Efficient Redistribution*

Corina Boar† Virgiliu Midrigan‡

April 2022

Abstract

We characterize the optimal shape of non-linear income and wealth taxes in a dynamic general equilibrium model with uninsurable idiosyncratic risk. Our analysis reproduces the distribution of income and wealth in the United States and takes into account the long-lived transition dynamics after policy reforms. We find that a uniform flat tax on capital and labor income combined with a lump-sum transfer is nearly optimal. The incremental welfare gains from steeper marginal income and wealth taxes are small, especially when the planner has a strong preference for redistribution, due to strong behavioral and general equilibrium effects.

Keywords: inequality, redistribution, tax policy.

JEL classifications: E2, E6, H2.

*We thank Federico Kochen and Audrey Azerot for superb research assistance, our discussant Dirk Krueger, as well as Marco Bassetto, V.V. Chari, Raquel Fernandez, Baris Kaymak, Albert Marcet, Abdou Ndiaye, Stefanie Stantcheva and numerous seminar participants for valuable comments. We gratefully acknowledge support from the National Science Foundation, Grant SES-1948119.

†New York University and NBER, corina.boar@nyu.edu.

‡New York University and NBER, virgiliu.midrigan@nyu.edu.

1 Introduction

Increased income and wealth inequality gave rise to numerous proposals for redistribution using higher top marginal income taxes and wealth taxes. Motivated by these policy discussions, our paper characterizes the optimal shape of non-linear income and wealth tax schedules in an unequal economy. Critically, and in contrast to existing studies[1], we study joint reforms whereby the government can simultaneously change the income and wealth tax schedules, as well as the amount of lump-sum transfers. We find that a flat uniform tax on income is nearly optimal in that the additional welfare gains from non-linear income and wealth taxation are relatively small. Interestingly, the incremental gains from a richer set of tax instruments are smaller, the stronger the planner's concern for redistribution.

We conduct our analysis in a dynamic general equilibrium economy with uninsurable idiosyncratic risk. We allow for non-linear income and wealth taxes, which, following the recommendation of Heathcote and Tsujiyama (2021b), are restricted to a simple parametric class. The revenue raised with these taxes is used to finance lump-sum transfers. Motivated by recent policy discussions, we consider once-and-for-all tax reforms and evaluate their welfare consequences taking into account the resulting transition dynamics. Optimal policy therefore balances the short-run desire to redistribute against long-run efficiency concerns.

Our economy is inhabited by households who face idiosyncratic shocks to their labor market ability and reproduces the inequality in wealth and income in the data. As in the United States, the income tax base includes both labor and capital income, so a wealth tax allows taxing capital and labor income at different rates. Motivated by Greulich et al. (2022), we consider several social welfare functions: average welfare (Benabou, 2000), which captures pure economic efficiency and disregards equity considerations, as well as objectives that place increasingly higher weight on the poor, such as utilitarian and Rawlsian welfare.

We first consider partial reforms and change the parameters of the tax schedules in isolation. We show that all instruments of redistribution – higher average marginal tax rates, steeper marginal income taxes, as well as wealth taxes – can increase the welfare of the poor. We then turn to the optimal tax experiments and consider joint policy reforms.

We proceed incrementally, by first restricting the planner to only use a flat income tax and then augmenting the set of instruments with non-linear income and wealth taxes. We find that an optimally chosen flat income tax delivers most of the welfare gains attainable with

[1] Guvenen et al. (2019), Kaymak and Poschke (2019), Kindermann and Krueger (2021), Bakis et al. (2015), Heathcote et al. (2017), Imrohoroglu et al. (2018), Brüggemann (2021), Ferriere et al. (2020).

more complex instruments. For example, a utilitarian planner restricted to a flat uniform tax sets it equal to 56%, raising the consumption-equivalent welfare by 7.4%. This represents 87% of the welfare gains attainable with optimally set upward sloping marginal income and wealth taxes. Interestingly, the incremental gains from a richer set of tax instruments are smaller, the stronger is the planner's concern for redistribution: a Rawlsian planner can achieve 97% of the maximum attainable welfare gains by using a flat uniform income tax.

At a first glance, the result that positively sloped marginal income and wealth taxes deliver small welfare gains seems to contradict the findings of our partial reform experiments. There is, in fact, no contradiction. Rather, the result reflects that starting from the optimal flat income tax that is high to begin with, steeper marginal income or wealth taxes generate little additional tax revenue that can be redistributed via lump-sum transfers.

Our finding that a uniform flat income tax is nearly optimal is robust to the details of the parameterization, including household preferences and distribution of ability. In all the experiments we considered, an optimally chosen flat income tax achieves between 72% and 96% of the welfare gains attainable with more complex tax instruments.

Related Work Our result that a flat tax is nearly optimal is reminiscent of Conesa and Krueger (2006) and Conesa et al. (2009). Conesa et al. (2009) allow for non-linear labor income taxes and a flat capital tax in an OLG environment and maximize long-run steady state welfare. They find that a flat tax on labor income combined with a sizable minimum deduction and a high tax on capital income are optimal. In their economy older agents supply less labor and have a higher labor supply elasticity, so it is optimal to tax their labor income at a lower rate. Absent age dependent tax instruments, a tax on capital mimics a labor income tax that falls with age. Relative to these papers, we study an infinite horizon economy and allow for a richer set of instruments, including non-linear wealth taxes. Crucially, we study the problem of a planner who maximizes welfare taking transition dynamics into account. In our framework, a planner concerned only with long-run welfare would subsidize wealth accumulation, ignoring the large welfare losses such policies would entail during the transition.

Two complementary papers that study optimal capital and labor income taxation are Dyrda and Pedroni (2021) and Acikgoz et al. (2021). These papers characterize the optimal path for flat capital and labor income taxes in an economy similar to ours. They find that it is optimal to tax capital initially at a very high rate and labor at a very low rate. Over time, the capital income tax is gradually reduced and the tax on labor increases. In contrast to this work, which restricts attention to flat taxes, we characterize the optimal shape of non-linear income and wealth tax schedules. Moreover, motivated by recent policy discussions on the desirability of increasing top marginal income taxes and introducing a wealth tax, we restrict attention to once-and-for-all tax reforms. Optimal policy in our setting therefore balances the short-run desire to redistribute against long-run efficiency concerns.

Relative to these papers, we argue that the incremental gains from differential taxation of these labor and capital income are very small. In our economy the optimal taxes on labor and capital income are 58% and 48%. Even though these rates are different, the welfare gains this policy entails are nearly as large as those achievable with uniform taxation.

Our paper is also related to the dynamic public finance literature (Farhi and Werning, 2013, Golosov et al., 2016, Stantcheva, 2017) which often finds that age-dependent linear taxes achieve the bulk of the welfare gains. In contrast to this research, usually set in partial equilibrium,[2] we study the general equilibrium consequences of tax reforms. We show that optimal policy would call for much steeper marginal wealth taxes in partial equilibrium, since the planner would not face the production consequences of depressed wealth accumulation.

2 Model

The economy is inhabited by a unit mass of households who face idiosyncratic shocks to labor market ability. We abstract from aggregate uncertainty and study the steady state of the model and transition dynamics after unanticipated optimal policy reforms.

2.1 Households

Households seek to maximize life-time utility given by

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{c_t^{1-\theta}}{1-\theta} - \frac{h_t^{1+\gamma}}{1+\gamma} \right), \tag{1}$$

where c_t is consumption and h_t is hours worked.[3] They derive income from two sources. Labor income $W_t e_t h_t$ depends on the wage rate W_t and the idiosyncratic ability e_t, which follows a Markov process. Asset income $r_{t-1} a_t$ depends on household wealth a_t and the return to savings r_{t-1}. Wealth a_t is the sum of government bonds and physical capital holdings. Absent aggregate uncertainty, the rates of return on these assets are equalized, so we only need to record total household wealth. Households cannot borrow, so $a_{t+1} \geq 0$.

[2]See Farhi and Werning (2012) who incorporate general equilibrium considerations.

[3]Though we refer to h_t as hours worked, we think of it as capturing a broader notion of labor effort.

The budget constraint is

$$(1+\tau_s)\,c_t + a_{t+1} = i_t - T^i\,(i_t) + a_t - T^a\,(a_t)\,, \tag{2}$$

where τ_s is a consumption tax. As in the United States, all income $i_t = W_t e_t h_t + r_{t-1} a_t$ is subject to a non-linear personal income tax schedule. We assume a modified HSV[4] tax function $T\,(i_t) = i_t - (1-\tau)\,\frac{i_t^{1-\xi}}{1-\xi} - \iota_t$, where ι_t is a lump-sum transfer. The parameter τ determines the average level of the marginal income tax and ξ determines its slope. As we show below, allowing for a lump-sum transfer is necessary to match the extent of redistribution in the United States. We note that even though many transfers in the data are income dependent, our assumption that transfers are lump-sum is not restrictive because what matters for households' incentives and welfare is the mapping from their pre-tax to their post-tax income. For example, a tax and transfer system with a phased out means-tested transfer can be equivalently recast as a system with a lump-sum transfer financed by a flat income tax.

We assume that the wealth tax $T^a\,(a_t)$ is zero, as in the United States. However, in computing optimal policy, we allow for non-linear wealth taxes, also of the HSV form $T^a\,(a_t) = a_t - \frac{1-\tau_a}{1-\xi_a} a_t^{1-\xi_a}$. Since all agents face the same rate of return on assets, a non-zero wealth tax simply allows for differential taxation of labor and capital income.

2.2 Technology

Firms produce a homogeneous good with technology $Y_t = K_t^{\alpha} L_t^{1-\alpha}$, where K_t and L_t are capital and labor. They rent capital at a rental rate R_t and hire labor at a wage rate W_t. No-arbitrage implies that $R_{t+1} = r_t + \delta$, where δ is the depreciation rate of capital. Output is used for consumption, investment and government spending.

2.3 Government

The government has an outstanding stock of debt B_t on which it pays the interest rate r_{t-1}. It finances exogenous spending G and collects taxes T_t on personal income, wealth and consumption. The budget constraint is $(1+r_{t-1})\,B_t + G = B_{t+1} + T_t$.

We formally define the equilibrium of this economy in the Appendix.

[4] Benabou (2002), Heathcote et al. (2017).

2.4 Tax Distortions

We next discuss the two distortions in the household saving and labor supply choices introduced by tax policies. Consider first labor supply and let $\tilde{\tau}_{it} = 1-(1-\tau)\left[r_{t-1}a_{it} + W_t e_{it} h_{it}\right]^{-\xi}$ denote the marginal income tax rate faced by household i. The income tax and the consumption tax distort household labor supply by reducing the marginal return to working. In particular, the labor supply choice is given by

$$h_{it}^{\gamma} = \frac{1-\tilde{\tau}_{it}}{1+\tau_s} c_{it}^{-\theta} W_t e_{it} = \frac{1}{\vartheta_{it}} c_{it}^{-\theta} W_t e_{it},$$

where the second equality implicitly defines the labor wedge ϑ_{it}.

Letting $\hat{c}_{it} = \frac{c_{it}}{C_t}$ denote the consumption share of household i and aggregating across households (see Berger et al., 2019 for details) gives the aggregate labor supply

$$L_t^{\gamma} = \frac{1}{\bar{\vartheta}_t} W_t C_t^{-\theta}, \quad \text{where} \quad \bar{\vartheta}_t = \left(\int \vartheta_{it}^{-\frac{1}{\gamma}} \hat{c}_{it}^{-\frac{\theta}{\gamma}} e_{it}^{1+\frac{1}{\gamma}} \mathrm{d}i \right)^{-\gamma}$$

is the aggregate labor wedge which depends on individual labor wedges and the covariance between consumption shares and labor market ability.

Consider next the household's savings choice and let $\tilde{\tau}_{it}^{a} = 1-(1-\tau_a)\, a_{it}^{-\xi_a}$ denote the marginal wealth tax faced by household i. The marginal income and wealth taxes distort the savings choice by lowering the marginal benefit of saving. In particular, the savings choice is

$$c_{it}^{-\theta} = \beta \mathbb{E}_t c_{it+1}^{-\theta} \left[1 - \tilde{\tau}_{it+1}^{a} + (1-\tilde{\tau}_{it+1})\, r_t + \chi_{it}\right] = \beta \mathbb{E}_t c_{it+1}^{-\theta} \frac{1+r_t}{\zeta_{it+1}}, \tag{3}$$

where χ_{it} is the multiplier on the no-borrowing constraint and the second equality implicitly defines the savings wedge ζ_{it}.

Aggregating across households yields the aggregate Euler equation

$$C_t^{-\theta} = \frac{1}{\bar{\zeta}_t} \beta C_{t+1}^{-\theta} (1+r_t), \quad \text{where} \quad \bar{\zeta}_t = \left(\int \mathbb{E}_t \left(\frac{\hat{c}_{it+1}}{\hat{c}_{it}} \right)^{-\theta} \frac{1}{\zeta_{it+1}} \mathrm{d}i \right)^{-1} \tag{4}$$

is the aggregate savings wedge which depends on individual savings wedges and the growth rates of consumption shares.

As is well understood, when the planner has access to a rich set of tax instruments, multiple tax policies can generate similar allocations, so wedges in the optimality conditions provide a more useful account of the distortions induced by the tax system.[5] For example, a wealth tax increases the savings wedge $\bar{\zeta}_t$ and therefore the equilibrium interest rate, reducing

[5] See Chari et al. (2020) for a recent illustration of this point.

the capital-labor ratio and the equilibrium wage. This effect is partly countered by a decline in the labor wedge stemming from wealth effects which encourage the labor supply of more productive households. A higher income tax worsens both the savings and labor wedge, and reduces labor supply, as well as the capital-labor ratio. Optimal policy balances the costs of these distortions against the benefits of insurance and redistribution.

3 Quantifying the Model

In this section we describe our calibration strategy and define measures of social welfare.

3.1 Calibration Strategy

We assume the economy is in a steady-state in 2013 and target statistics for this year.

Assigned Parameters A period is one year. We set the stock of government debt B equal to 100% of GDP, its value in 2013. We set the depreciation rate of capital $\delta = 0.06$, the capital elasticity $\alpha = 1/3$, the relative risk aversion $\theta = 1$, and the inverse of the Frisch elasticity of labor supply $\gamma = 2$, all conventional choices in the literature.

We set the wealth tax equal to zero in the initial steady state. We follow Bhandari and McGrattan (2020) and set the consumption tax $\tau_s = 0.065$. We assume that the unexpected capital gains generated from the reforms are taxed at a rate $\tau_k = 0.20$, consistent with the US tax code in 2013. We summarize these parameters in Panel B of Table 1.

Calibrated Parameters We parameterize the income tax function to replicate the degree of income redistribution in the United States. Specifically, we estimate the parameters ι, τ and ξ to match the CBO data on the shares of income before and after taxes and transfers for eight income groups: the first four quintiles, the 81^{st} to 90^{th} percentile, the 91^{st} to 95^{th} percentile, the 96^{th} to 99^{th} percentile, as well as the top 1 percent. The advantage of the CBO data is that it adjusts its estimates of means-tested transfers for survey under-reporting and thus provides a more accurate account of the transfers to low-income households.

We use data on the pre- and post-tax income shares of the various income groups to estimate the parameters of the tax function using non-linear least squares, weighting each group by its population share. The left panel of Figure 1 depicts both the data and the fitted values from our estimates. The tax function accounts almost perfectly for the extent of redistribution to the poorest quintiles and the degree of tax progressivity at the top. For

Table 1: Parameterization

A. Moments Used in Calibration

	Data	Model
Wealth to income ratio	6.6	6.6
Gini wealth	0.85	0.86
Gini income	0.64	0.65
Wealth share top 0.1%	0.22	0.22
Wealth share top 1%	0.35	0.34
Income share top 0.1%	0.14	0.14
Income share top 1%	0.22	0.22

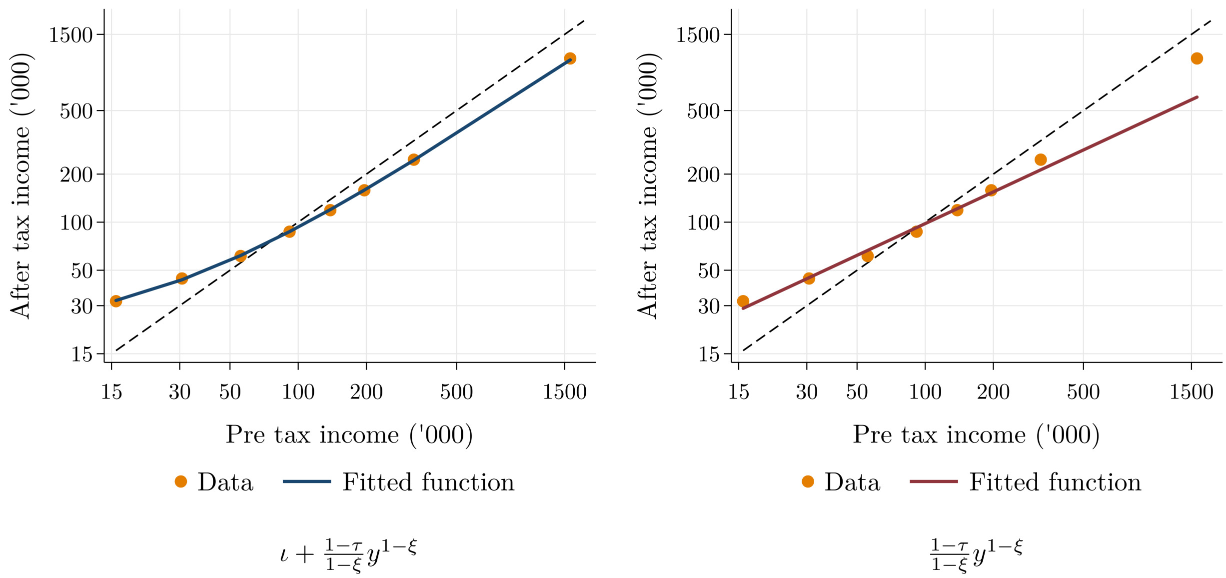
B. Parameter Values

Assigned			*Calibrated*		
θ	1	CRRA	β	0.975	discount factor
γ	2	inverse Frisch elasticity	ρ_e	0.982	autocorrelation e
α	1/3	capital elasticity	σ_e	0.200	std. dev. e shocks
δ	0.06	depreciation rate	p	2.2e-6	prob. enter super-star state
τ_a, ξ_a	0	wealth tax	q	0.990	prob. stay super-star state
τ_s	0.065	consumption tax	$\bar{e}$	504.3	ability super-star state, rel. to mean
τ_k	0.20	capital gains tax	ι	0.167	lump-sum transfer, rel. per-capita GDP
$\bar{B}$	1	government debt to GDP	τ	0.263	income tax schedule
			ξ	0.049	income tax schedule

comparison, we also estimated the tax function without lump-sum transfers. As the right panel shows, this function overstates the taxes paid by the richest households and understates the post-tax income of the poorest households, a point also made by Daruich and Fernández (2020).[6] Table 1 shows that our estimates of the parameters of the tax function are $\xi = 0.049$, $\iota = 0.167$ of the mean household income (or approximately \$14,000) and $\tau = 0.263$. These estimates imply that the marginal tax paid by the median household is equal to 26.7% and the marginal tax paid by a household at the 95^{th} percentile of the income distribution is equal to 34.5%. Our estimate of ξ is similar to that of Guner et al. (2014), but, owing to the presence of lump-sum transfers, lower than that of Heathcote et al. (2017). In our sensitivity analysis we show that our results are robust to alternative estimates of the tax function.

[6] We note that our estimates of the tax function imply a marginal income tax for the top 1% equal to 39%, similar to the 37% tax rate for the top income tax bracket in the US. In contrast, the estimated tax function without the lump-sum transfer implies a much higher marginal income tax for the top 1%, equal to 74.8%.

Figure 1: Tax Function



Notes: The figure plots the relationship between pre- and post-tax income under the assumption that post-tax income is equal to i) $\iota + \frac{1-\tau}{1-\xi} y^{1-\xi}$ in the left panel and ii) $\frac{1-\tau}{1-\xi} y^{1-\xi}$ in the right panel. The dashed line is the 45 degree line.

As is well known, matching the large degree of wealth and income inequality in an economy like ours requires departures from a Gaussian distribution of ability. We follow Castaneda et al. (2003) in assuming that ability can be either in a normal state or a super-star state. In the normal state it follows an AR(1) process with persistence ρ_e and volatility of Gaussian innovations σ_e. In the super-star state, ability is $\bar{e}$ times higher than the average. We assume that agents transit from the normal to the super-star state with a constant probability p and remain there with probability q. When agents return to the normal state, they draw a new ability from the ergodic distribution associated with the AR(1) process. In the sensitivity section we derive optimal policies for an alternative calibration with Gaussian ability shocks.

The discount factor and the parameters describing the ability process are jointly chosen to minimize the distance between moments in the model and in the data. We report the parameter values in Panel B of Table 1 and the moments in Panel A.

We target the average wealth to average income ratio, the wealth and income Gini coefficients, and the top 0.1% and 1% wealth and income shares in the 2013 SCF. The wealth to income ratio is 6.6 in both the data and the model. The model reproduces well the wealth and income Gini coefficients (0.85 vs. 0.86 and 0.64 vs. 0.65, respectively), the share of wealth held by the top 0.1% (0.22) and top 1% (0.35 vs. 0.34), as well as the share of income held by the top 0.1% and 1% (0.14 and 0.22). In the Appendix we show that the model reproduces additional untargeted moments of the wealth and income distribution.

Panel B of Table 1 reports the values of the calibrated parameters. The discount factor is $\beta = 0.975$. The process for ability in the normal state has persistence $\rho_e = 0.982$ and standard deviation $\sigma_e = 0.2$. Ability in the super-star state is $\bar{e} = 504$ times greater than the average. Households enter this state with probability $p = 2.2\text{e-}6$ and remain there with probability $q = 0.99$. This implies that 0.02% of households are in the super-star state at any point in time and that they earn 12% of all income. The autocorrelation and volatility of labor earnings in our model are similar to the estimates of Krueger et al. (2017) using the Panel Study of Income Dynamics: 0.98 (model) vs 0.94 (data) and 0.27 vs 0.29, respectively.

3.2 The Distribution of Household Welfare

We compute measures of household welfare using an approach similar to that of Benabou (2002) and Bakis et al. (2015). Specifically, we convert life-time utility into more interpretable units by calculating the constant consumption ω_i a household would need to receive every period, without working, in order to achieve the life-time utility V_i attained under the equilibrium allocations. That is, for a household i with life-time utility

$$V_i = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{c_{it}^{1-\theta}}{1-\theta} - \frac{h_{it}^{1+\gamma}}{1+\gamma} \right)$$

household welfare ω_i is the solution to

$$V_i = \sum_{t=0}^{\infty} \beta^t \frac{\omega_i^{1-\theta}}{1-\theta}.$$

This measure of welfare adjusts for risk, intertemporal substitution and mean-reversion in ability and, importantly, allows for interpersonal comparisons, a feature that is particularly useful in comparing the degree of redistribution that can be achieved by particular policies.

Table 2 shows that welfare inequality is substantially lower in our model compared to wealth and income inequality. For example, the share of wealth held by the top 1% is 34% and their share of pre-tax and post-tax income is 22% and 15%, respectively. Welfare is slightly less concentrated than post-tax income, owing to mean-reversion in labor ability, but is nevertheless unevenly distributed, with the top 1% receiving more than the bottom 25% combined (10% vs. 8%, respectively). We thus conclude that our economy is characterized by substantial inequality in welfare.

In our optimal policy exercise we need to take a stand on the objective of the planner. A parsimonious way of capturing alternative preferences for redistribution is to express the

Table 2: Dimensions of Inequality

	Welfare	Post-Tax Income	Pre-Tax Income	Wealth
Share top 1%	0.10	0.15	0.22	0.34
Share top 5%	0.20	0.29	0.40	0.59
Share top 10%	0.29	0.39	0.51	0.75
Share bottom 75%	0.52	0.40	0.27	0.06
Share bottom 50%	0.30	0.18	0.08	0.00
Share bottom 25%	0.08	0.06	0.02	0.00

social welfare function as

$$\text{social welfare function} = \left(\int \omega_i^{1-\Delta} \mathrm{d}i \right)^{\frac{1}{1-\Delta}},$$

where $\Delta \geq 0$ is a parameter that captures the desire to redistribute. We note that since the welfare objective increases in the utility of each agent, ω_i, the policy that maximizes this objective is Pareto optimal in the class of instruments we allow the planner to choose.

This specification captures a wide range of social welfare functions. For example, if $\Delta = 0$ the objective of the planner is to maximize average welfare. As pointed out by Benabou (2002), who refers to it as risk-adjusted GDP, this objective captures pure economic efficiency and disregards equity considerations in and of themselves. Alternatively, by setting $\Delta = \theta$, the households' coefficient of relative risk aversion, we recover the preferences of a utilitarian planner. More generally, a higher Δ implies a stronger preference for redistribution. In the limit, as $\Delta \to \infty$, this objective reduces to that of a Rawlsian planner.

4 Optimal Policy

We next study optimal tax reforms. To illustrate the role of each tax instrument, we first study partial reforms in which we vary one instrument at a time. We show that higher or steeper marginal income or wealth taxes can, in isolation, increase the welfare of poor households. We then characterize optimal policy, allowing the planner to use all instruments simultaneously. We show that flat income taxes deliver the bulk of the welfare gains that can be attained with non-linear income and wealth tax schedules.

4.1 Welfare Implications of Partial Reforms

We consider one-time, unanticipated and permanent increases in the parameters that determine the level of marginal income taxes, the slope of the marginal income tax schedule and the wealth tax.[7] For each experiment we adjust the lump-sum transfer ι_t to ensure that the government budget constraint is satisfied at all dates. As we vary each instrument, we keep all the other tax parameters and the amount of government debt unchanged.

Figure 2 reports the welfare implications of varying each instrument, taking into account the long-lived nature of the transitions and the general equilibrium implications of the reforms. We rank households according to their welfare ω_i in the initial steady state and report the utilitarian welfare change, as well as the welfare change for households in three groups of the distribution. Each column depicts the implications of changing a single tax instrument, τ, ξ and τ_a, respectively. The horizontal axes report the implied marginal tax rate at the 50^{th} and 95^{th} percentiles of the income distribution and the wealth tax, respectively.

We make two observations. First, higher wealth or income taxes increase the welfare of the poor at the expense of the rich. Intuitively, all these instruments allow the planner to collect additional revenue from the rich and finance transfers to the poor, either by increasing the lump-sum transfer or by reducing their marginal income tax. This increases utilitarian welfare as the insurance and redistributional value of transfers outweighs the distortions induced by higher taxes. For large enough tax increases the distortions dominate and welfare falls. Second, the instruments differ in how much redistribution they can achieve. For example, the planner can increase the welfare of the poorest third by as much as 30% by increasing the level of marginal income taxes, and by only 15% using wealth taxes.

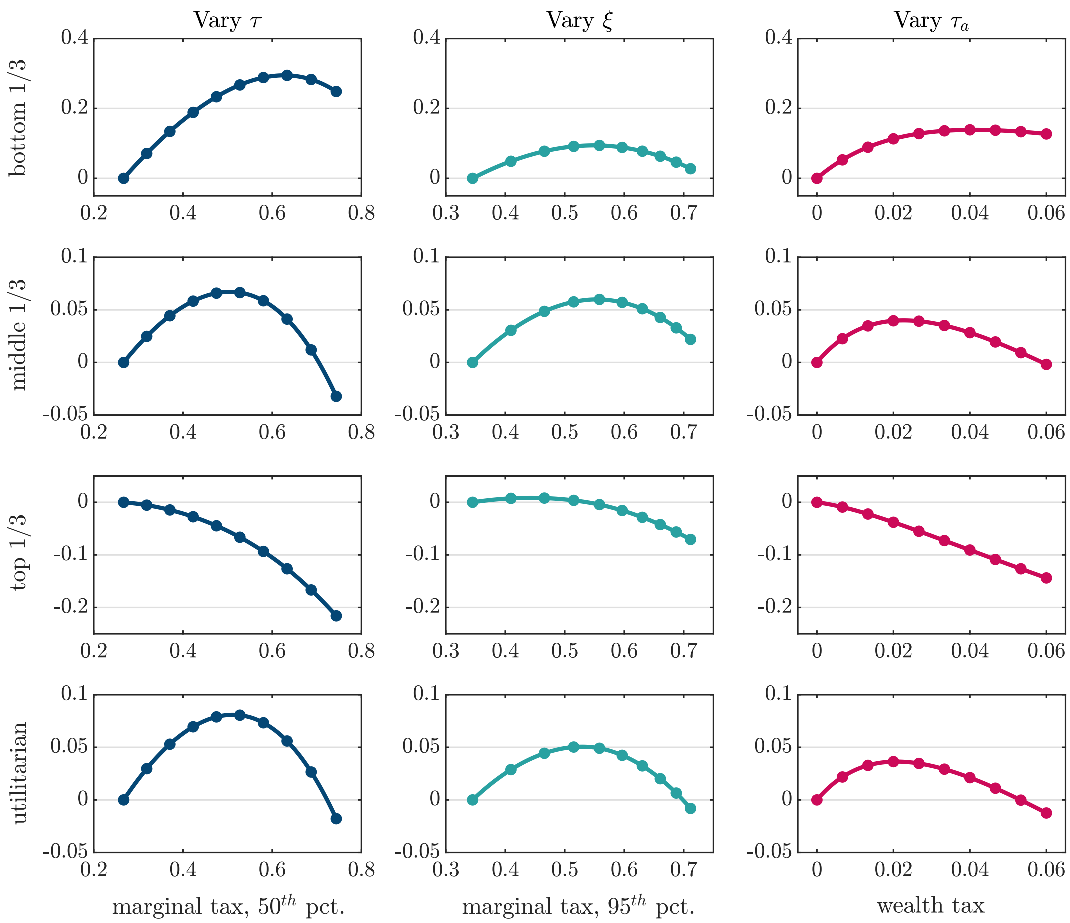
To understand the macroeconomic consequences of these partial reforms, Figure 3 illustrates the transition dynamics resulting from setting each of the three tax parameters at the values that maximize utilitarian welfare.[8] An increase in either the level or the slope of marginal income taxes worsens the labor wedge, while an increase in wealth taxes improves the labor wedge by stimulating labor supply by high-ability households due to wealth effects. All reforms worsen the savings wedge and depress the capital stock, labor and output. In turn, the interest rate increases and the wage rate falls.[9] These reforms have different implications for wealth inequality. While a steeper marginal income tax schedule reduces

[7]To conserve space, for this exercise we restrict attention to linear wealth taxes only and set $\xi_a = 0$.

[8]Specifically, we set $\tau = 0.5$, $\xi = 0.2$ and $\tau_a = 0.02$.

[9]Notice that the savings wedge increases to a level greater than one, implying that the interest rate increases above the rate of time preference, depressing capital relative to the modified golden rule.

Figure 2: Welfare Implications of Partial Reforms



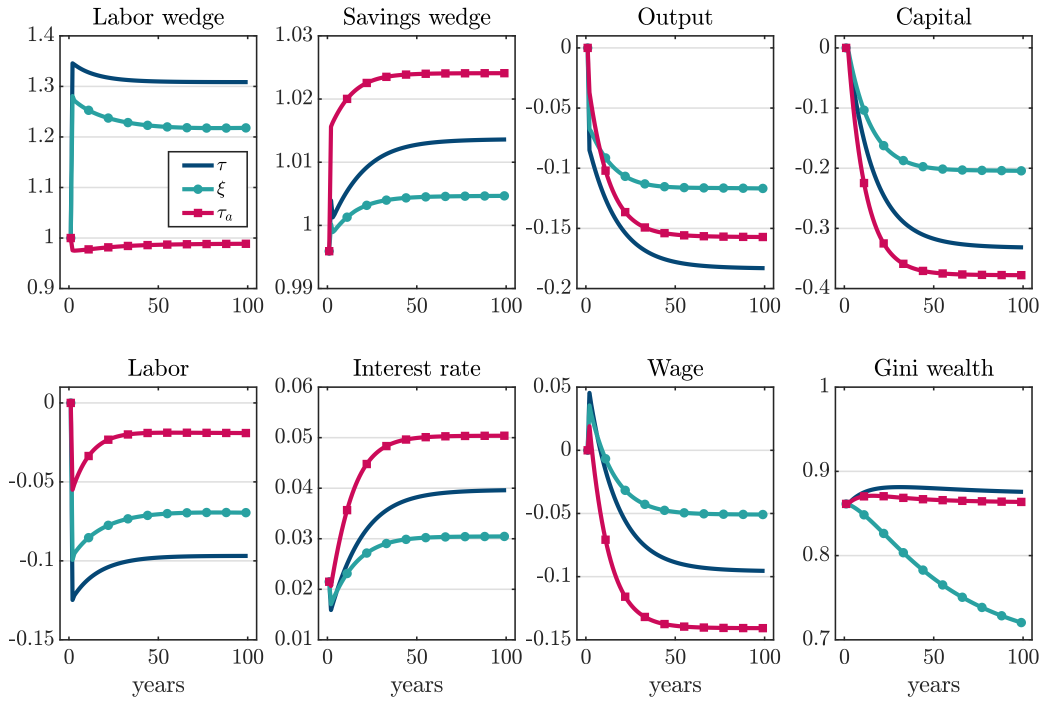
the wealth Gini coefficient, higher average income taxes and wealth taxes increase it because higher lump-sum transfers discourage labor supply and the precautionary savings by the poor. Since all of these policies increase the welfare of the poor, measures of wealth inequality provide a misleading account of the welfare consequences of tax reforms.

To summarize, increasing either wealth or income taxes can achieve redistribution. Previous work used this observation to argue in favor of a particular tax reform. However, most of this work allows the planner to use a single instrument at a time. In contrast to this research, the goal of our paper is to consider tax reforms that jointly change all of these instruments.

4.2 Optimal Tax Reforms

We next study optimal tax reforms. Specifically, we consider one-time, unanticipated, permanent changes in the parameters $\boldsymbol{\pi} = (\tau,\, \xi,\, \tau_a,\, \xi_a)$ that determine the income and wealth tax

Figure 3: Effect of Partial Reforms on Macro Aggregates



Notes: The labor wedge is expressed relative to its pre-reform value. The savings wedge, interest rate and Gini coefficient are expressed in levels. Changes in output, capital, labor and the wage are expressed relative to their pre-reform values.

schedules.[10] Our motivation for restricting the space of tax instruments comes from Heathcote and Tsujiyama (2021a), who show that the optimal income tax function in the HSV class approximates well the optimal Mirrlees policy in a static economy, and from Heathcote and Tsujiyama (2021b), who recommend using parametric families of tax functions when the distribution of ability is on a coarse grid.

Throughout, we maintain the assumption that government debt and the consumption tax are constant. The lump-sum transfer ι_t adjusts at every date to ensure that the government budget is balanced.[11] Since we consider once-and-for-all tax reforms (as do Domeij and Heathcote, 2004, Conesa et al., 2009, Guvenen et al., 2019), the planner balances the desire to tax wealth at date zero against the long-term distortions from capital income taxation.

[10]To ensure speed and stability, we modify the standard algorithm to compute transition dynamics in heterogeneous agents models along two dimensions. See the Appendix for details.

[11]We have experimented with allowing the planner to also choose debt optimally and found that raising government debt has similar implications to increasing the wealth tax. Both policies allow for a temporary increase in lump-sum transfers at the expense of a depressed capital stock. We found that the marginal gains from allowing the government to borrow more are small and therefore do not report these results for brevity.

We assume that the planner's objective is

$$\max_{\boldsymbol{\pi}} \left(\int \omega_i(\boldsymbol{\pi})^{1-\Delta} \mathrm{d}i \right)^{\frac{1}{1-\Delta}},$$

where $\omega_i(\boldsymbol{\pi})$ is the welfare of household i from a reform $\boldsymbol{\pi}$. We consider three values of Δ, corresponding to average, utilitarian and Rawlsian welfare.

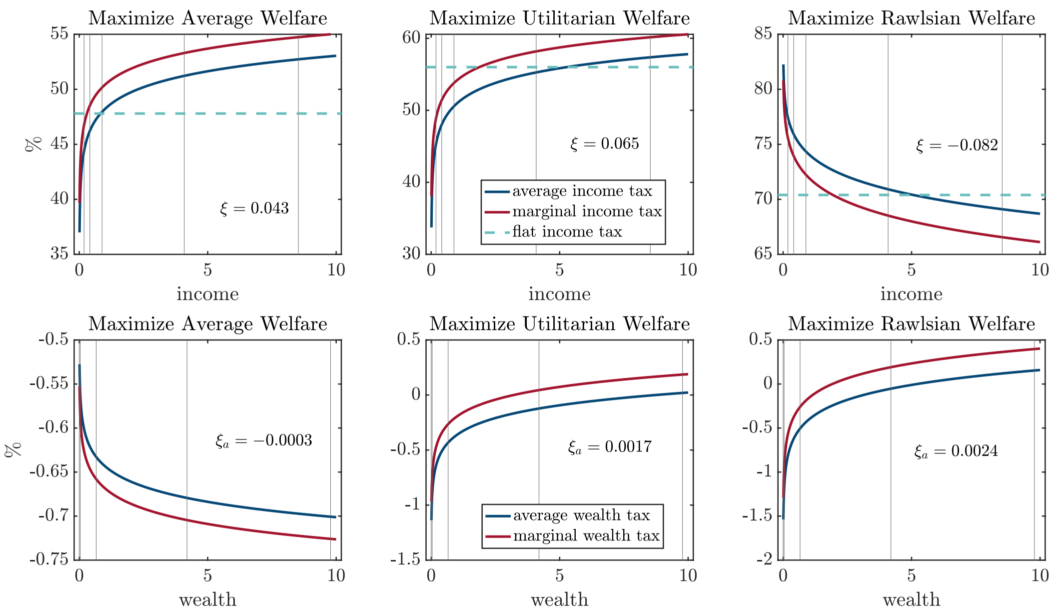
We proceed in two steps. First, we assume that the planner can only impose a flat tax on income and thus sets $\xi = \tau_a = \xi_a = 0$. We then compute the optimal value of τ. Second, we assume that the planner can use all tax instruments and compute the optimal values of τ, ξ, τ_a and ξ_a. Figure 4 displays the optimal tax schedules and Table 3 summarizes the welfare gains. For concreteness, we focus most of the discussion on a utilitarian objective.

When the utilitarian planner is restricted to using a flat tax on income, it sets it equal to 56%, which increases consumption-equivalent welfare by 7.4%. When the planner can use all instruments, it chooses positively sloped marginal income and wealth tax schedules. For example, the marginal income tax paid by the pre-reform median earner is equal to 50.5% and that paid by an earner at the 95^{th} percentile is 57.8%. The planner subsidizes wealth accumulation by the poor and taxes, albeit at a small rate, the wealth of the rich. Overall, the wealth tax does not change the steady state equilibrium interest rate relative to the economy with the optimally set flat income tax, suggesting that the savings wedge is unchanged.

With richer tax instruments utilitarian welfare increases by 8.5%, a modest gain relative to the 7.4% achievable with an optimally set flat income tax alone. Thus, the latter delivers 87% of the gains that can be achieved using all tax instruments. Interestingly, the households that benefit from a richer set of tax instruments are those in the middle of the distribution, not the poorest ones. Intuitively, with a richer set of tax instruments the planner can reduce the marginal income taxes paid by households in the middle to lower end of the distribution by cutting lump-sum transfers. Since low ability households prefer redistribution via lump-sum transfers, they lose from the lower marginal income taxes.

Our main conclusion is unchanged when considering alternative welfare objectives. As the first column of Figure 4 shows, a planner who seeks to maximize average welfare and therefore has no explicit concern for redistribution taxes income at a lower rate compared to a utilitarian planner, and subsidizes wealth accumulation. Nevertheless, such a planner also increases the welfare of the poor and of those in the middle class, as shown in Panel A of Table 3. Thus, even if the planner has no explicit concern for redistribution, it chooses a policy that greatly increases the welfare of the poor at the expense of the rich. We conclude

Figure 4: Optimal Income and Wealth Tax Schedules



Notes: The horizontal dashed line is the optimal flat income tax when the planner is restricted to only using a flat tax on income. The blue and red solid lines are the optimal average and marginal taxes when the set of instruments is unrestricted. Vertical bars represent the 25^{th}, 50^{th}, 75^{th}, 95^{th} and 99^{th} percentiles of the income and wealth distributions in the initial steady state. The x-axis reports income (wealth) relative to mean income (wealth) in the initial steady-state.

that measures of macroeconomic activity, such as output, which falls by 8% here, are poor indicators of how efficient a particular tax reform is. Our argument is thus distinct from that of Bowles and Gintis (1996) who argue that more redistributive policies may increase output. In contrast, the policies we consider here reduce output, but increase average welfare. Importantly, as in the case of a utilitarian objective, an optimally set flat income tax achieves most (81%) of the welfare gains attainable using more complex tax instruments.

The third column of Figure 4 shows that a Rawlsian planner sets much higher income taxes, subsidizes wealth accumulation by the poor and taxes the wealth of the rich. Moreover, such a planner chooses a negatively sloped marginal income tax schedule in order to increase lump-sum transfers, the main source of income for the poorest agent. Notice in Panel C of Table 3 that such a planner only increases the welfare of the bottom third of households by 29.4%, an increase only slightly larger than that achieved by a utilitarian planner. Importantly, we once again find that an optimally set flat income tax achieves 97% of the welfare gains attainable using all tax instruments.

Table 3 illustrates that the stronger is the planner's preference for redistribution, the smaller the additional welfare gains from deviating from a uniform flat income tax. Intuitively, a flat income tax generates nearly the same amount of (present value) tax revenue as richer tax instruments. This revenue is used to finance lump-sum transfers, which benefits the poor.

Table 3: Welfare Gains From Optimal Tax Policy

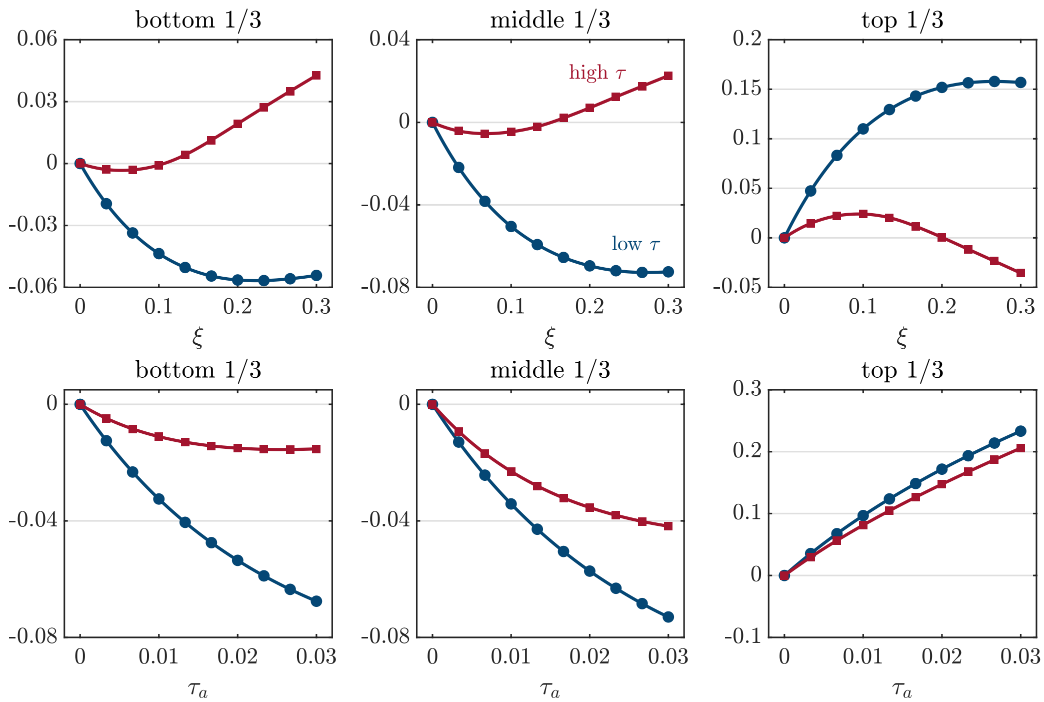
	Flat income tax	Non-linear income and wealth tax
A. Maximize Average Welfare		
average welfare gains	2.1	2.6
welfare gains, bottom 1/3	20.2	19.0
welfare gains, middle 1/3	4.8	4.7
welfare gains, top 1/3	−4.6	−3.1
B. Maximize Utilitarian Welfare		
utilitarian welfare gains	7.4	8.5
welfare gains, bottom 1/3	26.7	25.2
welfare gains, middle 1/3	5.7	7.0
welfare gains, top 1/3	−7.5	−4.8
C. Maximize Rawlsian Welfare		
Rawlsian welfare gains	65.0	66.9
welfare gains, bottom 1/3	29.4	29.4
welfare gains, middle 1/3	1.7	0.8
welfare gains, top 1/3	−16.3	−16.7

Notes: The first column restricts $\xi = \tau_a = \xi_a = 0$. The second column is the unrestricted optimum. The welfare gains are computed taking transitions into account. All numbers are expressed in percent.

4.3 Inspecting the Mechanism

We next explain why a richer set of tax instruments does not allow the planner to substantially increase welfare relative to an optimally chosen flat income tax. At a first glance, this result appears to contradict the conclusion of the partial reform exercises in Section 4.1, which showed that increasing wealth taxes or the slope of the marginal income tax schedule allows the planner to greatly increase social welfare. We show that there is, in fact, no contradiction. Rather, the result reflects that starting from an optimally chosen flat income tax, the marginal gains from additional instruments are small because steeper marginal income taxes or wealth taxes do not allow the government to raise much additional revenue. We illustrate this point by first considering the impact of increasing the slope of the marginal income tax schedule and then analyzing the effect of wealth taxes.

Figure 5: Differential Response of Tax Revenue to Varying ξ and τ_a



Notes: The tax bill is expressed relative to pre-reform GDP. We report the per capita, annuitized present value, discounted at the pre-reform equilibrium interest rate.

Steeper marginal income taxes The top panels of Figure 5 show the consequences of increasing the slope of the marginal income tax schedule ξ starting from two values of the parameter τ. The first value is $\tau = 0.26$, that under the status quo, and the second is $\tau = 0.56$, that optimally chosen by a utilitarian planner who uses flat income taxes only. For each of these values of τ, we calculate the transition dynamics from varying ξ, holding $\tau_a = 0$, and report the differential change in the present value of tax revenue obtained from a particular value of ξ relative to setting $\xi = 0$. The figure thus traces out the incremental effect of increasing the slope of the marginal tax schedule, conditional on a given value of τ.

The top panels of the figure show that when τ is high raising ξ does not increase the tax bill paid by the richest third of households.[12] Consequently, a steeper marginal income tax schedule does not reduce the tax bill of the poor. This is in contrast to what a steeper marginal income tax can achieve when τ is low. Intuitively, when τ is high, a steeper marginal income tax causes a large reduction in the income of the richest households, owing to both behavioral responses as well as equilibrium wage and output declines.[13] The large reduction

[12]Here, we consider all taxes paid by households, including those on consumption.
[13]We illustrate this with a numerical example in the Appendix.

in the income of the rich thus limits the amount of revenue the government can collect.

Wealth taxes The bottom panels of Figure 5 show the consequences of increasing the wealth tax τ_a starting from the same two values of τ as above, holding $\xi = 0$. Once again, the figure traces out the incremental effect of increasing the wealth tax, conditional on τ.

The overall tax bill paid by the richest one-third of households increases by a similar amount for the two values of τ. However, the higher revenues collected from the rich lead to much smaller declines in the taxes paid by the poor when τ is high, compared to when τ is low. The reason for this result is that when τ is high to begin with, a given wealth tax leads to a larger increase in the equilibrium interest rate. This increases the cost of servicing government debt, reducing the amount the planner can transfer to the poor.

4.4 Role of General Equilibrium Effects

We next study the role of general equilibrium effects in shaping optimal tax policy by considering a small open economy that takes the interest rate r as given. Since firms operate with a constant returns to scale technology, the equilibrium wage is also policy invariant.

Table 4 reports the welfare and aggregate outcomes. A utilitarian planner that can only use a flat income tax sets it equal to 60.1%, which generates welfare gains of 7.8%. When the planner can use the richer set of tax instruments, it chooses a positively sloped marginal income tax schedule, as in the general equilibrium setting, and a significantly steeper wealth tax schedule. For example, the marginal wealth tax paid by a household at the 50^{th} percentile of the pre-reform wealth distribution is -25.1%, while that paid by a household at the 95^{th} percentile is 34.9%. In effect, the planner immediately redistributes wealth from the rich to the poor, since it no longer faces the production consequences of depressing wealth accumulation. Notice that output falls by a much smaller amount across steady states in a small open economy, despite the fact that the wealth to income ratio falls much more. Also notice that even though the utilitarian welfare gains from using the richer set of tax instruments only increase by an additional 1.8 percentage points, the poor greatly benefit: the welfare gains of the bottom one-third increase from 29.4% to 40.5%. General equilibrium forces thus greatly constrain the planner's ability to use a wealth tax to achieve redistribution.

4.5 Role of Lump-Sum Transfer

We argue that lump-sum transfers are a potent means of redistribution, more so than upward-sloping marginal income taxes. To see this, we compute optimal policy restricting the planner to set the lump-sum transfer to zero. With flat income taxes and no lump-sum transfers utilitarian welfare falls by 28.1%. Unrestricted optimal policy calls for steep marginal income and wealth taxes ($\xi = 0.31$ and $\xi_a = 0.05$) and leads to modest welfare gains of 0.4% relative to the status quo. These results reflect that lump-sum transfers allow the government to redistribute to the poor who, owing to their low labor market ability, benefit much more from lump-sum transfers than from a reduction in marginal income taxes.

4.6 Separate Tax on Labor and Capital Income

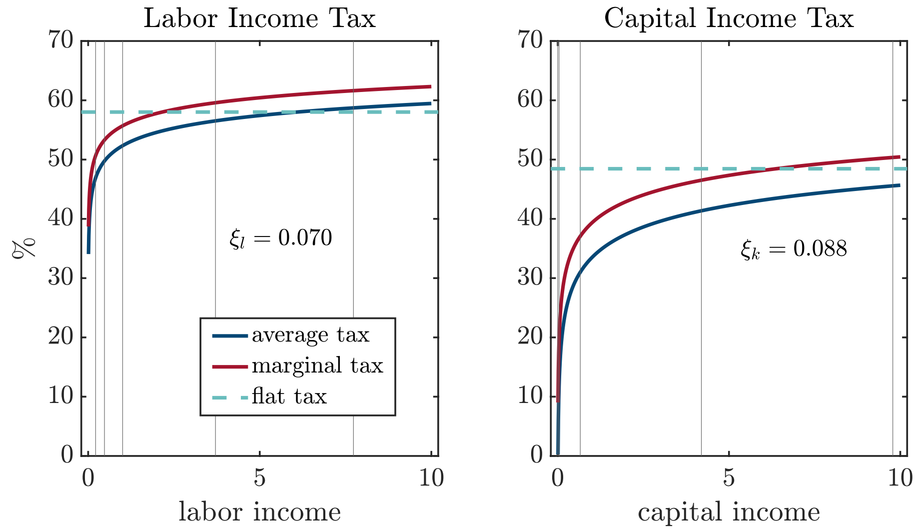
For comparison with related work, we also consider an alternative tax experiment in which the planner can separately tax labor and capital income. Specifically, we consider tax functions of the HSV form $T_l\left(W_t e_t h_t\right)$ and $T_k\left(r_{t-1} a_t\right)$ and optimize over the parameters τ_l, ξ_l, τ_k and ξ_k. As before, we proceed in two steps, first allowing for flat taxes only, and then optimizing over the richer set of parameters. For brevity, we report results for the utilitarian planner only. Figure 6 shows that when the planner is restricted to using flat income taxes only, it sets the labor income tax to 58% and the capital income tax to 48%.[14] Doing so generates welfare gains of 7.5%, similar to the 7.4% in the benchmark with a uniform tax on capital and labor income. With a richer set of instruments, the planner chooses upward sloping marginal income tax schedules ($\xi_l = 0.070$ and $\xi_k = 0.088$) and achieves welfare gains of 8.3%. These gains are slightly smaller than under the benchmark experiment, reflecting that the interest rate changes during the transition, which implies that capital income and wealth taxes are not exactly equivalent.

4.7 Capital Levy

In choosing time-invariant policies the planner balances the desire to tax the initial wealth against the distortions from depressing the capital stock. To see how the former shapes optimal policy, we consider an experiment where the planner levies a 100% wealth tax at date zero, which it redistributes lump-sum to all households. The column labeled 0 in Table

[14] Taxing capital in our economy is optimal for several reasons. First, as pointed out by Aiyagari (1995), our economy features capital over-accumulation relative to an economy with complete markets. Second, taxing capital prevents high ability households from accumulating wealth and leads them to supply more labor. Third, since the stock of wealth is inelastic in the short-run, taxing it generates government revenue.

Figure 6: Optimal Labor and Capital Income Tax Schedules



Notes: The horizontal dashed line is the optimal flat tax when the planner is restricted to only using flat taxes on labor and capital income. The blue and red solid lines are the optimal average and marginal taxes when the set of instruments is unrestricted. Vertical bars represent the 25^{th}, 50^{th}, 75^{th}, 95^{th} and 99^{th} percentiles of the labor and capital income distributions in the initial steady state. The x-axis reports labor (capital) income relative to mean labor (capital) income in the initial steady-state.

4 reports the effect of such a capital levy. The utilitarian welfare gains are nearly 15%, with the poorest one-third of households experiencing a welfare gain of 38.8%. We next allow the planner to also change the tax schedules, in addition to imposing a capital levy. We find that the planner no longer distorts capital accumulation, as evidenced by new steady state equilibrium interest rate of 2.2%, very similar to the 2.1% in the initial steady state. Once again, the marginal gains from deviating from uniform labor and capital taxation are small: the optimally set flat income tax achieves 86% of the maximum attainable welfare gains.

4.8 Higher Taxes on the Top 1%

We next assume piecewise linear income and wealth tax functions. Specifically, we assume that income (wealth) below a level $\bar{y}$ ($\bar{a}$) is taxed a flat tax rate τ_y^1 (τ_a^1) and every dollar of income (wealth) above this cutoff is taxed a flat tax rate τ_y^2 (τ_a^2). We set $\bar{y}$ and $\bar{a}$ equal to the 99^{th} percentile of the post-reform income and wealth distributions, thus allowing the planner to tax the top 1% at a higher rate than everyone else.

We find that it is optimal to tax the income of the bottom 99% at 52.5% and the income of the top 1% at 80.5%. The bottom 99% receive a 0.5% wealth subsidy and the top 1% pay a 3.4% wealth tax. The utilitarian welfare gains from this reform are 9.1%, reflecting a 25.9%, 6.8% and -3.4% increase in the welfare of the bottom, middle and top third of the

Table 4: Partial Equilibrium and Capital Levy, Utilitarian Planner

	Baseline		Partial Equilibrium		Capital Levy		
	I	II	I	II	0	I	II
A. Welfare Change							
utilitarian welfare gains	7.4	8.5	7.8	9.6	14.9	19.2	22.3
welfare gains, bottom 1/3	26.7	25.2	29.4	40.5	38.8	59.9	62.5
welfare gains, middle 1/3	5.7	7.0	6.8	13.8	20.5	25.3	29.3
welfare gains, top 1/3	−7.5	−4.8	−9.3	−17.8	−9.3	−15.4	−12.9
B. Aggregate Implications							
change in wage	−9.7	−9.8	0	0	0	−8.1	−0.5
interest rate	4.0	4.0	2.1	2.1	2.1	3.7	2.2
change in output	−19.1	−19.7	−9.3	−11.0	0	−16.2	−11.7
wealth to income	5.4	5.4	1.9	1.5	6.6	5.6	6.6

Notes: The columns labeled I refer to the optimal flat income tax reform. The columns labeled II refer to the optimal non-linear income and wealth tax reforms. The column labeled 0 reports the result of a one-time 100% wealth tax at date 0 that is rebated lump-sum to all households, with no additional changes in the parameters of the income and wealth tax schedules.

welfare distribution. The overall welfare gains are only slightly larger than the 8.5% under the HSV tax function, reflecting smaller losses at the top. Since these higher gains arise due the coarseness of the grid for the ability distribution, we focused our analysis on the more parsimonious HSV tax function, as recommended by Heathcote and Tsujiyama (2021b).

4.9 Sensitivity Analysis

It is well known that optimal tax policy is critically shaped by household preferences and the distribution of household ability. We next show that even though the size of optimal taxes indeed depends on these details of the model, our conclusion that a flat income tax achieves a large fraction of the gains attainable with non-linear wealth and income taxes is robust. We also show robustness to using alternative estimates of the tax schedule in the initial steady-state. For each experiment, we recalibrate the model to reproduce the same statistics as in the benchmark and revisit the optimal tax reforms for a utilitarian planner. Here, we briefly summarize the main findings and report more detailed results in the Appendix.

Preferences and Ability Distribution We consider three perturbations of the model.

First, we reduce the intertemporal elasticity of substitution to 0.5 by setting $\theta = 2$. The planner now desires more redistribution and the optimal flat income tax is 71.6%, larger than the 56% in our benchmark model. The unrestricted policy calls for positively-sloped marginal income and wealth tax schedules and imposes very high wealth taxes at the top: the marginal wealth tax paid by households at the 95^{th} percentile of the pre-reform wealth distribution is 21%. Overall, with an optimally set flat income tax the planner can achieve 72% of the welfare gains attainable with more instruments.

Second, we double the Frisch elasticity of labor supply by setting $\gamma = 1$. The optimal flat income tax is now smaller than in our benchmark and is equal to 49%. Once again, the planner prefers positively-sloped marginal income and wealth taxes. In contrast to the baseline parameterization, the planner taxes wealth at the top: the marginal wealth tax paid by households at the 95^{th} percentile of the pre-reform wealth distribution is 1%. Since labor is more elastic, the planner effectively taxes capital income at a higher rate than labor income. However, the gains from this flexibility are relatively small: a flat income tax achieves a substantial share (73%) of the welfare gains attainable with a richer set of tax instruments.

Lastly, we assume a Gaussian distribution of ability by eliminating the super-star state. As is well known, such an economy cannot reproduce the wealth dispersion in the data. Marginal income taxes now decrease with income (Saez, 2001, Mankiw et al., 2009). We find that the marginal income tax falls from 73.9% at the median to 57.9% at the 95^{th} percentile. In addition, the planner also finds it optimal to tax wealth at a decreasing rate: the marginal wealth tax falls from 0.5% at the median to 0.1% at the 95^{th} percentile. Once again, the incremental welfare gains of departing from a flat income tax are small: the flat income tax achieves 96% of the maximum attainable welfare gains.

Alternative Estimates of the Tax Function We consider two alternative of the tax function in the status quo. First, we re-estimate the tax function imposing the restriction $\xi = 0.181$, the estimate of Heathcote et al. (2017). Second, we halve the lump-sum transfer relative to the value estimated in our benchmark parameterization. Our results are robust to these alternatives. Specifically, under the HSV estimate of ξ we find that the utilitarian welfare gains from a flat uniform income tax are equal to 4.7%, only slightly smaller than the 6.3% attainable by the unrestricted wealth and income tax schedules. When we halve the lump-sum transfer ι, these gains are 12.3% and 13.4%, respectively.

5 Conclusions

Motivated by the large increase in wealth and income inequality in the United States, we characterize the optimal shape of income and wealth tax schedules in a dynamic general equilibrium model that reproduces the observed wealth and income inequality, as well as the extent of redistribution currently in place.

We find that taxing capital and labor income at a uniform flat rate is nearly optimal, in that the incremental gains from non-linear income and wealth tax schedules are relatively small. While steeper marginal income taxes or wealth taxes can greatly increase government revenue when average marginal income taxes are low to begin with, such policies lead to little or no increase in government revenue when marginal income taxes are already high due to a larger drop in the labor supply of high-ability households and the general equilibrium implications of depressed capital accumulation. Interestingly, the incremental gains from a richer set of tax instruments are smaller, the stronger the planner's concern for redistribution.

References

Acikgoz, Omer, Marcus Hagedorn, Hans Holter, and Yikai Wang, "The Optimum Quantity of Capital and Debt," Working Paper 2021.

Aiyagari, S. Rao, "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting," *Journal of Political Economy*, 1995, *103* (6), 1158–1175.

Bakis, Ozan, Baris Kaymak, and Markus Poschke, "Transitional Dynamics and the Optimal Progressivity of Income Redistribution," *Review of Economic Dynamics*, 2015, *18* (3), 679 – 693.

Benabou, Roland, "Unequal Societies: Income Distribution and the Social Contract," *American Economic Review*, March 2000, *90* (1), 96–129.

_ , "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?," *Econometrica*, 2002, *70* (2), 481–517.

Berger, David W, Luigi Bocola, and Alessandro Dovis, "Imperfect Risk-Sharing and the Business Cycle," Working Paper, National Bureau of Economic Research 2019.

Bhandari, Anmol and Ellen R McGrattan, "Sweat Equity in U.S. Private Business*," *The Quarterly Journal of Economics*, 12 2020, *136* (2), 727–781.

Bowles, Samuel and Herbert Gintis, "Efficient Redistribution: New Rules for Markets, States, and Communities," *Politics & Society*, 1996, *24* (4), 307–342.

Brüggemann, Bettina, "Higher Taxes at the Top: The Role of Entrepreneurs," *American Economic Review: Macroeconomics*, 2021, *13* (3), 1–36.

Castaneda, Ana, Javier Diaz-Gimenez, and Jose-Victor Rios-Rull, "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, 2003, *111* (4), 818–857.

Chari, V.V., Juan Pablo Nicolini, and Pedro Teles, "Optimal capital taxation revisited," *Journal of Monetary Economics*, 2020, *116*, 147–165.

Conesa, Juan Carlos and Dirk Krueger, "On the Optimal Progressivity of the Income Tax Code," *Journal of Monetary Economics*, 2006, *53*, 1425–1450.

_ , Sagiri Kitao, and Dirk Krueger, "Taxing Capital? Not a Bad Idea after All!," *American Economic Review*, March 2009, *99* (1), 25–48.

Daruich, Diego and Raquel Fernández, "Universal Basic Income: A Dynamic Assessment," Working Paper 27351, National Bureau of Economic Research June 2020.

Domeij, David and Jonathan Heathcote, "On the Distributional Effects of Reducing Capital Taxes," *International Economic Review*, 2004, *45* (2), 523–554.

Dyrda, Sebastian and Marcelo Pedroni, "Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Shocks," 2021. working paper.

Farhi, Emmanuel and Ivan Werning, "Capital Taxation: Quantitative Explorations of the Inverse Euler Equation," *Journal of Political Economy*, 2012, *120* (3), 398–445.

_ and _ , "Insurance and Taxation over the Life Cycle," *Review of Economic Studies*, 2013, *810* (2), 596–635.

Ferriere, Axelle, Philipp Grübener, Gaston Navarro, and Oliko Vardishvili, "Tax Progressivity, Credits, and Transfers in the United States," 2020. working paper.

Golosov, Mikhail, Maxim Troshkin, and Aleh Tsyvinski, "Redistribution and Social Insurance," *American Economic Review*, February 2016, *106* (2), 359–86.

Greulich, Katharina, Sarolta Laczo, and Albert Marcet, "Pareto-Improving Optimal Capital and Labor Taxes," Working Paper 951 2022.

Guner, Nezih, Remzi Kaygusuz, and Gustavo Ventura, "Income Taxation of U.S. Households: Facts and Parametric Estimates," *Review of Economic Dynamics*, 2014, *17* (4), 559–581.

Guvenen, Fatih, Gueorgui Kambourov, Burhan Kuruscu, Sergio Ocampo, and Daphne Chen, "Use It or Lose It: Efficiency Gains from Wealth Taxation," Working Paper 26284, National Bureau of Economic Research September 2019.

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

__ **and** __ , "Practical Optimal Income Taxation," 2021. working paper.

__ **, Kjetil Storesletten, and Giovanni L. Violante**, "Optimal Tax Progressivity: An Analytical Framework," *The Quarterly Journal of Economics*, 2017, *132* (4), 1693–1754.

Imrohoroglu, Ayse, Cagri S. Kumru, and Arm Nakornthab, "Revisiting Tax on Top Income," Technical Report 2018.

Kaymak, Barış and Markus Poschke, "The Macroeconomic and Distributional Effects of Progressive Wealth Taxes," 2019. working paper.

Kindermann, Fabian and Dirk Krueger, "High Marginal Tax Rates on the Top 1%? Lessons from a Life Cycle Model with Idiosyncratic Income Risk," *American Economic Journal: Macroeconomics, Forthcoming*, 2021.

Krueger, Dirk, Kurt Mitman, and Fabrizio Perri, "Macroeconomics and Household Heterogeneity," in John Taylor and Harald Uhlig, eds., *Handbook of Macroeconomics*, Vol. 2B, 2017.

Mankiw, N. Gregory, Matthew Weinzierl, and Danny Yagan, "Optimal Taxation in Theory and Practice," *Journal of Economic Perspectives*, December 2009, *23* (4), 147–74.

Rendahl, Pontus, "Inequality Constraints and Euler Equation-Based Solution Methods," *The Economic Journal*, 06 2014, *125* (585), 1110–1135.

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

Stantcheva, Stefanie, "Optimal Taxation and Human Capital Policies over the Life Cycle," *Journal of Political Economy*, 2017, *125* (6), 1931–1990.

Zhang, Junzi, Brendan O'Donoghue, and Stephen Boyd, "Globally Convergent Type-I Anderson Acceleration for Non-Smooth Fixed-Point Iterations," 2018.

Appendix

A Equilibrium Definition

A competitive equilibrium consists of: *(i)* aggregate prices W_t, R_t, r_t, *(ii)* consumption, saving and labor supply decisions of households $c_t(a,e)$, $a_{t+1}(a,e)$, $h_t(a,e)$, *(iii)* employment, capital and output choices of firms L_t, K_t, Y_t, and *(iv)* measures of households over their idiosyncratic states $n_t(a,e)$, such that

1. Given prices, households and firms solve their optimization problems.
2. The measure $n_t(a,e)$ evolves according to an equilibrium mapping dictated by the households' optimal choices and the stochastic process for labor market ability.
3. The budget constraint of the government is satisfied period by period.
4. Markets clear. The labor market clearing condition is

$$L_t = \int e h_t(a,e)\, dn_t(a,e).$$

The asset market clearing condition is

$$K_{t+1} + B_{t+1} = \int a_{t+1}(a,e)\, dn_t(a,e).$$

The goods market clears by Walras' Law.

B Computational Appendix

We solve the model by discretizing the distribution of labor market ability e using a Rouwenhorst method with 11 points. We approximate agents' value and consumption functions in the wealth space with 501 linear splines. We allow for a wide enough upper bound that ensures no extrapolation. Since finding the optimal policy requires computing the transition dynamics for many possible combinations of tax parameters, we require a fast, efficient and robust solution method. We therefore modify the standard algorithm used to compute transition dynamics in heterogeneous agents models along two dimensions.

First, we note that Euler equation iteration has a slow convergence rate. We therefore follow the approach proposed by Rendahl (2014) and apply Howard's improvement algorithm to iterate on the marginal valuation of wealth. Formally, we solve for the policy function

$a_{t+1}(a,e)$ by solving the following Euler equation for every grid point in the wealth and productivity space using Broyden's method

$$\frac{1}{1+\tau_s}c_t(a,e)^{-\theta} = \beta\mathbb{E}\frac{\partial V_{t+1}(a',e')}{\partial a'} + \varepsilon_t(a,e)$$

where $\varepsilon_t(a,e)$ is the multiplier on the $a_{t+1} \geq 0$ constraint. Even though the envelope condition implies that

$$\frac{\partial V_t(a,e)}{\partial a} = \frac{1}{1+\tau_s}c_t(a,e)^{-\theta}\left(1-\tilde{\tau}_{a,t}(a',e') + (1-\tilde{\tau}_t(a',e'))\, r_{t-1}\right)$$

at the optimum, the time iteration algorithm converges faster if we do not impose the envelope condition during the iterative process (when the envelope condition does not hold). We therefore iterate on

$$\begin{aligned}\frac{\partial V_t(a,e)}{\partial a} \quad &= \quad c_t(a,e)^{-\theta} c_{a,t}(a,e) + \left(\beta\mathbb{E}\frac{\partial V_{t+1}(a',e')}{\partial a'} + \varepsilon_t(a,e)\right)\times \\ &\quad \left(1-\tilde{\tau}_{a,t}(a',e') + (1-\tilde{\tau}_t(a',e'))\, r_t - (1+\tau_s)\, c_{a,t}(a,e)\right)\end{aligned}$$

using the Howard improvement algorithm.

Second, computing the transition dynamics requires iterating over a sequence of interest rates and lump-sum transfers during the transitions. This system is typically solved using a fixed point method, which is notoriously slow and unreliable. We achieved a significant speed gain and stability by employing the globally convergent Type-I Anderson acceleration algorithm in Zhang et al. (2018).

To ensure that we find a global maximum to the planner's problem, we first evaluate the social welfare function on a coarse grid for the values of tax parameters and used the best point in this grid to initialize our search. We also tested our results using several global optimization methods, including the genetic algorithm and particle swarm optimization method.

C Model Fit

In our calibration we only targeted the Gini coefficients of the wealth and income distributions and the shares of wealth and income held by the top 0.1% and 1%. Panels A and B of Table C.1 show that the model reproduces these distributions more broadly. For example, the wealthiest 10% of households hold 75% of wealth in both the data and the model. The richest 10% of households earn 51% of income in both the data and the model. The model also reproduces well the bottom of the wealth and income distribution. For example, households in the bottom half of the wealth distribution hold nearly no wealth in both the data and the model, while those in the bottom half of the income distribution earn 10% of income in the data and 8% the model.

Table C.1: Non-targeted moments

	Data	Model		Data	Model
A. Wealth Distribution			**B. Income Distribution**		
Share top 5%	0.63	0.59	Share top 5%	0.39	0.40
Share top 10%	0.75	0.75	Share top 10%	0.51	0.51
Share bottom 75%	0.09	0.06	Share bottom 75%	0.29	0.27
Share bottom 50%	0.01	0.00	Share bottom 50%	0.10	0.08
Share bottom 25%	-0.01	0.00	Share bottom 25%	0.02	0.02

Notes: The data moments are based on the 2013 SCF survey.

D Inspecting the Mechanism

D.1 Steeper Marginal Income Taxes

To understand why the government is unable to collect additional revenue from the rich by increasing the slope of the marginal income tax schedule when τ is high, let $T_t(\tau, \xi)$ denote the income taxes collected from the richest one-third of households in period t when the planner sets the parameters of the income tax schedule equal to τ and ξ. Similarly, let $I_t(\tau, \xi)$ denote the pre-tax income of these households under the same reform. Finally, let

$$\upsilon_t(\tau, \xi) \equiv \frac{T_t(\tau, \xi)}{I_t(\tau, \xi)}$$

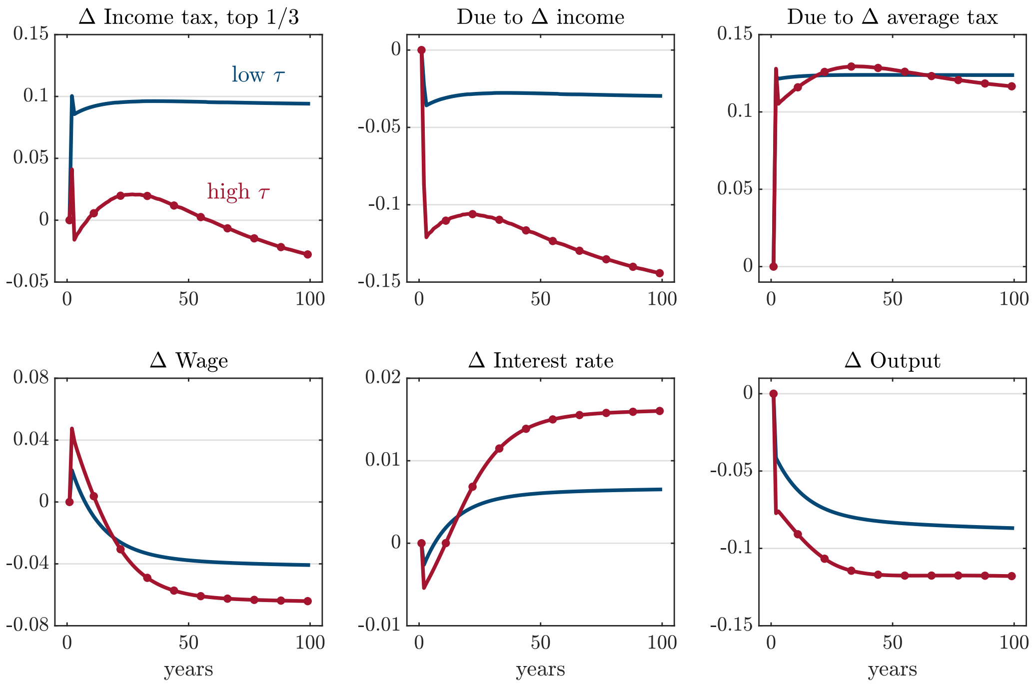
denote the implied average tax rate for this group. By definition, the differential change in the taxes paid by the richest third of households in an environment with $\xi > 0$ relative to an environment with $\xi = 0$ is equal to

$$T_t(\tau, \xi) - T_t(\tau, 0) = \upsilon_t(\tau, \xi)\left[I_t(\tau, \xi) - I_t(\tau, 0)\right] + \left[\upsilon_t(\tau, \xi) - \upsilon_t(\tau, 0)\right] I_t(\tau, 0). \qquad \text{(A1)}$$

Mechanically, the differential response of tax revenue is the sum of two components. The first one captures the behavioral response, the decline in income $I_t(\tau, \xi) - I_t(\tau, 0)$, brought about by steeper marginal income taxes. The second captures the change in average income taxes, $\upsilon_t(\tau, \xi) - \upsilon_t(\tau, 0)$.

We next zoom in on two reforms that increase the slope of the marginal income schedule starting from the two values of τ considered above. When $\tau = 0.56$, we set $\xi = 0.2$, which increases the average income tax paid by the richest one-third by 6.9% across steady states. Since the value of ξ on its own is not interpretable, to ensure comparability across the two reforms, we set $\xi = 0.079$ when $\tau = 0.26$. This leads to an identical steady-state increase in the average income tax paid by the richest one-third of 6.9%.

Figure A1: Transition Dynamics after Increasing ξ



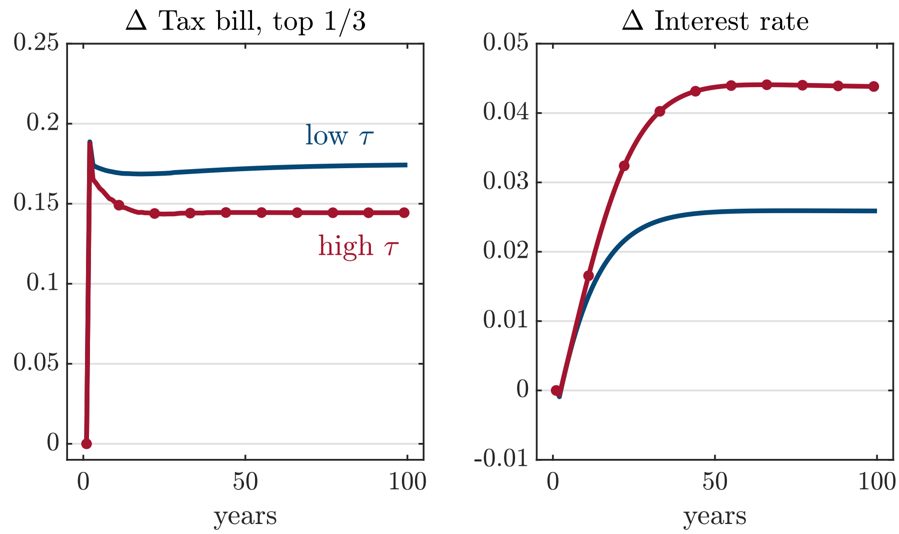
The top panel of Figure A1 traces out the differential response of income taxes collected from the richest one-third of households, as well as the two components in equation (A1). By construction, the second component is similar for high and low values of τ, which allows us to isolate the behavioral response. Notice that the income of the richest one-third of households falls by a much larger amount when increasing ξ starting from $\tau = 0.56$. This drop in income entirely offsets the gains from higher average income taxes, so the tax bill of the richest households changes little. The bottom panel of the figure shows that the equilibrium wage and output falls by a larger amount when increasing ξ starting from a high τ. Thus, increasing the slope of the marginal income tax schedule when τ is high generates little to no additional government revenue, but leads to large declines in output and wages and therefore reduces welfare for all households.

D.2 Wealth Taxes

Figure A2 shows that a given wealth tax leads to a larger increase in interest rates when income taxes are high. To understand why that is the case, note that the the steady-state equilibrium interest rate satisfies

$$\beta \int \mathbb{E}_t \left(\frac{\hat{c}_{it+1}}{\hat{c}_{it}} \right)^{-\theta} \big(1 - \tau^a + (1-\tau) r_t + \chi_{it}\big) \mathrm{d}i = 1, \tag{A2}$$

Figure A2: Transition Dynamics after Increasing τ_a



where we implicitly impose that wealth and income taxes are linear. Recall that χ_{it} is the multiplier on the no-borrowing constraint and $\hat{c}_{it}$ is the consumption share of household i. An increase in the wealth tax τ^a reduces the return to wealth and requires a countervailing increase in the after-tax interest rate $(1-\tau)r_t$ to ensure equation (A2) holds. Thus, the higher τ is, the larger the required increase in the interest rate r_t needed to clear the asset market. A given wealth tax is therefore more distortionary and yields much smaller welfare gains in an environment with high income taxes. We once again emphasize that in our economy income taxes apply to both capital and labor income, so the result that the gains from wealth taxation are small implies that the gains from taxing capital and labor income at different rates are low, not that there are no gains from capital income taxation.

E Sensitivity Analysis

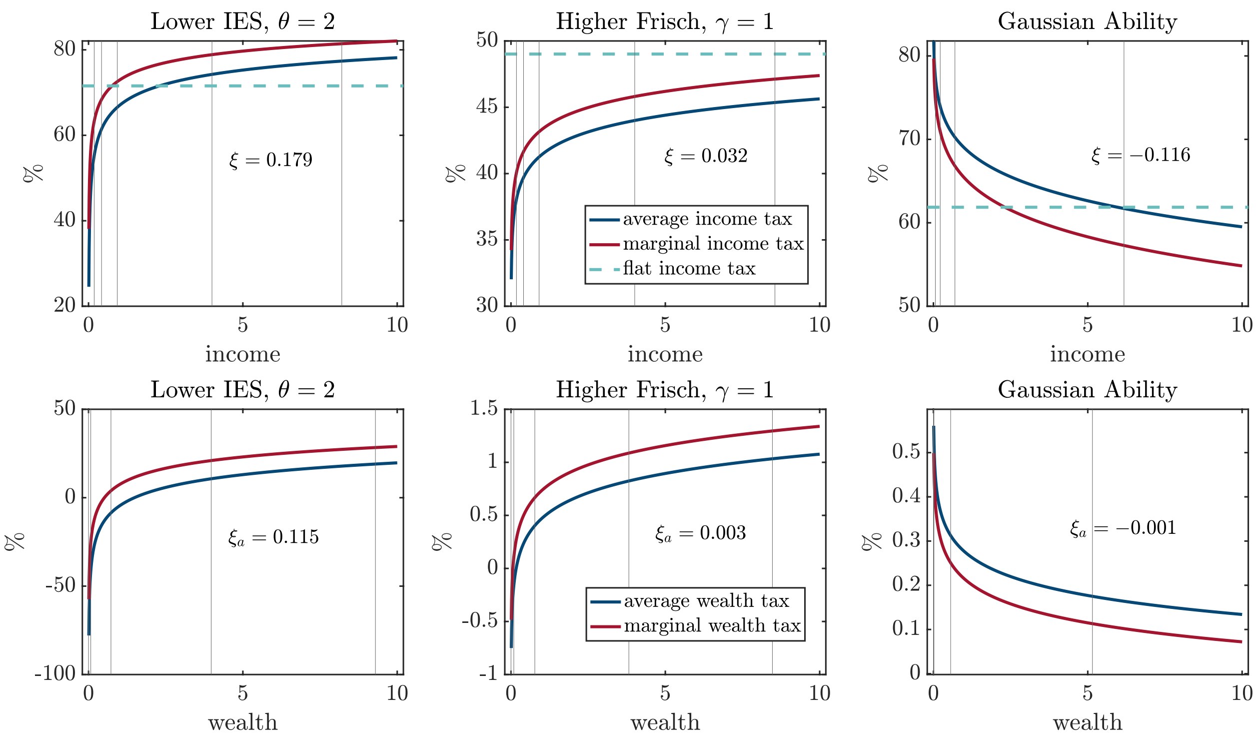
Parameterization. Table E.1 reports the parameter values under the different perturbations of the model in Panel A and the implied moments in Panel B. All these economies, with the exception of the economy with normally distributed labor ability, reproduce the wealth to income ratio, the Gini coefficients of wealth and income inequality, the top 0.1% and top 1% wealth and income shares. In contrast, as is well known, the economy in which labor ability is normally distributed cannot reproduce the top wealth and income shares and the fact that wealth is more concentrated than income.

Figure A1 plots the optimal tax schedules in the models with a higher elasticity of intertemporal substitution, a higher Frisch elasticity of labor supply and Gaussian ability,

Table E.1: Sensitivity: Parameterization

	Data	Lower IES $\theta = 2$	Higher Frisch $\gamma = 1$	Gaussian ability	$\xi = 0.18$	Halve ι
A. Parameter Values						
β, discount factor		0.958	0.970	0.968	0.978	0.974
ρ_e, autocorrelation e		0.981	0.963	0.979	0.984	0.983
σ_e, std. dev. e shocks		0.254	0.271	0.313	0.183	0.194
p, prob. enter super-star state		3.4e-6	3.0e-6	–	1.3e-5	2.7e-6
q, prob. stay super-star state		0.990	0.970	–	0.935	0.988
$\bar{e}$, ability super-star state, rel. mean		1147	790.1	–	532.9	507.4
B. Moments						
Wealth to income ratio	6.6	6.6	6.6	6.6	6.6	6.6
Gini wealth	0.85	0.85	0.84	0.87	0.82	0.85
Gini income	0.64	0.65	0.66	0.75	0.63	0.63
Wealth share top 0.1%	0.22	0.23	0.23	0.05	0.20	0.23
Wealth share top 1%	0.35	0.34	0.34	0.24	0.33	0.34
Income share top 0.1%	0.14	0.14	0.13	0.05	0.16	0.14
Income share top 1%	0.22	0.22	0.22	0.21	0.24	0.22

Figure A1: Optimal Tax Schedules, Alternative Preferences and Ability Distribution



Notes: The horizontal dashed line is the optimal flat income tax when the planner is restricted to only using a flat tax on income. The blue and red solid lines are the optimal average and marginal taxes when the set of instruments is unrestricted. Vertical bars represent the 25^{th}, 50^{th}, 75^{th}, 95^{th} and 99^{th} percentiles of the income and wealth distributions in the initial steady state. The x-axis reports income (wealth) relative to mean income (wealth) in the initial steady-state.

respectively.

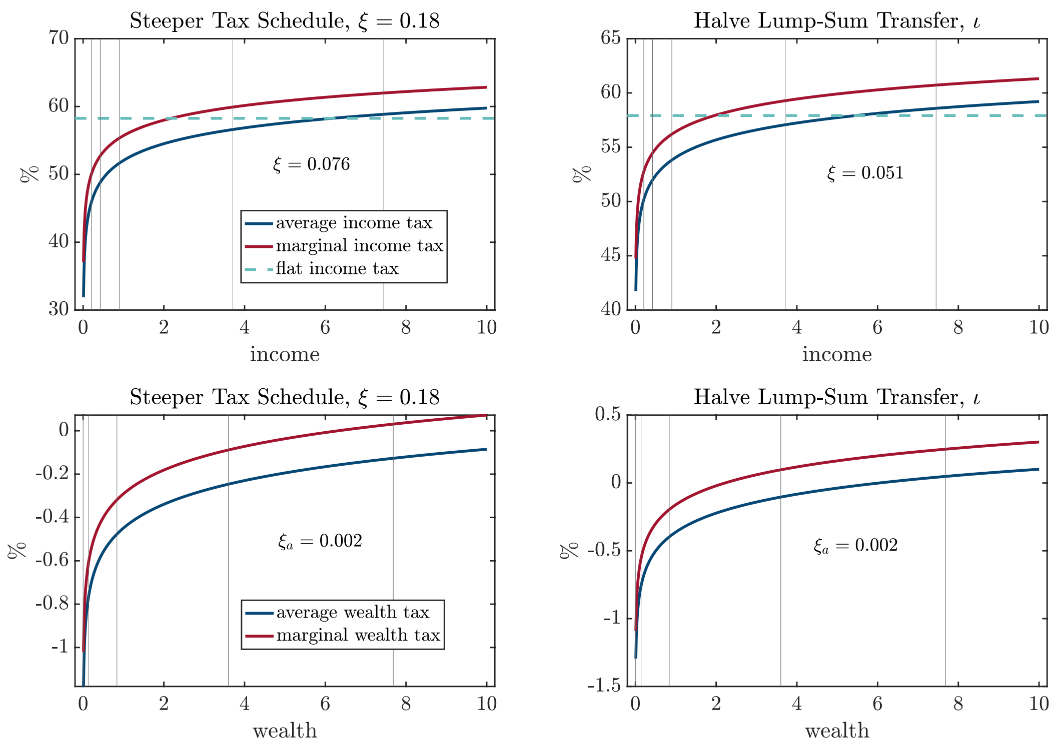
Figure A2 plots the optimal tax schedules in the models with alternative parameterizations of the HSV tax function.

Table E.2: Sensitivity: Welfare Gains, Utilitarian Planner

	Flat income tax	Non-linear income and wealth tax
A. Lower IES, $\theta = 2$		
bottom 1/3	57.5	75.1
middle 1/3	17.7	32.2
top 1/3	−4.3	−3.6
utilitarian	28.7	40.1
B. Higher Frisch, $\gamma = 1$		
bottom 1/3	14.3	15.1
middle 1/3	1.9	4.1
top 1/3	−4.4	−3.2
utilitarian	3.7	5.1
C. Gaussian Ability		
bottom 1/3	51.8	57.1
middle 1/3	19.4	20.2
top 1/3	−6.1	−7.9
utilitarian	19.4	20.2
D. Steeper Initial Income Tax Schedule, $\xi = 0.18$		
bottom 1/3	26.2	24.6
middle 1/3	1.4	3.3
top 1/3	−10.3	−6.6
utilitarian	4.7	6.3
E. Halve Initial Lump-Sum Transfer ι		
bottom 1/3	42.9	41.2
middle 1/3	8.6	10.0
top 1/3	−8.7	−6.0
utilitarian	12.3	13.4

Notes: The first column restricts $\xi = \tau_a = \xi_a = 0$. The second column is the unrestricted optimum. The welfare gains are computed taking transitions into account. All numbers are expressed in percent.

Figure A2: Optimal Tax Schedules, Alternative Estimates of the Tax Function



Notes: The horizontal dashed line is the optimal flat income tax when the planner is restricted to only using a flat tax on income. The blue and red solid lines are the optimal average and marginal taxes when the set of instruments is unrestricted. Vertical bars represent the 25^{th}, 50^{th}, 75^{th}, 95^{th} and 99^{th} percentiles of the income and wealth distributions in the initial steady state. The x-axis reports income (wealth) relative to mean income (wealth) in the initial steady-state.