Optimal Redistribution: Rising Inequality vs. Rising Living Standards

Axelle Ferriere
PSE & CNRS

Philipp Grübener
Goethe University Frankfurt

Dominik Sachs
University of St. Gallen*

March 24, 2024

Abstract

Over the last decades, the United States has experienced a large increase in, both, income inequality and living standards. The workhorse models of optimal income taxation call for more redistribution as inequality rises. By contrast, living standards play no role for taxes and transfers in these homothetic environments. This paper incorporates living standards into the optimal income tax problem by means of non-homothetic preferences. In a Mirrlees setup, we show that rising living standards alter both sides of the equity-efficiency trade-off. As an economy becomes richer, non-homotheticities imply a fall in the dispersion of marginal utilities, which weakens distributional concerns but has ambiguous effects on efficiency concerns. In a dynamic incomplete-market setup calibrated to the United States in 1950 and 2010, we quantify this new channel. Rising living standards dampen by at least 25% the desired increase in redistribution due to rising inequality.

Keywords: Fiscal Policy, Growth, Non-Homothetic Preferences, Redistribution
JEL Codes: E62, H21, H31, O23

*axelle.ferriere@psemail.eu, gruebener@econ.uni-frankfurt.de, dominik.sachs@unisg.ch. We thank Árpád Ábrahám, Timo Boppart, Georg Duernecker, Jonathan Heathcote, Jonas Loebbing, Ramon Marimon, Martí Mestieri, Raül Santaeulàlia-Llopis, Kjetil Storesletten, Gianluca Violante, and workshop and seminar participants at the EWMES 2020, BGSE Summer Forum EGF 2021, LAGV 2021, Oslo Macro Conference 2021, SED 2022, ESEM 2022, ZEW Public Finance Conference 2023, BSE Summer Forum Macroeconomics and Social Insurance 2023, ifo Conference on Macro and Survey Data 2023, NBER SI Inequality and Macroeconomics 2023, Sciences Po Workshop on Consumption and Saving in Macroeconomics 2023, the European University Institute, Goethe University Frankfurt, Lund University, and the University of Bristol for helpful comments and suggestions.

1 Introduction

Income inequality has been rising in the United States over the last decades, as documented in Piketty and Saez (2003), among others. As a result, fiscal redistribution has become a central topic in the policy debate, with popular calls for higher taxes and larger transfers. The literature on optimal income taxation characterizes the optimal tax-and-transfer ($t\&T$) system as a trade-off between equity and efficiency concerns. In the workhorse models, higher inequality indeed demands a more redistributive $t\&T$ system, as argued in Mankiw, Weinzierl, and Yagan (2009) and Diamond and Saez (2011).

Yet, in parallel to the rising income inequality, the United States has also experienced a very substantial increase in the standards of living. Mean income per capita has more than tripled since the 1950s, and the share of household expenditures spent on food has shrunk from more than 20% to less than 10%.[1] Standard models of optimal taxation feature homothetic preferences and cannot generate the observed heterogeneity in consumption baskets, both in the cross-section and over time. Loosely speaking, they cannot capture how being poor in the 1950s differs from being poor in the 2010s. Therefore, these models shed no light on how rising living standards affect efficiency and distribution concerns, and thus the optimal $t\&T$ system.

This paper incorporates living standards into the optimal income tax problem by means of non-homothetic preferences—that is, preferences featuring heterogeneous income elasticities of demand across multiple goods. First, we analytically show how changes in living standards affect the equity-efficiency trade-off in a static Mirrlees (1971) setup with fully flexible nonlinear taxes. Second, we quantify the relative effects of rising living standards and rising inequality from 1950 to 2010, using two complementary approaches: the Mirrlees framework; and a rich dynamic incomplete-market setup with flexible yet parametric nonlinear taxes. We consistently find that rising living standards reduce the desired increase in redistribution due to rising inequality by at least 25%, as measured by transfer-to-output ratios or by the difference in average income tax rates between the top- and bottom-income deciles.

Economic mechanisms. We mainly focus on the two recent state-of-the-art non-homothetic preference specifications in the structural change literature, namely Comin, Lashkari, and Mestieri (2021) and Alder, Boppart, and Müller (2022). These preferences imply heterogeneous income elasticities across goods, such that the marginal spending composition of an additional dollar depends on the level of income. As an economy grows, the share of expenditures spent on necessities falls, capturing the rising living standards. Heterogenous income elasticities across goods typically have two implications when living standards rise: the dispersion in marginal utilities across households falls,

[1] Data definitions are presented in Section 3.2.

and income effects weaken. These forces primarily affect the equity-efficiency trade-off in two ways. First, lower dispersion in marginal utilities reduces the gains from redistributing resources from rich to poor households. As such, the rising standard of living weakens distributional concerns—a force we refer to as the *distributional gains* channel of growth. Second, weaker income effects increase the efficiency costs of raising revenues but also decrease the efficiency costs of paying out transfers. As such, the rising standard of living has ambiguous effects on efficiency concerns—forces we refer to as the *efficiency costs* channel of growth. These dynamics arise as heterogeneous income elasticities typically imply decreasing relative risk aversion (DRRA), starting from a constant relative risk aversion (CRRA) for the homothetic counterpart of the utility function. Intuitively, richer households consume a smaller share of necessities, so that taking income risks is less costly. This property is also consistent with empirical micro evidence beyond the structural change literature.[2]

Two complementary approaches. The just described mechanisms are first formalized in a Mirrleesian setup. In particular, we consider fully nonlinear taxes in a static environment. We build on the analytical representation of optimal nonlinear taxes developed in Heathcote and Tsujiyama (2021) to formally decompose how living standards affect, both, efficiency costs and distributional gains of raising marginal tax rates along the income distribution. A calibration of this setup further allows us to quantify those different channels.

Second, we consider a quantitative dynamic incomplete-market setup in the tradition of Imrohoroğlu (1989), Huggett (1993), and Aiyagari (1994). In this richer setup, we follow a Ramsey approach and restrict the $t\&T$ system to belong to a flexible parametric class. While the Mirrleesian setup is powerful in imposing no restrictions on the $t\&T$ system, the dynamic environment allows to discipline preferences from intra- and intertemporal choices, with a meaningful notion of risk aversion. In addition, it allows to separate income from expenditure distributions, crucial to disentangle efficiency from distributional concerns. The dynamic setup is further used to discipline the calibration of the Mirrleesian setup mentioned above.

The logic of the quantitative exercise is as follows. First, we calibrate the model to the U.S. economy in 1950. We derive inverse optimum Pareto weights, which make the calibrated 1950 $t\&T$ system optimal. Keeping those weights constant, we then compute the optimal $t\&T$ system for two cases. First, we only account for the rise in inequality until 2010, as a benchmark comparable to the literature. Second, we compute the optimal $t\&T$ system when also accounting for rising living standards. We interpret the difference

[2] See Section 3.3.4 for a description of the empirical literature on risk aversion and the intertemporal elasticity of substitution (IES) over time and in the cross-section.

in the optimal $t\&T$ systems as the standard-of-livings channel. We now describe our calibration of the model and preview our quantitative results.

Quantification. We calibrate the dynamic model to be consistent with key micro- and macro-level developments of the U.S. economy from 1950 to 2010, with the non-homothetic CES (constant elasticity of substitution) preferences of Comin, Lashkari, and Mestieri (2021) as our benchmark preference specification. Key for distributional concerns, the model is consistent with the dynamics of inequality. Regarding the non-homotheticities, we use consumption and labor supply patterns in the cross-section and the time series to discipline preference parameters, which eventually govern the degree of DRRA in the calibrated economy. Furthermore, we derive an analytical relation between the degree of DRRA, wealth effects, and marginal propensities to consume (MPCs), and check consistency of model-implied wealth effects and MPCs with recent evidence. The implied degree of DRRA is modest, well within the range of plausible estimates from fields as diverse as portfolio choice, consumption Euler equation estimation, and development.

We then evaluate the effect of rising living standards on the optimal $t\&T$ system relative to the effect of rising inequality, using both approaches. In isolation, the large rise in inequality calls for a more redistributive $t\&T$ system, with the optimal transfer-to-output ratio going up by more than three percentage points in the dynamic Ramsey approach—and by more than six percentage points in the static Mirrlees framework. Accounting for the rise in living standards, the optimal $t\&T$ system still redistributes more in 2010 than in 1950, but to a lesser degree. Rising living standards dampen the optimal increase in transfer-to-output ratios by 30% and 40% in the Ramsey and Mirrlees frameworks. Rising living standards also dampen the optimal increase in the difference between top-10% and bottom-10% average $t\&T$ rates by about 25% in both frameworks.

We further use the Mirrlees setup for two purposes. First, we use the analytical income tax formula to quantify the different channels driving the effects of the rising living standards. We find that almost the whole effect stems from the *distributional gains* channel of growth. Second, we conduct a series of robustness checks, using alternative calibrations with the non-homothetic CES preferences, as well as the preference specification of Alder, Boppart, and Müller (2022). All experiments suggest effects of standards of living at least as large as in the benchmark.

Summing up, we consistently find that the rising living standards dampen by at least 25% the desired increase in redistribution due to rising inequality, and most of this effect comes from weakening distributional concerns.

Related literature Our work relates to both the public economics tradition, which studies optimal nonlinear income taxation (Diamond 1998; Heathcote and Tsujiyama 2021; Saez 2001), and the macroeconomic tradition, which focuses on restricted tax instruments in richer environments (Conesa and Krueger 2006; Heathcote, Storesletten, and Violante 2017; Holter, Krueger, and Stepanchuk 2019). We connect these approaches with the notion of standards of living by incorporating growth and non-homothetic preferences into the analysis (Alder, Boppart, and Müller 2022; Comin, Lashkari, and Mestieri 2021).

In doing so, we contribute to an emerging literature on optimal taxes with non-homothetic preferences. Oni (2023) analyzes the optimal progressivity of a loglinear income tax function in a static general equilibrium model with non-homotheticies. In that setup, lower progressivity increases demand for luxuries; the relative price of necessities thus falls, which is beneficial for the poor. As a result, optimal progressivity falls, from 0.07 with homothetic preferences to 0.03 with non-homothetic preferences. Jaravel and Olivi (2022) consider non-homothetic preferences in a Mirrleesian income taxation problem comparable to ours, and focus on the effects of heterogeneous inflation rates—that is, of changes in relative prices, with unequal incidence across the income distribution. In their setup, a rise in the price of necessities reduces optimal redistribution. Indeed, a rise in the price of a good reduces the value of a marginal dollar, as consumption baskets become more expensive. A rise in the price of necessities disproportionately affects the consumption baskets of the poor, reducing the value of redistributing a dollar from the rich to the poor. We adopt a different focus and analyze the effects of growth, modeled as a homogeneous fall in all prices.[3]

More related in terms of motivation are the works of de Magalhaes, Martorell, and Santaeulàlia-Llopis (2022) and Tsujiyama (2022), which explore optimal taxation as an economy develops using non-homothetic preferences but abstracting from heterogeneous income elasticities across goods.[4]

Finally, our paper complements the literature addressing to what extent the rise in inequality in the United States justifies an increase in tax progressivity. Considering the United States between 1980 and 2016, Heathcote, Storesletten, and Violante (2020) find that the inequality channel is neutralized by increasing efficiency costs of tax progressivity resulting from skill-biased technical change. Relatedly, Brinca, Duarte, Holter, and Oliveira (2022) reach a similar conclusion in a quantitative setup accounting for heterogeneous returns across occupations.[5] Our paper puts into perspective the focus on changes in inequality, i.e. on second moments of the income distribution, by accounting

[3]For calibration purposes, we also account for changes in relative prices in our quantitative exercise; Section 3.4 conducts a decomposition exercise to isolate the effects of relative price changes in our setup.

[4]de Magalhaes, Martorell, and Santaeulàlia-Llopis (2022) show that, when considering both private and public transfers, risk-sharing tends to be larger in developing economies than in rich countries, and further discuss optimality of this finding in a one-good model with Stone-Geary preferences. Tsujiyama (2022) considers how subsistence self-employment, which is more prevalent in developing economies, affects the equity-efficiency trade-off, also in a one-good environment.

[5]Considering a more recent period of 2004-2015, Jaravel and Olivi (2022) also question the typical implication of rising inequalities on optimal redistribution. The heterogenous inflation rates, which were higher for luxuries than for necessities, call for more regressive taxes and offset the inequality channel.

for concurrent changes in living standards, i.e. first moments of the income distribution. We show that first moments are key for the optimal $t\&T$ system.

2 Static Model: Theoretical Analysis

We consider a continuum of heterogeneous households with labor productivity θ, which is distributed according to a probability density function $f(\theta)$. Households supply labor n and earn gross income $y = \theta n$. This results in expenditure $e = y - \mathcal{T}(y)$, where $\mathcal{T}$ captures the $t\&T$ system. Households allocate their expenditures to J different goods. We denote as $c = (c_1, \ldots, c_J)$ the basket of consumption goods. We assume that utility is of the form

$$U(c) - B\frac{n^{1+\varphi}}{1+\varphi}.$$

Additive separability allows to separate the labor/expenditure choice from the consumption composition choice, and thus to decompose the optimization problem into two steps: Step 1 solves for the optimal labor/expenditure level, while Step 2 optimally allocates the expenditure across different goods.[6]

$$V(\theta; \mathcal{T}(\cdot), \Lambda, p) \equiv \max_{e,n} u(e; \Lambda, p) - B\frac{n^{1+\varphi}}{1+\varphi} \quad \text{s.t.} \quad e = n\theta - \mathcal{T}(n\theta), \tag{Step 1}$$

$$u(e; \Lambda, p) \equiv \max_{\{c_j\}_j} U(c) \quad \text{s.t.} \quad \sum_j \frac{p_j}{\Lambda} c_j = e. \tag{Step 2}$$

Let p denote the vector of relative prices—i.e. $p_j = 1$ w.l.o.g. in one sector j. We assume p to be constant. Instead we consider changes in Λ, which homogeneously scales the level of prices. As Λ grows by g, real expenditures grow by g for a given level of nominal expenditures—i.e. we model growth as a fall in prices. Thus, we refer to Λ as the *level* of the economy. With non-homothetic preferences, a higher Λ implies a shift of consumption baskets away from necessities. This is what we define as the rising living standards.

Importantly, Λ and p only affect the labor supply decision through their impact on $u_e(e; \Lambda, p)$. This insight is very useful to analyze the implications of Λ on the optimal $t\&T$ system in Section 2.3. Once we characterize the properties of $u(e; \Lambda, p)$, we can focus on (Step 1).

Section 2.1 introduces the two non-homothetic preferences we consider throughout the paper. Section 2.2 characterizes implications of the preferences on $u(e; \Lambda, p)$ when consistent with empirical regularities on consumption baskets. Section 2.3 analyzes how these properties alter the optimal $t\&T$ system.

[6] The additive separability also implies that the Atkinson-Stiglitz theorem holds in this environment. Hence, the optimal tax system implies uniform commodity taxes.

2.1 Heterogenous Expenditure Elasticities

We measure rising living standards by changes in consumption baskets as an economy grows. There is ample evidence that Engel curves, depicting how spending on different goods varies with income, are not linear, and consumption baskets are heterogeneous—both over time and in the cross-section.[7] In other words, expenditure elasticities of demand are heterogeneous across goods.[8] To capture the rising living standards, we thus consider utilities satisfying the following assumption.

Assumption 1. *Assume that $U(c)$ is such that expenditure elasticities are heterogeneous across goods. That is,*

$$\frac{\partial \log c_i}{\partial \log e} \neq \frac{\partial \log c_j}{\partial \log e} \text{ when } i \neq j.$$

There exist different functional forms for $U(c)$ that are consistent with this assumption. We focus on the two recent state-of-the-art non-homothetic preference specifications in the structural change literature, namely Comin, Lashkari, and Mestieri (2021) and Alder, Boppart, and Müller (2022).

Case 1. *(Non-homothetic CES Preferences)*

We first describe the non-homothetic CES preferences that go back to Hanoch (1975) and have been introduced into a multi-sector growth model by Comin, Lashkari, and Mestieri (2021). Non-homothetic CES preferences are defined over the basket of consumption goods c by:

$$U(c) = \frac{(\mathbb{C}(c))^{1-\gamma}}{1-\gamma},$$

where the consumption aggregator $\mathbb{C}(c)$ is implicitly defined by the following equation:

$$\sum_{j}^{J} \left(\Omega_j (\mathbb{C}(c))^{\varepsilon_j}\right)^{\frac{1}{\sigma}} c_j^{\frac{\sigma-1}{\sigma}} = 1, \tag{1}$$

with $\gamma > 0$, and $\sigma > 0$, $\Omega_j > 0 \; \forall j$, $\epsilon_j > 0$ ($\epsilon_j < 0$) if $\sigma < 1$ ($\sigma > 1$) $\forall j$. Preferences collapse to a homothetic CES when $\varepsilon_j = 1 - \sigma \; \forall j$.

For this utility function, one obtains the following elasticities of consumption w.r.t. expenditure:

$$\frac{\partial \log c_j}{\partial \log e} = \sigma + (1-\sigma)\frac{\varepsilon_j}{\bar{\varepsilon}},$$

[7] See Aguiar and Bils (2015), Boppart (2014), and Herrendorf, Rogerson, and Valentinyi (2014), among many others.

[8] We refrain from using the term income elasticities, which has been often used in this context. Indeed, after-tax incomes are relevant in a setting with income taxes. In the static model, after-tax income and expenditure are equivalent. For the dynamic version of the model, this is no longer the case. Thus, we prefer to refer to elasticities w.r.t. expenditure.

where $\bar{\varepsilon} = \sum_{j=1}^{J} \omega_j \varepsilon_j$ and ω_j are the expenditure shares of the different goods. Goods j with $\varepsilon_j < \bar{\varepsilon}$ are necessities, as their expenditure elasticities are lower than unity. Goods j with $\varepsilon_j > \bar{\varepsilon}$ are luxuries. As opposed to Stone-Geary preferences, non-homotheticities do not vanish: differences in $\partial \log c_j / \partial \log e$ also prevail as e keeps growing.[9]

Comin, Lashkari, and Mestieri (2021) show that one can express the expenditure function as

$$e = \left(\sum_j \Omega_j \mathcal{C}(e; \Lambda, p)^{\varepsilon_j} (p_j/\Lambda)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}, \tag{2}$$

where we denote $\mathcal{C}(e; \Lambda, p)$ the optimal consumption aggregator given an expenditure level e.

Throughout the rest of the paper, we focus on the case where $\sigma < 1$, appropriate to capture changes in consumption baskets across broad sectors reflecting rising living standards.

Case 2. *(Intertemporally Aggregable (IA) Preferences)*

The second state-of-the-art non-homothetic preferences are the IA preferences introduced by Alder, Boppart, and Müller (2022). These preferences are directly defined over expenditure:

$$u(e; p, \Lambda) = \frac{1}{1-\gamma} \left(\frac{1}{\mathbf{B}(p^*)} \left(e - \underbrace{\sum_j p_j^* \bar{c}_j}_{\mathbf{A}(p^*)} \right) \right)^{1-\gamma} - \mathbf{D}(p^*), \text{ with } p^* \equiv \frac{p}{\Lambda}, \tag{3}$$

where $\mathbf{B}(p^*) = \left(\sum_j \Omega_j \left(p_j^*\right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$ with $\sigma > 0, \sum_{j \in J} \Omega_j = 1$, and $\Omega_j \geq 0 \ \forall j$; $\mathbf{D}(p^*)$ homogenous of degree zero; and $\gamma \in (0, 1)$. IA preferences are homothetic when $\mathbf{A}(p^*) = 0$ and $\mathbf{D}(p^*) = 0$.

Alder, Boppart, and Müller (2022) show that these preferences allow for intertemporal aggregation, and nest both: generalized Stone-Geary (Herrendorf, Rogerson, and Valentinyi 2014), through the $\mathbf{A}$ term; and price independent generalized linearity (PIGL) preferences (Boppart 2014), through the $\mathbf{D}$ term.

2.2 Non-homotheticities and Marginal Utilities

We now derive implications of heterogeneous expenditure elasticities on the curvature of the indirect utility function $u(e; p, \Lambda)$, which contains all properties of non-homotheticities relevant for the optimal income taxation problem.

[9] Note that $\{\varepsilon_j\}$ can be rescaled, as shown in Comin, Lashkari, and Mestieri (2021). See Appendix A.1.1 for formal details in our setting.

Intratemporal allocations provide information on expenditure elasticities across goods. In the vein of Crossley and Low (2011), we argue that intratemporal allocations are informative of the curvature of u, and thus further impose restrictions on intertemporal allocations as well.[10] Crossley and Low (2011) prove in particular that the rich degree of heterogeneity in consumption baskets along the income dimension rules out the possibility that the IES is constant in income. Echoing a theoretical literature, we argue next that non-homothetic preferences typically imply DRRA.[11,12]

Non-homothetic CES preferences and DRRA. We start with Preferences 1. We denote by $\gamma(e; \Lambda, p)$ the coefficient of relative risk aversion at expenditure e. Risk aversion depends on both the consumption aggregator $\mathcal{C}$ and the curvature parameter γ:

$$\gamma(e; \Lambda, p) \equiv -\frac{u_{ee} e}{u_e} = \gamma \frac{\mathcal{C}_e(e; \Lambda, p) e}{\mathcal{C}(e; \Lambda, p)} - \frac{\mathcal{C}_{ee}(e; \Lambda, p) e}{\mathcal{C}_e(e; \Lambda, p)}, \tag{4}$$

where $\mathcal{C}_e$ and $\mathcal{C}_{ee}$ can be obtained through implicit differentiation of the expenditure function.

Intratemporal allocations of expenditure across goods discipline the parameters $\{\varepsilon_j\}$ and σ, thereby determining $\mathcal{C}$. Given $\mathcal{C}$, the parameter γ pins down the entire schedule of risk aversion—that is, both the level of risk aversion and how it varies with expenditures.

We now investigate the different parts of equation (4). The first term in equation (4) multiplies the curvature parameter γ with the elasticity of the consumption aggregator $\mathcal{C}$ with respect to expenditures. As stated below in Lemma 1, it is unambiguously decreasing in e, generating a force towards DRRA. The second term captures the elasticity of $\mathcal{C}_e$ with respect to expenditures, which may be increasing or decreasing in e. A larger curvature parameter γ not only increases the level of risk aversion, but also amplifies the first term which generates DRRA. Thus, DRRA is always satisfied for sufficiently high levels of risk aversion.

Lemma 1. *Preferences 1 satisfy DRRA at expenditure e—that is, $\gamma_e(e; \Lambda, p) < 0$—iff*

$$\gamma > \frac{\partial}{\partial e}\left(\frac{\mathcal{C}_{ee}(e; \Lambda, p) e}{\mathcal{C}_e(e; \Lambda, p)}\right)\left(\frac{\partial}{\partial e}\left(\frac{\mathcal{C}_e(e; \Lambda, p) e}{\mathcal{C}(e; \Lambda, p)}\right)\right)^{-1}.$$

[10] Quoting their conclusion: *"The importance of our result is in refuting the belief that properties of intertemporal allocations can be independent of the properties of within-period allocation. This belief underpins the use of the constant-IES assumption in much modern macroeconomics."* (Crossley and Low 2011, p.104).

[11] See for instance Browning and Crossley (2000): "luxuries are easier to postpone," implying an increasing IES. See also Hanoch (1977) and Stiglitz (1969) for additional discussions of how the shapes of Engel curves in the data are inconsistent with a constant IES. We review the empirical literature on DRRA in Section 3.3.4, when discussing the calibration of the quantitative model.

[12] We consider time-separable preferences for which the IES is the inverse of RA. Key to our analysis is to characterize how the curvature of static utility changes as expenditure grows.

Proof. The lemma follows from

$$\frac{\partial}{\partial e}\left(\frac{\mathcal{C}_e(e;\Lambda,p)e}{\mathcal{C}(e;\Lambda,p)}\right) < 0 \quad \forall e, \tag{5}$$

an inequality which is proved in Appendix A.1.1. □

Corollary 1. *Consider the homothetic CES parameterization of Preferences 1. Then, preferences feature CRRA:* $\gamma(e,\Lambda,p) = \gamma \ \forall e$.

We can characterize further the conditions for DRRA in two polar cases: the two-good case; and the case with a continuum of goods and a flexible distribution of parameters presented in Bohr, Mestieri, and Yavuz (2023).

Corollary 2. *Consider two polar cases of Preferences 1.*

- *Let* $J = 2$, *with* $\varepsilon_2 = 1$, $\varepsilon_1 = \varepsilon < 1$ *w.l.o.g. and* $\{\Omega_j = 1\}$. *Then, Preferences 1 satisfy DRRA if: (i) sufficient condition:* $\gamma > 2$*; (ii) necessary condition:* $\gamma > \varepsilon_1$.
- *Let there be a continuum of goods, and distributional assumptions as in Bohr, Mestieri, and Yavuz (2023). Then, Preferences 1 satisfy DRRA iff* $\gamma > 1$.

This corollary reveals further the role of the curvature parameter for the DRRA property. With a continuum of goods, γ is the only relevant statistic, with DRRA for $\gamma > 1$. Discreteness can obscure this relationship between γ and DRRA, creating tighter conditions for some levels of expenditures (sufficient condition larger than 1) and looser conditions for others (necessary condition weaker than 1). Yet, the main force remains: a larger γ always favors DRRA.

Given estimates for $\{\varepsilon_j\}$ and σ, the curvature parameter γ can be disciplined by alternative moments on the level of risk aversion or dynamics of labor supply. In particular, one can use long-run risk aversion, denoted $\bar{\gamma}$ and equal to

$$\bar{\gamma} \equiv (1-\sigma)\frac{\gamma - 1}{\varepsilon_J},$$

as shown in Appendix A.1.1. Easier to measure empirically, risk aversion at any point in time also uniquely pins down γ—this is the approach we follow in the quantitative model of Section 3, where we calibrate γ to match an average risk aversion of 1 in 2010. Finally, a large body of evidence has argued that aggregate labor supply falls with income, both over time and across countries.[13] Given estimates for $\{\varepsilon_j\}$ and σ, a level of γ translates into a certain fall in aggregate labor supply between two points in time.[14]

[13] See Section 3.3.2 for a summary of the literature.

[14] At a broader level, a literature has argued that aggregate labor supply falling with income over time or across countries provide direct support for non-homothetic preferences (Bick, Fuchs-Schündeln, and Lagakos 2018; Ohanian, Raffo, and Rogerson 2008; Restuccia and Vandenbroucke 2013).

We assume three goods in the quantification of Section 3, and borrow estimates of expenditure elasticities and the elasticity of substitution between goods from Comin, Lashkari, and Mestieri (2021). Targeting an average risk aversion of 1 in 2010, the model implies a fall in aggregate labor supply over time consistent with evidence, and satisfies DRRA as shown in Figure 1. More generally, DRRA is a property that consistently holds quantitatively.

IA preferences and DRRA.

Lemma 2. *Preferences 2 satisfy DRRA iff* $\mathbf{A}(p) > 0$*.*

Proof. See Appendix A.1.2. □

Corollary 3. *Preferences 2 satisfy CRRA under: (i) homothetic parameterizations; and (ii) PIGL parameterizations.*

For IA preferences, the DRRA property emerges from the subsistence term $\mathbf{A}$. Interestingly, $\mathbf{A}(p) > 0$ is a necessary condition for, both, DRRA and the fall in labor supply over time measured in the data. In the quantification of the IA preferences in Section 4.3, the curvature parameter yields the same long-run risk aversion as with the non-homothetic CES preferences, and the calibration of the $\{\bar{c}_j\}$ is such that $\mathbf{A}(p) > 0$, and thus, DRRA.

Taking stock. Non-homothetic preferences generally feature DRRA, a property which will be key to analyze how optimal $t\&T$ systems vary with living standards. As living standards will rise, the dispersion of marginal utilities will fall under DRRA, altering both distributional gains and efficiency costs of taxation.

2.3 Optimal Incomes Taxes

We consider a social planner that assigns Pareto weights $w(\theta)$ to households of type θ and optimally chooses a fully nonlinear $t\&T$ system $\mathcal{T}(\cdot;\Lambda)$ in the spirit of Mirrlees (1971). We derive an optimal income tax formula in this environment. We formally show that optimal taxes are independent of the level of the economy Λ when preferences are homothetic. Then, we analyze the effect of changes in Λ on optimal taxes when preferences are non-homothetic—holding fixed relative prices, the distribution of skills, and the Pareto weights. Throughout this section, we suppress the constant relative price vector p as an argument of tax and policy functions for readability.[15]

[15] In a similar formal environment, Jaravel and Olivi (2022) consider a different question, the effects of heterogenous inflation rates, for which they use a different setup. In particular, they abstract from the effects of price changes on u_{ee} for most of their analysis, which is why a homogenous fall in prices has no effect on optimal redistribution in their setup.

The government's problem is given by

$$\max_{\mathcal{T}(\cdot;\Lambda)} \int_{\underline{\theta}}^{\bar{\theta}} V(\theta;\mathcal{T}(\cdot;\Lambda),\Lambda)w(\theta)f(\theta)d\theta \ \text{ s.t. } \ \int_{\underline{\theta}}^{\bar{\theta}} \mathcal{T}(n(\theta;\mathcal{T}(\cdot;\Lambda),\Lambda)\theta;\Lambda)f(\theta)d\theta \geq G, \tag{6}$$

subject to optimal household behavior given the tax function:

$$n(\theta;\mathcal{T}(\cdot;\Lambda),\Lambda) \equiv \arg\max_{e,n} u(e;\Lambda) - B\frac{n^{1+\varphi}}{1+\varphi} \ \text{ s.t. } \ e = n\theta - \mathcal{T}\left(n\theta;\Lambda\right),$$

where G denotes exogenous spending and $V(\theta;\mathcal{T}(\cdot;\Lambda),\Lambda)$ is defined in (Step 1). To ease notation we replace $u_e(e(\theta,\Lambda),\Lambda)$ with $u_e(\theta,\Lambda)$, and omit $\mathcal{T}(\cdot;\Lambda)$ as arguments of household policy functions when possible.

We now state the solution to the optimal tax problem in the following lemma.

Lemma 3. *For each type θ^*, the optimal marginal tax rate $\mathcal{T}'(y(\theta^*;\Lambda);\Lambda)$ is characterized by:*

$$\underbrace{1 - \frac{1 - \frac{\mathcal{T}'(y(\theta^*;\Lambda);\Lambda)}{1-\mathcal{T}'(y(\theta^*;\Lambda);\Lambda)}\frac{1}{1+\varphi}\frac{\theta^* f(\theta^*)}{1-F(\theta^*)} + \int_{\theta^*}^{\bar{\theta}} \mathcal{T}'(y(\theta;\Lambda);\Lambda)\eta(\theta;\Lambda)\frac{dF(\theta)}{1-F(\theta^*)}}{1 + \int_{\underline{\theta}}^{\bar{\theta}} \mathcal{T}'(y(\theta;\Lambda);\Lambda)\eta(\theta;\Lambda)dF(\theta)}}_{E(\theta^*;\mathcal{T},\Lambda)} = \underbrace{1 - \frac{\int_{\theta^*}^{\bar{\theta}} u_e(\theta;\Lambda)\frac{w(\theta)}{f(\theta)}\frac{dF(\theta)}{1-F(\theta^*)}}{\int_{\underline{\theta}}^{\bar{\theta}} u_e(\theta;\Lambda)\frac{w(\theta)}{f(\theta)}dF(\theta)}}_{D(\theta^*;\mathcal{T},\Lambda)},$$

where

$$\eta(\theta;\Lambda) \equiv \frac{dy(\theta;\Lambda)}{d\mathcal{T}(0;\Lambda)} = \frac{\gamma(e;\Lambda)\frac{y(\theta;\Lambda)}{e(\theta;\Lambda)}}{\varphi + \gamma(e;\Lambda)\frac{y(\theta;\Lambda)}{e(\theta;\Lambda)}(1 - \mathcal{T}'\left(y(\theta;\Lambda);\Lambda\right)) - \frac{\mathcal{T}''(y(\theta;\Lambda);\Lambda)y(\theta;\Lambda)}{1-\mathcal{T}'(y(\theta;\Lambda);\Lambda)}} \tag{7}$$

denotes the income effect of type-θ worker.

Proof. See Appendix A.2.2. □

This derivation is a standard exercise. As in Heathcote and Tsujiyama (2021), we characterize the optimal marginal tax rate at income $y(\theta^*,\Lambda)$ as the one equalizing distributional gains $D(\theta^*;\mathcal{T},\Lambda)$ to efficiency costs $E(\theta^*;\mathcal{T},\Lambda)$.

Distributional gains. $D(\theta^*;\mathcal{T},\Lambda)$ captures the distributional gains from increasing the marginal tax at income $y(\theta^*,\Lambda)$ and redistributing the additional revenues lump-sum. The numerator in the fraction captures the utility loss from the higher taxes paid by workers of type $\theta \geq \theta^*$. The denominator in the fraction captures the utility gain from the larger lump-sum transfer to all workers.

When all workers are identical and Pareto weights are equalized, $D(\theta^*;\mathcal{T},\Lambda) = 0$: there is no gain from redistributing. Instead, with heterogeneous workers, the average marginal utility of workers above θ^*, in the numerator, is typically lower than the average marginal utility across the entire distribution, in the denominator. Thus, $D(\theta^*; \mathcal{T}, \Lambda) > 0$: there are positive gains from redistributing. The larger the dispersion in marginal utilities u_e, the larger the term $D(\theta^*; \mathcal{T}, \Lambda)$ becomes.

Efficiency costs. $E(\theta^*; \mathcal{T}, \Lambda)$ captures the efficiency costs from increasing the marginal tax at income $y(\theta^*; \Lambda)$ and redistributing the additional revenues lump-sum. The numerator in the fraction captures the distortionary effects on labor supply of increasing the marginal tax, which depends on two forces: (i) with positive Frisch elasticity $1/\varphi > 0$, workers with type θ^* reduce labor supply in response to the higher marginal tax;[16] (ii) with positive income effects $\eta(\cdot; \Lambda) > 0$, workers with type $\theta > \theta^*$ increase their labor supply in response to the higher average tax rate. The denominator in the fraction captures the distortionary income effects on labor supply of the larger lump-sum transfer to all workers.

With no income effects, the efficiency costs of increasing the marginal rate at $y(\theta^*; \Lambda)$ only arise from the workers with type θ^* decreasing labor. Assuming further inelastic labor supply $1/\varphi = 0$, there is no efficiency cost of taxation: $E(\theta^*; \mathcal{T}, \Lambda) = 0$. Instead, with elastic labor supply and positive income effects, the numerator in the fraction typically decreases while the denominator increases and $E(\theta^*; \mathcal{T}, \Lambda) > 0$: there are efficiency costs from redistributing. Larger income effects have ambiguous effects: they lower the efficiency cost of raising taxes in the numerator, but increase the efficiency costs of raising the lump-sum transfer in the denominator.

2.3.1 Benchmark: Homothetic Preferences

We first consider as a benchmark the homothetic parameterizations of Preferences 1 and 2. As discussed in Section 2.2, they satisfy CRRA. Proposition 1 formally states the irrelevance of the level of the economy for the optimal $t\&T$ system. We describe how the optimal allocation changes with growth and fully characterize the optimal tax reform that implements the new allocation.

Tax Reform. Consider a marginal increase in Λ by $d\Lambda$, and let $g \equiv d\Lambda/\Lambda$. We denote the optimal tax reform by

$$\forall y: \; d\mathcal{T}(y; \Lambda) \equiv \lim_{g \to 0} \frac{1}{g} \left\{ \mathcal{T}(y; \Lambda(1+g)) - \mathcal{T}(y; \Lambda) \right\}.$$

For any variable $v(\theta; \mathcal{T}, \Lambda)$, we denote its growth rate under a given tax reform $d\mathcal{T}$ as

[16] As shown in Appendix A.2.2, one can express the tax formula in terms of the distribution of income instead of types. When doing so, the compensated labor supply elasticity appears explicitly in the formula, and it does change with Λ. However, the density of income also changes with Λ, so that the two effects cancel out. This is why only the constant Frisch elasticity φ^{-1} appears in the formula with types that we use in Lemma (3).

$$\hat{v}(\theta;\mathcal{T},\Lambda,d\mathcal{T}) \equiv \lim_{g\to 0}\frac{1}{g}\frac{v\big(\theta;\mathcal{T}(\cdot;\Lambda)+g\times d\mathcal{T}(\cdot;\Lambda),\Lambda(1+g)\big)}{v\big(\theta;\mathcal{T}(\cdot;\Lambda),\Lambda\big)}-1.$$

Proposition 1. *Assume preferences $u(e;\Lambda)$ satisfy CRRA in Preferences 1 and 2. The optimal tax reform, which we denote $d\tilde{\mathcal{T}}$, to a marginal change in Λ is characterized by:*

$$\forall y:\ d\tilde{\mathcal{T}}(y;\Lambda) = (\mathcal{T}(y;\Lambda)-\mathcal{T}'(y;\Lambda)y)\,\alpha, \tag{8}$$

where $\alpha \equiv (1-\gamma)/(\varphi+\gamma)$. The resulting allocation is such that:

1. *Expenditures and incomes grow at constant rate α $\forall\theta$:*

$$\forall\theta:\ \hat{y}(\theta;\mathcal{T},\Lambda,d\tilde{\mathcal{T}}) = \hat{e}(\theta;\mathcal{T},\Lambda,d\tilde{\mathcal{T}}) = \alpha.$$

2. *The optimal marginal and average tax rate of a type-θ household do not change:*

$$\forall\theta:\ \mathcal{T}'(y(\theta;\Lambda(1+g));\Lambda(1+g)) = \mathcal{T}'(y(\theta;\Lambda);\Lambda)$$

and

$$\forall\theta:\ \frac{\mathcal{T}(y(\theta;\Lambda(1+g)),\Lambda(1+g))}{y(\theta;\Lambda(1+g))} = \frac{\mathcal{T}(y(\theta;\Lambda),\Lambda)}{y(\theta;\Lambda)}.$$

Proof. See Appendix A.2.3. □

Proposition 1 characterizes the optimal tax reform (8) in response to growth. At the new $t\&T$ system, households' incomes and expenditures optimally grow at a constant rate α, which can be positive or negative depending on the relative strength of income and substitution effects. Given the households' optimal policies, both marginal and average optimal $t\&T$ rates remain constant for each type θ. Pre-tax and after-tax income inequality remain constant.

To provide intuition on the irrelevance of growth for the optimal $t\&T$ system, we build on the optimal tax formula in Lemma 3. We first show that both distributional gains and efficiency costs are unchanged with growth when income and expenditure policies grow at a constant rate.

We start with distributional gains. As expenditures grow at the same rate $\forall\theta$, with CRRA preferences marginal utilities also grow at a constant rate $\forall\theta$. Thus, the welfare gains from redistributing from workers above θ^* to those below θ^* are unaffected: $\hat{D}(\theta^*;\mathcal{T},\Lambda,d\tilde{\mathcal{T}}) = 0\ \forall\theta^*$.

We turn to efficiency costs. In principle, efficiency costs may change with Λ as income effects defined in equation (7) do depend on Λ. Yet, income effects become independent of growth under the optimal tax reform. Income effects are a function of: (i) income-over-expenditure ratios, which are independent of Λ as both terms grow at the same rate;

(ii) marginal rates, which are constant at the optimal tax reform for each θ; and (iii) $\mathcal{T}'' \times y$, which is shown in Appendix A.2.3 to be constant as Λ increases. Thus, η is independent of Λ under the optimal tax reform:

$$\eta\big(\theta; \mathcal{T}(\cdot;\Lambda), \Lambda\big) = \eta\big(\theta; \mathcal{T}(\cdot;\Lambda) + g \times d\tilde{\mathcal{T}}(\cdot;\Lambda), \Lambda(1+g)\big) \ \forall\theta,$$

and thus efficiency costs are also unchanged with growth: $\hat{E}(\theta^*; \mathcal{T}, \Lambda, d\tilde{\mathcal{T}}) = 0 \ \forall\theta^*$.

As a consequence, the tax reform described in Proposition 1 is optimal, to the extent that it is consistent with household variables growing at the same rate α for all θ. Completing the proof requires to show that given the optimal tax reform at $\Lambda(1+g)$, household optimal policies adjust by a constant $\alpha = (1-\gamma)/(\varphi+\gamma)$ independent of θ—which we do in Appendix A.2.3.

Summing up, in the homothetic parameterizations of Preferences 1 and 2 satisfying CRRA, growth is not associated with a rise in living standards and leaves both marginal and average $t\&T$ rates unchanged.

2.3.2 Accounting for living standards

We now consider the non-homothetic parameterizations of Preferences 1 and 2 which satisfy the DRRA property. A change in the level of the economy Λ alters the optimal $t\&T$ system through three channels: (i) a *distributional gains* channel; (ii) an *efficiency costs* channel; and (iii) an *income distribution* channel.

To isolate these channels, we build on Proposition 1 in two steps. First, we assume that, as with homothetic preferences, incomes and expenditures grow at the same rate α for all θ, so that, relative to their mean, income and expenditure distributions do not change with growth. In that setup, we show how growth alters both distributional gains and efficiency costs of taxation. Then, we further show how the distributions of income and expenditure change with growth.

Step 1: Holding fixed the income distribution.

Proposition 2 (Distributional gains channel)**.** *Assume preferences $u(e;\Lambda)$ satisfy DRRA in Preferences 1 and 2. Consider the tax reform $d\tilde{\mathcal{T}}$, and assume incomes and expenditures grow at a constant rate α $\forall\theta$. Distributional gains decrease with growth:*

$$\forall\theta^*: \ \hat{D}(\theta^*; \mathcal{T}, \Lambda, d\tilde{\mathcal{T}}) = (1+\alpha)\Big[\mathbb{E}^u\left[\gamma(\theta;\Lambda)\right] - \mathbb{E}^u\left[\gamma(\theta;\Lambda)|\theta < \theta^*\right]\Big] < 0,$$

where $\mathbb{E}^u$ denotes the expectation using $f^u(\theta) \equiv w(\theta)u_e(\theta;\Lambda)f(\theta)/\int_\theta w(\theta)u_e(\theta;\Lambda)f(\theta)d\theta$, and $\gamma(\theta;\Lambda)$ denotes the relative risk aversion of a worker θ for a given Λ.

Proof. See Appendix A.2.4. □

Formally, $\hat{D}$ is negative as risk aversion decreases in expenditure, and thus in θ: average risk aversion over the entire population is smaller than the risk aversion of workers with $\theta < \theta^*$, $\forall \theta^*$. Intuitively, when preferences feature DRRA, the ratio of marginal utilities is no longer independent of growth, even with expenditures growing at the same rate across workers. As the economy grows, the dispersion in marginal utilities decreases, and thus distributional gains from redistributing from the rich to the poor decrease.

Proposition 3 (Efficiency costs channel)**.** *Assume preferences $u(e;\Lambda)$ satisfy DRRA in Preferences 1 and 2. Consider the tax reform $d\tilde{\mathcal{T}}$, and assume incomes and expenditures grow at a constant rate α $\forall\theta$. Income effects decrease with growth:*

$$\forall \theta: \ \eta(\theta;\mathcal{T} + g \times d\tilde{\mathcal{T}}, \Lambda(1+g)) < \eta(\theta;\mathcal{T},\Lambda), \tag{9}$$

which affects efficiency costs of taxation in two ways: (i) the efficiency costs of raising tax revenue increase; (ii) the efficiency costs of distributing a lump sum decrease. As such, $\hat{E}(\theta^*;\mathcal{T},\Lambda,d\tilde{\mathcal{T}})$ can be positive or negative.

Proof. See Appendix A.2.5. □

In contrast to the CRRA benchmark, income effects weaken with growth. This implies that the efficiency cost of raising revenue increases, but also that the efficiency cost of paying out transfers decreases. Thus, the effect of growth on the efficiency costs of taxation is ambiguous.

Step 2: Accounting for changes in the income distribution. When non-homothetic preferences in Preferences 1 and 2 feature DRRA, growth not only changes distributional gains and efficiency costs of taxes, but also directly changes income and expenditure distributions. Indeed, the growth rate of income now depends on θ.

Proposition 4 (Income distribution channel)**.** *Assume preferences $u(e;\Lambda)$ satisfy DRRA in Preferences 1 and 2. Consider the tax reform $d\tilde{\mathcal{T}}$. Then, the change in income is given by*

$$\forall \theta: \ \hat{y}(\theta;\mathcal{T},\Lambda,d\tilde{\mathcal{T}}) = \frac{1-\gamma}{\varphi+\gamma} + (\gamma - \gamma(\theta;\Lambda))d(\theta;\mathcal{T},\Lambda)$$

where $d(\theta;\mathcal{T},\Lambda) > 0$ is made explicit in Appendix A.2.6.

Proof. See Appendix A.2.6. □

Proposition 4 nests the CRRA case: when $\gamma(\theta;\Lambda) = \gamma$, the change in income is independent of θ. With DRRA, income changes depend on both risk aversion and $d(\cdot)$, where the latter is a function of $\mathcal{T}''$. We show in Appendix A.2.6 that $\hat{y}(\theta;\mathcal{T},\Lambda,d\tilde{\mathcal{T}})$ is unambigously increasing in θ under the standard loglinear tax function used in Feldstein

(1969), Benabou (2002), and Heathcote, Storesletten, and Violante (2017). Changes in income distribution are typically primarily driven by the other term, $\gamma - \gamma(\theta; \Lambda)$, which generates a force towards an increase in income inequality: with growth, labor income increases by less (or decreases by more) for low-θ than for high-θ workers.

Taking stock. To sum up, we have identified three channels of an increase in the standards of living on the optimal $t\&T$ system:

1. **Distributional gains channel** (Proposition 2). Higher standards of living lower the distributional gains of taxes. This channel calls for less redistribution as an economy grows.

2. **Efficiency costs channel** (Proposition 3). Higher standards of living raise the efficiency costs of raising revenue but lower the efficiency costs of distributing tax revenue back in form of lump-sum transfers. This channel has an ambiguous effect on optimal redistribution as an economy grows.

3. **Income distribution channel** (Proposition 4). Higher standards of living typically increase income inequality, and thus expenditure inequality, as labor supply of low-θ workers falls by more (or decreases by less) with growth. This channel calls for more redistribution as an economy grows.

Which of these effects dominates is a quantitative question that we explore in detail next. Anticipating the results, we will find that the distributional gains channel dominates: the optimal $t\&T$ system becomes less redistributive with growth.

3 Quantitative Models

We now present the frameworks to quantify the effects of rising living standards on the optimal $t\&T$ system. To put these effects in perspective, the frameworks also account for rising inequality. We use two complementary approaches.

We start with a Ramsey approach and describe optimal parametric $t\&T$ systems in a rich dynamic incomplete-market version of the model in the tradition of Imrohoroğlu (1989), Huggett (1993), and Aiyagari (1994). A dynamic model offers two main advantages. First, precautionary savings endogenously generate a distribution of expenditure given the observed distribution of income, which is crucial to disentangle efficiency from distribution concerns. Second, a model with savings generates dynamic moments such as MPCs and wealth effects, for which we have empirical counterparts, and which are intrinsically related to risk aversion.[17] Thus, a dynamic model is useful to discipline the DRRA property arising from non-homothetic preferences, alongside consumption composition.

We then use a Mirrlees approach in a static setup. This approach offers two advantages. First, it allows to check that the results are not driven by the specific $t\&T$ functional forms assumed in the Ramsey exercise. Second, it allows to build on the optimal tax formula in Lemma 3 and decompose the relative importance of the three channels of growth identified in Section 2.

Section 3.1 formally introduces the dynamic model. Section 3.2 describes the calibration of preferences, growth, and changes in inequality in the dynamic model to the U.S. economy from 1950 to 2010. Section 3.3 checks the model implications of non-homotheticities on static decisions—that is, consumption and labor patterns—as well as on dynamic decisions—that is, wealth effects and MPCs; we also compare the implied level of DRRA in the model to empirical estimates provided in the literature. Finally, Section 3.4 addresses the calibration of the static model.

3.1 Dynamic Model: Setup

The dynamic model is a standard incomplete-market setup in the tradition of Imrohoroğlu (1989), Huggett (1993), and Aiyagari (1994). Households are characterized by their productivity θ and holdings of a risk-free bond a. The household problem reads as follows:

$$\begin{aligned} V(a,\theta) &= \max_{e,a',n} \left\{ u(e;\Lambda,p) - B\frac{n^{1+\varphi}}{1+\varphi} + \beta \mathbb{E}_{\theta'}\left[V(a',\theta')\,|\theta\right] \right\} \\ \text{s.t.}\quad & e + a' \leq \theta n + (1+r)a - \mathcal{T}(\theta n), \quad a' \geq 0, \end{aligned} \tag{10}$$

where the utility function u will be non-homothetic. Households discount the future with discount factor β and face a no-borrowing constraint. Productivity θ follows a stochastic process. The $t\&T$ system $\mathcal{T}(\cdot)$ is modeled as a parametric function of labor income. We describe all functional forms in the calibration section. The problem is cast in partial equilibrium, with the interest r and the vector of prices taken as exogenous.

3.2 Dynamic Model: Calibration

We calibrate the model in two points in time: 1950 and 2010. We use non-homothetic CES preferences with three sectors: agriculture/food, manufacturing/goods, and services. Rising living standards result from falling prices, disciplined by GDP per capita growth and changes in relative prices. Rising income inequality results from changes in the distribution of idiosyncratic productivity shocks. Taxes and transfers describe the U.S.

[17] We derive an explicit relationship between MPCs, wealth effects, and risk aversion in equation (12) in Section 3.3.3.

fiscal system in 1950 and 2010. Table 1 presents all parameter values while Table 2 summarizes all targets.

3.2.1 Preferences

Non-homothetic CES. The benchmark preference specification uses non-homothetic CES preferences. We rely on the estimates of Comin, Lashkari, and Mestieri (2021), based on micro data from the CEX, for the parameters ε_j, governing the expenditure elasticities of demand, and σ, governing the substitutability of the different commodities. We set $\sigma = 0.3$, $\varepsilon_A = 0.1$, $\varepsilon_G = 1.0$, and $\varepsilon_S = 1.8$. As such, agricultural products are the necessities, with a low expenditure elasticity of demand, whereas services are the luxury, with a high expenditure elasticity of demand. We set the parameters Ω_j of the non-homothetic CES to match aggregate sector shares in 2010, based on Herrendorf, Rogerson, and Valentinyi (2013): 8% for agriculture, 26% for goods, and 67% for services.[18]

Other preference parameters. We set the discount factor β to match a wealth-to-income ratio of 4 in 2010 (Piketty and Zucman 2014). We fix the Frisch elasticity at a standard value with $1/\varphi = 0.5$ and the labor disutility parameter B such that average labor supply in 2010 is 0.3. We target an average relative risk aversion in 2010 of 1, a standard value in the literature that often relies on log utility.[19] This procedure yields a curvature parameter γ of 0.75—lower than the value required by the condition in Lemma 1, which delivers $\gamma > 1.9$ for our set of parameters to guarantee DRRA at *any* level of expenditures. Still, with an implied long-run relative risk aversion of 0.9, this calibration delivers DRRA at all relevant expenditure levels in our computations, as shown in Figure 1. We discuss further the magnitude of DRRA in Section 3.3.4.

3.2.2 Growth and Prices

We fix the interest rate at 2% for both years. As is standard with this class of model, we can normalize the price vector in one period.[20] We calibrate the three prices in the other period to match three moments: aggregate growth in GDP per capita from 1950 to 2010; and changing relative prices of agriculture and services to goods over the same time period.

Accounting for changes in relative prices is necessary to compare the model-implied rise in living standards, as captured by changes in sectoral expenditure shares, to its

[18]We follow the final expenditure rather than value added approach in Herrendorf, Rogerson, and Valentinyi (2013) since we are modeling household expenditure behavior rather than production.

[19]See for instance Guner, Kaygusuz, and Ventura (2023), Heathcote, Storesletten, and Violante (2017), and Saez (2001).

[20]See for instance Buera, Kaboski, Rogerson, and Vizcaino (2022) for an explanation of the price normalization.

Table 1: Parameter Values

Parameter	Interpretation	Value
Preferences		
β	Discount factor	0.957
γ	Curvature utility	0.750
$1/\varphi$	Frisch elasticity	0.500
B	Labor disutility	8.340
σ	Non-homothetic CES parameter	0.300
ε_A	Non-homothetic CES parameter	0.100
ε_G	Non-homothetic CES parameter	1.000
ε_S	Non-homothetic CES parameter	1.800
Ω_A	Non-homothetic CES parameter	0.057
Ω_G	Non-homothetic CES parameter	1.000
Ω_S	Non-homothetic CES parameter	10.303
Prices		
p_A^{1950}	Price agriculture 1950	3.032
p_G^{1950}	Price goods 1950	5.669
p_S^{1950}	Price services 1950	1.794
p_A^{2010}	Price agriculture 2010	1.000
p_G^{2010}	Price goods 2010	1.000
p_S^{2010}	Price services 2010	1.000
r	Interest rate	0.020
Inequality		
ρ_θ	Persistence productivity	0.900
σ_θ^{1950}	Std. dev. productivity innovation 1950	0.273
σ_θ^{2010}	Std. dev. productivity innovation 2010	0.302
α_θ^{1950}	Pareto tail parameter 1950	2.200
α_θ^{2010}	Pareto tail parameter 2010	1.650
Government		
λ^{1950}	Tax function level 1950	0.146
τ^{1950}	Tax function progressivity 1950	0.134
T^{1950}	Transfer 1950	0.005
G^{1950}	Government spending 1950	0.070
λ^{2010}	Tax function level 2010	0.172
τ^{2010}	Tax function progressivity 2010	0.073
T^{2010}	Transfer 2010	0.020
G^{2010}	Government spending 2010	0.078

Notes: Table 1 summarizes the parameter values.

Table 2: Targeted Data and Model Moments

Moment	Source	Data	Model
Moments related to Preferences			
Wealth-to-income ratio 2010	Piketty et al. (2014)	4	4
Avg. RRA in 2010	Standard value (log utility)	1.00	0.99
Avg. labor supply 2010	Normalization	-	0.3
Agg. agriculture share 2010	Herrendorf et al. (2013)	7.5%	7.5%
Agg. goods share 2010	Herrendorf et al. (2013)	25.6%	25.6%
Agg. services share 2010	Herrendorf et al. (2013)	66.9%	66.9%
Moments related to Prices			
Change rel. price A to G	Herrendorf et al. (2013)	1.87	1.87
Change rel. price S to G	Herrendorf et al. (2013)	3.16	3.16
GDP per capita growth	NIPA	3.34	3.34
Moments related to Inequality			
Variance log income 1950	SCF+	0.57	0.55
Variance log income 2010	SCF+	0.78	0.78
Moments related to Government			
T/Y 1950	OMB	1.1%	1.1%
G/Y 1950	OMB / const. spending	14.0%	14.0%
AMTR difference 1950	Mertens et al. (2018)	12.86%	12.86%
T/Y 2010	OMB	3.6%	3.6%
G/Y 2010	OMB / const. spending	14.0%	14.0%
AMTR difference 2010	Mertens et al. (2018)	8.67%	8.67%

Notes: Table 2 summarizes data moments and their model counterparts.

empirical counterpart which is measured in nominal terms.[21] As discussed in Jaravel and Olivi (2022), changes in relative prices may also have efficiency and distributional implications on the optimal $t\&T$ system because heterogeneous households consume heterogeneous baskets of goods. We quantitatively isolate this force in a decomposition exercise in Section 4.2.

We compute aggregate growth in GDP per capita from 1950 to 2010 from National Income and Product Accounts (NIPA) to be equal to 3.3. We compute changes in relative prices based on Herrendorf, Rogerson, and Valentinyi (2013). From 1950 to 2010, the relative price of agriculture (food) rises by a factor of 1.87 relative to goods, and the relative price of services rises by a factor of 3.16. These targets translate into falling prices for all commodities from 1950 to 2010, with the largest fall in goods and the smallest fall in services.

3.2.3 Inequality Dynamics

Household productivity follows a log AR(1) process, to which a Pareto tail is appended, with a time-varying Pareto tail parameter α set to 2.2 in 1950 and 1.65 in 2010 (Aoki and Nirei 2017). We fix the persistence of the productivity process ρ_θ to 0.9 and set the standard deviation of the innovation each period to match the variance of log income in 1950 (0.57) and 2010 (0.78) computed in the extended Survey of Consumer Finances (SCF+) of Kuhn, Schularick, and Steins (2020).[22]

The variance of log income is targeted explicitly, but the model provides a good fit for income inequality along the entire income distribution. Table 4 shows income shares by quintile. As in the data, the income share of the bottom quintile falls by a third in the model, and the share of income going to the top quintile strongly increases.

3.2.4 Government

We restrict the analysis to a parametric but flexible functional form, following Ferriere, Grübener, Navarro, and Vardishvili (2023). The tax payment is given by

$$\mathcal{T}(y) = \exp\left[\log(\lambda)\left(y^{-2\tau}\right)\right] y - T. \tag{11}$$

The first part of the equation describes a two-parameter tax function, with parameter λ governing the level of taxes and parameter τ governing the progressivity, and T is a lump-sum transfer. As compared to the widely used loglinear tax function popularized by Feldstein (1969), Benabou (2002), and Heathcote, Storesletten, and Violante (2017), our functional form allows to better jointly match the bottom and the top of the tax

[21] See Section 3.3.1 for a discussion of the changes in sectoral expenditure shares implied by this calibration.

[22] See Appendix B.1 for details on the SCF+ data.

distribution (Ferriere, Grübener, Navarro, and Vardishvili 2023). Loosely speaking, T is disciplined by average tax-net-of-transfer rates and τ by the marginal tax rates at the top.

We set T and τ for the years 1950 and 2010 to match the transfer-to-output ratio and the difference in average marginal tax rates (AMTRs) between the top-10% and the bottom-90% of the income distribution. We measure transfers as income security programs, amounting to 1.1% of GDP in 1950 and 3.6% of GDP in 2010.[23] We use data from Mertens and Montiel Olea (2018) to compute the difference in AMTRs, equal to 13% in 1950 and 9% in 2010. We set exogenous spending G to match a spending-to-output ratio of 14% in both years.[24] The parameter λ of the tax function is determined by the restriction that the government budget has to clear period by period.

The flexible functional form with transfers modeled separately from progressive taxes allows to capture two key developments of the U.S. $t\&T$ system over the last decades. First, marginal tax rates have become less progressive, reflected in a lower progressivity parameter in 2010 than in 1950 (Ferriere and Navarro 2023). Second, transfers have risen significantly over this time period, such that average $t\&T$ rates have become more progressive (Heathcote, Storesletten, and Violante 2020; Splinter 2020).

3.3 Dynamic Model: Validation

We now validate the calibration. We investigate expenditure and labor supply patterns, both over time and in the cross-section. We also verify dynamic decisions with wealth effects and MPCs. We end this section with a comparison of the implied degree of DRRA to estimates in the literature.

3.3.1 Expenditures

Aggregate expenditure shares over time. We investigate the (untargeted) change in aggregate sector shares between 1950 and 2010, to validate the rising living standards in the model. As shown in Table 3, the model captures well the structural change out of agriculture towards services, with an agricultural sector share of 17% (data: 22%), goods share of 49% (39%), and services share of 34% (39%) in 1950.

[23] Income security programs consist of general retirement and disability insurance (excluding social security), federal employee retirement and disability, unemployment compensation, housing assistance, food and nutrition assistance, and other income security (White House Office of Management & Budget, OMB).

[24] Spending has risen over time in the data, but this increase has been largely deficit financed, which we do not model in our stationary setup—see for instance https://fred.stlouisfed.org/series/FYFSDFYGDP. We keep the spending-to-output ratio constant, as changing spending requirements would introduce further dynamics in the optimal level of tax progressivity—see Heathcote and Tsujiyama (2021) and Ayaz, Fricke, Fuest, and Sachs (2023) for discussions of how "fiscal pressure" influences optimal tax progressivity.

Table 3: Untargeted Data and Model Moments

Moment	Source	Data	Model
Agg. agriculture share 1950	Herrendorf et al. (2013)	21.5%	16.7%
Agg. goods share 1950	Herrendorf et al. (2013)	39.2%	49.1%
Agg. services share 1950	Herrendorf et al. (2013)	39.2%	34.2%
Wealth-to-income ratio 1950	Piketty et al. (2014)	3.5	3.0
Agg. fall in labor supply	Ramey et al. (2009)	5-7%	7%

Notes: Table 3 summarizes a subset of untargeted data moments and their model counterparts.

Table 4: Income and Wealth Distributions

1950	**Income** Share by Quintile				
Model	6%	11%	13%	21%	49%
Data (SCF+)	6%	11%	15%	21%	48%
2010	**Income** Share by Quintile				
Model	4%	9%	11%	19%	56%
Data (SCF+)	4%	9%	13%	21%	53%
1950	**Wealth** Share by Quintile				
Model	0%	2%	6%	17%	76%
Data (SCF+)	0%	1%	4%	11%	84%
2010	**Wealth** Share by Quintile				
Model	0%	1%	5%	13%	81%
Data (SCF+)	-1%	1%	3%	10%	87%

Notes: Table 4 compares income and wealth shares by quintile of the respective distribution in model and data. Data comes from the SCF+.

Expenditure inequality. We investigate the (untargeted) change in expenditure inequality between 1950 and 2010, to validate the change in distributional concerns in the model. Expenditure inequality in the model is the result of income inequality, which we match, and private savings decisions.

In line with evidence, the expenditure distribution is more equal than the income distribution. In 2010, the variance of log expenditure in the model is 0.46, close to the number of 0.36 reported in Attanasio and Pistaferri (2014) using CEX data. The discrepancy traces back to our assumption of a Pareto tail in the income distribution, while the CEX does not oversample high-income households. The model also matches well the distribution of wealth by quintile, as reported in Table 4.

There is no evidence on the distribution of expenditure in 1950. Yet, the model generates a reasonable wealth-to-income ratio (Table 3) and distribution of wealth by quintile (Table 4), which is informative of the capacity of the model to also generate a reasonable distribution of log expenditure. We obtain a variance of log expenditure of 0.33 in 1950 in the model, thus smaller than in 2010.[25]

Expenditure shares in the cross-section. We investigate cross-sectional heterogeneity in expenditure sector shares in 2010 in the model to validate further the preference parameters $\{\varepsilon_j\}$ estimated in Comin, Lashkari, and Mestieri (2021). The agriculture expenditure share is 8.8 percentage points larger in the bottom than in the top expenditure quintile in 2017 (Meyer and Sullivan 2023), to be compared to 11.6 percentage points in the model. Instead, the services expenditure share is 10 percentage points smaller in the bottom than in the top income quintile in 2010 (Boppart 2014), to be compared to 12.5 percentage points in the model. Overall, the model captures well cross-sectional heterogeneity in expenditure sector shares in 2010. Again, there is no cross-sectional evidence for 1950.

3.3.2 Labor Supply

Recent literature has documented key patterns of labor supply over time, across countries, and in the cross-section within a country. Boppart and Krusell (2020) find as a robust pattern of labor supply over time across countries a steady fall in hours worked by roughly 0.5% per year. For the postwar United States, McGrattan and Rogerson (2004) and

[25]There is no consensus in the literature on dynamics of consumption inequality over time. Krueger and Perri (2006) or Heathcote, Perri, Violante, and Zhang (2023), among others, report stable or moderately increasing consumption inequality. Accounting for measurement error, Attanasio, Battistin, and Ichimura (2007), Attanasio, Hurst, and Pistaferri (2014), or Aguiar and Bils (2015) find instead that consumption inequality has increased as much as income inequality since the 1980s. Meyer and Sullivan (2023) focus on well-measured consumption only in the Consumer Expenditure Survey and find that consumption inequality rose less than income inequality between 1961 and 2017. See this paper also for a more complete review of the different approaches and results in the literature.

Ramey and Francis (2009) find a fall in hours per worker of 5-7%.[26] Cross-sectional patterns of labor supply have also changed over time. Before the 1970s, low-wage workers worked more hours than high-wage workers, a pattern which has reversed since then (Costa 2000; Heathcote, Perri, and Violante 2010; Heathcote, Perri, Violante, and Zhang 2023; Mantovani 2023).

We compute aggregate and cross-sectional changes in labor supply in the model. Aggregate labor supply falls by 7% over time, a number which is somewhat high but consistent with the estimates in the literature. The correlation between hours worked and hourly wage increases by 12 points from 1950 to 2010—Heathcote, Perri, Violante, and Zhang (2023) compute an increase of 22 points for men and 7 points for women from 1967 to 2021. A larger curvature in utility would increase the change in the correlation, at the expense of a larger fall in labor supply. Overall, this parsimonious model captures fairly well the effects of growth on labor supply, both in the aggregate and the cross-section.

3.3.3 Wealth Effects and MPCs

We exploit the dynamic dimension in the model to further validate the calibration of preferences and the implied degree of DRRA. In particular, we link RRA to concepts that are better measurable in the data, such as MPCs and wealth effects. To do so, we derive the following expression from the households' budget constraint and savings decisions:

$$\eta\left(\varphi\frac{e}{\theta n}+\frac{e\mathcal{T}''(\theta n)}{\mathcal{T}'(\theta n)}\right)=\text{MPC}\times\text{RRA}, \tag{12}$$

where η denotes the wealth effect.[27]

In 2010, the calibrated model produces MPCs and wealth effects well in line with available evidence.[28] For MPCs, we compute the expenditure response to a \$500 increase in wealth. On average, the model produces an MPC of 18%. While this is relatively low compared to most of the available evidence (Fagereng, Holm, and Natvik 2021; Johnson, Parker, and Souleles 2006; Kaplan and Violante 2022), we consider it a success in a model with just one asset calibrated to the entire stock of wealth. For wealth effects, we compare the model response to a one-time unanticipated wealth shock to the evidence provided by Golosov, Graber, Mogstad, and Novgorodsky (2023), who measure the earnings response to lottery winnings using the universe of U.S. taxpayers. The model captures very well that earnings fall by \$2.3 in response to a \$100 wealth shock.

[26]In terms of total hours this is compensated by rising female labor force participation, a pattern we abstract from in the model.

[27]Details of the derivation are presented in Appendix C.1.1.

[28]See Appendix C.1.2 for more details on the computations of the MPCs and the wealth effects in the model.

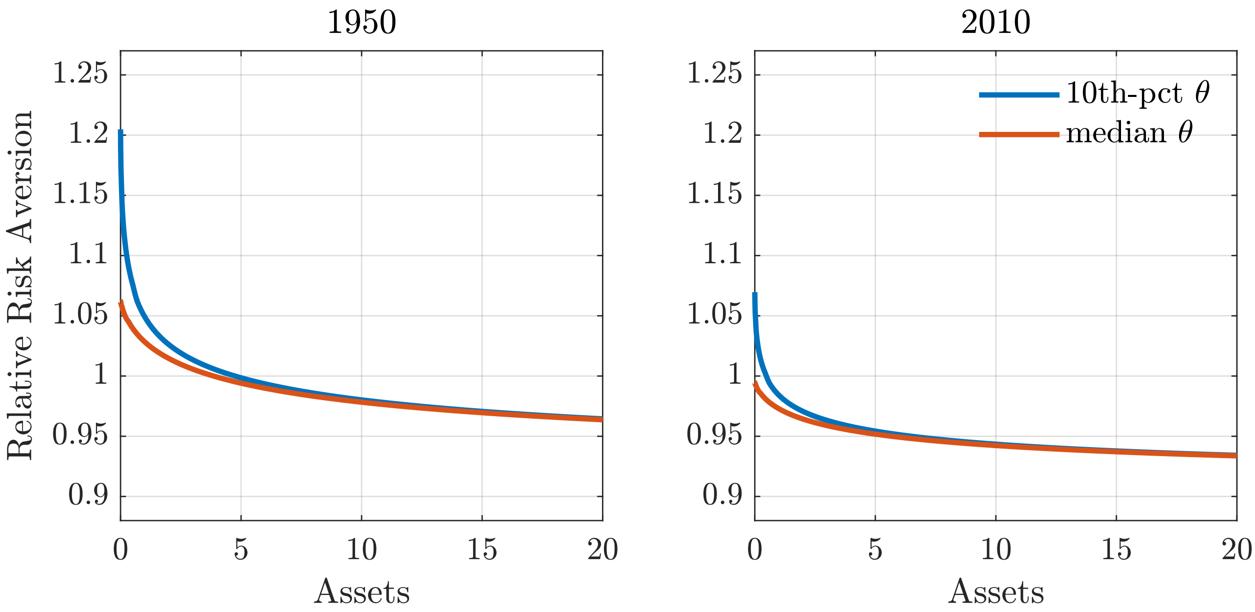

Figure 1: Relative Risk Aversion

Notes: Figure 1 plots dispersion in relative risk aversion in the calibrated model in 1950 (left panel) and 2010 (right panel). Wealth is normalized by mean wealth.

3.3.4 DRRA: Relation to the Literature

Finally, we directly compare the model implied degree of DRRA to available evidence from the literature, attained using a vast variety of different approaches. The calibrated model implies a modest degree of DRRA. From 1950 to 2010, average relative risk aversion falls from 1.07 to 0.99. Similarly, cross-sectional dispersion in risk aversion is small, as shown in Figure 1. This degree of DRRA is small relative to available evidence, as we describe next.

Evidence on DRRA first stems from direct empirical estimates of varying RRA or IES. Ogaki and Zhang (2001) and Zhang and Ogaki (2004) reject the hypothesis of CRRA in favor of DRRA using consumption data from Pakistani and Indian villages. Atkeson and Ogaki (1996) estimate the IES both using Indian panel data and in the aggregate time series for India and the United States. They find that the IES of the richest households in India is 60% higher than the one of the poorest households. The ratio of the IES between the United States and India is roughly 1.5. In the U.S. time series they estimate an increase in the IES from 0.38 to 0.41 from 1929 to 1988. Blundell, Browning, and Meghir (1994) and Attanasio and Browning (1995) also estimate an IES increasing in consumption using UK data. Finally, Blundell, Browning, and Meghir (1994) who report variation in the IES from the 10th to the 90th percentile ranging from 0.66 to 1.10 or 0.96 to 2.8, depending on the specification. Relative to this body of evidence, the variation in our model is modest.

In addition to these direct estimates, DRRA is an important feature in making theory consistent with data in a variety of fields. In consumption theory, Straub (2019) employs non-homothetic preferences that imply risk aversion decreasing in income and wealth to account for consumption responses to permanent income changes. In finance, DRRA typically helps in matching portfolios across the wealth distribution (Cioffi 2021; Meeuwis 2022; Wachter and Yogo 2010) and in mitigating the equity premium puzzle (Ait-Sahalia, Parker, and Yogo 2004). More closely related to us, in the development literature, Donovan (2021) argues that the heterogenous income elasticities coming from non-homothetic preferences also imply DRRA, a feature which is important in accounting for aggregate productivity differences across countries.[29]

3.4 Static Model: Calibration

Finally, we briefly describe the quantification of the Mirrlees static model. We partly use the dynamic model to quantitatively discipline the static model, as we explain next. Table C.1 summarizes all parameters.

Preference parameters $\{\varepsilon_j; \sigma\}$ rely on the micro estimates from Comin, Lashkari, and Mestieri (2021), while the parameters $\{\Omega_j\}$ are set to match aggregate sector shares. We also keep the other parameters of the utility function, γ and φ, as in the dynamic model. Prices are set to replicate aggregate growth and changes in relative prices over time. We set government parameters to match transfer-to-output ratios, spending-to-output ratios, and the difference in AMTRs between the top-10% and bottom-90% of the distribution.

In contrast to the dynamic model, there is no distinction between after-tax income and expenditure in this environment. For the calibration of this static model, we therefore follow a partial insurance approach and calibrate productivities such that the model is consistent with inequality in expenditures. Specifically, we calibrate the distribution of skills as an exponentially modified Gaussian distribution (EMG), as in Heathcote and Tsujiyama (2021). The Pareto tail of the expenditure distribution is thinner than that of the income distribution. For 2010, we follow the literature and set the tail parameter to 3.3, twice larger than the one for the income distribution (Aoki and Nirei 2017; Toda and Walsh 2015). For 1950 there are no available estimates of the tail parameter for the expenditure distribution. Thus, we assume constant the relation between income and consumption tail parameters. Based on a tail parameter for the income distribution of 2.2, we set the parameter to 4.4 for the expenditure distribution. Next, we set the variance of the normal shock in 2010 to match a variance of log expenditure of 0.37—a number well in line with the literature (Attanasio and Pistaferri 2014; Heathcote, Perri,

[29] The mechanism is that lower intermediate input usage in agriculture in developing countries is driven by the combination of idiosyncratic shocks, incomplete markets, and subsistence requirements, where the latter imply DRRA, because lower intermediate input usage limits exposure to uninsurable shocks.

and Violante 2010). The variance of log expenditure is thus about 40% the variance of log income in 2010. We maintain this ratio and set the variance of the income shock to calibrate a variance of log expenditure of about 0.26 in 1950.[30]

We validate the calibration of the static model by examining labor supply behavior over time and in the cross-section. Over time, aggregate labor supply is falling by 5%. Labor supply is monotonically decreasing in productivity in 1950 and monotonically increasing in productivity in 2010, as in the data. Average risk aversion amounts to numbers which are very comparable to the dynamic model, at 1.08 in 1950 and 0.99 in 2010.

4 Optimal Policies

We now quantify the effects on the optimal $t\&T$ system of the rising living standards relative to the rising inequality. Section 4.1 follows a Ramsey approach in the dynamic model and computes the optimal fiscal system in 2010 within the class of $t\&T$ functions described in Section 3.2. Section 4.2 complements the analysis with the optimal fully nonlinear $t\&T$ system in the static model, and further uses the Mirrlees setup to decompose the three effects of growth isolated in Section 2.3. Section 4.3 presents various robustness exercises.

4.1 Ramsey Analysis in Dynamic Model

We start with the Ramsey analysis in the dynamic model. We proceed in three steps. First, we find inverse optimum Pareto weights that make the observed $t\&T$ system in 1950 optimal. Second, we add the change in inequality from 1950 to 2010. Third, we also account for rising income levels with the fall in prices.

Pareto Weights. We start from the 1950 calibrated $t\&T$ system and find the Pareto weights under which it is optimal. When evaluating the optimal $t\&T$ system in 2010, we will then assume social preferences fixed over time.[31]

In static Mirrlees models, it is natural to make welfare weights a function of productivity. In the dynamic model, heterogeneity is two-dimensional, with households differing both in productivity and wealth. A one-dimensional measure, capturing how well-off a household is, is expenditure.[32]

[30] We show that the expenditure distribution by quintiles are comparable across the static and the dynamic models; see Table C.2 in the appendix.

[31] In Section 4.3, we also present robustness with a Utilitarian social welfare function.

[32] Chang, Chang, and Kim (2018) also use an inverse optimum approach conditioning Pareto weights on expenditures. As a robustness, we have also made weights a function of productivity only and obtained similar results as those reported here. See also Le Grand and Ragot (2023) for an example of inverse optimum Pareto weights on productivity in a similar setup.

The $t\&T$ system in 1950 is characterized by two parameters, T and τ. Hence, we use a two-parameter function for the Pareto weights, which we assume of the following form:

$$\omega\left(p_i\right)=\mu+p_i(e_i)^{\nu},$$

where we loosely think of ν to relate to the progressivity and μ to the lump-sum transfer. The Pareto weight ω depends on the percentile p_i of the expenditure distribution, to avoid to mechanically increase Pareto weights on the rich as inequality increases. Appendix C.4 reports more details on the Pareto weights.

Figure 2 reports the calibrated (and optimal) marginal and average $t\&T$ rates in 1950. In 1950, average rates are only very modestly negative at the bottom, given the small transfer.

Rising inequality. Starting from the 1950 economy, we first adjust only inequality to 2010 levels and compute the optimal $t\&T$ system. To do so, we modify both the Pareto tail parameter α and the variance of the innovation of the AR(1) process governing the dispersion in productivity, but we keep prices constant at their 1950 level.

As shown in Figure 2, the $t\&T$ system becomes more redistributive when inequality rises. Taxes become more progressive, as marginal rates rise across most of the income distribution and especially so at the top. The government raises more revenues and redistributes through a larger lump-sum transfer, amounting to 4.4% of output. Overall, the 2010 optimal $t\&T$ system provides much more redistribution, with average $t\&T$ rates that go as low as -40% for the poorest households. This result echoes the typical finding in the literature that rising inequality calls for more redistribution.

Rising living standards. The third scenario in Figure 2 accounts for rising living standards due to growth, in addition to rising inequality. To do so, we adjust prices to their 2010 level.

When also accounting for rising living standards, marginal tax rates do increase, as compared to the 1950 $t\&T$ system, but not as much as when only accounting for rising inequality. The optimal lump-sum transfer amounts to only 3.3% of output—again larger than the 0.9% of the 1950 $t\&T$ system, but smaller than the 4.4% obtained when only accounting for rising inequality. In fact, the optimal transfer-to-output ratio comes close to its data counterpart of 3.6%.

Thus, rising inequality is the quantitatively dominant force, but rising living standards reduce the desired increase in the transfer-to-output ratio by more than 30%. Rising living standards also reduce the desired increase in top 10% average rates minus bottom 10% average $t\&T$ rates by around 30%. The average $t\&T$ rate now only amounts to -30% for the poorest households.

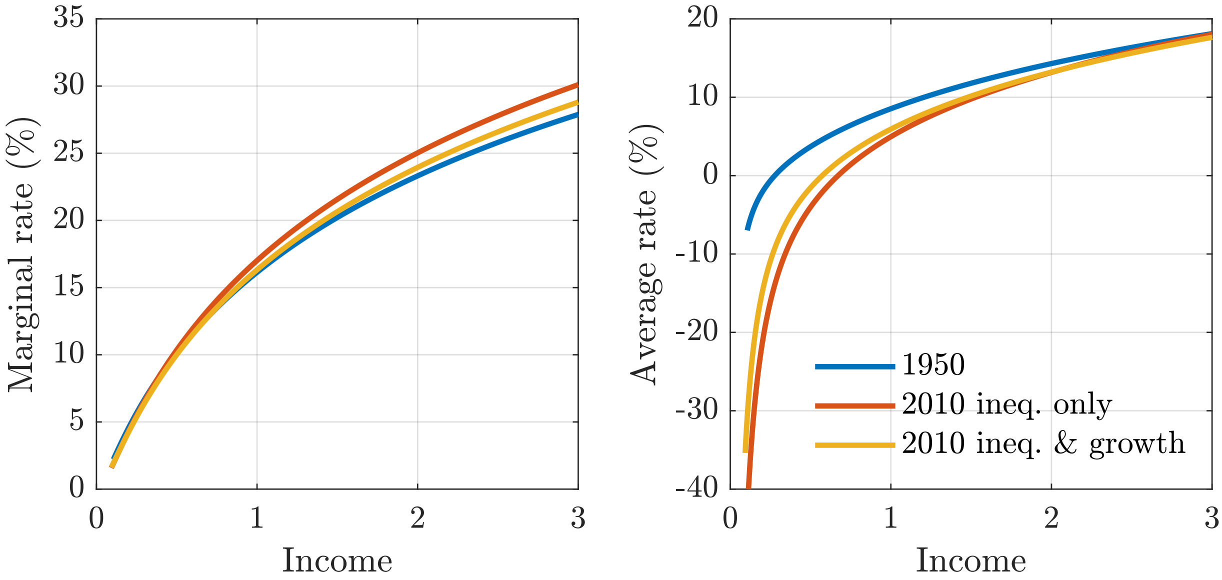

Figure 2: Optimal $t\&T$ Rates in the Dynamic Model

Notes: Figure 2 shows the optimal marginal (left panel) and average (right panel) $t\&T$ rates schedule in the dynamic model for three cases: (1) the 1950 inverse optimum; (2) a counterfactual economy with only the rise in inequality from 1950 to 2010, called "2010 ineq. only"; and (3) the 2010 economy with rising inequality and falling prices, called "2010 ineq. & growth". Income is normalized by mean income.

4.2 Mirrlees Analysis in Static Model

We now turn to the optimal policy analysis in the static Mirrlees model with unrestricted nonlinear income taxes. We follow the same approach as with the dynamic model. We start with finding inverse optimum weights making the 1950 $t\&T$ system optimal. In this framework, we can find a unique set of Pareto weights as a function of productivity, which in this environment captures inequality in earnings potential (Bourguignon and Spadaro 2012; Hendren 2020; Jacobs, Jongen, and Zoutman 2017; Lockwood and Weinzierl 2016). As before, over time we keep these weights constant for the two following scenarios as functions of the position in the distribution.

Figure 3 shows the optimal marginal and average $t\&T$ rates in the static model for the three cases: (1) the 1950 calibrated $t\&T$ function; (2) the optimal 2010 $t\&T$ system when accounting only for rising inequality; and (3) the optimal 2010 $t\&T$ system when also accounting for rising living standards. Results are comparable to those in the dynamic model. In 1950 marginal tax rates are monotonically increasing, as imposed by the calibration. When only accounting for rising inequality, marginal tax rates rise across most of the distribution except at the very bottom. The transfer-to-output ratio, which equates 1.2% in the calibration of the 1950 economy, rises to 6.7%—a slightly larger increase than in the dynamic model. When also accounting for rising living standards,

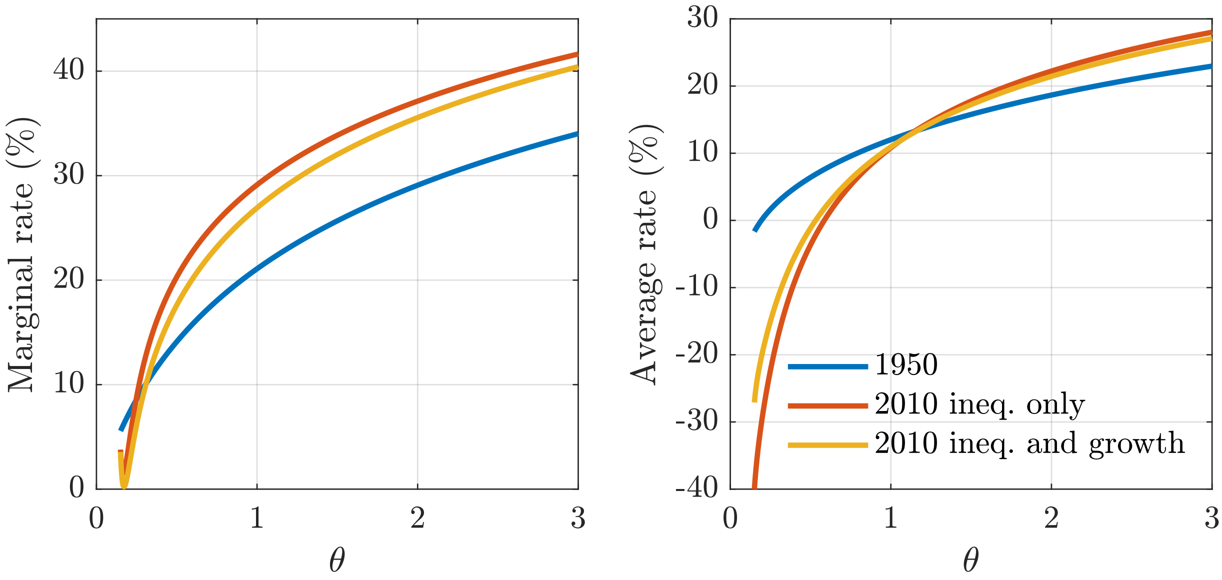

Figure 3: Optimal $t\&T$ Rates in the Static Model

Notes: Figure 3 shows the optimal marginal (left panel) and average (right panel) $t\&T$ rates schedule in the Mirrlees setup for three cases: (1) the 1950 inverse optimum; (2) a counterfactual economy with only the rise in inequality from 1950 to 2010, called "2010 ineq. only"; and (3) the 2010 economy with rising inequality and and falling prices, called "2010 ineq. & growth". Income is normalized by mean income.

the optimal $t\&T$ system provides more redistribution than in 1950, but less so than when only accounting for rising inequality: The transfer-to-output ratio increases to only 4.5%. Hence, by this metric, growth reduces the desired increase in the transfer-to-output ratio by 40%. As in the dynamic model, rising living standards also reduce the increase in top-10% average rates minus bottom-10% average rates by almost 30%. Overall, the rising living standards dampen the desired increase in redistribution due to rising inequality, and this result is robust across the Ramsey and Mirrlees approaches.

We now use the quantified version of the static model to conduct two decompositions.

Decomposition of the growth effect: three channels. The first decomposition builds on the optimal tax formula in Lemma 3 and the comparative statics in Propositions 2-4 to quantify the main drivers of the effect of growth on optimal taxes. For this purpose, we abstract from the rising inequality and keep the distribution of θ fixed. Recall that changes in Λ affect the tax formula through three channels: the *distributional gains* channel (Proposition 2), the *efficiency costs* channel (Proposition 3), and the *income distribution* channel (Proposition 4). The left panel of Figure 4 presents the decomposition.

We start from the 1950 calibrated $t\&T$ system. First, we account for the total effect of growth on optimal taxes—that is, we evaluate the tax formula keeping the θ distribution as in 1950 but using 2010 prices. Marginal taxes fall across the board. Hence, fewer

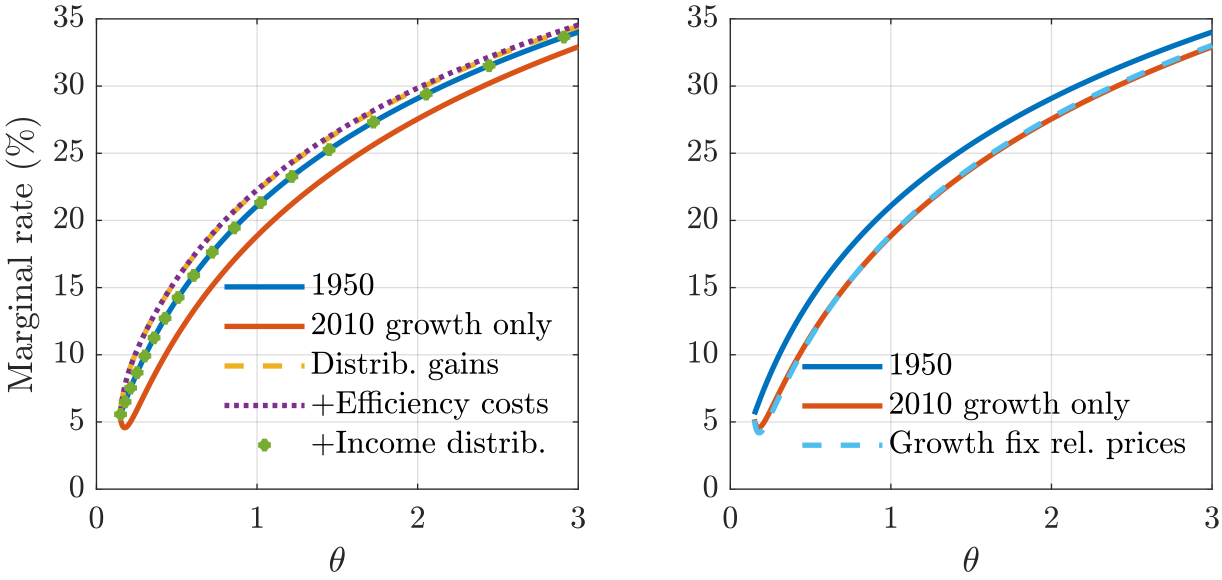

Figure 4: Decompositions of the Effect of Growth in the Static Model

Notes: Figure 4 presents two decompositions of the effects of growth in the Mirrlees setup. Both panels present marginal rates for two cases: (1) the 1950 inverse optimum; (2) a counterfactual economy with only falling prices from 1950 to 2010, called "2010 growth only". The left panel decomposes the effects of growth into three channels: the *distributional gains*, the *efficiency costs*, and the *income distribution* channels. The right panel decomposes the effects of growth into two steps: a homogeneous fall in prices fixing relative prices, followed by an adjustment in marginal prices.

revenues are raised and the small lump-sum transfer turns into a lump-sum tax, with the transfer-to-output ratio decreasing from 1.2% to -0.7%. This result illustrates again that rising living standards call for less redistribution.

We then reverse-engineer this overall effect of growth in three steps. To isolate the *distributional gains* channel, we derive the optimal $t\&T$ system with marginal utilities computed under 1950 prices, but income effects and hours worked computed under the 2010 prices. With distributional gains as in 1950, redistribution increases strongly, with marginal rates increasing by 3 to 5 percentage points across the board; the transfer-to-output ratio rises to 2.4%. To also account for the *efficiency costs* channel, we derive the optimal $t\&T$ system with also income effects computed under the 1950 prices. Theoretically, larger income effects have ambiguous effects. Quantitatively, the opposing effects essentially cancel out. Marginal tax rates and the transfer-to-output ratio barely change relative to the previous scenario. Finally, to also account for the *income distribution* channel, we compute optimal hours using the 1950 prices as well—which retrieves exactly the 1950 calibrated $t\&T$ system. This last step further lowers marginal rates as, given constant skill inequality, income inequality is lower at 1950 than at 2010 prices: the variance of log expenditure decreases moderately, from 0.28 to 0.26. Overall, the first effect dominates quantitatively: rising living standards decrease optimal redistribution mostly due to lower distributional gains.

Relative prices. The second decomposition disentangles the effect of g, which drives the aggregate fall in prices, from the effect of changes in relative prices. Indeed, from 1950 to 2010 all prices fall, but prices in agriculture and manufacturing fall by more than prices in services.

Again, we start from the 1950 calibrated $t\&T$ system, and we first account for the total effect of growth—that is, we evaluate the tax formula keeping the θ distribution as in 1950 but using 2010 prices, accounting for both the aggregate fall in prices and the change in relative prices. We then isolate the effect of the aggregate fall in prices. To do so, we compute a counterfactual where relative prices in 2010 remain as in 1950 but all prices fall to generate the same growth in GDP per capita as in the data. As shown in the right panel of Figure 4, the effect of changes in relative prices is very modest in our setup.

4.3 Robustness

To conclude the analysis, we conduct three robustness exercises. First, we recalibrate the economy assuming a larger degree of RRA. Second, we derive the optimal $t\&T$ system of a utilitarian planner. Third, we replicate the benchmark exercise using the other non-homothetic preferences generally used in the literature, the IA preferences of Alder, Boppart, and Müller (2022). We conduct these exercises in the Mirrlees environment, which is quantitatively more tractable.

Higher risk aversion. The right panel of Figure 5 features marginal $t\&T$ rates in an economy calibrated to higher risk-aversion, with the curvature parameter γ moving from 0.75 to 1.5.[33] Average risk aversion now amounts to 1.37 in 2010, and to 1.6 in 1950 as higher levels of risk aversion also amplify DRRA. This moderate increase in the level of risk aversion amplifies significantly the effects of growth on the optimal $t\&T$ system. From 1.1% of GDP in 1950, the optimal transfers increase to 7.2% in 2010 when only accounting for rising inequality, but only to 1.8% when also accounting for rising living standards. By this metric rising living standards thus reduce the desired increase in redistribution by almost 90%. The difference in average $t\&T$ rates between top-10% and bottom-10% also decreases by almost two-thirds when accounting for rising living standards. Yet, this alternative calibration with higher risk aversion also generates a counterfactually large fall in aggregate labor supply.

[33] Table C.1 summarizes the parameters for this calibration.

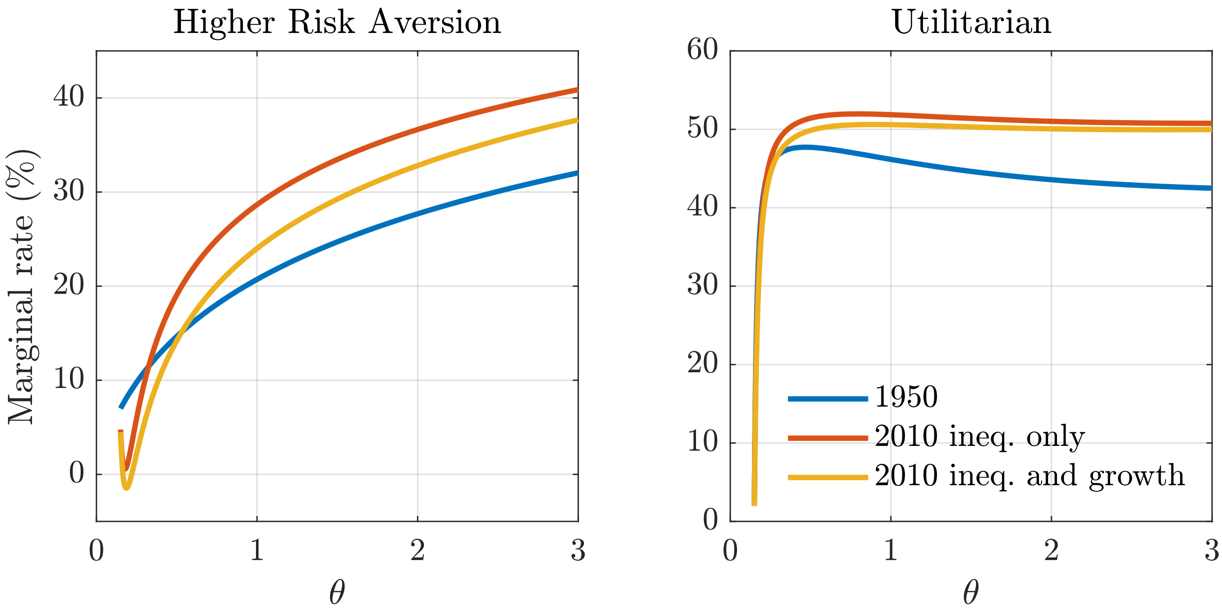

Figure 5: Optimal $t\&T$ Rates in the Static Model: Robustness

Notes: Figure 5 shows the optimal marginal $t\&T$ rates schedule in the static model for three cases: (1) the 1950 economy; (2) a counterfactual economy with only a rise in inequality; and (3) the 2010 economy with rising inequality and growth. The left panel assumes a larger risk aversion; the right panel assumes a Utilitarian planner instead of the inverse-optimum Pareto weights.

Utilitarian planner. The left panel of Figure 5 features marginal $t\&T$ rates for a utilitarian planner. Marginal rates are much higher across all scenarios—a common finding in the literature when assuming a Utilitarian planner (Heathcote and Tsujiyama 2021; Saez 2001). In 1950, the optimal transfer amounts to 25.2% of GDP. With only rising inequality, optimal marginal rates increase and the transfer reaches 29.2% of GDP. Adding rising living standards, optimal marginal rates increase by less and finance a lower transfer, at only 27.6% of GDP. By this metric rising living standards thus reduce the desired increase in redistribution by almost 40%. Using the alternative metric of the difference in average $t\&T$ rates between top-10% and bottom-10%, rising living standards reduce the optimal increase in redistribution by almost 10%.

IA preferences. As a last robustness, we perform the analysis replacing the non-homothetic CES preferences with the IA preferences of Alder, Boppart, and Müller (2022), calibrated to match the same targets as with the non-homothetic CES. We follow the functional form for $\mathbf{D}$ presented in Alder, Boppart, and Müller (2022):

$$\mathbf{D}(p^*) = \frac{\nu}{\eta}\left(\left[\frac{B(p^*)}{\tilde{D}(p^*)}\right]^{\eta} - 1\right), \quad \tilde{D}(p^*) = \left(\sum_{j\in J} \theta_j p_j^{*1-\iota}\right)^{\frac{1}{1-\iota}},$$

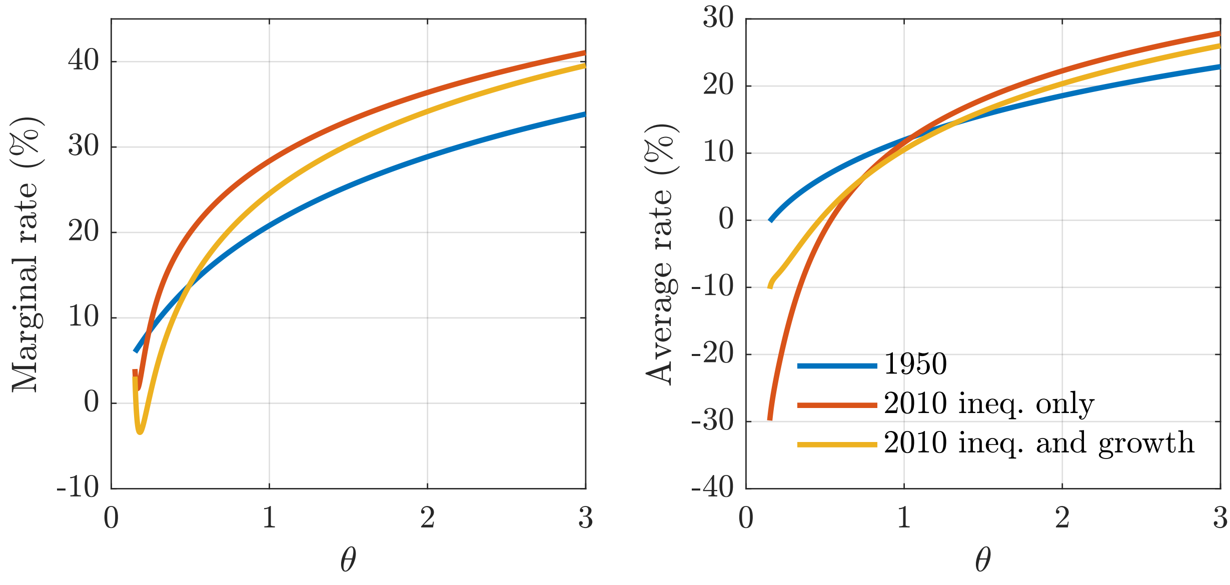

Figure 6: Optimal $t\&T$ Rates in the Static Model: IA Preferences

Notes: Figure 6 shows the optimal marginal and average $t\&T$ rates schedule with IA preferences in the static model for three cases: (1) the 1950 economy; (2) a counterfactual economy with only a rise in inequality; and (3) the 2010 economy with rising inequality and growth.

with $\nu \geq 0$, $\eta \in (0,1)$, $\iota > 0$, $\sum_{j\in J} \theta_j = 1$, and $\theta_j \geq 0 \; \forall j$. Key to the calibration are the levels of $\{\bar{c}_j\}$, which are such that the generalized Stone-Geary term A is positive—a necessary condition for the IA preferences to generate a fall in labor supply, and which also results in DRRA as explained in Lemma 2.[34]

As compared to the benchmark with non-homothetic CES, the effects of growth are larger. The transfer-to-output ratio moves from 1.1% in 1950 to 5.6% when only accounting for rising income inequality, but to only 2.0% when also accounting for growth. Hence, rising living standards reduce the optimal rise in the transfer-to-output ratio by more than 80%. Rising living standards also reduce the optimal increase in the difference in average $t\&T$ rates between top-10% and bottom-10% by close to half.

5 Conclusion

In this paper, we explore the impact of rising living standards on the optimal design of the tax-and-transfer system. We do so by considering growth in the optimal income tax problem with non-homothetic preferences over heterogeneous goods. This allows to examine how households' consumption baskets shift from necessities to luxuries as they become more affluent and how these changes influence optimal redistribution.

[34]The parameters for the IA preferences are the following: $\gamma = 1 - \eta = 0.9$, , $\bar{c}_A = 0.03$, $\bar{c}_G = 0.00$, $\bar{c}_S = 0.005$, $\sigma = 0.001$, $\Omega_A = 0.06$, $\Omega_G = 0.4$, $\nu = 15$, $\iota = 2$, $\theta_A = 0.22$, $\theta_G = 0.62$

We theoretically show that a crucial property of such preferences is that they feature decreasing relative risk aversion (DRRA). This property implies that as living standards rise, equity concerns in the economy unequivocally decrease, while the effects on efficiency concerns remain theoretically ambiguous. Quantitatively, we employ both a Ramsey approach within a dynamic incomplete-market framework and a Mirrlees approach in a static environment. Our analysis reveals that accounting for the impact of rising living standards reduces by at least 25% the desired increase in redistribution attributed to rising inequality in the United States from 1950 to 2010.

References

Aguiar, Mark and Mark Bils (2015). “Has consumption inequality mirrored income inequality?” *The American Economic Review* 105.9, pp. 2725–56.

Ait-Sahalia, Yacine, Jonathan A. Parker, and Motohiro Yogo (2004). “Luxury goods and the equity premium.” *The Journal of Finance* 59.6, pp. 2959–3004.

Aiyagari, S. Rao (1994). “Uninsured idiosyncratic risk and aggregate saving.” *The Quarterly Journal of Economics* 109.3, pp. 659–684.

Alder, Simon, Timo Boppart, and Andreas Müller (2022). “A theory of structural change that can fit the data.” *American Economic Journal: Macroeconomics* 14.2, pp. 160–206.

Aoki, Shuhei and Makoto Nirei (2017). “Zipf’s Law, Pareto’s Law, and the Evolution of Top Incomes in the United States.” *American Economic Journal: Macroeconomics* 9.3, pp. 36–71.

Atkeson, Andrew and Masao Ogaki (1996). “Wealth-varying intertemporal elasticities of substitution: Evidence from panel and aggregate data.” *Journal of Monetary Economics* 38.3, pp. 507–534.

Attanasio, Orazio, Erich Battistin, and Hidehiko Ichimura (2007). “What Really Happened to Consumption Inequality in the United States?” *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches*. University of Chicago Press, pp. 515–543.

Attanasio, Orazio, Erik Hurst, and Luigi Pistaferri (2014). “The evolution of income, consumption, and leisure inequality in the United States, 1980–2010.” *Improving the measurement of consumer expenditures*. University of Chicago Press, pp. 100–140.

Attanasio, Orazio P. and Martin Browning (1995). “Consumption over the Life Cycle and over the Business Cycle.” *The American Economic Review*, pp. 1118–1137.

Attanasio, Orazio P. and Luigi Pistaferri (2014). “Consumption inequality over the last half century: some evidence using the new PSID consumption measure.” *The American Economic Review* 104.5, pp. 122–126.

Ayaz, Mehmet, Lea Fricke, Clemens Fuest, and Dominik Sachs (2023). “Who should bear the burden of COVID-19 related fiscal pressure? An optimal income taxation perspective.” *European Economic Review* 153, p. 104381.

Benabou, Roland (2002). “Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?” *Econometrica* 70.2, pp. 481–517.

Bick, Alexander, Nicola Fuchs-Schündeln, and David Lagakos (2018). "How do hours worked vary with income? Cross-country evidence and implications." *The American Economic Review* 108.1, pp. 170–99.

Blundell, Richard, Martin Browning, and Costas Meghir (1994). "Consumer demand and the life-cycle allocation of household expenditures." *The Review of Economic Studies* 61.1, pp. 57–80.

Bohr, Clement E., Martí Mestieri, and Emre Enes Yavuz (2023). "Aggregation and Closed-Form Results for Nonhomothetic CES Preferences." *Working Paper.*

Boppart, Timo (2014). "Structural change and the Kaldor facts in a growth model with relative price effects and non-Gorman preferences." *Econometrica* 82.6, pp. 2167–2196.

Boppart, Timo and Per Krusell (2020). "Labor supply in the past, present, and future: a balanced-growth perspective." *Journal of Political Economy* 128.1, pp. 118–157.

Bourguignon, François and Amedeo Spadaro (2012). "Tax–benefit revealed social preferences." *The Journal of Economic Inequality* 10.1, pp. 75–108.

Brinca, Pedro, João B. Duarte, Hans Aasnes Holter, and João Henrique Barata Gouveia de Oliveira (2022). "Technological Change and Earnings Inequality in the US: Implications for Optimal Taxation." *Working Paper.*

Browning, Martin and Thomas F. Crossley (2000). "Luxuries are easier to postpone: A proof." *Journal of Political Economy* 108.5, pp. 1022–1026.

Buera, Francisco J., Joseph P. Kaboski, Richard Rogerson, and Juan I. Vizcaino (2022). "Skill-biased structural change." *The Review of Economic Studies* 89.2, pp. 592–625.

Chang, Bo Hyun, Yongsung Chang, and Sun-Bin Kim (2018). "Pareto weights in practice: A quantitative analysis across 32 OECD countries." *Review of Economic Dynamics* 28, pp. 181–204.

Cioffi, Riccardo A. (2021). "Heterogeneous Risk Exposure and the Dynamics of Wealth Inequality." *Working Paper.*

Comin, Diego, Danial Lashkari, and Martí Mestieri (2021). "Structural change with long-run income and price effects." *Econometrica* 89.1, pp. 311–374.

Conesa, Juan Carlos and Dirk Krueger (2006). "On the optimal progressivity of the income tax code." *Journal of Monetary Economics* 53.7, pp. 1425–1450.

Costa, Dora L. (2000). "The Wage and the Length of the Work Day: From the 1890s to 1991." *Journal of Labor Economics* 18.1, pp. 156–181.

Crossley, Thomas F. and Hamish W. Low (2011). “Is the Elasticity of Intertemporal Substitution Constant?” *Journal of the European Economic Association* 9.1, pp. 87–105.

de Magalhaes, Leandro, Enric Martorell, and Raül Santaeulàlia-Llopis (2022). “Progressivity and Development.” *Working Paper.*

Diamond, Peter A. (1998). “Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates.” *The American Economic Review*, pp. 83–95.

Diamond, Peter A. and Emmanuel Saez (2011). “The case for a progressive tax: From basic research to policy recommendation.” *Journal of Economic Perspectives* 25.4, pp. 165–190.

Donovan, Kevin (2021). “The equilibrium impact of agricultural risk on intermediate inputs and aggregate productivity.” *The Review of Economic Studies* 88.5, pp. 2275–2307.

Fagereng, Andreas, Martin B. Holm, and Gisle J. Natvik (2021). “MPC heterogeneity and household balance sheets.” *American Economic Journal: Macroeconomics* 13.4, pp. 1–54.

Feldstein, Martin S. (1969). “The effects of taxation on risk taking.” *Journal of Political Economy* 77.5, pp. 755–764.

Ferriere, Axelle, Philipp Grübener, Gaston Navarro, and Oliko Vardishvili (2023). “On the optimal design of transfers and income tax progressivity.” *Journal of Political Economy Macroeconomics* 1.2, pp. 276–333.

Ferriere, Axelle and Gaston Navarro (2023). “The Heterogeneous Effects of Government Spending: It’s All About Taxes.” *Working Paper.*

Golosov, Mikhail, Michael Graber, Magne Mogstad, and David Novgorodsky (2023). “How Americans respond to idiosyncratic and exogenous changes in household wealth and unearned income.” *Forthcoming in the Quarterly Journal of Economics.*

Guner, Nezih, Remzi Kaygusuz, and Gustavo Ventura (2023). “Rethinking the welfare state.” *Econometrica* 91.6, pp. 2261–2294.

Hanoch, Giora (1975). “Production and demand models with direct or indirect implicit additivity.” *Econometrica* 43.3, pp. 395–419.

— (1977). “Risk aversion and consumer preferences.” *Econometrica*, pp. 413–426.

Heathcote, Jonathan, Fabrizio Perri, and Giovanni L. Violante (2010). “Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006.” *Review of Economic Dynamics* 13.1, pp. 15–51.

Heathcote, Jonathan, Fabrizio Perri, Giovanni L. Violante, and Lichen Zhang (2023). "More unequal we stand? Inequality dynamics in the United States, 1967–2021." *Review of Economic Dynamics* 50, pp. 235–266.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante (2017). "Optimal tax progressivity: An analytical framework." *The Quarterly Journal of Economics* 132.4, pp. 1693–1754.

— (2020). "Presidential Address 2019: How Should Tax Progressivity Respond to Rising Income Inequality?" *Journal of the European Economic Association* 18.6, pp. 2715–2754.

Heathcote, Jonathan and Hitoshi Tsujiyama (2021). "Optimal income taxation: Mirrlees meets Ramsey." *Journal of Political Economy* 129.11, pp. 3141–3184.

Hendren, Nathaniel (2020). "Measuring economic efficiency using inverse-optimum weights." *Journal of Public Economics* 187, p. 104198.

Herrendorf, Berthold, Richard Rogerson, and Akos Valentinyi (2013). "Two perspectives on preferences and structural transformation." *The American Economic Review* 103.7, pp. 2752–2789.

— (2014). "Growth and structural transformation." *Handbook of Economic Growth*. Vol. 2. Elsevier, pp. 855–941.

Holter, Hans A., Dirk Krueger, and Serhiy Stepanchuk (2019). "How do tax progressivity and household heterogeneity affect Laffer curves?" *Quantitative Economics* 10.4, pp. 1317–1356.

Huggett, Mark (1993). "The risk-free rate in heterogeneous-agent incomplete-insurance economies." *Journal of Economic Dynamics and Control* 17.5, pp. 953–969.

Imrohoroğlu, Ayşe (1989). "Cost of business cycles with indivisibilities and liquidity constraints." *Journal of Political Economy* 97.6, pp. 1364–1383.

Jacobs, Bas, Egbert L. W. Jongen, and Floris T. Zoutman (2017). "Revealed social preferences of Dutch political parties." *Journal of Public Economics* 156, pp. 81–100.

Jaravel, Xavier and Alan Olivi (2022). "Prices, Non-homotheticities, and Optimal Taxation." *Working Paper*.

Johnson, David S., Jonathan A. Parker, and Nicholas S. Souleles (2006). "Household expenditure and the income tax rebates of 2001." *The American Economic Review* 96.5, pp. 1589–1610.

Kaplan, Greg and Giovanni L. Violante (2022). "The marginal propensity to consume in heterogeneous agent models." *Annual Review of Economics* 14, pp. 747–775.

Krueger, Dirk and Fabrizio Perri (2006). “Does income inequality lead to consumption inequality? Evidence and theory.” *The Review of Economic Studies* 73.1, pp. 163–193.

Kuhn, Moritz, Moritz Schularick, and Ulrike I. Steins (2020). “Income and wealth inequality in America, 1949–2016.” *Journal of Political Economy* 128.9, pp. 3469–3519.

Le Grand, François and Xavier Ragot (2023). “Optimal Fiscal Policy With Heterogeneous Agents and Capital: Should We Increase or Decrease Public Debt and Capital Taxes?” *Working Paper*.

Lockwood, Benjamin B. and Matthew Weinzierl (2016). “Positive and normative judgments implicit in US tax policy, and the costs of unequal growth and recessions.” *Journal of Monetary Economics* 77, pp. 30–47.

Mankiw, N. Gregory, Matthew Weinzierl, and Danny Yagan (2009). “Optimal taxation in theory and practice.” *Journal of Economic Perspectives* 23.4, pp. 147–74.

Mantovani, Cristiano (2023). “Hours-Biased Technological Change.” *Working Paper*.

McGrattan, Ellen R. and Richard Rogerson (2004). “Changes in hours worked, 1950-2000.” *Federal Reserve Bank of Minneapolis Quarterly Review* 28.1, pp. 14–33.

Meeuwis, Maarten (2022). “Wealth fluctuations and risk preferences: Evidence from US investor portfolios.” *Working Paper*.

Mertens, Karel and José Luis Montiel Olea (2018). “Marginal tax rates and income: New time series evidence.” *The Quarterly Journal of Economics* 133.4, pp. 1803–1884.

Meyer, Bruce D. and James X. Sullivan (2023). “Consumption and Income Inequality in the United States since the 1960s.” *Journal of Political Economy* 131.2, pp. 247–284.

Mirrlees, James A. (1971). “An exploration in the theory of optimum income taxation.” *The Review of Economic Studies* 38.2, pp. 175–208.

Ogaki, Masao and Qiang Zhang (2001). “Decreasing relative risk aversion and tests of risk sharing.” *Econometrica* 69.2, pp. 515–526.

Ohanian, Lee, Andrea Raffo, and Richard Rogerson (2008). “Long-term changes in labor supply and taxes: Evidence from OECD countries, 1956–2004.” *Journal of Monetary Economics* 55.8, pp. 1353–1362.

Oni, Mehedi Hasan (2023). “Progressive income taxation and consumption baskets of rich and poor.” *Journal of Economic Dynamics and Control* 157, p. 104758.

Piketty, Thomas and Emmanuel Saez (2003). “Income inequality in the United States, 1913–1998.” *The Quarterly Journal of Economics* 118.1, pp. 1–41.

Piketty, Thomas and Gabriel Zucman (2014). “Capital is back: Wealth-income ratios in rich countries 1700–2010.” *The Quarterly Journal of Economics* 129.3, pp. 1255–1310.

Ramey, Valerie A. and Neville Francis (2009). "A century of work and leisure." *American Economic Journal: Macroeconomics* 1.2, pp. 189–224.

Restuccia, Diego and Guillaume Vandenbroucke (2013). "A century of human capital and hours." *Economic Inquiry* 51.3, pp. 1849–1866.

Saez, Emmanuel (2001). "Using elasticities to derive optimal income tax rates." *The Review of Economic Studies* 68.1, pp. 205–229.

Splinter, David (2020). "US Tax Progressivity and Redistribution." *National Tax Journal* 73.4, pp. 1005–1024.

Stiglitz, Joseph E. (1969). "Behavior towards risk with many commodities." *Econometrica*, pp. 660–667.

Straub, Ludwig (2019). "Consumption, Savings, and the Distribution of Permanent Income." *Working Paper*.

Toda, Alexis Akira and Kieran Walsh (2015). "The double power law in consumption and implications for testing Euler equations." *Journal of Political Economy* 123.5, pp. 1177–1200.

Tsujiyama, Hitoshi (2022). "Optimal Taxation Along the Development Spectrum." *Working Paper*.

Wachter, Jessica A. and Motohiro Yogo (2010). "Why do household portfolio shares rise in wealth?" *The Review of Financial Studies* 23.11, pp. 3929–3965.

Zhang, Qiang and Masao Ogaki (2004). "Decreasing relative risk aversion, risk sharing, and the permanent income hypothesis." *Journal of Business & Economic Statistics* 22.4, pp. 421–430.

A Theory

A.1 Heterogenous Expenditure Elasticities

A.1.1 Non-homothetic CES Preferences

We abuse notation for the following proofs and use $\mathcal{C}(e) = \mathcal{C}(e; \Lambda, p)$.

Risk aversion. Differentiating the expenditure function (2), denoting $\Omega_j^* \equiv \left(p_j^*\right)^{1-\sigma} \Omega_j$, one obtains

$$\mathcal{C}_e(e) = (1-\sigma)e^{-\sigma} \left(\sum_j \Omega_j^* \varepsilon_j \mathcal{C}(e)^{\varepsilon_j - 1} \right)^{-1} \tag{13}$$

$$\mathcal{C}_{ee}(e) = -\frac{\mathcal{C}_e(e)}{e} \left(\sigma + \mathcal{C}_e(e) e \frac{\sum_j \Omega_j^* \varepsilon_j (\varepsilon_j - 1) \mathcal{C}(e)^{\varepsilon_j - 2}}{\sum_j \Omega_j^* \varepsilon_j \mathcal{C}(e)^{\varepsilon_j - 1}} \right)$$

where $\mathcal{C}_e(e) > 0 \; \forall e$. Rearranging, this yields to risk aversion equal to

$$\gamma(e) = \sigma + (1-\sigma) \frac{1}{\chi(\mathcal{C}(e))} \left(\gamma - 1 + \zeta(\mathcal{C}(e)) \right),$$

where

$$\chi(C) \equiv \frac{\sum_j \Omega_j^* \varepsilon_j C^{\varepsilon_j}}{\sum_j \Omega_j^* C^{\varepsilon_j}} \quad \text{and} \quad \zeta(C) \equiv \frac{\sum_j \Omega_j^* \varepsilon_j^2 C^{\varepsilon_j}}{\sum_j \Omega_j^* \varepsilon_j C^{\varepsilon_j}}.$$

We characterize long-run risk aversion as the risk aversion level when $e \to \infty$, that is, when $C \to \infty$. Notice that $\lim_{C \to \infty} \chi(C) = \lim_{C \to \infty} \chi(C) = \varepsilon_J$. Thus, we get

$$\bar{\gamma} = \sigma + (1-\sigma) \frac{1}{\varepsilon_J} \left(\gamma - 1 + \varepsilon_J \right) = (1-\sigma)(\gamma - 1) \frac{1}{\varepsilon_J}. \tag{14}$$

Proof. (Lemma 1) We prove inequality (5). Using equation (13), one can rewrite the first term in the risk aversion formula (4) as

$$\gamma \frac{\mathcal{C}_e(e) e}{\mathcal{C}(e)} = \gamma (1-\sigma) \frac{\sum_j \Omega_j^* \mathcal{C}(e)^{\varepsilon_j}}{\sum_j \Omega_j^* \varepsilon_j \mathcal{C}(e)^{\varepsilon_j}}. \tag{15}$$

As $\mathcal{C}(e)$ is strictly increasing in e, and $\gamma(1-\sigma) > 0$, we only need to show that the fraction in (15) is decreasing in C, or, equivalently, that the inverse of the fraction in (15) is increasing in C:

$$\frac{\partial}{\partial C} \left[\frac{\sum_j \Omega_j^* \varepsilon_j C^{\varepsilon_j}}{\sum_j \Omega_j^* C^{\varepsilon_j}} \right] = \left(\sum_j \Omega_j^* C^{\varepsilon_j} \right)^{-2} \frac{1}{C} \frac{1}{2} \sum_k \sum_j \Omega_k^* \Omega_j^* (\varepsilon_k - \varepsilon_j)^2 C^{\varepsilon_k + \varepsilon_j} > 0,$$

which completes the proof.

□

Proof. (Corollary 2) We start with $J = 2$. We assume $\varepsilon_1 = \varepsilon < 1$ and $\varepsilon_2 = 1$ w.l.o.g. as well as $\Omega_j^* = 1\ \forall j$. Abusing notation, we characterize risk aversion as a function of C, which we have shown is increasing in e:

$$\omega(C) = \sigma + (1-\sigma)\frac{\Omega C^{\varepsilon-1}+1}{\Omega\varepsilon C^{\varepsilon-1}+1}\left(\gamma + (\varepsilon-1)\frac{\Omega\varepsilon C^{\varepsilon-1}}{\Omega\varepsilon C^{\varepsilon-1}+1}\right).$$

We define $y \equiv C^{\varepsilon-1}$. Note that y is a decreasing function of C as $\varepsilon < 1$, thus a decreasing function of e. Therefore, to prove that $\omega'(e) < 0$, we need to show that $f'(y) > 0$, where $f(y)$ is defined as

$$f(y) \equiv \frac{y+1}{\varepsilon y+1}\left(\gamma - (1-\varepsilon)\frac{\varepsilon y}{\varepsilon y+1}\right) \quad \forall y > 0.$$

Some algebra yields

$$f'(y) = \frac{1-\varepsilon}{(\varepsilon y+1)^3}\left[\gamma(\varepsilon y+1) - 2\varepsilon y + \varepsilon^2 y - \varepsilon\right]. \tag{16}$$

The fraction in (16) is strictly positive, so we have DRRA as long as

$$g(y) \equiv (\gamma - 2 + \varepsilon)\varepsilon y + \gamma - \varepsilon > 0, \quad \text{where } y > 0.$$

Thus, $g(y) > 0\ \forall y$ when $\gamma > 2$, while $g(y) < 0\ \forall y$ when $\gamma < \varepsilon$, which completes the proof.

We now turn to the case with a continuum of goods. Bohr, Mestieri, and Yavuz (2023) make the following set of assumptions:

1. The price parameters $\{p_i^*\}_{i\in[0,1]}$ and taste parameters $\{\Omega_i\}_{i\in[0,1]}$ have a log-linear relationship with $\{\varepsilon_i\}_{i\in[0,1]}$, with a regularity condition regarding the intercept.
2. $\{\varepsilon_i\}_{i\in[0,1]}$ follow a gamma distribution: $\varepsilon_i \sim Gamma(\alpha, \beta)$, with $\alpha > 0$ and $\beta > 0$.

Then, they obtain a closed-form relationship between e and $\mathcal{C}(e)$ as shown in equation (7) of their paper:

$$\log \mathcal{C}(e) = \hat{Y} - \frac{\Psi}{1-\sigma}e^{-\frac{1-\sigma}{\alpha}},$$

where $\hat{Y}$ and Ψ are positive scalars which depend on distribution parameters. As such, we obtain the following closed forms for the derivatives:

$$\mathcal{C}_e(e) = \frac{\Psi}{\alpha}e^{-\frac{1-\sigma}{\alpha}-1}\mathcal{C}(e), \quad \mathcal{C}_{ee}(e) = \frac{\Psi}{\alpha}e^{-\frac{1-\sigma}{\alpha}-1}\mathcal{C}_e(e) - \left(\frac{1-\sigma}{\alpha}+1\right)\frac{\mathcal{C}_e(e)}{e}.$$

Thus, we can express risk-aversion as, abstraction from Λ and p:

$$\begin{aligned}\gamma(e) &= \underbrace{\gamma \times \frac{\Psi}{\alpha} e^{-\frac{1-\sigma}{\alpha}}}_{\text{first term}} + \underbrace{-\frac{\Psi}{\alpha} e^{-\frac{1-\sigma}{\alpha}} + \left(\frac{1-\sigma}{\alpha} + 1\right)}_{\text{second term}} \\ &= (\gamma - 1)\frac{\Psi}{\alpha} e^{-\frac{1-\sigma}{\alpha}} + \left(\frac{1-\sigma}{\alpha} + 1\right).\end{aligned}$$

As $\sigma < 1$, it follows that $\gamma'(e) < 0$ iff $\gamma > 1$.

□

Rescaling Comin, Lashkari, and Mestieri (2021) show that all ε_j can be multiplied by a positive scalar without implications on intratemporal consumption allocations. We show next that the rescaling irrelevance extends to (i) risk aversion, (ii) labor supply, when γ and B are rescaled appropriately.

When multiplying all ε_j by a scalar ι, one needs to rescale $1-\gamma$ by that same scalar—that is, $\gamma_\iota \equiv 1 - \iota(1-\gamma)$, where x_ι defines the rescaled version of variable x. Rescaling both ϵ_J and $(1-\gamma)$ will leave long-run risk aversion unchanged in equation (14). More generally, the expenditure function defines a new $\mathcal{C}_\iota(e)$ which appears in both numerators and denominators of χ and ζ. Thus, we have $\chi_\iota = \iota\chi$ and $\zeta_\iota = \iota\zeta$, and risk aversion becomes

$$\begin{aligned}\gamma_\iota(e) &= \sigma + (1-\sigma)\frac{1}{\iota\chi(\mathcal{C}(e))}\left(\gamma_\iota - 1 + \iota\zeta(\mathcal{C}(e))\right) \\ &= \sigma + (1-\sigma)\frac{1}{\iota\chi(\mathcal{C}(e))}\left(\iota(\gamma-1) + \iota\zeta(\mathcal{C}(e))\right) = \gamma(e).\end{aligned}$$

Rescaling the curvature parameter γ as defined above leaves risk aversion unchanged.

We turn to labor supply. When multiplying all ε_j by a scalar ι, one needs to rescale the labor disutility parameter such that $B_\iota \equiv B/\iota$. Abstracting from taxes w.l.o.g. the first-order condition reads

$$\mathcal{C}_\iota\left(e\right)^{-\gamma_\iota} \mathcal{C}_{e\iota}\left(e\right) = B_\iota e^{\varphi}.$$

Using $\mathcal{C}_\iota\left(e\right) = \mathcal{C}\left(e\right)^{1/\iota}$ and $\mathcal{C}_{e\iota}\left(e\right) = (1/\iota)\,\mathcal{C}\left(e\right)^{1/\iota - 1}\mathcal{C}_{e\iota}\left(e\right)$ we obtain

$$\mathcal{C}_\iota\left(e\right)^{-\gamma_\iota} \mathcal{C}_{e\iota}\left(e\right) = \mathcal{C}^{-\frac{1}{\iota}[\iota(\gamma-1)+1]}\frac{1}{\iota}\mathcal{C}\left(e\right)^{\frac{1}{\iota}-1}\mathcal{C}_e\left(e\right) = \mathcal{C}\left(e\right)^{-\gamma}\mathcal{C}_e\left(e\right)\frac{1}{\iota},$$

such that the first order condition coincides after rescaling the disutility parameter.

A.1.2 IA Preferences

Proof. (Lemma 2) Differentiating the IA indirect utility function (3) yields

$$u_e\left(e;p,\Lambda\right)=\left(\frac{1}{\mathbf{B}(p^*)}\right)^{1-\gamma}\left(e-\mathbf{A}\left(p^*\right)\right)^{-\gamma},\quad u_{ee}\left(e;p,\Lambda\right)=-\gamma\left(\frac{1}{\mathbf{B}(p^*)}\right)^{1-\gamma}\left(e-\mathbf{A}\left(p^*\right)\right)^{-\gamma-1}.$$

The coefficient of relative risk aversion follows as

$$\gamma\left(e;p,\Lambda\right)=\gamma\frac{e}{e-\mathbf{A}\left(p^*\right)},$$

and thus

$$\frac{\partial\gamma\left(e;p,\Lambda\right)}{\partial e}=-\gamma\frac{\mathbf{A}\left(p^*\right)}{\left[\left(e-\mathbf{A}\left(p^*\right)\right)\right]^2},$$

which is negative for $\mathbf{A}\left(p^*\right)>0$. □

A.2 Optimal Income Taxes

A.2.1 Household Behavior Given Taxes

In this part of the appendix, we describe the household behavior given taxes and derive some results that are used in the proofs that follow afterwards. Recall the household problem:

$$V(\theta;\mathcal{T}(\cdot),\Lambda,p)\equiv\max_{e,n}u(e;\Lambda,p)-Bn^{1+\varphi}/(1+\varphi)\quad\text{s.t.}\quad e=n\theta-\mathcal{T}\left(n\theta\right)\qquad\text{(Step 1)}$$

$$u(e;\Lambda,p)\equiv\max_{\{c_j\}_j}\;U(c)\quad\text{s.t.}\quad\sum_j\frac{p_j}{\Lambda}c_j=e.\qquad\text{(Step 2)}$$

In the following, we suppress dependence of variables on the price vector p for simplicity.

Relation between u_e and u_Λ. The Lagrangian associated with problem (Step 2) is

$$\mathcal{L}=U(c)+\mu\left(e-\sum_j\frac{p_j}{\Lambda}c_j\right).$$

Application of the envelope theorem yields (omitting arguments)

$$u_e=\mu\quad\text{and}\quad u_\Lambda=\mu\sum_j\frac{p_j}{\Lambda^2}c_j=\mu\frac{e}{\Lambda},$$

which together implies

$$u_e=\frac{u_\Lambda\Lambda}{e}.$$

Since this relation holds for each e, we can take derivatives, which yields

$$u_{ee} = \frac{u_{\Lambda e}\Lambda e - u_\Lambda \Lambda}{e^2} = \Lambda \left(\frac{u_{\Lambda e}}{e} - \frac{u_\Lambda}{e^2} \right).$$

Hence

$$u_{e\Lambda} = \frac{e}{\Lambda} u_{ee} + u_\Lambda \frac{1}{e} = \frac{e}{\Lambda} u_{ee} + u_e \frac{1}{\Lambda}. \tag{17}$$

Labor supply decision. The first-order condition (FOC) of (Step 1) reads as:

$$-Bn^\varphi + u_e(e;\Lambda)(1-\mathcal{T}')\theta = 0.$$

The second-order condition (SOC) is

$$-B\varphi n^{\varphi-1} + u_{ee}(e;\Lambda)\left((1-\mathcal{T}')\theta\right)^2 - u_e(e;\Lambda)\mathcal{T}''\theta^2 < 0, \tag{18}$$

which can be written as

$$-B\varphi n^\varphi \frac{1}{\theta\,(1-\mathcal{T}')} + u_{ee}(e;\Lambda)\left((1-\mathcal{T}')\theta\right) n - u_e(e;\Lambda)\frac{\mathcal{T}''}{1-\mathcal{T}'} n\theta < 0.$$

Now divide by u_e and use the FOC to rewrite the first term and obtain:

$$-\varphi + \frac{u_{ee}(e;\Lambda)}{u_e(e;\Lambda)}\left((1-\mathcal{T}')n\theta\right) - \rho(n\theta) < 0,$$

where $\rho(y) = -\frac{d\log(1-T'(y))}{d\log(y)} = \frac{T''(y)y}{(1-T'(y))}$. The SOC is fulfilled if the tax function is not too concave:

$$\rho(n\theta) > -\varphi - \gamma(e;\Lambda)\frac{(1-\mathcal{T}')n\theta}{e}.$$

Wealth effect on labor supply. First, we derive the responsiveness to an increase in the lump-sum element of the tax $\mathcal{T}(0;\Lambda)$. Implicit differentiation yields:

$$\frac{\partial n}{\partial \mathcal{T}(0;\Lambda)} = -\frac{-u_{ee}(e;\Lambda)(1-\mathcal{T}')\theta}{-\varphi B n^{\varphi-1} + u_{ee}(e;\Lambda)(1-\mathcal{T}')^2\theta^2 - u_e(e;\Lambda)\mathcal{T}''\theta^2},$$

which can be written as (using the FOC $u_e = \frac{Bn^\varphi}{(1-\mathcal{T}')\theta}$):

$$\frac{\partial n}{\partial \mathcal{T}(0;\Lambda)} = -\frac{-\frac{u_{ee}(e;\Lambda)}{u_e(e;\Lambda)}(1-\mathcal{T}')\theta}{-\varphi B n^{\varphi-1}\frac{(1-\mathcal{T}')\theta}{Bn^\varphi} + \frac{u_{ee}(e;\Lambda)}{u_e(e;\Lambda)}(1-\mathcal{T}')^2\theta^2 - \mathcal{T}''\theta^2}$$

and hence

$$\frac{\partial n}{\partial \mathcal{T}(0;\Lambda)} = -\frac{\gamma(e;\Lambda)\frac{1}{e}(1-\mathcal{T}')\theta}{-\varphi\frac{(1-\mathcal{T}')\theta}{n} - \gamma(e;\Lambda)\frac{1}{e}(1-\mathcal{T}')^2\theta^2 - \mathcal{T}''\theta^2}.$$

Some further steps reveal that

$$\frac{\partial n}{\partial \mathcal{T}(0;\Lambda)} = \frac{1}{\varphi}\frac{\gamma(e;\Lambda)\frac{n}{e}}{1+\frac{\gamma(e;\Lambda)}{\varphi}\frac{n}{e}(1-\mathcal{T}')\theta+\frac{1}{\varphi}\frac{\mathcal{T}''}{1-\mathcal{T}'}\theta n}$$

Finally, we define (and use that $\rho(y) = \frac{\mathcal{T}''}{1-\mathcal{T}'}y$) to obtain:

$$\eta(\theta;\mathcal{T},\Lambda,p) \equiv -\theta\frac{\partial n(\theta;\mathcal{T},\Lambda,p)}{\partial \mathcal{T}(0;\Lambda)} = -\frac{1}{\varphi}\times\frac{\gamma(e;\Lambda)\frac{y}{e}}{1+\frac{\gamma(e;\Lambda)}{\varphi}\frac{y}{e}(1-\mathcal{T}')+\frac{\rho(y)}{\varphi}}.$$

Elasticity w.r.t. type. Implicit differentiation of the individual FOC yields:

$$\begin{aligned}
\frac{\partial n}{\partial \theta} &= -\frac{u_{ee}(e;\Lambda)(1-\mathcal{T}')^2\theta n + u_e(e;\Lambda)(1-\mathcal{T}') - u_e(e;\Lambda)\mathcal{T}''\theta n}{-\varphi B n^{\varphi-1} + u_{ee}(e;\Lambda)(1-\mathcal{T}')^2\theta^2 - u_e(e;\Lambda)\mathcal{T}''\theta^2}\\
&= -\frac{\frac{u_{ee}(e;\Lambda)}{u_e(e;\Lambda)}(1-\mathcal{T}')\theta n + 1 - \frac{\mathcal{T}''}{1-\mathcal{T}'}\theta n}{-\varphi\frac{\theta}{n}+\frac{u_{ee}(e;\Lambda)}{u_e(e;\Lambda)}(1-\mathcal{T}')\theta^2-\frac{\mathcal{T}''}{1-\mathcal{T}'}\theta^2}\\
&= \frac{1}{\varphi}\frac{-\gamma(e;\Lambda)\frac{(1-\mathcal{T}')\theta n}{e}+1-\frac{\mathcal{T}''}{1-\mathcal{T}'}y}{\frac{\theta}{n}+\frac{\gamma(e;\Lambda)}{\varphi}\frac{(1-\mathcal{T}')\theta^2}{e}+\frac{1}{\varphi}\frac{\mathcal{T}''}{1-\mathcal{T}'}\theta^2}.
\end{aligned}$$

The elasticity follows as

$$\begin{aligned}
\varepsilon_{n,\theta} &= \frac{1}{\varphi}\frac{\frac{\theta}{n}\left(1-\gamma(e;\Lambda)\frac{(1-\mathcal{T}')\theta n}{e}-\rho(y)\right)}{\frac{\theta}{n}+\frac{\gamma(e;\Lambda)}{\varphi}\frac{(1-\mathcal{T}')\theta^2}{e}+\frac{1}{\varphi}\frac{\mathcal{T}''}{1-\mathcal{T}'}\theta^2}\\
&= \frac{1}{\varphi}\frac{\left(1-\gamma(e;\Lambda)\frac{(1-\mathcal{T}')\theta n}{e}-\rho(y)\right)}{1+\frac{\gamma(e;\Lambda)}{\varphi}\frac{(1-\mathcal{T}')\theta n}{e}+\frac{\rho(y)}{\varphi}}.
\end{aligned}$$

Elasticity w.r.t. to $1-T'$. Implicit differentiation of the individual FOC yields:

$$\begin{aligned}
\frac{\partial n}{\partial(1-\mathcal{T}')} &= -\frac{u_e(e;\Lambda)\theta}{-\varphi B n^{\varphi-1}+u_{ee}(e;\Lambda)(1-\mathcal{T}')^2\theta^2-u_e(e;\Lambda)\mathcal{T}''\theta^2}\\
&= \frac{1}{\varphi}\frac{1}{\frac{1-\mathcal{T}'}{n}-\frac{u_{ee}(e;\Lambda)}{u_e}(1-\mathcal{T}')^2\theta\frac{1}{\varphi}+\frac{1}{\varphi}\mathcal{T}''\theta}\\
&= \frac{1}{\varphi}\times\frac{1}{\frac{1-\mathcal{T}'}{n}+\frac{\gamma(e;\Lambda)}{\varphi}\frac{(1-\mathcal{T}')^2\theta}{e}+\frac{1}{\varphi}\mathcal{T}''\theta}
\end{aligned}$$

The elasticity follows as

$$\varepsilon_{n,1-\mathcal{T}'} = \frac{1}{\varphi}\times\frac{1}{1+\frac{\gamma(e;\Lambda)}{\varphi}\frac{(1-\mathcal{T}')\theta n}{e}+\frac{\rho(y)}{\varphi}}.$$

Ratio of $\frac{\varepsilon_{y,1-\mathcal{T}'}}{\varepsilon_{y,\theta}}$. First of all note that

$$\varepsilon_{y,1-\mathcal{T}'} = \varepsilon_{n,1-\mathcal{T}'} \quad \text{and} \quad \varepsilon_{y,\theta} = 1 + \varepsilon_{n,\theta}.$$

Intuitively, y depends directly on θ and indirectly through n. The ratio $\frac{\varepsilon_{y,1-\mathcal{T}'}}{\varepsilon_{y,\theta}}$ is then given by:

$$\begin{aligned}
\frac{\varepsilon_{y,1-\mathcal{T}'}}{\varepsilon_{y,\theta}} &= \frac{\frac{1}{\varphi}\frac{1}{1+\frac{\gamma(e;\Lambda)}{\varphi}\frac{(1-\mathcal{T}')\theta n}{e}+\frac{\rho(y)}{\varphi}}}{1+\frac{1}{\varphi}\frac{\left(1-\gamma(e;\Lambda)\frac{(1-\mathcal{T}')\theta n}{e}-\rho(y)\right)}{1+\frac{\gamma(e;\Lambda)}{\varphi}\frac{(1-\mathcal{T}')\theta n}{e}+\frac{\rho(y)}{\varphi}}} \\
&= \frac{1}{\varphi}\frac{1}{1+\frac{\gamma(e;\Lambda)}{\varphi}\frac{(1-\mathcal{T}')\theta n}{e}+\frac{\rho(y)}{\varphi}+\frac{1}{\varphi}\left(1-\gamma(e;\Lambda)\frac{(1-\mathcal{T}')\theta n}{e}-\rho(y)\right)} = \frac{1}{1+\varphi}.
\end{aligned} \tag{19}$$

A.2.2 Optimal Income Tax Formula

Proof. (Lemma 3) The Lagrangian of the government's problem (6) is:

$$\begin{aligned}
\mathcal{L} = &\int_{\underline{\theta}}^{\overline{\theta}} \left(u\left[n(\theta;\mathcal{T}(\cdot,\Lambda),\Lambda,p)\theta - \mathcal{T}\left(n(\theta;\mathcal{T}(\cdot,\Lambda),\Lambda,p)\theta\right);\Lambda,p\right] - Bn^{1+\varphi}/(1+\varphi)\right)\tilde{f}(\theta)d\theta \\
&+ \lambda\left(\int_{\underline{\theta}}^{\overline{\theta}} \mathcal{T}(n(\theta;\mathcal{T}(\cdot,\Lambda),\Lambda,p)\theta)f(\theta)d\theta - G\right),
\end{aligned}$$

where $\tilde{f}(\theta) = f(\theta)w(\theta)$.

We follow the heuristic approach going back to Saez (2001) for deriving the optimality condition for the marginal tax rate. Consider an increase in the marginal tax rate by $d\mathcal{T}'$ within a small interval $[(y(\theta^*;\mathcal{T}(\cdot,\Lambda),\Lambda), y(\theta^*;\mathcal{T}(\cdot,\Lambda),\Lambda)+dy]$. The mass of people affected by this increase in the marginal tax rate is approximately given by:

$$h(y(\theta^*;\mathcal{T},\Lambda)) \times dy.$$

Note that the density function h is endogenous w.r.t. $(\mathcal{T},\Lambda)$, which we suppress for ease of notation. To express this mass in terms of primitives, note that

$$F(\theta^*) = H(y(\theta^*;\mathcal{T},\Lambda)) \quad \text{and hence} \quad f(\theta^*) = h(y(\theta^*;\mathcal{T},\Lambda))y_\theta(\theta^*;\mathcal{T},\Lambda),$$

which implies that the mass of households that face an increase in the marginal tax rate is given by

$$h(y(\theta^*;\mathcal{T},\Lambda)) \times dy = \frac{f(\theta^*)dy}{y_\theta(\theta^*;\mathcal{T},\Lambda)}.$$

Note that each person affected by the increase in the marginal tax rate changes their earnings by (for simplicy we denote $A(\theta)$ instead of $A(\theta^*;\mathcal{T},\Lambda)$ for individual choice variables)

$$\frac{\partial y(\theta^*)}{\partial T'}d\mathcal{T}' = -\varepsilon_{y,1-\mathcal{T}'}(\theta^*;\Lambda)\frac{y(\theta^*)}{1-\mathcal{T}'(y(\theta^*);\Lambda)}d\mathcal{T}'.$$

Hence, the fiscal externality through the substitution effect is given by

$$\begin{aligned}
dS(\theta^*;\mathcal{T},\Lambda) &= -\lambda\mathcal{T}'(y(\theta^*);\Lambda)\varepsilon_{y,1-\mathcal{T}'}(\theta^*;\Lambda)\frac{y(\theta^*)}{1-\mathcal{T}'(y(\theta^*);\Lambda)}d\mathcal{T}'\frac{f(\theta^*)}{y_\theta(\theta^*)}dy \\
&= -\lambda\frac{\mathcal{T}'(y(\theta^*);\Lambda)}{1-\mathcal{T}'(y(\theta^*);\Lambda)}\frac{\varepsilon_{y,1-\mathcal{T}'}(\theta^*;\Lambda)}{\varepsilon_{y,\theta}(\theta^*;\Lambda)}\theta^* d\mathcal{T}' f(\theta^*)dy \\
&= \lambda\frac{\mathcal{T}'(y(\theta^*);\Lambda)}{1-\mathcal{T}'(y(\theta^*);\Lambda)}\frac{1}{\varphi+1}\theta^* d\mathcal{T}' f(\theta^*)dy,
\end{aligned}$$

where the last equality uses equation (19) from Appendix A.2.1.

Next, there is a mechanical effect: all households with $\theta > \theta^*$ pay $d\mathcal{T}'dy$ more in taxes:

$$dM(\theta^*;\mathcal{T},\Lambda) = d\mathcal{T}'dy \times \int_{\theta^*}^{\overline{\theta}}\left(\lambda - u_e(\theta;\Lambda)\frac{\tilde{f}(\theta)}{f(\theta)}\right)f(\theta)d\theta,$$

where $u_e(x;\Lambda) = u_e(e(x;\Lambda);\Lambda)$.

Finally, there is an income effect: all households with $\theta > \theta^*$ now get poorer by $d\mathcal{T}'dy$ and change their income, which has a tax revenue effect:

$$dI(\theta^*;\mathcal{T},\Lambda) = d\mathcal{T}'dy \times \lambda\int_{\theta^*}^{\overline{\theta}}\mathcal{T}'(y(\theta);\Lambda)\eta(\theta;\Lambda)f(\theta)d\theta,$$

where $\eta(x,\Lambda) = -\frac{\partial y(\theta)}{\partial \mathcal{T}(0)}$.

If the tax schedule is optimal, all welfare effects have to add up to zero: $dS(\theta^*;\mathcal{T},\Lambda))+dM(\theta^*;\mathcal{T},\Lambda)+dI(\theta^*;\mathcal{T},\Lambda) = 0$. This then yields the optimality condition as in Lemma 3. □

A.2.3 Homothetic Benchmark

Proof. (Proposition 1) We proceed in 4 steps:

1. We show that income choices grow at rate $\alpha = \frac{1-\gamma}{\varphi+\gamma}$ due to growth $g \to 0$ and tax reform (8).
2. We show that this implies that expenditure levels also grow at this rate.
3. We show that this implies that marginal and average tax rates for different types θ stay constant.
4. Based on 1.-3., we show that tax reform (8) indeed is the optimal tax reform.

Step 1. In the first step, we show that tax reform (8) in combination with a small increase in Λ by $d\Lambda$, where $g \equiv \frac{d\Lambda}{\Lambda} \to 0$ indeed implies that all incomes grow at rate $\frac{1-\gamma}{\varphi+\gamma}$. To derive the comparative statics on labor supply, note first of all that tax reform (8) implies a marginal change in the absolute level of the tax payment for income level y by:

$$d\tilde{\mathcal{T}}(y;\Lambda) = \frac{1-\gamma}{\varphi+\gamma}\left(\mathcal{T}(y;\Lambda) - \mathcal{T}'(y;\Lambda)y\right). \tag{20}$$

The implied change in the marginal tax rate is then given by

$$d\tilde{\mathcal{T}}'(y;\Lambda) = \frac{1-\gamma}{\varphi+\gamma}\left(\mathcal{T}'(y;\Lambda) - \mathcal{T}'(y;\Lambda) - \mathcal{T}''(y;\Lambda)y\right) = -\frac{1-\gamma}{\varphi+\gamma}\mathcal{T}''(y;\Lambda)y. \tag{21}$$

Now consider a small perturbation of the individual first-order conditions by g and associated change in the absolute tax level as defined in (20) and associated change in the marginal tax rate as defined in (21). The adjustment of labor supply dn such that the FOC still holds is then defined by:

$$\begin{aligned} &SOC dn + u_{e\Lambda}(e;\Lambda)(1-\mathcal{T}')\theta\Lambda - u_e(e;\Lambda)\theta\left(-\frac{1-\gamma}{\varphi+\gamma}\mathcal{T}''(y)y\right) \\ &- u_{ee}(e;\Lambda)(1-\mathcal{T}')\theta\frac{1-\gamma}{\varphi+\gamma}\left(\mathcal{T}(y) - \mathcal{T}'(y)y\right) = 0, \end{aligned}$$

where

$$SOC = -B\varphi n^{\varphi-1} + u_{ee}(e;\Lambda)\left((1-\mathcal{T}')\theta\right)^2 - u_e(e;\Lambda)\mathcal{T}''\theta^2$$

is the second-order condition, see Appendix A.2.1. Solving for dn and using $u_{e\Lambda}(e;\Lambda) = u_{ee}(e;\Lambda)\frac{e}{\Lambda} + \frac{u_e(e;\Lambda)}{\Lambda}$ (see Appendix A.2.1 above, equation 17) yields (omitting arguments)

$$\begin{aligned} dn =& \frac{-\left(u_{ee}e + u_e\right)(1-\mathcal{T}')\theta - u_e\theta\frac{1-\gamma}{\varphi+\gamma}\mathcal{T}''y}{-B\varphi n^{\varphi-1} + u_{ee}\left((1-\mathcal{T}')\theta\right)^2 - u_e\mathcal{T}''\theta^2} \\ &+ \frac{u_{ee}(1-\mathcal{T}')\theta\frac{1-\gamma}{\varphi+\gamma}\left(\mathcal{T} - \mathcal{T}'y\right)}{-B\varphi n^{\varphi-1} + u_{ee}\left((1-\mathcal{T}')\theta\right)^2 - u_e(e;\Lambda)\mathcal{T}''\theta^2}. \end{aligned}$$

Rearranging yields:

$$dn = \frac{-u_{ee}(1-\mathcal{T}')\theta\left(e - \frac{1-\gamma}{\varphi+\gamma}\left(\mathcal{T}(y) - \mathcal{T}'(y)y\right)\right) - u_e\theta\left((1-\mathcal{T}') + \frac{1-\gamma}{\varphi+\gamma}\mathcal{T}''(y)y\right)}{-B\varphi n^{\varphi-1} + u_{ee}\left((1-\mathcal{T}')\theta\right)^2 - u_e\mathcal{T}''\theta^2}.$$

Divide by $u_e\theta(1-\mathcal{T}') = Bn^{\varphi}$:

$$dn = \frac{\frac{-u_{ee}e}{u_e}\left(1 - \frac{1-\gamma}{\varphi+\gamma}\left(\mathcal{T}(y) - \mathcal{T}'(y)y\right)\frac{1}{e}\right) - \left(1 + \frac{1-\gamma}{\varphi+\gamma}\frac{\mathcal{T}''(y)y}{1-\mathcal{T}'(y)}\right)}{-\frac{\varphi}{n} - \frac{-u_{ee}e}{u_e}\frac{((1-\mathcal{T}')\theta)}{e} - \frac{\mathcal{T}''}{1-\mathcal{T}'(y)}\theta}.$$

Turn this absolute labor supply change into a relative change:

$$\frac{dn}{n}=\frac{\gamma(e;\Lambda)\left(1-\frac{1-\gamma}{\varphi+\gamma}\left(\mathcal{T}(y)-\mathcal{T}'(y)y\right)\frac{1}{e}\right)-\left(1+\frac{1-\gamma}{\varphi+\gamma}\frac{\mathcal{T}''(y)}{1-\mathcal{T}'(y)}y\right)}{-\varphi-\gamma(e;\Lambda)\frac{((1-\mathcal{T}')\theta n)}{e}-\frac{\mathcal{T}''}{1-\mathcal{T}'(y)}y}.$$

In what follows, we isolate $\frac{1-\gamma}{\varphi+\gamma}$ on the RHS:

$$\frac{dn}{n}=\frac{1-\gamma}{\varphi+\gamma}\frac{\gamma(e;\Lambda)\left(\frac{\varphi+\gamma}{1-\gamma}-\frac{1}{e}\left(\mathcal{T}(y)-\mathcal{T}'(y)y\right)\right)-\left(\frac{\varphi+\gamma}{1-\gamma}+\frac{\mathcal{T}''(y)}{1-\mathcal{T}'(y)}y\right)}{-\varphi-\gamma(e;\Lambda)\frac{((1-\mathcal{T}')\theta n)}{e}-\frac{\mathcal{T}''}{1-\mathcal{T}'(y)}y}.$$

Hence:

$$\frac{dn}{n}=\frac{1-\gamma}{\varphi+\gamma}\frac{\frac{\varphi+\gamma}{1-\gamma}\left(\gamma(e;\Lambda)-1\right)-\gamma(e;\Lambda)\frac{1}{e}\left(\mathcal{T}(y)-\mathcal{T}'(y)y\right)-\frac{\mathcal{T}''(y)}{1-\mathcal{T}'(y)}y}{-\varphi-\gamma(e;\Lambda)\frac{((1-\mathcal{T}')\theta n)}{e}-\frac{\mathcal{T}''}{1-\mathcal{T}'(y)}y}.$$

Next we add and subtract $\varphi+\gamma(e;\Lambda)$ in the numerator:

$$\frac{dn}{n}=\frac{1-\gamma}{\varphi+\gamma}\frac{\frac{\varphi+\gamma}{1-\gamma}\left(\gamma(e;\Lambda)-1\right)+\varphi-\varphi+\gamma(e;\Lambda)-\gamma(e;\Lambda)-\gamma(e;\Lambda)\left(\frac{1}{e}\left(\mathcal{T}(y)-\mathcal{T}'(y)y\right)\right)-\frac{\mathcal{T}''(y)}{1-\mathcal{T}'(y)}y}{-\varphi-\gamma(e;\Lambda)\frac{((1-\mathcal{T}')\theta n)}{e}-\frac{\mathcal{T}''}{1-\mathcal{T}'(y)}y}.$$

Now use $y=e+\mathcal{T}$ and hence $1+\frac{\mathcal{T}}{e}=\frac{y}{e}$ to obtain

$$\frac{dn}{n}=\frac{1-\gamma}{\varphi+\gamma}\left(1+\frac{\frac{\varphi+\gamma}{1-\gamma}\left(1-\gamma(e;\Lambda)\right)-\gamma(e;\Lambda)-\varphi}{\varphi+\gamma(e;\Lambda)\frac{((1-\mathcal{T}')\theta n)}{e}+\frac{\mathcal{T}''}{1-\mathcal{T}'(y)}y}\right). \tag{22}$$

This implies that

$$\frac{dy}{y}=\frac{1-\gamma}{\varphi+\gamma}$$

in the homothetic case with $\gamma(e;\Lambda)=\gamma$ for each expenditure level and each level of economic development.

Step 2. Next, we show that expenditures also grow at rate $\frac{1-\gamma}{\varphi+\gamma}$. Therefore, note that expenditure and income are related through:

$$e=y-\mathcal{T}(y;\Lambda).$$

The change in expenditure is given by

$$\begin{aligned}de=&dy\left(1-\mathcal{T}'(y;\Lambda)\right)-d\tilde{\mathcal{T}}(y;\Lambda)=y\frac{1-\gamma}{\varphi+\gamma}\left(1-\mathcal{T}'(y;\Lambda)\right)-\frac{1-\gamma}{\varphi+\gamma}\left(\mathcal{T}(y)-\mathcal{T}'(y)y\right)\\=&\frac{1-\gamma}{\varphi+\gamma}\left(y-\mathcal{T}(y)\right)=\frac{1-\gamma}{\varphi+\gamma}e.\end{aligned}$$

which implies that expenditures grow at the same rate as income:

$$\frac{de}{e} = \frac{1-\gamma}{\varphi+\gamma}.$$

Step 3. We need to show that average and marginal tax rates do not change for a given type θ. For the average tax rate, this follows immediately. Recall the definition of the average tax rate

$$\frac{\mathcal{T}(y;\Lambda)}{y} = \frac{y-e}{y}.$$

Since both y and e grow at the same rate (see Step 1 and Step 2), the average tax rate stays constant.

Next, we turn to the marginal tax rate. Use first of all that the new tax schedule is defined as:

$$\lim_{g\to 0} \mathcal{T}(y;\Lambda\,(1+g)) = \mathcal{T}(y;\Lambda) + \lim_{g\to 0} g\alpha\,(\mathcal{T}(y;\Lambda) - \mathcal{T}'(y;\Lambda)y)$$

and hence

$$\lim_{g\to 0} \mathcal{T}'(y;\Lambda\,(1+g)) = \mathcal{T}'(y;\Lambda) - \lim_{g\to 0} g\alpha\mathcal{T}''(y;\Lambda)y.$$

We want to show that

$$\lim_{g\to 0} \mathcal{T}'\left(y\,(1+\alpha g)\,;\Lambda\,(1+g)\right) = \mathcal{T}'(y;\Lambda).$$

Now use the above and evaluate it at $y(1+\alpha g)$

$$\begin{aligned}
\lim_{g\to 0} \mathcal{T}'(y\,(1+\alpha g)\,;\Lambda\,(1+g)) =& \lim_{g\to 0} \left(\mathcal{T}'(y\,(1+\alpha g)\,;\Lambda) - g\alpha\mathcal{T}''(y\,(1+\alpha g))y\,(1+\alpha g)\right)\\
=& \mathcal{T}'(y;\Lambda) + \lim_{g\to 0} \left(g\alpha y\mathcal{T}''(y\,(1+\alpha)\,;\Lambda) - g\alpha\mathcal{T}''(y\,(1+\alpha g))y\,(1+\alpha g)\right)\\
=& \mathcal{T}'(y;g) - \lim_{g\to 0} g^2\alpha^2\mathcal{T}''(y\,(1+\alpha))y\\
=& \mathcal{T}'(y;g).
\end{aligned}$$

Step 4. Now we show that the tax reform (8) also satisfies the governments optimality condition at the allocation that it implements. We start with the distributional gains term. Recall that it is given by

$$D(\theta^*;\mathcal{T},\Lambda) = 1 - \frac{\int_{\theta^*}^{\bar{\theta}} u_e(x;\Lambda)\frac{w(x)}{f(x)}\frac{dF(x)}{1-F(\theta^*)}}{\int_{\underline{\theta}}^{\bar{\theta}} u_e(x;\Lambda)\frac{w(x)}{f(x)}dF(x)}.$$

We have shown already that $e(\theta;\Lambda)$ grows at rate α for all θ. It therefore follows:

$$\hat{D}(\theta^*;\mathcal{T},\Lambda,d\tilde{\mathcal{T}}) = \frac{\int_{\underline{\theta}}^{\bar{\theta}} \left(u_{ee}(x;\Lambda)\alpha e(x;\Lambda) + u_{e\Lambda}(x;\Lambda)\Lambda\right) \frac{w(x)}{f(x)} \frac{dF(x)}{1-F(\theta^*)}}{\int_{\underline{\theta}}^{\bar{\theta}} u_e(x;\Lambda) \frac{w(x)}{f(x)} \frac{dF(x)}{1-F(\theta^*)}} - \frac{\int_{\theta^*}^{\bar{\theta}} \left(u_{ee}(x;\Lambda)\alpha e(x;\Lambda) + u_{e\Lambda}(x;\Lambda)\Lambda\right) \frac{w(x)}{f(x)} \frac{dF(x)}{1-F(\theta^*)}}{\int_{\theta^*}^{\bar{\theta}} u_e(x;\Lambda) \frac{w(x)}{f(x)} \frac{dF(x)}{1-F(\theta^*)}}.$$

Now note that (see (17))

$$u_{e\Lambda} = \frac{e}{\Lambda} u_{ee} + u_e \frac{1}{\Lambda}. \tag{23}$$

Hence we obtain

$$\hat{D}(\theta^*;\mathcal{T},\Lambda,d\tilde{\mathcal{T}}) = \frac{\int_{\underline{\theta}}^{\bar{\theta}} \left(u_{ee}(x;\Lambda)\left(1+\alpha\right) e(x;\Lambda) + u_e(x;\Lambda)\right) \frac{w(x)}{f(x)} \frac{dF(x)}{1-F(\theta^*)}}{\int_{\underline{\theta}}^{\bar{\theta}} u_e(x;\Lambda) \frac{w(x)}{f(x)} \frac{dF(x)}{1-F(\theta^*)}} - \frac{\int_{\theta^*}^{\bar{\theta}} \left(u_{ee}(x;\Lambda)\left(1+\alpha\right) e(x;\Lambda) + u_e(x;\Lambda)\right) \frac{w(x)}{f(x)} \frac{dF(x)}{1-F(\theta^*)}}{\int_{\theta^*}^{\bar{\theta}} u_e(x;\Lambda) \frac{w(x)}{f(x)} \frac{dF(x)}{1-F(\theta^*)}},$$

and

$$\hat{D}(\theta^*;\mathcal{T},\Lambda,d\tilde{\mathcal{T}}) = \frac{\int_{\underline{\theta}}^{\bar{\theta}} \left(\gamma(\theta;\Lambda) u_e(x;\Lambda)\left(1+\alpha\right) + u_e(x;\Lambda)\right) \frac{w(x)}{f(x)} \frac{dF(x)}{1-F(\theta^*)}}{\int_{\underline{\theta}}^{\bar{\theta}} u_e(x;\Lambda) \frac{w(x)}{f(x)} \frac{dF(x)}{1-F(\theta^*)}} - \frac{\int_{\theta^*}^{\bar{\theta}} \left(\gamma(\theta;\Lambda) u_e(x;\Lambda)\left(1+\alpha\right) + u_e(x;\Lambda)\right) \frac{w(x)}{f(x)} \frac{dF(x)}{1-F(\theta^*)}}{\int_{\theta^*}^{\bar{\theta}} u_e(x;\Lambda) \frac{w(x)}{f(x)} \frac{dF(x)}{1-F(\theta^*)}}.$$

For the homothetic case where $\gamma(\theta;\Lambda) = \gamma$ for all θ, this simplifies to

$$\hat{D}(\theta^*;\mathcal{T},\Lambda,d\tilde{\mathcal{T}}) = \left(1+\gamma\left(1+\alpha\right)\right) \left(\frac{\int_{\underline{\theta}}^{\bar{\theta}} u_e(x;\Lambda) \frac{w(x)}{f(x)} \frac{dF(x)}{1-F(\theta^*)}}{\int_{\underline{\theta}}^{\bar{\theta}} u_e(x;\Lambda) \frac{w(x)}{f(x)} \frac{dF(x)}{1-F(\theta^*)}} - \frac{\int_{\theta^*}^{\bar{\theta}} u_e(x;\Lambda) \frac{w(x)}{f(x)} \frac{dF(x)}{1-F(\theta^*)}}{\int_{\theta^*}^{\bar{\theta}} u_e(x;\Lambda) \frac{w(x)}{f(x)} \frac{dF(x)}{1-F(\theta^*)}} \right) = 0.$$

The final step concerns showing that the efficiency cost term is unchanged. Hence we need to show:

$$\eta(\theta;\mathcal{T},\Lambda) = \eta(\theta;\mathcal{T} + g \times d\tilde{\mathcal{T}}, \Lambda\left(1+g\right)).$$

To show that, consider the RHS

$$\eta(\theta;\mathcal{T}+g\times d\tilde{\mathcal{T}},\Lambda\left(1+g\right))=-\frac{1}{\varphi}\times\frac{\gamma\frac{y(1+\alpha g)}{e(1+\alpha g)}}{1+\frac{\gamma}{\varphi}\frac{y(1+\alpha g)}{e(1+\alpha g)}(1-\mathcal{T}'\left((1+\alpha g)y;\Lambda\left(1+g\right)\right))+\frac{\rho(y(1+\alpha g);\Lambda(1+g))}{\varphi}}$$
$$-\frac{1}{\varphi}\times\frac{\gamma\frac{y}{e}}{1+\frac{\gamma}{\varphi}\frac{y}{e}(1-\mathcal{T}'\left(y;\Lambda\right))+\frac{\rho(y(1+\alpha g);\Lambda(1+g))}{\varphi}}.$$

Hence, $\eta(\theta;\mathcal{T},\Lambda)=\eta(\theta;\mathcal{T}+g\times d\tilde{\mathcal{T}},\Lambda\left(1+g\right))$ requires that $\rho(y\left(1+\alpha g\right);\Lambda\left(1+g\right))=\rho(y;\Lambda)$. Hence, consider

$$\mathcal{T}'(y\left(1+\alpha g\right);\Lambda\left(1+g\right))=\mathcal{T}'(y;\Lambda).$$

Since this has to hold for each value of y, this implies

$$\mathcal{T}''(y\left(1+\alpha g\right);\Lambda\left(1+g\right))\left(1+\alpha g\right)=\mathcal{T}''(y;\Lambda)$$

Now turn to ρ:

$$\rho(y\left(1+\alpha g\right);\Lambda\left(1+g\right))=\frac{\mathcal{T}''(y\left(1+\alpha g\right);\Lambda\left(1+g\right))y\left(1+\alpha g\right)}{1-\mathcal{T}'(y\left(1+\alpha g\right);\Lambda\left(1+g\right))}=\frac{\mathcal{T}''(y;\Lambda)y}{1-\mathcal{T}'(y;\Lambda)}=\rho(y;\Lambda),$$

which in turn implies

$$\eta(\theta;\mathcal{T},\Lambda)=\eta(\theta;\mathcal{T}+g\times d\tilde{\mathcal{T}},\Lambda\left(1+g\right)).$$

This completes the proof.

□

A.2.4 Non-homothetic Preferences: Distributional Gains Channel

Proof. (Proposition 2) Recall from Appendix A.2.3 that

$$\hat{D}(\theta^{*};\mathcal{T},\Lambda,d\tilde{\mathcal{T}})=\frac{\int_{\underline{\theta}}^{\bar{\theta}}\left(\gamma(\theta;\Lambda)u_{e}(x;\Lambda)\left(1+\alpha\right)+u_{e}(x;\Lambda)\right)\frac{w(x)}{f(x)}\frac{dF(x)}{1-F(\theta^{*})}}{\int_{\underline{\theta}}^{\bar{\theta}}u_{e}(x;\Lambda)\frac{w(x)}{f(x)}\frac{dF(x)}{1-F(\theta^{*})}}$$
$$-\frac{\int_{\theta^{*}}^{\bar{\theta}}\left(\gamma(\theta;\Lambda)u_{e}(x;\Lambda)\left(1+\alpha\right)+u_{e}(x;\Lambda)\right)\frac{w(x)}{f(x)}\frac{dF(x)}{1-F(\theta^{*})}}{\int_{\theta^{*}}^{\bar{\theta}}u_{e}(x;\Lambda)\frac{w(x)}{f(x)}\frac{dF(x)}{1-F(\theta^{*})}}.$$

This implies

$$\hat{D}(\theta^*;\mathcal{T},\Lambda,d\tilde{\mathcal{T}}) = (1+\alpha)\frac{\int_{\underline{\theta}}^{\bar{\theta}}\left(\gamma(\theta;\Lambda)u_e(x;\Lambda)\right)\frac{w(x)}{f(x)}\frac{dF(x)}{1-F(\theta^*)}}{\int_{\underline{\theta}}^{\bar{\theta}}u_e(x;\Lambda)\frac{w(x)}{f(x)}\frac{dF(x)}{1-F(\theta^*)}} - (1+\alpha)\frac{\int_{\theta^*}^{\bar{\theta}}\left(\gamma(\theta;\Lambda)u_e(x;\Lambda)\right)\frac{w(x)}{f(x)}\frac{dF(x)}{1-F(\theta^*)}}{\int_{\theta^*}^{\bar{\theta}}u_e(x;\Lambda)\frac{w(x)}{f(x)}\frac{dF(x)}{1-F(\theta^*)}}.$$

which yields the result in Proposition 2. □

A.2.5 Non-homothetic Preferences: Efficiency Costs Channel

Proof. (Proposition 3)

$$\eta(\theta;\mathcal{T},\Lambda) > \eta(\theta;\mathcal{T}+g\times d\tilde{\mathcal{T}},\Lambda\,(1+g)).$$

Recall from Appendix A.2.3, that the RHS is (but now account for relative risk aversion not to be constant):

$$\eta(\theta;\mathcal{T}+g\times d\tilde{\mathcal{T}},\Lambda\,(1+g)) = -\frac{1}{\varphi}\times\frac{\gamma\left(\theta;\Lambda\,(1+g)\right)\frac{y}{e}}{1+\frac{\gamma(\theta;\Lambda(1+g))}{\varphi}\frac{y}{e}(1-\mathcal{T}'\left(y;\Lambda\,(1+g)\right))+\frac{\rho(y;\Lambda)}{\varphi}}.$$

Hence

$$\begin{aligned}
&\lim_{g\to 0}\left(\eta(\theta;\mathcal{T}+g\times d\tilde{\mathcal{T}},\Lambda\,(1+g))-\eta(\theta;\mathcal{T},\Lambda)\right)\\
&=\lim_{g\to 0}\frac{1}{\varphi}\frac{\gamma_\Lambda\frac{y}{e}\left(\varphi+\frac{\gamma(\theta;\Lambda)}{\varphi}\frac{y}{e}(1-\mathcal{T}'\left(y;\Lambda\right))+\frac{\rho(y;\Lambda)}{\varphi}\right)-\frac{\gamma_\Lambda(\theta;\Lambda)}{\varphi}\frac{y}{e}(1-\mathcal{T}'\left(y;\Lambda\right))\gamma\left(\theta;\Lambda\right)\frac{y}{e}}{\left(1+\frac{\gamma(\theta;\Lambda)}{\varphi}\frac{y}{e}(1-\mathcal{T}'\left(y;\Lambda\right))+\frac{\rho(y;\Lambda)}{\varphi}\right)^2}\\
&=\lim_{g\to 0}\frac{1}{\varphi}\frac{\gamma_\Lambda\frac{y}{e}\left(1+\frac{\rho(y;\Lambda)}{\varphi}\right)}{\left(1+\frac{\gamma(\theta;\Lambda)}{\varphi}\frac{y}{e}(1-\mathcal{T}'\left(y;\Lambda\right))+\frac{\rho(y;\Lambda)}{\varphi}\right)^2}<0.
\end{aligned}$$

□

A.2.6 Non-homothetic Preferences: Income Distribution Channel

Proof. (Proposition 4) In Appendix A.2.3 we derived the relative change in income due to tax reform (8) and a small growth in the level Λ by $g\to 0$

$$\frac{dy}{y}=\frac{dn}{n}=\frac{1-\gamma}{\varphi+\gamma}\left(1+\frac{\gamma(e;\Lambda)\left(-1-\frac{\varphi+\gamma}{1-\gamma}\right)+\frac{\varphi+\gamma}{1-\gamma}-\varphi}{\varphi+\gamma(e;\Lambda)\frac{((1-\mathcal{T}')\theta n)}{e}+\frac{\mathcal{T}''}{1-\mathcal{T}'(y)}y}\right)$$

$$\forall \theta: \ d(\theta; \mathcal{T}, \Lambda) \equiv \frac{1+\varphi}{\varphi+\gamma} \left(\varphi + \gamma(\theta;\Lambda) \frac{y(\theta;\Lambda)}{e(\theta;\Lambda)} (1-\mathcal{T}') + \frac{\mathcal{T}''}{1-\mathcal{T}'} y(\theta;\Lambda) \right)^{-1}.$$

We now consider the sign of $\frac{d\left(\frac{dy}{y}\right)}{\theta}$ as this term is no longer constant if risk aversion $\gamma(e;\Lambda)$ is not constant. If $\frac{d\left(\frac{dy}{y}\right)}{\theta} > 0$ for all θ, then tax reform (8) would imply an increase in inequality. Signing this in general is hard. We therefore consider a special illustrative case to gain intuition and a clear benchmark. We consider the following tax function:

$$\mathcal{T}(y) = y - (1-\lambda)\, y^{1-\tau},$$

which has first derivative

$$1 - \mathcal{T}'(y) = (1-\lambda)\,(1-\tau)\, y^{-\tau}$$

and second derivative

$$\mathcal{T}''(y) = \tau\,(1-\lambda)\,(1-\tau)\, y^{-\tau-1}.$$

This implies

$$\frac{\mathcal{T}''}{1-\mathcal{T}'(y)} y = \tau$$

and

$$\frac{((1-\mathcal{T}')\theta n)}{e} = \frac{(1-\lambda)\,(1-\tau)\, y^{-\tau+1}}{(1-\lambda)\, y^{1-\tau}}.$$

In this case, the relative change in income simplifies to:

$$\frac{dy}{y} = \frac{dn}{n} = \frac{1-\gamma}{\varphi+\gamma} \left(1 + \frac{\gamma(e;\Lambda)\left(-1 - \frac{\varphi+\gamma}{1-\gamma}\right) + \frac{\varphi+\gamma}{1-\gamma} - \varphi}{\varphi + \gamma(e;\Lambda)\,(1-\tau) + \tau} \right)$$

and the derivative w.r.t. to type is given by

$$\frac{d\left(\frac{dy}{y}\right)}{\theta} = \frac{1-\gamma}{\varphi+\gamma} \left(\frac{\gamma_e(e;\Lambda) e_\theta \left(-1 - \frac{\varphi+\gamma}{1-\gamma}\right)(\varphi + \gamma(e;\Lambda)\,(1-\tau)+\tau)}{(\varphi + \gamma(e;\Lambda)\,(1-\tau)+\tau)^2} \right.$$
$$\left. \frac{\gamma_e(e;\Lambda) e_\theta\,(1-\tau)\left(\gamma(e;\Lambda)\left(-1 - \frac{\varphi+\gamma}{1-\gamma}\right) + \frac{\varphi+\gamma}{1-\gamma} - \varphi\right)}{(\varphi + \gamma(e;\Lambda)\,(1-\tau)+\tau)^2} \right)$$

.

Focus on the numerator:

$$\begin{aligned}
&\frac{1-\gamma}{\varphi+\gamma}\gamma_e(e;\Lambda)e_\theta\Bigg[\left(-1-\frac{\varphi+\gamma}{1-\gamma}\right)(\varphi+\gamma(e;\Lambda)(1-\tau)+\tau)\\
&-(1-\tau)\left(\gamma(e;\Lambda)\left(-1-\frac{\varphi+\gamma}{1-\gamma}\right)+\frac{\varphi+\gamma}{1-\gamma}-\varphi\right)\Bigg]\\
=&\frac{1-\gamma}{\varphi+\gamma}\gamma_e(e;\Lambda)e_\theta\left[\left(-1-\frac{\varphi+\gamma}{1-\gamma}\right)(\varphi+\tau)-(1-\tau)\frac{\varphi+\gamma}{1-\gamma}+(1-\tau)\varphi\right]\\
&\frac{1-\gamma}{\varphi+\gamma}\gamma_e(e;\Lambda)e_\theta\left[\left(-1-\frac{\varphi+\gamma}{1-\gamma}\right)(\varphi+\tau+(1-\tau))+(1-\tau)(1+\varphi)\right]\\
&\frac{1-\gamma}{\varphi+\gamma}\gamma_e(e;\Lambda)e_\theta\left[\left(-\frac{1+\varphi}{1-\gamma}\right)(\varphi+1)+(1-\tau)(1+\varphi)\right]
\end{aligned}$$

Case 1: $\gamma < 1$. Then this is > 0 if

$$\left(-\frac{1+\varphi}{1-\gamma}\right)(\varphi+1)+(1-\tau)(1+\varphi)<0$$

hence

$$-1-\varphi+(1-\tau)(1-\gamma)<0$$

$$-\varphi-\tau-(1-\tau)\gamma<0$$

Case 2: $\gamma > 1$. Then this is > 0 if

$$\left(-\frac{1+\varphi}{1-\gamma}\right)(\varphi+1)+(1-\tau)(1+\varphi)>0$$

Rearranging yields

$$-\varphi-\tau-(1-\tau)\gamma<0$$

which is the same condition. Note that this condition is basically not a constraint: it is the SOC if the tax schedule is of this form and if there is a risk aversion parameter γ. So a sufficient condition (which would hold for any $\gamma > 0$ – note that here, γ is a tax reform parameter) is

$$\tau > -\varphi$$

which is the condition on the SOC being fulfilled in case of no risk aversion being equal to zero. □

B Data

In this section, we describe the dataset we use to compute income and wealth distributions in 1950 and 2010.

B.1 SCF+

The SCF+ provides long-run data on income and wealth inequality in the United States It is compiled by Kuhn, Schularick, and Steins (2020), based on historical waves of the SCF. The covered time period is from 1949 to 2016.

As income components in the data, we use wages and salaries, income from professional practice and self-employment, and business and farm income. We exclude rental income, interest, dividends, and transfers, as we model asset income and transfers separately from the labor income process.

For wealth, we compute net worth as the sum of all assets minus the sum of all debts. Assets include liquid assets (checking, savings, call/money market accounts, certificates of deposit), housing and other real estate, bonds, stocks and business equity, mutual funds, cash value of life insurance, defined-contribution retirement plans, and cars. Debt consists of housing debt (debt on owner-occupied homes, home equity loans and lines of credit) and other debt (car loans, education loans, consumer loans).

We restrict the sample to the working age population, i.e. household heads aged 25 to 60. We impose that minimum household income is \$5,000 in 2010 (in 2016 dollars). In 1950, we choose the cutoff such that the ratio of minimum income to median income is the same as in 2010, which results in a cutoff of \$2,700 (in 2016 dollars).

C Quantitative Models

C.1 Risk Aversion, Wealth Effects and MPCs

C.1.1 Relationship between Risk Aversion, Wealth Effects and MPCs

First-order intratemporal condition in the household's optimization problem (10) gives:

$$v'(n) - u_e(e)\theta(1 - \mathcal{T}'(\theta n)) = 0, \quad \text{with } v'(n) = Bn^{\varphi}.$$

We implicitly differentiate this equation to obtain

$$
\begin{aligned}
v''(n)\frac{\partial n}{\partial T} - u_{ee}(e)\frac{\partial e}{\partial T}\theta(1-\mathcal{T}'(\theta n)) + u_e(e)\theta^2\mathcal{T}''(\theta n)\frac{\partial n}{\partial T} &= 0 \\
-\frac{v''(n)}{\theta}\eta - u_{ee}(e)\theta(1-\mathcal{T}'(\theta n)) \times \text{MPC} - u_e(e)\theta\mathcal{T}''(\theta n)\eta &= 0 \\
-\eta\left(\frac{1}{\theta}v''(n) + u_e(e)\theta\mathcal{T}''(\theta n)\right)e - u_{ee}(e)\frac{v'(n)}{u_e}e \times \text{MPC} &= 0 \\
-\eta\left(\frac{1}{\theta}\frac{v''(n)}{v'(n)} + \frac{\theta u_e(e)}{v'(n)}\mathcal{T}''(\theta n)\right)e + \text{RRA} \times \text{MPC} &= 0
\end{aligned}
$$

which delivers equation (12).

C.1.2 Dynamic Model: Measurement of Wealth Effects and MPCs

Wealth effects. We compare model-implied wealth effects on household earnings with evidence provided by Golosov, Graber, Mogstad, and Novgorodsky (2023). They merge data from lottery winnings with earnings data covering the universe of U.S. taxpayers. Our preferred measure of comparison is the average reduction in per-adult total labor earnings in the five years following a lottery win. They report a reduction of labor earnings by \$2.3 per \$100 of lottery wealth.

In the model, we expose households to a wealth shock corresponding to the average post-tax win size reported by Golosov, Graber, Mogstad, and Novgorodsky (2023). This win size is \$181, 200 in 2016 dollars, which we adjust to 2010 dollars using the consumer price index. Then, we simulate two panels of households, one of which is exposed to this wealth shock and one of which is not. We compute the average difference between the two groups and obtain a drop of labor earnings of \$2.3 per \$100 of additional wealth.

MPCs. The empirical literature typically computes MPCs as the consumption response to a small windfall gain. To be comparable with that approach, we expose households in the model calibrated to the year 2010 to a one-time wealth shock of \$500. We compute MPCs as the differences in the expenditure after the wealth shock relative to a counterfactual in which no such shock occurs. We report the population average.

C.2 Mirrlees Parameterizations

Table C.1 summarizes the calibrated parameters of the Mirrlees setups we use in Section 4.

C.3 Quantitative Models: Expenditure Distributions

Table C.2 shows the expenditure distributions in the static and the dynamic model.

Table C.1: Mirrlees Parameterization

Parameter	Interpretation	Baseline	Higher RA
Preferences			
γ	Curvature utility	0.750	1.500
$1/\varphi$	Frisch elasticity	0.500	0.500
B	Labor disutility	13.000	52.000
σ	Non-homothetic CES parameter	0.300	0.300
ε_A	Non-homothetic CES parameter	0.100	0.100
ε_G	Non-homothetic CES parameter	1.000	1.000
ε_S	Non-homothetic CES parameter	1.800	1.800
Ω_A	Non-homothetic CES parameter	0.093	0.093
Ω_G	Non-homothetic CES parameter	1.000	1.000
Ω_S	Non-homothetic CES parameter	2.400	2.600
Prices			
p_A^{1950}	Price agriculture 1950	1.000	1.000
p_G^{1950}	Price goods 1950	1.000	1.000
p_S^{1950}	Price services 1950	1.000	1.000
p_A^{2010}	Price agriculture 2010	0.274	0.229
p_G^{2010}	Price goods 2010	0.147	0.123
p_S^{2010}	Price services 2010	0.464	0.387
Inequality			
α^{1950}	Pareto tail 1950	4.400	4.400
α^{2010}	Pareto tail 2010	3.300	3.300
σ_a^{1950}	EMG parameter 1950	0.350	0.481
σ_a^{2010}	EMG parameter 2010	0.480	0.640
Government			
λ^{1950}	Tax function level 1950	0.242	0.243
τ^{1950}	Tax function progressivity 1950	0.160	0.160
T^{1950}	Transfer 1950	0.003	0.003
G^{1950}	Government spending 1950	0.043	0.041
λ^{2010}	Tax function level 2010	0.225	0.238
τ^{2010}	Tax function progressivity 2010	0.095	0.095
T^{2010}	Transfer 2010	0.011	0.009
G^{2010}	Government spending 2010	0.042	0.033

Notes: Table C.1 summarizes the calibrated parameters of the Mirrlees setup, for both the benchmark calibration and the robustness calibration with higher risk aversion.

Table C.2: Expenditure Distribution in the Dynamic and the Static Model

1950	**Expenditure** Share by Quintile				
Dynamic model	8%	13%	17%	23%	39%
Static model	9%	13%	17%	23%	38%
2010	**Expenditure** Share by Quintile				
Dynamic model	7%	11%	16%	21%	45%
Static model	7%	12%	16%	23%	43%

Notes: Table C.2 compares the expenditure distributions in the static and the dynamic model.

C.4 Pareto Weights in the Dynamic Model

The inverse optimum weights that make the 1950 $t\&T$ system optimal are high on the top expenditure households relative to the rest of the distribution. With parameters $\mu = 0.05$ and $\nu = 116.4$, the weights are flat up to the 95th percentile of the expenditure distribution, but then strongly increase in expenditure up to roughly 20 times the weight at the bottom. Though there is no guarantee to match exactly the calibrated system, we come very close to matching the observed tax system, with an optimal progressivity of 0.15 and a transfer-to-output ratio of 0.9%, relative to the calibrated values of 0.13 and 1.1%.