065)	(0.067)	(0.059)
Respondents in subsample	1,304	1,251	1,456	1,369

Notes: Each cell shows the average simulated elasticity or wage-tax elasticity wedge in elasticities for the subsample specified in the column header. To compute the elasticities and wedge, we first estimate the extended model from Table 4 row 4, and then simulate the elasticities and wedges for the indicated subsample at the reported (for non-tax variables) or tax-simulated values of each respondent (one for each respondent), and then averaged over respondents. Standard errors computed by delta method in parentheses. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$.

Appendix Table A6: Wage-tax elasticity wedge, by attitudes towards government

	Wage-tax elasticity wedge, simulated for each level of the specified heterogeneity variable					
	Marshallian		Hicksian		Frisch	
	Het=0 (1)	Het=1 (2)	Het=0 (3)	Het=1 (4)	Het=0 (5)	Het=1 (6)
Panel A: Political affiliation						
Respondent is Democrat	0.20***	0.23***	0.54***	0.61***	0.44***	0.58***
	(0.023)	(0.019)	(0.053)	(0.046)	(0.077)	(0.067)
Panel B: Attitudes towards the government						
General opinion of government	0.22***	0.27***	0.56***	0.68***	0.48***	0.74***
	(0.019)	(0.025)	(0.043)	(0.057)	(0.050)	(0.10)
Importance of government	0.25***	0.24***	0.61***	0.61***	0.57***	0.59***
	(0.029)	(0.019)	(0.065)	(0.044)	(0.090)	(0.063)
Trust in government	0.23***	0.28***	0.58***	0.74***	0.54***	0.82***
	(0.019)	(0.031)	(0.044)	(0.075)	(0.057)	(0.15)
Programs benefit people like me	0.22***	0.27***	0.58***	0.69***	0.53***	0.70***
	(0.021)	(0.024)	(0.048)	(0.059)	(0.066)	(0.091)
Government revenue allocation is fair	0.21***	0.26***	0.54***	0.65***	0.47***	0.67***
	(0.018)	(0.022)	(0.043)	(0.052)	(0.052)	(0.083)

Notes: Each cell reports the simulated wage-tax elasticity wedge based on the model used in Table 5 Panels A and B. Columns 1–2, 3–4, and 5–6 correspond to Table 5 columns 1, 2, and 3, respectively. Table 5 Panel C is omitted because its heterogeneity variable is continuous. To construct the wedge by heterogeneity status, we estimate the model from Table 5 and simulate outcomes under counterfactuals in which all respondents are assigned Het=0 (odd-numbered columns) or Het=1 (even-numbered columns). Standard errors computed by delta method in parentheses. Asterisks denote significance: * $p < .10$, ** $p < .05$, *** $p < .01$.

B Derivations

B.1 Micro-foundation of our main model

In this section, we elaborate the utility function to explicitly include social preferences from the tax contribution of other people. We show that the model simplifies to our main model used in Section 2 given an assumption that is likely to hold in our vignette experiment.

In any time period, a person with initial assets level a_i chooses consumption c_i, next period assets a_i', and labor supply h_i to solve the utility maximization problem

$$V\left(a_i, w_i, \tau, N_i\right) = \max_{c_i, a_i', h_i} \tilde{U}\left(c_i, h_i, g_i, g_{-i}, G\right) + \beta E\left[V\left(a_i', w_i', \tau', N_i'\right)\right] \tag{B1}$$

$$\text{s.t. } c_i + \frac{1}{1+r} a_i' = a_i + \left(1-\tau\right) w_i h_i + N_i, \tag{B2}$$

$$g_i = \tau w_i h_i, \tag{B3}$$

$$g_{-i} = \tau Y_{-i}, \tag{B4}$$

$$G = G_0 + \tau Y_{-i} + g_i, \tag{B5}$$

where w_i is the person's wage, τ is the tax rate on earnings from work $w_i h_i$, N_i is non-labor income, β is an impatience parameter, r is the interest rate, $V\left(\cdot\right)$ is the indirect utility function, and $E\left[\cdot\right]$ is the expectations operator. Equation (B2) is a usual budget constraint which specifies that consumption c_i and savings $\frac{1}{1+r} a_i' - a_i$ equal post-tax earnings and non-labor income.

Compared to the usual labor supply model, per-period utility $\tilde{U}\left(\cdot\right)$ in Equation (B1) depends on three additional factors: the individual's own tax contribution g_i (reflecting warm glow from own contribution), the other people's tax contribution g_{-i} (reflecting warm glow from others' contribution), and the total government expenditure G (reflecting pure altruism, which by definition would not distinguish between a contribution by self or by others). g_i is equivalent to the tax rate applied to the individual's total earnings from work by Equation (B3), and g_{-i} is equivalent to the same income tax rate applied to the total earnings of everyone else Y_{-i} by Equation (B4). Equation (B5) computes total government expenditure as the sum of non-income tax revenue G_0, g_i and g_{-i}. Notice that $\tilde{U}\left(\cdot\right)$ depends (through g_{-i} and G) on the common tax rate τ. Hence, when tax rates increase, the individual gets social preferences utility not just from her own contribution but also the increased contributions of others. This modeling aspect reflects the fact that tax changes in practice often apply to broad groups of people, while wage changes are often perceived by the individual to be specific to herself.

While conceptually easy to describe, the maximization problem (B1) does not yield equations that can be easily estimated. This is because estimating the problem directly requires data on the individual's expectations of Y_{-i}, which is difficult to elicit. However, notice that substituting Equation (B4) and (B5) into the utility function yields $\tilde{U}\left(c, h, g_i, \tau Y_{-i}, G_0 + g_i + \tau Y_{-i}\right)$, with G_0 and Y_{-i} exogenous to the individual. Hence, treating these as parameters in the utility function, we can rewrite utility as

$$U\left(c_i, h_i, g_i, \tau\right) = \tilde{U}\left(c, h, g_i, \tau Y_{-i}, G_0 + g_i + \tau Y_{-i}\right). \tag{B6}$$

The maximization problem then becomes

$$V\left(a_i, w_i, \tau, N_i\right) = \max_{c_i, a_i', h_i} U\left(c_i, h_i, g_i, \tau\right) + \beta E\left[V\left(a_i', w_i', \tau', N_i'\right)\right] \tag{B7}$$

subject to Equations (B2) and (B3). This formulation yields estimating equations involving τ, which are more easily observed in the data.

For our empirical analysis, we specialize this utility function to

$$U\left(c_i, h_i, g_i, \tau\right) = u\left(c_i, h_i\right) + v\left(g_i, \tau\right) \tag{B8}$$

so that we have an easily interpretable social preferences term $v\left(g_i, \tau\right)$. Furthermore, since taxes paid are highly correlated with tax rates, we expect the model misspecification that comes from ignoring the dependence of $v\left(g_i, \tau\right)$ on τ to be relatively benign. In that case, we can further simplify to

$$U\left(c_i, h_i, g_i, \tau\right) = u\left(c_i, h_i\right) + \tilde{v}\left(g_i\right). \tag{B9}$$

This is equivalent to the utility function that we use in our main analysis, as described in Section 2. The first derivative of $\tilde{v}\left(g_i\right)$ will be positive if people had positive social preferences — broadly defined, since social preferences in our model also includes the tax money that comes from other individuals g_{-i} (in addition to the usual discussion about pure altruism and warm glow from the individual's own contribution).

B.2 Derivation of the solution for the main model

The utility maximization problem from Section 2, slightly rewritten, is

$$V\left(a_i, w_i, \tau, N_i\right) = \max_{c_i, a_i', h_i} U\left(c_i, h_i, g_i\right) + \beta E\left[V\left(a_i', w_i', \tau', N_i'\right)\right] \tag{B10}$$

$$\text{s.t. } c_i = A_i + \left(1 - \tau\right) w_i h_i + N_i, \tag{B11}$$

$$A_i = a_i - \frac{1}{1+r} a_i', \tag{B12}$$

$$g_i = \tau w_i h_i, \tag{B13}$$

where utility $U(\cdot)$ depends on consumption c_i, labor supply h_i, and tax-funded government expenditure g_i, the state variables for the maximization problem are asset level a_i, wage w_i, tax rate on earnings τ, and non-labor income N_i, β is an impatience parameter, r is the interest rate, $V(\cdot)$ is the indirect utility function, $E[\cdot]$ is the expectations operator, and the prime symbol ($'$) denotes the next period. A_i is the intertemporal resource allocation amount, useful for connecting this to static models.

Let λ_i be the Lagrangian multiplier on Equation (B11). The (intratemporal) first order conditions for c_i and h_i are, respectively,

$$U_c(c_i, h_i, g_i) = \lambda_i \tag{B14}$$

and

$$U_h(c_i, h_i, g_i) + \tau w_i U_g(c_i, h_i, g_i) + \lambda_i (1-\tau) w_i = 0. \tag{B15}$$

Substituting Equation (B14) into Equation (B15) and rewriting in matrix form (more convenient for derivation of the Marshallian and Hicksian elasticities), we have

$$\mathcal{M}(h_i; c_i, A_i, a_i, w_i, \tau, N_i) \equiv DU(c_i, h_i, g_i) \begin{bmatrix} (1-\tau) w_i \\ 1 \\ \tau w_i \end{bmatrix} = 0, \tag{B16}$$

where D is the directional derivative operator so that

$$DU(c_i, h_i, g_i) \equiv \begin{bmatrix} U_c(c_i, h_i, g_i) & U_h(c_i, h_i, g_i) & U_g(c_i, h_i, g_i) \end{bmatrix}$$

is the Jacobian matrix of $U(c_i, h_i, g_i)$, and the optimal-choice c_i and g_i satisfy Equations (B11) and (B13) respectively.

B.2.1 Marshallian elasticities

In a life-cycle framework, the Marshallian elasticity holds constant the intertemporal allocation of resources A_i (see Blundell and MaCurdy 1999, Section 4.2.2). Implicit differentiation of Equation (B16) gives

$$\epsilon_w^M \equiv \left. \frac{\partial \log h_i}{\partial \log w_i} \right|_{A_i} = -\frac{w_i \frac{\partial \mathcal{M}(h_i)}{\partial w_i}}{h_i \frac{\partial \mathcal{M}(h_i)}{\partial h_i}} \tag{B17}$$

and

$$\epsilon_{1-\tau}^{M} \equiv \left.\frac{\partial \log h_i}{\partial \log\left(1-\tau\right)}\right|_{A_i} = -\frac{\left(1-\tau\right)\frac{\partial \mathcal{M}(h_i)}{\partial(1-\tau)}}{h\frac{\partial \mathcal{M}(h_i)}{\partial h_i}}, \tag{B18}$$

where we have suppressed the dependence of the elasticities on i to reduce notational burden. Similarly suppressing the dependence of derivatives $U\left(c_i, h_i, g_i\right)$ on their arguments and on i, the derivatives of the $\mathcal{M}\left(h_i\right)$ terms in Equations (B17) and (B18) are

$$\frac{\partial \mathcal{M}\left(h_i\right)}{\partial h_i} = \left[\begin{array}{ccc} \left(1-\tau\right)w_i & 1 & \tau w_i \end{array}\right] D^2U \left[\begin{array}{c} \left(1-\tau\right)w_i \\ 1 \\ \tau w_i \end{array}\right],$$

$$\frac{\partial \mathcal{M}\left(h_i\right)}{\partial w_i} = \left[\begin{array}{ccc} \left(1-\tau\right)h_i & 0 & \tau h_i \end{array}\right] D^2U \left[\begin{array}{c} \left(1-\tau\right)w_i \\ 1 \\ \tau w_i \end{array}\right] + DU \left[\begin{array}{c} \left(1-\tau\right) \\ 0 \\ \tau \end{array}\right],$$

and

$$\frac{\partial \mathcal{M}\left(h_i\right)}{\partial\left(1-\tau\right)} = \left[\begin{array}{ccc} w_ih_i & 0 & -w_ih_i \end{array}\right] D^2U \left[\begin{array}{c} \left(1-\tau\right)w_i \\ 1 \\ \tau w_i \end{array}\right] + DU \left[\begin{array}{c} w_i \\ 0 \\ -w_i \end{array}\right],$$

where D^2U is the Hessian matrix of $U\left(c_i, h_i, g_i\right)$. Specializing to the separable social preferences utility in Equation (4) and evaluating (we use the Python symbolic mathematics package SymPy), Equations (B17) and (B18) simplify to Equations (8) and (9).

B.2.2 Hicksian elasticities

The expenditure minimization problem is

$$e\left(a_i, w_i, \tau, \bar{U}_i\right) = \min_{c_i, a_i', h_i} c_i - A_i - \left(1-\tau\right)w_ih_i \tag{B19}$$

$$\text{s.t. } U\left(c_i, h_i, g_i\right) = \bar{U}_i \tag{B20}$$

and Equations (B12) and (B13). Let $h^H\left(a_i, w_i, \tau, \bar{U}_i\right)$ denote the solution. Duality gives

$$h^H\left(a_i, w_i, \tau, \bar{U}_i\right) = h\left(a_i, w_i, \tau, e\left(a_i, w_i, \tau, \bar{U}_i\right)\right), \tag{B21}$$

where $h(\cdot)$ is the Marshallian labor supply function, so that holding A_i constant and differentiating with respect to w_i and $(1-\tau)$ gives, respectively,

$$\begin{aligned}\left.\frac{\partial h_i^H}{\partial w_i}\right|_{A_i,\bar{U}_i} &= \left.\frac{\partial h_i}{\partial w_i}\right|_{A_i} + \left.\frac{\partial h_i}{\partial N_i}\right|_{A_i} \left.\frac{\partial e\left(a_i, w_i, \tau, \bar{U}_i\right)}{\partial w_i}\right|_{A_i,\bar{U}_i} \\ &= \left.\frac{\partial h_i}{\partial w_i}\right|_{A_i} - (1-\tau)\, h_i \left.\frac{\partial h_i}{\partial N_i}\right|_{A_i} \left(1 + \frac{\tau U_g}{(1-\tau)\, U_c}\right) \end{aligned} \tag{B22}$$

and

$$\begin{aligned}\left.\frac{\partial h_i^H}{\partial (1-\tau)}\right|_{A_i,\bar{U}_i} &= \left.\frac{\partial h_i}{\partial (1-\tau)}\right|_{A_i} + \left.\frac{\partial h_i}{\partial N_i}\right|_{A_i} \left.\frac{\partial e\left(a_i, w_i, \tau, \bar{U}_i\right)}{\partial (1-\tau)}\right|_{A_i,\bar{U}_i} \\ &= \left.\frac{\partial h_i}{\partial (1-\tau)}\right|_{A_i} - w_i h_i \left.\frac{\partial h_i}{\partial N_i}\right|_{A_i} \left(1 - \frac{U_g}{U_c}\right), \end{aligned} \tag{B23}$$

where we have used the envelope theorem and the first order condition with respect to c_i in the second equality. The elasticities are thus

$$\epsilon_w^H \equiv \left.\frac{\partial \log h_i^H}{\partial \log w_i}\right|_{A_i,\bar{U}_i} = \epsilon_w^M - (1-\tau)\, w_i \left.\frac{\partial h_i}{\partial N_i}\right|_{A_i} \left(1 + \frac{\tau U_g}{(1-\tau)\, U_c}\right) \tag{B24}$$

and

$$\epsilon_{1-\tau}^H \equiv \left.\frac{\partial \log h_i^H}{\partial \log (1-\tau)}\right|_{A_i,\bar{U}_i} = \epsilon_{1-\tau}^M - (1-\tau)\, w_i \left.\frac{\partial h_i}{\partial N_i}\right|_{A_i} \left(1 - \frac{U_g}{U_c}\right), \tag{B25}$$

with $\left.\frac{\partial h_i}{\partial N_i}\right|_{A_i}$ obtained by implicit differentiation of Equation (B16):

$$\left.\frac{\partial h_i}{\partial N_i}\right|_{A_i} = -\frac{\frac{\partial \mathcal{M}(h_i)}{\partial N_i}}{\frac{\partial \mathcal{M}(h_i)}{\partial h_i}} = -\frac{\begin{bmatrix} 1 & 0 & 0 \end{bmatrix} D^2 U \begin{bmatrix} (1-\tau)\, w_i \\ 1 \\ \tau w_i \end{bmatrix}}{\frac{\partial \mathcal{M}(h_i)}{\partial h_i}}. \tag{B26}$$

Substituting in the expressions from Equations (B17) and (B18), and simplifying, we get

$$\epsilon_w^H = \frac{\text{Num}_w^H}{h_i \frac{\partial \mathcal{M}(h_i)}{\partial h_i}}, \tag{B27}$$

$$\text{Num}_w^H = \begin{bmatrix} \frac{U_g}{U_c}\tau w_i h_i & 0 & -\tau w_i h_i \end{bmatrix} D^2 U \begin{bmatrix} (1-\tau)\, w_i \\ 1 \\ \tau w_i \end{bmatrix} + DU \begin{bmatrix} -(1-\tau)\, w_i \\ 0 \\ -\tau w_i \end{bmatrix},$$

and

$$\epsilon_{1-\tau}^{H} = \frac{\text{Num}_{1-\tau}^{H}}{h_i \frac{\partial \mathcal{M}(h_i)}{\partial h_i}}, \tag{B28}$$

$$\text{Num}_{1-\tau}^{H} = \left[\begin{array}{ccc} -\frac{U_g}{U_c}(1-\tau) w_i h_i & 0 & (1-\tau) w_i h_i \end{array} \right] D^2 U \left[\begin{array}{c} (1-\tau) w_i \\ 1 \\ \tau w_i \end{array} \right] + DU \left[\begin{array}{c} -(1-\tau) w_i \\ 0 \\ (1-\tau) w_i \end{array} \right].$$

Finally, specializing to the separable social preferences utility in Equation (4) gives Equation (11) and (12).

B.2.3 Frisch elasticities

The Frisch elasticities hold constant the marginal utility of wealth λ_i. To do this, rewrite Equations (B14) and (B15) in matrix form to get

$$\boldsymbol{\mathcal{F}}(c_i, h_i) \equiv \left[\begin{array}{c} U_c(c_i, h_i, g_i) \\ U_h(c_i, h_i, g_i) + \tau w_i U_g(c_i, h_i, g_i) \end{array} \right] + \left[\begin{array}{c} -\lambda_i \\ \lambda_i (1-\tau) w_i \end{array} \right] = \mathbf{0}.$$

This is a system of two equations in c_i and h, with parameters τ, w_i, and λ_i. (g_i is not a parameter since it is pinned down by Equation (B13)). Implicit differentiation gives

$$\left[\begin{array}{c} \frac{\partial c_i}{\partial w_i}\Big|_{\lambda_i} \\ \frac{\partial h_i}{\partial w_i}\Big|_{\lambda_i} \end{array} \right] = -\left(D\boldsymbol{\mathcal{F}}(c_i, h_i)\right)^{-1} \frac{\partial \boldsymbol{\mathcal{F}}(c_i, h_i)}{\partial w_i}\Big|_{\lambda_i}, \tag{B29}$$

and

$$\left[\begin{array}{c} \frac{\partial c_i}{\partial (1-\tau)}\Big|_{\lambda_i} \\ \frac{\partial h_i}{\partial (1-\tau)}\Big|_{\lambda_i} \end{array} \right] = -\left(D\boldsymbol{\mathcal{F}}(c_i, h_i)\right)^{-1} \frac{\partial \boldsymbol{\mathcal{F}}(c_i, h_i)}{\partial (1-\tau)}\Big|_{\lambda_i}, \tag{B30}$$

where

$$\frac{\partial \boldsymbol{\mathcal{F}}(c_i, h_i)}{\partial w_i}\Big|_{\lambda_i} = \left[\begin{array}{c} \tau h_i U_{cg} \\ \tau h_i U_{hg} + \tau U_g + \tau^2 w_i h_i U_{gg} + (1-\tau)\lambda_i \end{array} \right],$$

$$\frac{\partial \boldsymbol{\mathcal{F}}(c_i, h_i)}{\partial (1-\tau)}\Big|_{\lambda_i} = \left[\begin{array}{c} -w_i h_i U_{cg} \\ -w_i h_i U_{hg} - w_i U_g - \tau w_i^2 h_i U_{gg} + w_i \lambda_i \end{array} \right],$$

and

$$D\boldsymbol{\mathcal{F}}(c_i, h_i) = \left[\begin{array}{cc} U_{cc} & U_{ch} + \tau w_i U_{cg} \\ U_{ch} + \tau w_i U_{cg} & U_{hh} + 2\tau w_i U_{hg} + \tau^2 w_i^2 U_{gg} \end{array} \right].$$

The Frisch elasticities are the lower components of the vectors in Equations (B29) and (B30), multiplied by w_i/h_i and $(1-\tau)/h_i$ respectively. Specializing to the separable social preferences utility in Equation (4) gives Equations (5) and (6).

B.3 Solution for model with w_i and τ in the utility function

At the end of Section 2, we discuss what happens when we allow utility to have warm glow preferences with regards to the wage. In Section 4.3, we discuss a robustness check that explicitly includes the common tax rate τ in the utility function along with g_i, shown in Equation B8. (The results are presented in Table 4.) In this section, we show the functional form of elasticities when we incorporate w_i and τ in the social preference term.

The expanded utility function is

$$V(a_i, w_i, \tau, N_i) = \max_{c_i, a_i', h_i} U(c_i, h_i, g_i, w_i, \tau) + \beta E\left[V(a_i', w_i', \tau', N_i')\right] \tag{B31}$$

$$\text{s.t. } c_i = A_i + (1-\tau) w_i h_i + N_i, \tag{B32}$$

$$A_i = a_i - \frac{1}{1+r} a_i', \tag{B33}$$

$$g_i = \tau w_i h_i, \tag{B34}$$

where terms follow Section B.2. Derivations of the solution proceed analogous to Section B.2, except with four components in the matrices rather than three. Specializing to $U(c_i, h_i, g_i, w_i, \tau) = u(c_i, h_i) + v(g_i, w_i, \tau)$, the elasticities obtained are

$$\epsilon_w^M = \frac{-w_i(1-\tau)\left\{\left[U_c + \frac{\tau}{1-\tau}(U_g + U_{gg} g_i + w_i U_{gw})\right] + h_i\left[(1-\tau) w_i U_{cc} + U_{ch}\right]\right\}}{h_i\left[(1-\tau)^2 w_i^2 U_{cc} + 2(1-\tau) w_i U_{ch} + U_{hh} + \tau^2 w_i^2 U_{gg}\right]}, \tag{B35}$$

$$\epsilon_{1-\tau}^M = \frac{-w_i(1-\tau)\left\{\left[U_c - (U_g + U_{gg} g_i + \tau U_{g\tau})\right] + h_i\left[(1-\tau) w_i U_{cc} + U_{ch}\right]\right\}}{h_i\left[(1-\tau)^2 w_i^2 U_{cc} + 2(1-\tau) w_i U_{ch} + U_{hh} + \tau^2 w_i^2 U_{gg}\right]}, \tag{B36}$$

$$\epsilon_w^H = \frac{-w_i(1-\tau)\left\{\left[U_c + \frac{\tau}{1-\tau}(U_g + U_{gg} g_i + w_i U_{gw})\right] - \left(\frac{\tau}{1-\tau} h_i \frac{U_g}{U_c} + \frac{U_w}{(1-\tau) U_c}\right)\left[(1-\tau) w_i U_{cc} + U_{ch}\right]\right\}}{h_i\left[(1-\tau)^2 w_i^2 U_{cc} + 2(1-\tau) w_i U_{ch} + U_{hh} + \tau^2 w_i^2 U_{gg}\right]}, \tag{B37}$$

$$\epsilon_{1-\tau}^H = \frac{-w_i(1-\tau)\left\{\left[U_c - (U_g + U_{gg} g_i + \tau U_{g\tau})\right] + \left(h_i \frac{U_g}{U_c} + \frac{U_\tau}{w_i U_c}\right)\left[(1-\tau) w_i U_{cc} + U_{ch}\right]\right\}}{h_i\left[(1-\tau)^2 w_i^2 U_{cc} + 2(1-\tau) w_i U_{ch} + U_{hh} + \tau^2 w_i^2 U_{gg}\right]}, \tag{B38}$$

$$\epsilon_w^F = \frac{-w_i(1-\tau) U_{cc}\left[U_c + \frac{\tau}{1-\tau}(U_g + U_{gg} g_i + w_i U_{gw})\right]}{h_i\left(U_{cc} U_{hh} - U_{ch}^2 + \tau^2 w_i^2 U_{cc} U_{gg}\right)}, \tag{B39}$$

and

$$\epsilon^F_{1-\tau} = \frac{-w_i\left(1-\tau\right)U_{cc}\left[U_c - \left(U_g + U_{gg}g_i + \tau U_{g\tau}\right)\right]}{h_i\left(U_{cc}U_{hh} - U_{ch}^2 + \tau^2 w_i^2 U_{cc}U_{gg}\right)}. \tag{B40}$$

Compared with the main-model analogs in Equations (5) to (13), the wage elasticities ϵ^M_w,ϵ^H_w and ϵ^F_w contain an additional U_{gw} term while the net-of-tax rate elasticities $\epsilon^M_{1-\tau}$, $\epsilon^H_{1-\tau}$ and $\epsilon^F_{1-\tau}$ contain an only one additional $U_{g\tau}$ term. Interestingly, the marginal utilities of wage and tax rate, U_w and U_τ, appear once in the Hicksian elasticities (and not in other elasticities) and the curvature terms U_{ww} and $U_{\tau\tau}$ do not appear at all. The interpretation here is that the direct marginal utilities of wage and tax rate do not affect the labor-leisure substitution (captured by the first term in the numerator comprising U_c, $U_g + U_{gg}g_i$, and $U_{gw}/U_{g\tau}$). This occurs because, unlike g_i which depends on how much labor the individual decides to supply via Equation (B3), the wage or tax rate is exogenous to the individual. Focusing on the tax rate, intuitively, increasing labor supply increases taxes paid which provide a direct utility benefit (or cost) beyond the effect on disposable income. In contrast, the tax rate is exogenous (working more would not change the tax rate), and hence working more does not provide any first-order social preferences benefit once g_i and disposable income are accounted for. Consequently, only the complementarity or substitutability in utility between w or τ and other variables matters for labor-leisure substitution. Finally, in the Hicksian elasticity, U_w and U_τ appear in front of the income effect term because wages and taxes directly affect utility levels and the income effect compensates for that.

B.4 Solution for program-specific taxes paid

Again, the solution proceeds analogous to Section B.2. After specializing to the separable utility function

$$U\left(c_i, h_i, g_i, s_i\right) = u\left(c_i, h_i\right) + v\left(g, s_i\right),$$

the elasticities obtained are

$$\epsilon^M_w = \frac{-w_i\left(1-\tau-\sigma\right)\left(\mathfrak{S}_w + \mathfrak{I}\right)}{h_i\left(w_i^2\left(1-\tau-\sigma\right)^2 U_{cc} + 2U_{ch}w_i\left(1-\tau-\sigma\right) + U_{hh} + U_{gg}\tau^2 w_i^2 + 2U_{gs}\tau\sigma w_i^2 + U_{ss}\sigma^2 w_i^2\right)},$$

$$\epsilon^M_{1-\sigma} = \frac{-w_i\left(1-\sigma\right)\left(\mathfrak{S}_{1-\sigma} + \mathfrak{I}\right)}{h_i\left(w_i^2\left(1-\tau-\sigma\right)^2 U_{cc} + 2U_{ch}w_i\left(1-\tau-\sigma\right) + U_{hh} + U_{gg}\tau^2 w_i^2 + 2U_{gs}\tau\sigma w_i^2 + U_{ss}\sigma^2 w_i^2\right)},$$

$$\epsilon^H_w = \frac{-w_i\left(1-\tau-\sigma\right)\left(\mathfrak{S}_w - \left(\frac{\tau}{1-\tau-\sigma}\frac{U_g}{U_c} + \frac{\sigma}{1-\tau-\sigma}\frac{U_s}{U_c}\right)\mathfrak{I}\right)}{h_i\left(w_i^2\left(1-\tau-\sigma\right)^2 U_{cc} + 2U_{ch}w_i\left(1-\tau-\sigma\right) + U_{hh} + U_{gg}\tau^2 w_i^2 + 2U_{gs}\tau\sigma w_i^2 + U_{ss}\sigma^2 w_i^2\right)},$$

$$\epsilon^H_{1-\sigma} = \frac{-w_i\left(1-\sigma\right)\left(\mathfrak{S}_{1-\sigma} + \frac{U_s}{U_c}\mathfrak{I}\right)}{h_i\left(w_i^2\left(1-\tau-\sigma\right)^2 U_{cc} + 2U_{ch}w_i\left(1-\tau-\sigma\right) + U_{hh} + U_{gg}\tau^2 w_i^2 + 2U_{gs}\tau\sigma w_i^2 + U_{ss}\sigma^2 w_i^2\right)},$$

$$\epsilon_w^F = \frac{-w_i\left(1-\tau-\sigma\right)U_{cc}\mathfrak{S}_w}{h_i\left(U_{cc}U_{hh}+U_{cc}U_{gg}\tau^2w_i^2+2U_{cc}U_{gs}\tau\sigma w_i^2+U_{cc}U_{ss}\sigma^2w_i^2-U_{ch}^2\right)},$$

and

$$\epsilon_{1-\sigma}^F = \frac{-w_i\left(1-\sigma\right)U_{cc}\mathfrak{S}_{1-\sigma}}{h_i\left(U_{cc}U_{hh}+U_{cc}U_{gg}\tau^2w_i^2+2U_{cc}U_{gs}\tau\sigma w_i^2+U_{cc}U_{ss}\sigma^2w_i^2-U_{ch}^2\right)},$$

where

$$\mathfrak{S}_w = U_c + \frac{\tau}{1-\tau-\sigma}\left(U_g+U_{gg}g_i\right) + \frac{\sigma}{1-\tau-\sigma}\left(U_s+U_{ss}s_i\right) + 2\frac{\tau\sigma}{1-\tau-\sigma}U_{gs}h_iw_i,$$

$$\mathfrak{S}_{1-\sigma} = U_c - U_s - U_{ss}s_i - U_{gs}g_i,$$

and

$$\mathfrak{I} = U_{cc}h_iw_i\left(1-\tau-\sigma\right) + U_{ch}h_i.$$

B.5 Optimal income taxation

In this subsection, we derive implications of government-regarding social preferences for optimal income taxation. We examine two main cases: the optimal uniform tax rate and the optimal nonlinear tax schedule. Derivations below follow the steps in Piketty and Saez (2013).

B.5.1 Optimal uniform income taxation

Consider a uniform tax rate τ used to fund a demogrant R (a transfer of equal value for all individuals). The government chooses τ to maximize[28]

$$SWF = \int_i \omega_i \Gamma\left(\hat{V}_i\left(\tau,R\right)\right)d\nu\left(i\right), \tag{B41}$$

where the integration is over a measure-one population of individuals in the economy (indexed by i) with distribution $\nu\left(i\right)$, ω_i is a set of Pareto weights, $\Gamma\left(\cdot\right)$ is increasing and concave. $\hat{V}_i\left(\tau,R\right)$ is the i-specific intratemporal value function from Section B.2:

$$\begin{aligned}\hat{V}_i\left(\tau,R\right) &\equiv V\left(a_i,w_i,\tau,N_i+R\right)|_{A_i} \\ &= U\left(A_i+\left(1-\tau\right)w_ih_i^*+N_i+R,h_i^*,\tau w_ih_i^*\right)+\beta E\left[\cdot\right],\end{aligned} \tag{B42}$$

[28] Following the literature, we introduce the i notation to emphasize possible heterogeneity in the population. All other notational elements are chosen to maintain continuity with previous sections.

where the $\beta E[\cdot]$ term is a constant once we hold A_i constant, and $h_i^* = h_i(a_i, w_i, \tau, N_i + R)|_{A_i}$. The demogrant R is evenly distributed to everybody, so

$$R = \tau \int_i z_i^* d\nu(i) - E \tag{B43}$$

where E is some government non-transfer spending and $z_i^* \equiv w_i h_i^*$ is total pre-tax income, introduced to save on some notation. Note that h_i^* (and hence z_i^*) is a function of τ, so

$$\begin{aligned}\frac{dR}{d\tau} &= \int_i z_i^* d\nu(i) + \tau \int_i w_i \frac{\partial h_i^*}{\partial \tau} d\nu(i) \\ &= \int_i z_i^* d\nu(i) - \frac{\tau}{1-\tau} \epsilon_{1-\tau}^M \int_i z_i^* d\nu(i)\end{aligned} \tag{B44}$$

after some derivations. As is well known in the literature, when the tax rate increases, the demogrant increases mechanically (the first term) but not as much as expected because labor supply decreases (the second term), the latter of which is governed by the Marshallian elasticity $\epsilon_{1-\tau}^M$.

The first-order condition for the government, computed by differentiating Equation (B41) subject to Equation (B43), is

$$\int_i \omega_i \Gamma'\left(\hat{V}_i(\tau, R)\right)\left[U_{c,i}\left[-z_i^* + \frac{dR}{d\tau}\right] + U_g z_i^*\right] d\nu(i) = 0. \tag{B45}$$

Note that $U_g z_i^*$ does not envelope out since individuals cannot choose the amount of government expenditure. For comparative statics, it is convenient to use the "usual" social marginal welfare weight on person i (usual because it does not incorporate any U_g term):

$$\gamma_i \equiv \frac{\omega_i \Gamma'\left(\hat{V}_i(\tau, R)\right) U_c}{\int_j \omega_j \Gamma'\left(\hat{V}_j(\tau, R)\right) U_{c,j}}. \tag{B46}$$

Substituting Equations (B44) and (B46) into Equation (B45), and after some manipulation, we get

$$\tau = \frac{1 - \bar{\gamma}}{1 - \bar{\gamma} + \epsilon_{1-\tau}^M} \tag{B47}$$

where

$$\bar{\gamma} \equiv \frac{\int_i \gamma_i \left(1 - \frac{U_g}{U_c}\right) z_i^* d\nu(i)}{\int_i z_i^* d\nu(i)} \tag{B48}$$

is a penalized income-weighted average of the social marginal welfare weight, with penalty due to government-regarding social preferences.

Compared to a world without social preferences, the only difference is that $\bar{\gamma}$ in Equation

(B48) now has $\left(1-\frac{U_g}{U_c}\right)$ instead of 1 in the numerator. Hence, assuming that $U_g \geq 0$ for all individuals, the existence of government-regarding social preferences leads to smaller $\bar{\gamma}$.[29] From Equation (B47), this means that the optimal tax rate would be larger in the presence of positive social preferences.

B.5.2 Optimal nonlinear tax schedule

Now the government's problem is to choose a nonlinear tax schedule $T(z)$ to maximize

$$SWF = \int_i \omega_i \Gamma\left(\hat{V}_i(T,R)\right) d\nu(i) \tag{B49}$$

$$\text{s.t. } R = \int_i T(z_i^*)\, d\nu(i) - E$$

where

$$\hat{V}_i(T,R) = V(a_i, w_i, \tau_i, N_i + R)\,|_{A_i} = U(A_i + z_i^* - T(z_i^*) + N_i + R, h_i^*, T(z_i^*)) + \beta E[\cdot].$$

$T(z)$ gives the total amount of taxes paid, such that $T'(z)$ is the marginal tax rate at any income z, and $\hat{V}_i(\cdot)$ is a higher order function for notational simplicity. We assume away income effects, and follow the graphical argument around Piketty and Saez (2013, Figure 6) to consider the mechanical, behavioral, and welfare loss effects of a $d\tau$ increase in the marginal tax rate between z and dz.

The first two effects do not depend on the utility function and retain their standard forms (well known in the literature). The mechanical effect is the total increase in tax rate $dzd\tau$ applied to all incomes above z, hence:

$$dM = dzd\tau\,[1 - H(z)], \tag{B50}$$

[29] $U_g \geq 0$ for all individuals is of course an overly-conservative bound, since the presence of some individuals with negative social preferences need not necessarily turn the whole average negative. The precise bound can be obtained by rewriting Equation (B48) as

$$\bar{\gamma} = E\left[\gamma_i \frac{z_i^*}{E[z_i^*]}\right]\left(1 - \left(E\left[\frac{U_g}{U_c}\right] + \frac{Cov\left[\gamma_i \frac{z_i^*}{E[z_i^*]}, \frac{U_g}{U_c}\right]}{E\left[\gamma_i \frac{z_i^*}{E[z_i^*]}\right]}\right)\right).$$

The first term $E\left[\gamma_i \frac{z_i^*}{E[z_i^*]}\right]$ is the non-penalized weighted average social marginal welfare weight (i.e. the $\bar{\gamma}$ in a world without social preferences). Hence, $\bar{\gamma}$ decreases relative to a world without social preferences if $E\left[\frac{U_g}{U_c}\right] + \frac{Cov\left[\gamma_i \frac{z_i^*}{E[z_i^*]}, \frac{U_g}{U_c}\right]}{E\left[\gamma_i \frac{z_i^*}{E[z_i^*]}\right]} > 0$.

where $H(z)$ is the (endogenous) cumulative distribution function of income with density $h(z)$. The behavioral effect only affects the $h(z)\,dz$ individuals with incomes between z and dz since all other incomes face no change in the marginal tax rate and we have assumed away income effects. An affected individual i changes her income by $w_i \frac{\partial h_i^*}{\partial \tau} d\tau = -z_i^* \frac{d\tau}{1-T'(z)} \epsilon_{1-\tau,i}$, where $\epsilon_{1-\tau,i}$ is the individual-specific ETI, yielding tax revenue consequences of $-z_i^* \frac{d\tau}{1-T'(z)} \epsilon_{1-\tau,i} T'(z)$. Put together, the behavioral effect is

$$dB = -dz d\tau h(z) \epsilon_{1-\tau} z \frac{T'(z)}{1-T'(z)}. \tag{B51}$$

The welfare loss of the change is obtained by considering the effect on the social welfare function in Equation (B49), restricted to the people with incomes above z who pay $dz d\tau$ more in taxes, and noting that any labor supply effects envelope out when considering the effects on welfare. Using the marginal value of public funds for the government p to convert from utility to dollars, the money-metric welfare loss is (after some algebraic manipulation)

$$dW = -dz d\tau \int_{z_i > z} \gamma_i \left(1 - \frac{U_g}{U_c}\right) dH(z_i) \tag{B52}$$

$$= -dz d\tau \left[1 - H(z)\right] \gamma^+(z), \tag{B53}$$

where $\gamma_i \equiv \omega_i \Gamma' \left(\hat{V}_i(T, R)\right) U_c / p$ in Equation (B52) is again the "usual" social marginal welfare weight on person i, and in Equation (B53),

$$\gamma^+(z) \equiv \frac{1}{1 - H(z)} \int_{z_i > z} \gamma_i \left(1 - \frac{U_g}{U_c}\right) dH(z_i) \tag{B54}$$

is the average social marginal welfare weight for individuals with income above z, penalized by government-regarding social preferences.

The optimal marginal tax rate at any z is obtained by setting the sum of the effects in Equations (B50), (B51), and (B53) to zero. This yields

$$T'(z) = \frac{1 - \gamma^+(z)}{1 - \gamma^+(z) + \alpha(z) \epsilon_{1-\tau}}$$

where $\alpha(z) \equiv \frac{z h(z)}{[1 - H(z)]}$ is a parameter that is related to the Pareto distribution.

Again, compared to a world without social preferences, the only difference is that $\gamma^+(z)$ in Equation (B54) incorporates a social preferences penalization term. On top of the interpretation from Section B.5.1 — positive social preferences allows the government to levy higher income taxes on average — Equation (B54) reveals that heterogeneity in social preferences can also affect the marginal tax rate at different points of the income distribution.

C Survey and vignette experiment details

A live version of the survey is available at https://nus.syd1.qualtrics.com/jfe/form/SV_cAszYDfScYSrEUu. The full questionnaire is available at https://www.pinchuanong.com/wp-content/uploads/2024/07/appx-d.pdf. Some of our background and political view questions were adapted from Stantcheva (2022) and Doherty et al. (2015). Below, we provide details on specific elements of the survey.

C.1 Ensuring high-quality response

We maximized response quality in several ways.

In our survey landing page (consent page), we followed a good practice in the literature and warned participants that responding without adequate effort might result in their responses being marked as low quality and not used (see e.g. Stantcheva 2023). Also following the literature, we appealed to their sense of social responsibility by emphasizing that their responses might form the basis for policy recommendations, and requested that they read the questions carefully and answered honestly. To avoid selection based on survey topics, our landing page provided limited information about what the survey was about. (Specifically, "We are a group of independent non-partisan researchers studying how people around the world make decisions about their jobs.") Throughout the survey, we used simple and neutral language to avoid giving respondents the impression that we had a hypothesis in mind.

We took several steps to ensure accurate measurement of employment variables. For ease of answering, when asking respondents about their earnings, we adapted our questions to the respondents by offering the option of entering an annual, monthly or weekly figure. Because the hourly wage (earnings divided by work hours) can be difficult to measure (see e.g. Borjas 1980), we programmed a dynamic check of the wage in the survey instrument. Specifically, immediately after the respondent entered her earnings and work hours, we showed her the computed wage and asked her if it sounded "about right". If the answer was "No", we provided the computation formula and requested that she checked the figures again. The key employment variables (work hours, earnings, and verified wage) were all asked on the same webpage to make internal consistency easy for the respondent (since there was no need to click the "back" button to change numbers).

Our vignette experiment was necessarily more complex than other parts of our survey which comprised simple-to-answer questions. We implemented four features to ensure the highest possible quality answers. First, before the vignette experiment, we required respondents to go through an animated slideshow that walked them through the instructions. In total, there were ten slides that explained the hypothetical scenarios, the assumptions (as

described in Section 3.2), and the choices they would be asked to make. Sentences on the slides appeared at 250 words per minute (between the usual reading and speaking speed), and the next slide button was grayed out (i.e., not responsive) until all the text had appeared. Hence, unlike an instructional video, participants could not simply start the slideshow and return after some time to resume the survey — they had to advance the slides a total of ten times throughout the slideshow.

Second, we focused the respondents' attention on the question at hand by hiding all non-relevant information in the questionnaire matrix. For example, when the respondent had to choose whether to work or not in the first scenario, we hid the second scenario and the question on which scenario he preferred. When he had to make the analogous choice for the second scenario, we similarly hid the first scenario, including the previous response to minimize anchoring effects. We revealed all scenarios and choices only when the respondent was to choose between the two scenarios.

Third, on the first pair of scenarios, we ensured that respondents understood the layout and how to answer by showing an interactive "guided website tour" through each component (see Appendix Figure A2). Every time we introduced a change in the vignette layout afterwards (i.e., pairs 5 and 9), we showed a new interactive guided tour to explain the change and to maintain their attention.

Fourth, we programmed a check for answers exhibiting intransitive preferences — choices to work in a scenario of a pair and not the other, but yet the respondent preferred the latter scenario to the former — and showed a non-intrusive prompt asking respondents to double-check their answers when an intransitive choice was made. Without the prompt, the worry is that some (more attentive) respondents may notice that we allow inconsistent choices, think that the survey was badly designed, and lose motivation. We kept track of when the prompt was observed, and investigate robustness to observation of the prompt in Section 4.3.

When eliciting opinions about government spending on specific programs, we used a more complex sampling mechanism (of the programs) to balance multiple objectives. Our primary goal is to study how opinions of a program affect the program-specific wage-tax elasticity wedge in pairs 9 and 10 of the vignettes (one program per pair). These two programs were selected from a list of ten using a randomization procedure that weighted dissimilar programs more. Specifically, each pair of programs (e.g. Social Security and Medicare) had a probability of selection that was proportional to the Euclidean distance in liking score between the programs of the pair.[30] Later on in the survey when we asked respondents how much they liked/disliked their tax money being used to fund specific programs, each

[30]The Euclidean distances were computed based on a pilot dataset. Interpretation of these distances as a measure of dissimilarity is frequently used in data clustering algorithms.

respondent was asked to rate five programs including the two they saw in pairs 9 and 10 of the vignette and three other randomly chosen programs. We added the three randomly chosen programs to better mask our objective for asking about the opinion of government spending and because these opinions are of general interest in and of themselves. (We did not present the full list of ten programs to every respondent to reduce response burden.) Liking was measured on a 5-point Likert scale, with the two extremes corresponding to "dislike a lot" and "like a lot" (we randomized the scale direction between respondents). To ensure that respondents understood the programs, we included a short description in both the last two pairs of vignettes and the liking questions.

We conducted the survey when there was some worry about bots being used to answer surveys (Goodrich et al. 2023). Thus, we followed good practices in the literature to ensure that our respondents were legitimate and verified (Stantcheva 2023). Besides distributing our survey through a reliable company which also performed checks of their participants, we incorporated Qualtric's bot detection capabilities and added a "honey pot" question on our landing page. The question "Check this box if it applies to you. [I am a piece of software/I am not a piece of software]" appeared for one second and was displaced downwards by one screen height, so that a human participant saw a blank screen for the first second (consistent with the webpage still loading). Among all participants who started the survey, this question only detected one bot that passed upstream bot checks. (We screened the respondent out.)

We kept track of three measures of response quality in the survey. First, we followed the literature (see e.g. Mas and Pallais 2017) to build two attention checks into the survey. The first check (which we refer to as the pre-vignette attention check) doubles as a comprehension check and asked participants a simple question about the instructions of the vignette following the slideshow. Specifically, respondents were asked whether the income tax rate in the hypothetical scenarios referred to federal, state or local income tax rate (the correct answer being federal). The second check (the in-vignette attention check) was built into pair 7 of our vignette experiment. In this pair, by design, one of the two scenarios was strictly dominating: both scenarios showed the same tax rate and work hours but one had a strictly higher hourly wage rate. We would expect attentive participants to prefer the strictly dominating scenario. We chose pair 7 as our in-vignette attention check since vignette fatigue would likely affect respondents the most around this pair.

Second, we exploited the interactive nature of our instructional slideshow to identify "impatient" respondents. Since respondents could not advance the slides until all the text was displayed — and the graying out of the "next" button made this obvious — respondents who were thinking about the instructions carefully would only click "next" a total of ten times. Hence, the number of times this "next" button was clicked when it was still grayed

out serves as a natural measure of response quality. In practice, we labeled respondents who clicked more than twenty times (twice as many as needed) as impatient respondents.

Finally, the time spent by respondents on the survey as a whole, as well as on the vignette experiment and other questions, allows us to identify respondents who spent too little time answering questions. This is a standard measure used in the survey research and practitioner literature (Zhang and Conrad 2014; Greszki et al. 2015; Leiner 2019). The literature and practitioners often use thresholds relative to the median time (see e.g. Greszki et al. 2015); we follow this practice and label respondents who responded faster than half the median time as "too fast". We considered three time intervals when setting this measure: the overall time taken, the time taken on the vignette experiment (our main focus), and the time taken for the government expenditure liking questions (to measure quality near the end of the survey, and because we use a quality sample when investigating the impact of liking the program that taxes fund).

C.2 Income tax rate simulation

To avoid showing tax rates that differed too much from the respondent's actual situation, our vignette experiment randomly selected tax rates around the respondent's actual tax rate. The actual tax rate was computed by simulation, both to simplify the details that respondents should focus on (e.g. we simply needed them to focus on the average tax rate for our exercise, ignoring the complications that come with marginal tax rates) and because respondents might not recall the tax rates that they were paying. We did not ask respondents directly about their taxes paid or ask them to verify the simulated tax rate to avoid emphasizing taxes beyond what was necessary. To avoid the internet speed reduction and risk of connection failure that comes with querying an external website, we elected to program the simulation directly in the survey, instead of obtaining the tax rate from an external source like the National Bureau of Economic Research TAXSIM.

We modeled three filing categories: married filing jointly if the respondent was married, single-filer if the respondent was non-married without any children, and head of household if the respondent was non-married with children. Total household income comprised wage income, interest and dividend income, capital gains income, and business income.[31] Wage income was directly collected in our survey (for both the respondent and their spouse if present). We imputed the other types of income separately for married and single respondents. To do this, we ran a regression of each income type on several demographic characteristics using the 2015–2019 Current Population Survey data. We then used these

[31] All numbers in the income tax rate simulation were adjusted to be in 2022 dollars.

models and the demographic information collected in our survey to predict the interest and dividend income (set to zero if the resulting prediction was negative), capital gains income and business income for each respondent.

We modeled deductions as a weighted average of the standard deduction and an itemized deduction which varied across household income and filing category. The weights and itemized deduction amount were estimated based on the Internal Revenue Service's Statistics of Income (Internal Revenue Service 2021, Table 1.2). We computed taxable income as total household income less deductions, and applied the 2022 federal income tax brackets to obtain an income tax before credits amount for each respondent.

Next, we modeled two large tax credits, the Child Tax Credit and the Earned Income Tax Credit, using the federal tax credit formulas which consider income, number of children, and marital status. We chose these two credits to model because they were large credits that were likely to apply to our respondent sample and did not require additional questions that deviated from our main research question. We then computed the simulated income tax for each respondent as the income tax before credits less the imputed tax credits. Finally, we divided the income tax after credits by total household income to obtain the average tax rate for each respondent.

C.3 Sample

Our target sample comprised US residents aged 25–64 who were working at the time of the survey. Full-time students who might have worked part-time were excluded from the survey since they might be less familiar with the labor market implications of their choices. We also excluded the self-employed (including unpaid family workers) as they might find it more difficult to relate to the scenarios in our vignette which were about finding a new job with similar conditions to the participant's current job. We asked Bilendi (previously called Respondi), our survey company, to target participants such that our sample is representative along the age, gender, and income group dimensions. (We provided target proportions for different groups based on the working population in the 2022 CPS data.)

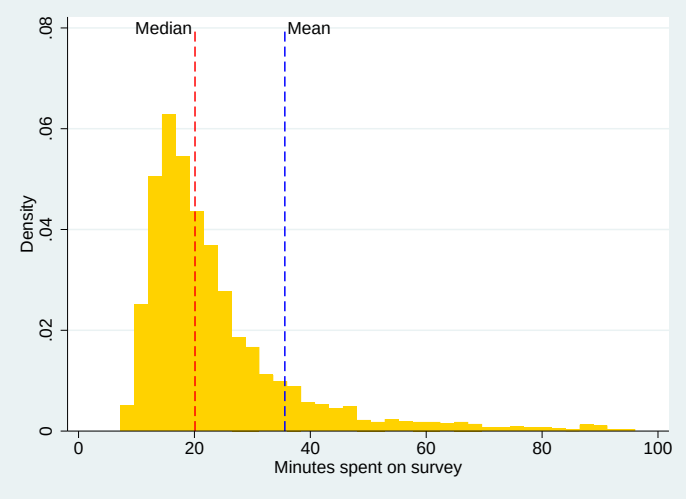
Our survey started with demographic questions to determine participants' eligibility: those who did not meet our sample definition were screened out of the survey. 12,164 respondents started the survey and were not screened out. Appendix Table C1 describes, step-by-step, the sample size reductions for each of our data cleaning steps to arrive at our main sample. Appendix Figure C1 shows the distribution of the time respondents in our main sample spent on completing the survey. The mean and median durations were 35.6 and 20.1 minutes, respectively.

Appendix Table C1: Sample size and data-cleaning steps

	Number of Respondents
Started the survey and not screened out	12,164
Got to just before the instructional slideshow	8,338
Passed the instructional slideshow	7,856
Completed the vignette experiment	7,158
Completed the entire survey	7,122
Excluded:	
Work hours, consumption, or tax amounts below 5th or above 95th percentiles	1,611
Vignette consumption not imputable	90
Skipped slide show	45
Excluded: Failed second-order condition in main specification	1
Final Sample	5,439

Notes: This table shows the number of respondents who started our survey and the sample size reduction for each of our data-cleaning steps. Respondents may be excluded for more than one of the three reasons stated.

Appendix Figure C1: Time Spent on Survey



Notes: The figure shows the distribution of the time respondents in the main analysis sample spent on the survey (truncated at 100 minutes). The mean duration is 35.6 minutes, the median 20.1, and the 25th and 75th percentile are 15.5 and 28.6, respectively.

Appendix Table C2: Inattentive, impatient or too fast respondents

	Number of respondents	Percent of final sample
Inattentive in pre-vignette attention check	380	7.0
Inattentive in in-vignette attention check	446	8.2
Inattentive in either attention check	725	13.3
Clicked next slide >20 times during vignette instruction slideshow	909	16.7
Time taken for any of overall, vignette, and gov. expenditure liking questions is less than half the median time	1,286	23.6

Notes: This table shows the number of respondents and the proportion of the main sample that are "inattentive" (failed our pre-vignette or in-vignette attention checks), "impatient" (clicked next more than 20 times during the vignette instructional slideshow), and "too fast" (those with time spent on any of the overall, vignette and government expenditure liking questions that was less than half of the median time).

We flag respondents with potentially low-quality answers and exclude them in some of our analyses. We did this using the three measures — "inattentive", "impatient" and "too fast" — described in Appendix C.1. Appendix Table C2 reports the number of these respondents and their proportion of our main sample.

C.4 Preregistration and compliance details

This subsection reports on compliance with the preregistration AEARCTR-0011052.

1. Trial information The aim of the preregistered experiment was to investigate whether individuals' labor supply responses differ between tax changes and wage changes using the vignette experiment. The preregistration was submitted on March 7, 2023. The survey commenced March 9, 2023. The last respondent who completed the entire survey submitted the response on June 6, 2023, before the planned end date of June 30, 2023. Before preregistration, we conducted a pilot survey from February 17–25, 2023, focused on logistical issues such as the instructional slideshow and vignette-screen formatting. Pilot responses are not used in the analysis.

2. Experimental details The preregistered outcomes were labor supply responses to wages and taxes in the vignette experiment, and views on income tax and the government. The preregistered experimental design specified multiple pairs of scenarios for a total of 20 per respondent, with respondents choosing whether to take up the job in each scenario and choosing their preferred scenario within each pair.

3. Sample As preregistered, we planned to recruit 5,000 respondents. Bilendi fielded the survey to more respondents to account for non-completion and ensure that the target was met. After applying the data-cleaning steps described in Appendix C.3, our final sample comprises 5,439 respondents. The cleaning steps were not specified in the preregistration because we did not anticipate the influence of extreme values in work hours, consumption, and tax amounts on the simulated wedges and standard errors. We show robustness to respondent quality and alternative treatment of extreme values in Table 4.

4. IRB As declared in the preregistration, the study was approved by the Faculty Ethics Review Committee, NUS Business School, on January 18, 2023, protocol SP-23-0103.

5. Comparison between the study and the preregistration The experiment was conducted in compliance with the preregistration. We summarize the main implementation details that differ from, or are not fully described in, the preregistration.

1. Sample size as explained above.

2. While the total number of scenarios was implemented as preregistered, the preregistration omitted the breakdown of the three scenario layouts.

3. The preregistered hypothesis was nondirectional. The implemented survey and its analysis strictly followed this nondirectional hypothesis. Some terminology may appear directional because it follows the existing literature, but the empirical specifications are two-sided.

References

Blundell, Richard and Thomas MaCurdy, "Labor supply: A review of alternative approaches," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 3, Amsterdam, The Netherlands: Elsevier, 1999, chapter 27, pp. 1559–1695.

Borjas, George J, "The relationship between wages and weekly hours of work: The role of division bias," *Journal of Human Resources*, 1980, *15* (3), 409–423.

Doherty, Carroll, Jocelyn Kiley, Alec Tyson, and Bridget Jameson, "Beyond distrust: How Americans view their government," Technical Report, Pew Research Center 2015.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren, and Michael Westberry, "Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]," 2022. Minneapolis, MN: IPUMS, Retrieved November 7, 2022, from https://doi.org/10.18128/D030.V10.0.

Gallup, "Party Affiliation," 2023. Retrieved October 29, 2023, from https://news.gallup.com/poll/15370/party-affiliation.aspx.

Goodrich, Brittney, Marieke Fenton, Jerrod Penn, John Bovay, and Travis Mountain, "Battling bots: Experiences and strategies to mitigate fraudulent responses in online surveys," *Applied Economic Perspectives and Policy*, 2023, *45* (2), 762–784.

Greszki, Robert, Marco Meyer, and Harald Schoen, "Exploring the effects of removing "too fast" responses and respondents from web surveys," *Public Opinion Quarterly*, 2015, *79* (2), 471–503.

Internal Revenue Service, "Statistics of income — 2019 individual income tax returns," Publication 1304 (Rev. 12-2021), Department of the Treasury Internal Revenue Service, Washington, D.C. 2021. Retrieved September 27, 2022, from https://www.irs.gov/statistics/soi-tax-stats-individual-income-tax-returns-complete-report-publication-1304.

Leiner, Dominik Johannes, "Too fast, too straight, too weird: Non-reactive indicators for meaningless data in internet surveys," *Survey Research Methods*, 2019, *13* (3), 229–248.

Leip, David, "Atlas of U.S. Presidential Elections," 2023. Retrieved October 29, 2023, from https://uselectionatlas.org/.

Mas, Alexandre and Amanda Pallais, "Valuing alternative work arrangements," *American Economic Review*, December 2017, *107* (12), 3722–3759.

Piketty, Thomas and Emmanuel Saez, "Optimal labor income taxation," in "Handbook of Public Economics," Vol. 5, Elsevier, 2013, pp. 391–474.

Stantcheva, Stefanie, "Understanding of trade," Working Paper 30040, National Bureau of Economic Research 2022.

_ , "How to run surveys: A guide to creating your own identifying variation and revealing the invisible," *Annual Review of Economics*, 2023, *15*, 205–234.

Zhang, Chan and Frederick Conrad, "Speeding in web surveys: The tendency to answer very fast and its association with straightlining," *Survey Research Methods*, 2014, *8* (2), 127–135.