Redistribution and Investment

Wendy Morrison[*,†]

March 31st, 2026

Click here for current version

Abstract

This paper studies the trade-offs associated with income redistribution in an overlapping generations model in which marginal propensities to save increase with permanent income. Transferring permanent income from high savers to low savers lowers aggregate savings and depresses investment. If more capital is welfare improving, the government faces a trade-off between redistribution and investment whose size depends on the distribution of marginal propensities to save across the income distribution. I characterize this trade-off theoretically and quantify it. In a simple model, I show that the welfare costs of labor income redistribution can be decomposed into the familiar equity-efficiency trade-off and the permanent income redistribution-investment channel. I calibrate a quantitative model to match empirical estimates of marginal propensities to save in the U.S. In a back-of-the-envelope calculation using the simple model, this channel is between 1/4 and 1/3 of the size of the labor distortion channel. In the quantitative model, this ratio rises to between 44 and 47 percent.

Keywords: Redistribution, Non-Homothetic Preferences, Optimal Capital Accumulation

JEL codes: E21, E22, H21, D63

[*]I would like to thank Jennifer La'O and Émilien Gouin-Bonenfant for their continuous support and excellent feedback. I would like to also thank Hassan Afrouzi, Stephen Zeldes, Mariacristina De Nardi, Patrick Kehoe, and seminar participants at Columbia University, the 2024 Summer Meeting of the Society for Economic Dynamics, and the 2024 NBER Summer Institute for many helpful conversations.

[†]Duke University. Email: wendy.morrison@duke.edu, Phone: 610-283-6532

1 Introduction

What is the optimal amount of income redistribution? The existing literature has primarily focused on the trade-off between equity and efficiency: weighing the welfare benefits of a more equal distribution of resources against the efficiency costs associated with distortionary redistributive taxation (Mirrlees, 1971, Piketty and Saez, 2013b, Werning, 2007). In this paper, I argue that this framing misses a quantitatively important additional cost. Several studies have documented that marginal propensities to save (MPS) out of *permanent income* rise sharply over the income distribution (Dynan et al. (2004), Straub (2019)).[1] I show that this pattern introduces an additional trade-off between *intra*-generational equality and *inter*-generational equality.

Intuitively, if savings increase with permanent income, then all redistributive policies – including non-distortionary lump-sum redistribution – move permanent income from high-MPS households to low-MPS households, which lowers aggregate savings and increases borrowing costs (Straub, 2019, Mian et al., 2021).[2] In the long run, a lower savings rate curbs firms' investment in capital, implying fewer resources set aside for the consumption of future generations in dynamically efficient economies. This trade-off is distinct from the classic equity-efficiency trade-off, which states that a more equal distribution of resources requires movement *inside* the Pareto Frontier. In this paper, I argue that when high-income households have higher MPS out of permanent income, movements in the distribution of permanent income generate movements *along* the Pareto Frontier – from allocations rationalized by greater weight on future generations towards allocations rationalized by lower weight on future generations.[3]

In this paper, I characterize this permanent income (PI) redistribution-investment trade-off theoretically and quantify its size. In the first section, I begin with a simple overlapping generations (OLG) model with exogenous labor supply, ex-ante inequal-

[1] Dynan et al. (2004) find a strong positive relationship between saving rates and permanent income, as well as a positive relationship between the marginal propensity to save and permanent income in the PSID. Straub (2019) uses updated PSID data with measures of consumption to estimate the elasticity of consumption with respect to permanent income. His key finding is a consumption elasticity of approximately 0.7, significantly below 1, implying that marginal propensities to save increase with permanent income.

[2] As Straub (2019) notes, this channel is absent in canonical heterogeneous agent models, which generate heterogeneous marginal propensities to consume and save out of *transitory shocks* but not out of *permanent income*.

[3] When MPS out of permanent income increase over the income distribution, lump-sum redistribution has the same effect on aggregate savings as government debt or pay-as-you-go social security.

ity in labor productivity, and MPS out of PI that increase with permanent income. The fiscal authority has access to capital taxes or subsidies, limited inter-generational transfers, and *type-dependent lump-sum* transfers between households within a generation.[4] In this context, I present sufficient conditions for the existence of the PI redistribution-investment trade-off. I show that for such a trade-off to exist, it is sufficient that both high-income households have higher MPS out of permanent income and for an economy to be below the Golden Rule capital stock.[5] Intuitively, it must both be the case that PI redistribution alters the long-run level of savings, and that, even after exhausting available non-distortionary tools to boost the savings supply, the economy is not over-saving.

In this simple model, I perform a series of comparative static exercises that explore what determines the size of the trade-off. Importantly, whether differences in MPS stem from agents' *type* (e.g., type-dependent discount factors) or are simply a function of agents' income (e.g., non-homothetic preferences) determines whether the distribution of MPS compresses as the income distribution becomes more equal, implying a smaller impact of additional redistribution on aggregate savings.[6] I then show that the size of this channel also depends not only on the distribution of MPS out of permanent income, but also on the *interest rate elasticities* of firms and savers. If firms are slow to adjust their capital stock to increases in borrowing costs, permanent income redistribution will simply drive up interest rates and crowd out household debt with little effect on aggregate capital. Indeed this is the mechanism in both Straub (2019) and Mian et al. (2021). Similarly, if the domestic economy is open and small relative to the rest of the world, a decline in aggregate domestic savings has little impact on the total supply of savings available to domestic firms, muting the size of the PI redistribution-investment trade-off.

While allowing for lump-sum redistribution is useful to illustrate the PI redistribution-investment trade-off, this trade-off applies to all policies that redistribute permanent income. To examine how this channel operates in a more realistic policy setting, and to assess its magnitude relative to the classic equity-efficiency trade-off, I introduce

[4]Note that the inter-generational transfers are equivalent to debt policy with a uniform lump-sum tax/transfer (Atkinson and Sandmo, 1980).

[5]Importantly, the existence of this trade-off does not depend on *why* high-income households have higher MPS out of PI. I consider multiple possible micro-foundations.

[6]If savings are simply a function of income, then as the income distribution becomes more equal, the MPS of the rich decreases and the MPS of the non-rich increases, decreasing the size of the trade-off.

endogenous labor into the simple model and replace the lump-sum redistribution with a standard linear labor income tax that funds a uniform lump-sum transfer. I show that the welfare effects of this policy can be decomposed into the familiar equity-efficiency trade-off and a new term corresponding to the PI redistribution-investment trade-off. In a back-of-envelope numerical illustration using this decomposition, the PI redistribution channel is about 26 percent of the size of the welfare cost of the labor income distortion channel, depending on welfare weights and the assumed labor earnings elasticity. This magnitude is large enough to meaningfully affect the optimal setting of labor income tax rates.

Finally, I solve a quantitative OLG model with a more realistic earnings life cycle and fiscal policy, as well as uninsurable idiosyncratic risk, and calibrate it to match empirical estimates of the MPS distribution from U.S. household panel data. I find that a one percentage point increase in the labor income tax rate lowers steady-state capital by 1.01 percent; roughly one third of this decline is attributable to the direct effect of the change in the expected permanent income distribution, holding labor supply fixed. I show that the size of the decline depends on whether heterogeneous MPS are generated by type-dependent preferences or scale-dependent (non-homothetic) preferences

I then perform the same decomposition as in the simple model, isolating the welfare impact of the PI-redistribution channel from the labor income distortion channel. In the quantitative model, the PI redistribution channel accounts for between 44 and 47 percent of the welfare cost of the labor distortion channel across a wide range of inter-generational welfare weights and Pareto weights on types.

These findings have direct implications for the design of fiscal policy. Holding existing inter-generational transfer schemes and debt levels constant, this paper implies that optimal labor income tax rates are likely substantially lower than formulas based on models with homogeneous MPS would suggest. When the government has access to a broader set of fiscal instruments, however, the analysis points toward a specific policy mix: greater income redistribution should be paired with pro-investment reforms — a transition from pay-as-you-go to fully funded social security and reductions in government debt for example — that offset the negative effect of redistribution on aggregate investment. That is, the permanent income redistribution channel does not necessarily imply that redistribution is undesirable; but rather, that policy makers should take into account the likely impact of redistribution on savings rates and pair more progressive redistribution with policies that restore the long-run capital stock.

Related Literature This paper is related to the substantial literature on redistributive taxation. In their review of the optimal labor income tax literature, Piketty and Saez (2013a) note that researchers typically focus on ' the classical trade-off between equity and efficiency which is at the core of the optimal labor income tax problem.' Similarly, Piketty and Saez (2013b) analyze the optimal inheritance tax through the lens of an equity-efficiency trade-off, noting that their results are orthogonal to concerns over optimal capital accumulation.[7] Werning (2007) considers the equity-efficiency trade-off in a dynamic economy subject to aggregate shocks.

Golosov et al. (2016) focuses on the trade-offs between efficiency and both equity and insurance in a model with idiosyncratic household labor income shocks. Heathcote et al. (2017) focus on the trade-off between the benefits of equity and insurance and the costs of labor supply distortions and disincentives to invest in skills. Imrohoroglu et al. (2018) study the trade-off between greater equity through taxing top earners and entrepreneurial activity. I depart from much of the literature in considering the *non-distortionary* effects of redistributing permanent income on optimal capital accumulation.

This paper is certainly not the first to consider a trade-off between capital accumulation and taxation (Atkinson and Sandmo 1980; Hamada 1972, Pizzo 2023). However, this paper is one of only a few that analyze a trade-off between redistribution and capital accumulation while abstracting away from inefficiency concerns (notably Pestieau and Possen (1978) and Okuno and Yakita (1981)).

A growing literature studies the effect of heterogeneous savings behavior on optimal tax policy in various settings. Golosov et al. (2013) solve a static model with preference heterogeneity. Pestieau and Possen (1978) and Judd (1985) consider a 'two-class' model with capitalists and workers. Sheshinski (1976) considers a model with infinitely lived agents. Several papers take a Mirrleesian approach (Saez and Stantcheva (2018); Gerritsen et al. (2020); Schulz (2021)). This literature primarily studies the effect of *distortionary* taxation in infinitely lived models. In this paper, I highlight a new channel through which even *non-distortionary* permanent income redistribution creates a trade-off between intra-generational and inter-generational

[7]Note that if the economy is dynamically efficient, less capital may be sub-optimal for a *given* set of Pareto weights, but is not *inefficient*. That is, one can find Pareto weights such that a lower level of capital is optimal, namely by weighting current generations more.

equality in life-cycle models.

This paper also contributes modestly to the empirical literature studying the relationship between permanent household income and savings. Following Dynan et al. (2004), I use the Panel Study of Income Dynamics (PSID) to estimate savings rates by permanent income quintile, taking advantage of the new consumption data added to the PSID in 1999 to construct a *direct* measure of *active* savings (income less consumption). I then combine these updated savings rate estimates with the permanent income elasticity of savings implied by Straub (2019)—approximately 1.3—to recover marginal propensities to save by quintile. The resulting estimates confirm a steep positive gradient in marginal propensities to save across the permanent income distribution.

Using Norwegian administrative data, Fagereng et al. (2019) find that active saving rates — excluding capital gains — are approximately flat across the wealth distribution; it is only gross saving rates, which include capital gains, that rise sharply with wealth. In contrast, Hubmer et al. (2026), also using Norwegian data, find that higher saving rates among the wealthy account for roughly a third of the wealth gap between the top 0.1 percent and median-wealth households. Whether these findings extend to the United States is unclear, given differences in demographics and fiscal policies between the two countries. The estimates in this paper are constructed directly from PSID data, conditioning on permanent income rather than wealth, and focusing on active saving — the margin most relevant for the capital accumulation channel studied here.

Finally, this paper contributes to a small recent literature on the macroeconomic effects of heterogeneous household savings behavior. Straub (2019) shows that non-homothetic savings behavior and increased inequality can explain falling interest rates. Blanco and Diz (2021) study the effects of non-homothetic preferences on optimal monetary policy. Mian et al. (2021) show how non-homothetic preferences have contributed to increased indebtedness and dampened aggregate demand in the long run. Doerr et al. (2023) show that because high-income households save relatively more in stocks and bonds, this raises relative borrowing costs for bank-dependent firms and lowers their employment share.

Layout. The rest of the paper proceeds as follows. In Section 2, I lay out the baseline overlapping generations model with exogenous labor and establish its key properties. I derive sufficient conditions for a welfare trade-off between permanent income redistribution and capital accumulation. In Section 3, I introduce endogenous labor into the model, and decompose the welfare effects of a simple labor income redistribution policy into equity, efficiency, and the new channel. In Section 4, I present the quantitative model and results. Section 5 concludes.

2 Simple Model.

2.1 Environment

I begin with a variant of the canonical two-generation overlapping generations closed-economy model with fixed labor supply (Diamond, 1965) with ex-ante heterogeneity in labor productivity and time preferences. Time is discrete. Agents have perfect foresight over future variables and there is no uncertainty. The government is able to impose a capital income tax which funds age-dependent lump-sum transfers subject to a political constraint which limits transfers from the current old to the current young. The government can also impose *type-dependent* lump-sum transfers.

Households. There is a unit mass of households who each live for 2 periods, $j \in \{y, o\}$ and have heterogeneous labor productivity types, θ_i for $i \in \{L, H\}$ where $\theta_L < \theta_H$. There is a constant fraction π of households of each productivity type-i born in period t. While young, households supply a single unit of labor to firms and receive $w_t\theta_i$ in labor income. The weighted sum of labor productivity is normalized to 1. Households can borrow and save at net rate of return r_t subject to the linear tax rate τ_{kt}. Capital depreciates at rate δ. Households receive a type-specific lump-sum tax (transfer) T^y_{it} when young as well as a uniform transfer when old, T^o_t. A type-i household born in year t has lifetime utility given by equation (1).

$$U(c^y_{it}, c^o_{it+1}) = \frac{(c^y_{it})^{1-\sigma_y}}{1-\sigma_y} + \beta_i \frac{(c^o_{it+1})^{1-\sigma_o}}{1-\sigma_o} \tag{1}$$

Note that the discount factor, β_i may be type-specific and that the parameters σ_y and σ_o may differ from one another. Households choose consumption when young to maximize (1) subject to their lifetime budget constraint:

$$c^y_{it} + \frac{c^o_{i,t+1}}{(1-\tau_{kt+1})(1+r_{t+1})} = w_t\theta_i + T^y_{it} + \frac{T^o_{t+1}}{(1-\tau_{kt+1})(1+r_{t+1})} = PI_{it} \tag{2}$$

I define the right hand side of equation (2) as the household's permanent income, PI_{it}. Denote a_{it} as the savings of type-i households.

Discussion of preferences. The parameters σ_y and σ_o govern the elasticity of substitution between consumption over the life-cycle. I make the following assumption about these parameters and the discount factor, β_i.

Assumption. Let $\sigma_o \leq 1$, $\sigma_y \geq \sigma_o$, and $\beta_H \geq \beta_L$.

The above assumption allows for the possibility that households with higher incomes have a higher propensity to save out of their lifetime income. When any of the above inequalities are strict, as long as the high productivity types have higher permanent income, the marginal propensity to save out of permanent income for high-productivity type households, $\frac{\partial s_{Ht}}{\partial PI_H}$ is greater than for low-productivity households.

When all elasticity parameters are equal and discount factors are uniform across types, the marginal propensity to save out of permanent income is constant over types. In this case, any lump-sum transfer from the high-types to the low-types will have no effect on aggregate savings or the interest rate. These results are summarized in the following Lemma.

Lemma 1 *If either $\sigma_y > \sigma_o$ or $\beta_H > \beta_L$, and $PI_{Ht} > PI_{Lt}$, the marginal propensity to save out of permanent income is higher for high-productivity types, $\frac{\partial s_{Ht}}{\partial PI_{Ht}} > \frac{\partial s_{Lt}}{\partial PI_{Lt}}$. For a proof, see Appendix A.1.*

Lemma 1 shows that this simple life-cycle model nests several major explanations for non-homothetic savings behavior. When $\sigma_y > \sigma_o$, consumption later in life is considered a luxury, and households consume a greater share in the second period as their lifetime income increases (Straub, 2019) and therefore save a greater share of their permanent income.[8] I also allow for the possibility that high-productivity households are simply more patient, which may explain some of the observed differences in savings rates over the income distribution (De Nardi and Fella, 2017). These *type-dependent* preferences can capture other type-dependent channels – for example, type-dependent rates of return – in a simple way. I explicitly consider type-dependent returns in an extension in the Appendix.

Firms. There is a continuum of perfectly competitive firms who rent capital and labor from households and produce output in order to maximize profit subject to a

[8]For example, more lavish retirements, private school for children, and out-of-pocket medical expenses are all luxury goods purchased later in life.

constant elasticity of substitution production function (3).

$$f(k_t, \ell_t) = \left(\alpha_k k_t^{\zeta} + \alpha_\ell \ell_t^{\zeta}\right)^{\frac{1}{\zeta}} \tag{3}$$

Here, I assume that $\alpha_\ell + \alpha_k = 1$.

Government. The government runs a balanced budget each period.[9] The government budget constraint is given by the following expression.

$$\pi \sum_{i \in I} T_{it}^y + \pi T_t^o = (1 + r_t) k_t \tau_{kt}$$

I consider the possibility that the government is subject to *a political constraint* on lump-sum transfers. First, net transfers from the young to the old must remain above a lower limit, $\underline{\mathrm{T}}$.[10] For simplicity, I consider two cases: (1) where $\underline{\mathrm{T}} = -\infty$ and transfers are unlimited, and (2) $\underline{\mathrm{T}} = 0$, which rules out a *reverse* pay-as-you-go social security scheme in which resources are transferred from the current old to the current young.[11]

$$T_t^o - \sum_{i \in I} T_{it}^y \geq \underline{\mathrm{T}} \tag{4}$$

$$T_{Lt}^y - T_{Ht}^y \leq (\theta_H - \theta_L) w_t \tag{5}$$

Second, net transfers within a generation from the high-productivity households to the low-productivity households are capped to ensure that high-productivity households have permanent income that is at least as high as that of low-productivity households. That is, the government can redistribute to fully equalize the income distribution, but cannot generate *reverse* inequality in which so much is transferred from the high-productivity types to the low-productivity types that the latter becomes richer than the former.

[9] However, the lump-sum inter-generational transfers ensure that the government is able to achieve any allocation available with debt policy. Therefore, this assumption is without loss.

[10] If age-dependent transfers were replaced with debt policy, this would be equivalent to assuming an upper-bound on the government surplus.

[11] This assumption is a natural one, given that the initial old would be made worse off.

2.2 Equilibrium.

2.2.1 Equilibrium Definition.

I define an allocation $\mathcal{A} \equiv \{\{c^y_{it}, c^o_{it}\}_{i\in I}, k_t\}_{t\geq 0}$. An equilibrium, $\mathcal{X}$ is an allocation, a sequence of financial positions, $\{a_{it}\}_{i\in I, t\geq 0}$, a sequence of prices, $\{r_t, w_t\}_{t\geq 0}$, and policies $T \equiv \{\{T_{it}\}_{i\in I}, T^o_t, \tau_{kt}\}_{t\geq 0}$ such that the household first order conditions and budget constraint, the firms' first order conditions, and the government's budget constraint are satisfied, the labor market clears ($\ell_t = 1$), and the resource constraint (6) and asset market clearing condition (7) are satisfied.

$$\sum_I \pi(c^y_{it} + c^o_{it}) + k_{t+1} = F(k_t, 1) + (1-\delta)k_t \tag{6}$$

$$k_{t+1} = \sum_I \pi a_{it} \tag{7}$$

I define the set of all *feasible* allocations, $\mathcal{X}^f$ as the set of allocations that satisfy the resource constraint (6). I define the set of all *implementable* allocations, $\mathcal{X}^I$ as the set of allocations for which prices and policies exist that implement the allocation as an equilibrium. Let $\mathcal{X}^f_s$ denote the set of all feasible steady state allocations and $\mathcal{X}^I_s$ be the set of all implementable steady state allocations.

2.2.2 Characterization.

In Appendix A.2, I establish several properties of equilibrium in the simple model. In particular, I show that the equilibrium capital each period can be expressed as a unique function of the previous periods' capital and policy. When the sequence of policy converges, the equilibrium converges to a unique steady state. Furthermore, equilibrium capital is a non-increasing function of net transfers from the high-productivity households to the low-productivity households and a non-increasing function of net transfers from the young to the old. These properties are summarized in Lemma 2.

Lemma 2 *Equilibrium capital in period t+1 is a unique function,* $k^E_{t+1}(k_t, T^y_{Lt}, T^y_{Ht}, T^o_{t+1}, \tau_{kt+1})$, *which is non-increasing in* $T^y_{Lt} - T^y_{Ht}$ *and non-increasing in* T^o_{t+1}. *Furthermore, if the sequence of policy converges to* $\{T^y_L, T^y_H, T^o, \tau_k\}$, *then the equilibrium converges to a unique steady state.*

For a proof, see Appendix A.2.

Intuitively, marginal propensities to save are weakly higher for type-H households, implying that savings and investment are decreasing in net transfers from type-H to type-L households. Old households do not save, and therefore budget-balanced transfers to the old also decrease the savings rate, and therefore investment in next period's capital.

2.3 The first-best benchmark.

Here, I consider the problem of an unconstrained planner, who chooses among all feasible allocations, $\mathcal{X}^f$ to maximize the following utilitarian social welfare function with Pareto weights $\gamma^t\lambda_i$ on type-i households born in year t for some $\gamma \in (0,1)$. The solution to this problem is the first-best allocation, which characterizes the ideal degree of intra-generational equality and the modified Golden Rule capital stock.

2.3.1 Unconstrained planner's problem.

The planner's objective function is given by equation (8).

$$\sum_{t=0}^{\infty}\gamma^t\sum_{i\in I}\lambda_i\pi\left(\frac{(c_{it}^y)^{1-\sigma_y}}{1-\sigma_y}+\beta_i\frac{(c_{it+1}^o)^{1-\sigma_o}}{1-\sigma_o}\right)+\gamma^{-1}\sum_{i\in I}\beta_i\lambda_i\pi\left(\frac{(c_{i0}^o)^{1-\sigma_o}}{1-\sigma_o}\right) \tag{8}$$

The unconstrained planner's problem is to choose an allocation, defined as a sequence $\mathcal{A} \equiv \{\{c_{it}^y, c_{it}^o\}_{i\in I}, k_{t+1}\}_{t\geq 0}$ in order to maximize (8) subject to the resource constraints (9) and (10), non-negativity constraints, as well as an initial capital stock, k_0.

$$k_{t+1} = k_t(1-\delta) + y(k_t, \ell_t) - \pi\sum_{i\in I}(c_{it}^y + c_{it}^o) \tag{9}$$

$$k_{t+1} = \pi\sum_{i\in I} a_{it} \tag{10}$$

$$c_{it}^y \geq 0, \;\; c_{it}^o \geq 0, \;\; k_{t+1} \geq 0$$

Assumption. Let $\lambda_H \geq \lambda_L$ and $\gamma \in (0,1)$.

The above assumption on the Pareto weights ensures that the planner's social welfare function is bounded and – for simplicity – rules out a planner that prefers an unequal distribution of resources that 'reverses' the inequality present in the laissez-faire competitive equilibrium. That is, the planner can desire for the type-L households to be *as* well off as the type-H households, but not *more* well off. The solution to the unconstrained planner's problem is given in the following Proposition.

Proposition 1 *(The first-best allocation)*
Let $\mathcal{A}_u \equiv \{\{\bar{c}^y_{it}, \bar{c}^o_{it}\}_{i\in I}, \bar{k}_{t+1}\}_{t\geq 0}$ *be the solution to the unconstrained planner's problem. This allocation is characterized by the first best level of intra-generational equality (11), and the first best sequence of capital (12) for all* $t \geq 0$.

$$\left(\frac{\bar{c}^y_{Ht}}{\bar{c}^y_{Lt}}\right)^{\sigma_y} = \frac{\lambda_H}{\lambda_L} \tag{11}$$

$$\frac{1}{\gamma}(\bar{c}^y_{it})^{-\sigma_y} = (\bar{c}^y_{it+1})^{-\sigma_y}(f_k(\bar{k}_{t+1}, 1) + (1-\delta)) \tag{12}$$

For a proof, see Appendix A.3.

At the first-best allocation, consumption shares are set so that the ratio of marginal utility between the high and low types equals the ratio of their Pareto weights. That is, the marginal benefit of redistributing to one type is exactly offset by the cost of redistributing away from the other type. Note that the steady state version of equation (12) implicitly defines the Modified Golden Rule capital stock, k^{gr}.

$$\frac{1}{\gamma} = 1 - \delta + f_k(k^{gr}, 1) \tag{13}$$

Corollary 1 *If the government faces no political constraints on inter-generational transfers, then the first best allocation is implementable,* $\mathcal{A}_u \in \mathcal{X}^I$.
For a proof, see Appendix A.4.

When the government has the ability to transfer resources from the current old to the current young in an unrestricted manner, then they can use intra-generational transfers to achieve the first-best allocation within a generation and inter-generational transfers to achieve the first-best sequence of capital. However, when inter-generational transfers are constrained, a redistribution-investment trade-off emerges.

2.4 The redistribution-investment trade-off.

2.4.1 The constrained optimal allocation.

In this section, I consider the problem of the constrained fiscal authority who chooses among the implementable allocation, $\mathcal{X}^I$ in order to maximize social welfare (8). Using the results from Lemma 2, I can write the fiscal authority's problem in the following way.

Definition. The constrained fiscal authority's problem is:

$$\max_{\{c_{it}^y, c_{it}^o, k_{t+1}\}_{t\geq 0}} \sum_{t=0}^{\infty} \gamma^t \sum_{i\in I} \lambda_i \pi \left(\frac{(c_{it}^y)^{1-\sigma_y}}{1-\sigma_y} + \beta_i \frac{(c_{it+1}^o)^{1-\sigma_o}}{1-\sigma_o} \right) + \gamma^{-1} \sum_{i\in I} \beta_i \lambda_i \pi \left(\frac{(c_{i0}^o)^{1-\sigma_o}}{1-\sigma_o} \right)$$

subject to the political constraints (4) and (5), and implementability, $\mathcal{A} \in \mathcal{X}^I$.

The solution to the constrained planner's problem is characterized in Proposition 2. This Proposition provides sufficient conditions for the existence of a trade-off between redistribution and investment. That is, a trade-off between the ideal inter- and intra-generational income distributions.

Proposition 2 *(Redistribution investment trade-off)*
Define the constrained optimal allocation, $\mathcal{A}^* \equiv \{\{c_{it}^{y*}, c_{it}^{o*}\}_{i\in I}, k_{t+1}^*\}_{t\geq 0}$ as the solution to the constrained fiscal authority's problem and the full equality steady state level of capital, k_f as the steady state level of capital where T_L^y and T_H^y are set such that $\left(\frac{c_H^y}{c_L^y}\right)^{\sigma_y} = \frac{\lambda_H}{\lambda_L}$, $\tau_k = 0$, and T^o is set at the political constraint.

(1) If $\beta_L = \beta_H$ and $\sigma_y = \sigma_o$, then $\frac{(c_{Ht}^{y*})^{\sigma_y}}{(c_{Lt}^{y*})^{\sigma_y}} = \frac{\lambda_H}{\lambda_L}$ for all $t \geq 0$.

(2) If either (a) $\beta_H > \beta_L$ or (b) $\sigma_y > \sigma_o$, then if

i) $k^{gr} > k_f$, $\exists\ \tau$ s.t. $\frac{(c_{Ht}^{y*})^{\sigma_y}}{(c_{Lt}^{y*})^{\sigma_y}} > \frac{\lambda_H}{\lambda_L}$ for $t \geq \tau$

ii) If $k^{gr} > k_f$ and $k_0 < k^{gr}$, then $\frac{(c_{Ht}^{y*})^{\sigma_y}}{(c_{Lt}^{y*})^{\sigma_y}} > \frac{\lambda_H}{\lambda_L}$ for $t \geq 0$

For a proof, see Appendix A.5

Part (1) of Proposition 2 states that if discount factors are uniform and preferences are homothetic, then the fiscal authority should set intra-generational redistribution policy to achieve an allocation with the ideal level of inequality. Intuitively, because

MPS are uniform, intra-generational redistribution cannot be used to alter the capital stock, even if the planner would wish to do so.

If however, either high-productivity households are more patient or preferences are non-homothetic, the planner is *able to* trade off intra-generational equality for a higher capital stock. Whether doing so is *desirable* depends on whether trying to achieve the optimal *intra-generational* income distribution results in the capital stock being below the Golden Rule level, even after the planner has exhausted their ability to boost capital in a non-distortionary way through inter-generational transfers.[12]

Consider a planner who implements the first-best level of intra-generational equality — setting $\left(\frac{c^y_{Ht}}{c^y_{Lt}}\right)^{\sigma_y} = \frac{\lambda_H}{\lambda_L}$ in every period — using only lump-sum transfers and no capital taxes. In the Appendix, I show that this implies that the economy will eventually converge to the full-equality steady state. If after pushing transfers from younger savers to old spenders as low as possible, k_f is still lower than the Golden Rule capital stock, k^{gr}, then the planner will eventually find it optimal to sacrifice some intra-generational equality (along with inter-temporal distortions) in order to gain additional investment.

Part (2) of Proposition 2 states that if capital in the full equality steady state is below the Golden Rule capital stock, then there exists a finite period τ such that the constrained optimal allocation will feature more intra-generational inequality than the first best allocation for all $t > \tau$. Furthermore, if the *initial* capital stock is also below the Golden Rule level, then the constrained optimal allocation will feature more inequality than first best in every period. Intuitively, the fiscal authority is balancing the benefits of redistribution against the costs of lower investment and the costs of inter-temporal distortions generated by the capital tax/subsidy.

2.4.2 Numerical Illustration.

Consider the following simple numerical example in which $\theta_H = 3$, $\theta_L = 1$, $\alpha = .3$, $\sigma_y = 2.5$, $\sigma_o = 1.5$, $\beta_L = \beta_H = .96$, $\delta = .12$ and $\zeta = 0$ (Cobb-Douglas production). Figure 1 plots the Pareto Frontier for the steady state of this model. On the y-axis, I plot the implied Pareto weight on future generations that rationalizes a particular level of steady state capital using equation (13). On the x-axis I plot the implied ratio of Pareto weights on high and low types that rationalizes a particular division

[12] By transferring resources lump-sum from the current old to the current young, the planner can increase the savings rate without resorting to distortionary taxes.

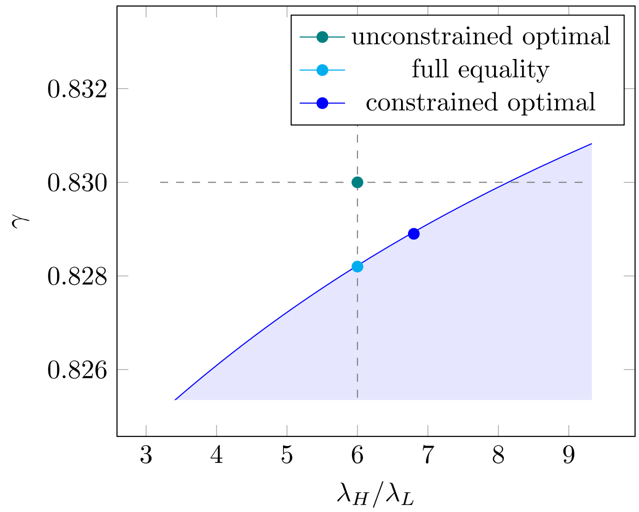

Figure 1: The Pareto Frontier

of steady state resources between types. I choose an arbitrary $\gamma = .83$ and $\lambda_H/\lambda_L = 6$ for the social welfare function, and plot the implied unconstrained optimal allocation.

The shaded blue region corresponds to the portion of the Pareto Frontier that is implementable using non-distortionary policy.[13] The boundary of this region corresponds to the case in which net transfers between the young and old are pushed to the political constraint: inter-generational transfers are set to 0. As the degree of within-generation redistribution increases, aggregate savings and steady state capital declines, implying a lower γ. For a given degree of intra-generational inequality, any allocation with a *lower* steady state capital stock – and therefore a lower implied γ – is possible. To see this, note that lump-sum transfers from the current young to the current old – which lower the capital stock – are unrestricted.

Figure 1 presents a case corresponding to part (2) of Proposition 2 in which the steady state capital stock that results from implementing the first-best level of inequality (the 'full-equality' steady state) is below the golden rule capital stock. In order to plot the solution in two-dimensions, I solve numerically for and report the constrained optimal allocation for a case in which τ_k is restricted to be 0. In this case, the constrained government trades off intra-generational equality with investment in the steady state capital stock. The true solution would involve a three-way tradeoff between equality, capital accumulation, and inter-temporal distortions.[14]

[13]Being on the Pareto frontier rules out inter-temporal distortions, implying $\tau_k = 0$.

[14]Note that the intertemporal distortions introduced by capital taxes/subsidies mean that they

2.5 What determines the size of this trade-off?

The size of the PI redistribution-investment trade-off depends on what microfoundations are used to generate heterogeneous MPS out of permanent income and on the general equilibrium impact of the decline in savings supply on capital investment.

2.5.1 Micro-foundations for heterogeneous MPS.

Type vs. scale dependence. In the simple model presented above, I show that differences in MPS out of permanent income can be generated by both heterogeneous time-preferences and by non-homothetic preferences. While the *existence* of the redistribution-investment trade-off does not depend on the micro-foundation, the *size* of the trade-off does.

If preferences are homothetic and high-productivity households simply have higher discount factors, β_i, then their high MPS are less sensitive to changes in the distribution of after-tax permanent income. In this case, the marginal impact of additional redistribution on savings remains roughly constant. If instead discount rates were uniform, but preferences were non-homothetic, then high-productivity types would have higher MPS out of PI only *because their after-tax income is high.* Every additional dollar of redistribution compresses the income distribution, and therefore also compresses the distribution of MPS. This implies that the marginal impact of additional redistribution on aggregate savings decreases as the degree of redistribution increases, dampening the size of the trade-off.

This point is illustrated in Figure 2. In a numerical example, I consider two versions of the simple model. Both models are calibrated to match the same steady state aggregate capital stock and the same marginal propensities to save out of permanent income for the high and low types at some arbitrary initial steady state. The first model generates these marginal propensities to save using heterogeneous β_i, while the second uses non-homothetic preferences ($\sigma_y > \sigma_o$).

As in Figure 1, Figure 2 plots the portion of the Pareto Frontier that is implementable with fiscal policy. As before, the boundary of this region corresponds to the set of allocations that can be implemented by setting the inter-generational transfers to the political limit while varying the degree of intra-generational transfers. Starting at the initial allocation (around $\lambda_H/\lambda_L = 9$) the distribution of MPS are

cannot be used to resolve the redistribution-investment trade-off.

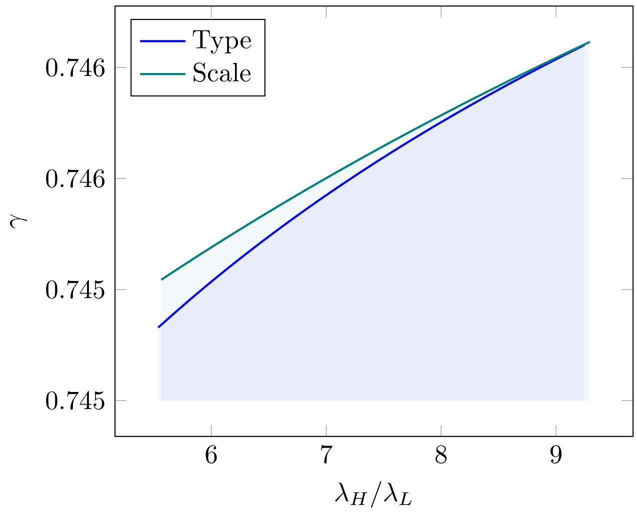

Figure 2: The Pareto Frontier: Type vs. Scale Dependence

identical by construction, and therefore the size of the local redistribution-investment trade-off are the same. Moving leftward, as the degree of redistribution increases, the distribution of MPS compresses in the scale-dependent model relative to the type-dependent model. Intuitively, in the type-dependent model, high income households save more because they are permanently more patient. In the scale-dependent model, high income households save more *because they have more income.* Greater redistribution therefore increases the MPS of the low-productivity types and decreases the MPS of the high-productivity types. This generates a relatively weaker trade-off with each additional degree of redistribution. This exercise suggests that pinning down the relative importance of type and scale dependent factors will be important when calibrating the quantitative model.

2.5.2 General equilibrium feedback.

The ultimate size of the trade-off does not just depend on the direct impact of redistribution on savings, but also on *what would have been with those savings.* That is, when aggregate savings fall and borrowing costs increase as a result, does this primarily crowd out capital investment or household debt? Equivalently, does the rise in borrowing costs crowd in household savings via the *substitution effect*?

To understand what features of the economic environment determine the general equilibrium feedback effects of redistribution on investment, it is useful to examine the total derivative of capital with respect to net transfer from the high to low productivity types. This total derivative is reported in Lemma 3.

Lemma 3 *The general equilibrium derivative of equilibrium capital with respect to net transfers between type-H and type-L households is given by equation (14).*

$$\frac{dk_{t+1}}{dT_{Lt}} = \frac{\pi MPS_{Lt} - \pi MPS_{Ht}}{1 - (k_{t+1}^{r_{t+1}})^{-1}\mathcal{A}_t^{r_{t+1}}} \tag{14}$$

where MPS_{Lt} and MPS_{Ht} are the marginal propensities to save out of permanent income for type-L and type-H households respectively, $\mathcal{A}_t^{r_{t+1}} = \sum_I \frac{a_{it}}{k_{t+1}} \frac{d\log a_{it}}{d\log r_{t+1}}$ is the asset-weighted interest rate elasticity of household savings and $k_{t+1}^{r_{t+1}} = \left(\frac{d\log r_{t+1}}{d\log k_{t+1}}\right)^{-1}$ is the partial equilibrium elasticity of firms' capital demand with respect to the interest rate.

For a proof, see Appendix A.6.1.

Lemma 3 shows that anything that lowers the interest rate elasticity of firm demand for capital will decrease the size of the trade-off. Intuitively, if firms are slow to adjust their capital stock in response to rising borrowing costs, the decline in aggregate savings following redistribution is absorbed primarily through higher interest rates rather than lower investment, leaving the capital stock largely unchanged. At the same time, if the interest rate elasticity of household savings increases, this also decreases the size of the trade-off. This is the mechanism at work in Mian et al. (2021). In their model as in mine, higher marginal propensities to save out of permanent income imply that greater income inequality generates an increase in savings that pushes down rates of return. In their baseline model however, there is no capital so the decline in interest rates simply generates more indebtedness among lower income households through the substitution effect. From equation (14) we can see that the ultimate general equilibrium effect of changing the income distribution depends on the relative strength of the interest rate elasticity of households compared to that of firms.

In Appendix A.6.3 I show that the effect of capital-skill complementarity on the PI-redistribution channel are ambiguous. If low-skill households are more substitutable with capital but have higher social welfare weights, the welfare losses from a given decline in steady state capital are less severe. However, because high-income households are complementary with capital, redistribution will generate a larger equilibrium loss in steady state capital. The ultimate effect depends on the relative magnitude of these two forces. Furthermore, I examine how the size of the trade-off changes when differences in MPS are generated by type-dependent rates of return in Appendix A.6.5.

Open economy. Consider an open economy variant of the simple model. Let $\rho \in (0,1)$ be the domestic population share of the total world population. Define the foreign asset share, $\frac{f_t}{k_t}$ and the interest rate elasticity of foreign savings, $\mathcal{A}_r^f$. In this case, equation (14) becomes the following.[15]

$$\frac{dk_{t+1}}{dT_{Lt}} = \frac{\rho\Bigg(\pi MPS_{Lt} - \pi MPS_{Ht}\Bigg)}{1-(k_{t+1}^{r_{t+1}})^{-1}\left(\frac{(1-f_{t+1})}{k_{t+1}}\mathcal{A}_t^{r_{t+1}} - \frac{f_{t+1}}{k_{t+1}}\mathcal{A}_r^f\right)}$$

In this case, as the interest rate elasticity of foreign savings increases, the trade-off between redistribution and investment decreases. Intuitively, the decline in domestic savings following an increase in redistribution pushes up borrowing costs, but this spurs additional investment from abroad, dampening the impact on domestic capital. At the same time, as the relative size of the domestic economy, ρ decreases, the trade-off vanishes as domestic savings become less and less relevant for the domestic capital stock. In the small open economy limit ($\rho = 0$), the trade-off is eliminated.

3 Endogenous Labor

While the PI redistribution-investment trade-off is easiest to see in a model with non-distortionary lump-sum redistribution, this trade-off will be present for any policy that redistributes permanent income, whenever marginal propensities to save differ across the income distribution. To illustrate how this channel impacts the welfare costs of *distortionary* taxation, I alter the simple model to incorporate endogenous labor and replace the lump-sum redistribution policy with a linear tax on labor that funds a uniform lump-sum transfer. In this setting, I show that the marginal impact on welfare of the labor income redistribution policy can be decomposed into the familiar equity-efficiency trade-off, and, when MPS out of permanent income are higher

[15] This follows from modifying the asset market clearing condition to include foreign savings, and re-deriving the total derivative of capital with respect to T_{Lt}.

for high-income households, a new term capturing the PI redistribution-investment trade-off.

3.1 Environment.

Households. The household sector is the same as before, but household lifetime utility is now given by equation (15).

$$u(c^y_{it}, c^o_{it+1}, \ell_{it}) = \frac{(c^y_{it})^{1-\sigma_y}}{1-\sigma_y} - \psi_i \frac{(\ell_{it})^{1+\nu_i}}{1+\nu_i} + \beta_i \frac{(c^o_{it+1})^{1-\sigma_o}}{1-\sigma_o} \tag{15}$$

Note that the parameters of the labor disutility term are type-specific. The household period budget constraints are now given by the following.

$$c^y_{it} = w_t\theta_i(1-\tau_\ell)\ell_{it} + T_t + \tau_\ell w_t \ell_t - a_{it}$$
$$c^o_{it} = a_{it-1}(1+r_t)$$

Government. The fiscal policy from Section 2 is replaced with a linear labor income tax, $\tau_{\ell t}$ that funds a uniform lump-sum transfer, T_t. The government's budget constraint is now simply:

$$\tau_{\ell t} w_t \ell_t = T_t$$

The firm sector is unchanged from Section 2.

As in the case with exogenous labor, I can repeat the exercise in Lemma 3 and solve for the general equilibrium impact of an increase in the linear labor income tax, here on *steady state* aggregate capital.

Lemma 4 *Define $\epsilon_{\tau\ell}$ as the general equilibrium elasticity of aggregate labor supply with respect to τ_ℓ and $\epsilon_{i\tau_\ell}$ is the general equilibrium elasticity of type-i labor supply. Then the semi-elasticity of steady state capital with respect to τ_ℓ, $\frac{dk}{d\tau_\ell}\frac{1}{k}$ is given by the* *following expression.*

$$\frac{dk}{kd\tau_\ell} \propto \Bigg(\underbrace{\sum_I MPS_i\Big(1-\frac{\theta_i\ell_i}{\ell}\Big)}_{\text{Effect of PI Redistribution } (\mathcal{K}_D)} + \underbrace{\bar{MPS}\cdot\epsilon_{\tau_\ell} + Cov\Big(MPS_i, \epsilon_{i\tau_\ell}\Big) - A_r w_\ell \epsilon_{\tau_\ell}}_{\text{Feedback Effect of Distorted Labor } (\mathcal{K}_I)} \Bigg) \cdot \Big(1 - A_r k_r^{-1} - A_w w_k\Big)^{-1} \tag{16}$$

Here, $\bar{MPS}$ is the average marginal propensity to save out of permanent income. I denote x_y as the elasticity of variable x with respect to y.

For a proof, see Appendix A.7.1.

Relative to the exogenous labor case, the effect of redistribution on capital investment consists of both a direct effect, $\mathcal{K}_D$, as well as an indirect effect on savings resulting from changes in labor supply. The direct effect captures the redistributive effect of the tax and transfer, holding labor supply constant. The labor income tax and uniform transfer redistributes income from households who earn above-average labor income to households who earn below-average labor income. If households with above average labor income have higher marginal propensities to save, the direct effect implies a decline in capital investment. Put differently, the direct effect reflects the change in capital that would have occurred if households faced a lump-sum tax or transfer of income equivalent to the direct effect of the labor income tax and uniform transfer, keeping their labor supply fixed. As in the exogenous labor case, the size of this decline is determined by the interest rate elasticity of household savings and firm investment demand, and now by the elasticity of steady state wages to capital.

The indirect effect captures the effect of changes in the supply of savings as a result of distorted labor supply. A one percentage point increase in τ_ℓ decreases aggregate labor income by ϵ_{τ_ℓ} percent, generating a decline in savings proportional to the average MPS. If those with higher MPS have a greater labor income response to taxation, this effect is amplified. Furthermore, a decline in aggregate labor lowers the steady state wage, which reduces labor income and therefore savings further.

3.2 Welfare impact of labor income redistribution.

Consider the following *steady state* utilitarian social welfare function given in equation (17). This social welfare function weights the period utility of two representative generations – the current young and the current old – discounting the welfare of the current young by γ as in Section 2.

$$SW \equiv \sum_{i \in I} \pi \lambda_i \left(\gamma \left(\frac{(c_i^y)^{1-\sigma_y}}{1-\sigma_y} - \psi_i \frac{\ell_i^{1+\nu_i}}{1+\nu_i} \right) + \beta_i \frac{(c_i^o)^{1-\sigma_o}}{1-\sigma_o} \right) \tag{17}$$

Proposition 3 presents the marginal impact on steady state social welfare of a small (permanent) increase in labor income redistribution, τ_ℓ.

Proposition 3 *Let the elasticity of wages with respect to capital and labor as $\epsilon_{w,k}$ and $\epsilon_{w,\ell}$. Define a type-i household's social welfare weight, $\kappa_i \equiv \lambda_i (c_i^y)^{-\sigma_y}$. Assuming that the economy is in steady state and that the λ_i are scaled such that $\sum_I \kappa_i = 1$, then the marginal change in social welfare from a permanent increase in τ_ℓ is given by equation (18).*

$$\frac{dSW}{d\tau_\ell} \propto \underbrace{\sum_I \kappa_i \left(1 - \frac{\theta_i \ell_i}{\ell}\right)}_{\text{Equality}} + \underbrace{\epsilon_\ell}_{\text{Labor Distor.}} + \underbrace{\Delta \epsilon_{w,k} \mathcal{K}_D}_{\text{PI Redistr.}} \tag{18}$$

where Δ is defined as in equation (19).

$$\Delta = \sum_I \kappa_i \left((1-\tau_\ell) \frac{\theta_i \ell_i}{\ell} + \tau_\ell - \frac{a_i}{\gamma k (1+r)} \right) \tag{19}$$

Here, $\mathcal{K}_D$ is the direct effect of permanent income redistribution on investment derived in Lemma 4 and ϵ_ℓ is the product of the elasticity of social welfare weighted total after-tax income with respect to aggregate labor and the elasticity of aggregate labor with respect to the tax.[16] *For a proof and the full expression for ϵ_ℓ, see Appendix A.7.*

From equation (18), we see that the marginal impact on social welfare can be decomposed into three components. The first two correspond to the 'classic' equity-efficiency trade-off. The first term captures the direct redistributive effect of the tax

[16]The complete expression for ϵ_ℓ is given in Appendix equation (A.27).

and transfer. Households face the same marginal tax rate, but differing average tax rates, which implies the tax transfers resources from households with higher than average labor earnings to households with lower than average labor earnings. These transfers are weighted by the social welfare weight of each type.

The second term includes the welfare impact of the distortionary tax through changes in after-tax labor earnings.[17] This term corresponds to the elasticity used in Piketty and Saez (2013a), and captures both the weighted change in labor income through the change in wages induced by falling labor supply, as well as the decline in the lump-sum transfer induced by the shrinking tax base. In a static model as in Sheshinski (1972) or Piketty and Saez (2013a) with uniform marginal propensities to consume out of permanent income and no capital, these two channels characterize the full trade-off.[18] When capital is added, labor supply distortions also indirectly affect rates of return as well as savings behavior and the capital stock.

Finally, when $\beta_H > \beta_L$ or $\sigma_y > \sigma_o$ – and therefore MPS are higher for high-productivity households – an additional term emerges. This term is the product of (i) the direct impact of a change in the permanent income distribution on capital, $\mathcal{K}_D$ and (ii) the welfare impact of additional capital, $\Delta\epsilon_{w,k}$. Intuitively, an increase in steady state capital boosts the after-tax wage while decreasing the rate of return, shifting income towards younger wage-earning households, particularly those who earn a relatively larger share of their income in the labor market, while also raising lump-sum transfers for all young households.[19] Because the households who benefit from the increase in wages and transfers are members of the next generation, this increase in welfare is discounted by γ.

It is instructive to examine the Δ term for a special case of the model: a steady state without ex-ante heterogeneity and in which the current $\tau_\ell = 0$. In this case, Δ collapses to a modified golden rule capital stock condition as in Atkinson and Sandmo (1980). In this special case, if $1 + f_k - \delta > \frac{1}{\gamma}$, then $\Delta > 0$. Intuitively, when the marginal product of capital is sufficiently large relative to its depreciation rate, more capital increases the size of the aggregate consumption "pie" for future generations. In a model with ex-ante heterogeneity and redistributive taxes, Δ captures this effect,

[17] Note that there are no direct welfare effects of changing households' individual labor supply, as households are already optimizing.

[18] Indeed, the redistribution term and ϵ_ℓ correspond exactly to the terms in the simple model in Piketty and Saez (2013a).

[19] This is exactly the same redistribution force present in Dávila et al. (2012).

along with the additional redistributive effects of capital through changes in factor prices and T. This insight reflects the same phenomenon at work in Proposition 2. In the simple model, for a trade-off to exist between permanent income redistribution and capital, it was sufficient for the full-equality steady state to be dynamically efficient. Similarly here, if the current steady state is dynamically efficient, then more capital is *desirable*, ensuring that the permanent income redistribution channel is welfare-reducing and a trade-off between redistribution and investment exists.

3.3 Numerical illustration.

A natural question is whether the permanent income redistribution channel is large relative to the labor supply distortion channel, as the latter has been the main focus of the literature on optimal labor income taxation. Proposition 3 can be used to generate a back-of-the-envelope estimate of the relative importance of the PI redistribution channel for small increases in labor income redistribution. Given the difficulty in estimating the indirect effects of the labor income tax on capital accumulation and rates of return, I focus on comparing the permanent income redistribution channel to the *direct* effect of the labor supply distortion on after-tax labor income. This will allow me to make a direct comparison between this channel and the efficiency costs of linear labor income redistribution from other optimal-tax formula papers. From equation (18), we see that the size of the channel depends on both steady state moments and elasticities as well as *subjective* welfare weights, κ_i and γ. In Table 1, I report the values of the model moments and elasticities used in my illustration. I then report the relative size of the permanent income redistribution channel to the labor earnings distortion channel for various values of the subjective welfare weights.

3.3.1 Values used in numerical exercise.

For several moments, I choose standard values common in the literature. In particular, I assume a real net interest rate of 5 percent and a capital-to-output ratio of 2. I set $\tau_\ell = .35$ as in Piketty and Saez (2013a). For the interest rate elasticity of firm investment, I consider estimates ranging from 0 to -0.65. The latter is the long run elasticity implied by a Cobb-Douglas production function with a labor share of .65. For the elasticity of wages with respect to capital and labor, I use .35 and -.35 respectively, the values implied by a Cobb-Douglas production function. For ϵ_ℓ, I follow Saez et al. (2012) and consider a range between .1 and .25.

Income Distributions and welfare weights. For the back-of-the-envelope measure of the distribution of capital and labor income, I use analysis from the Congressional Budget Office (CBO) on the composition of income by quintile for the United States in 2019. The CBO uses data from the Statistics of Income, a nationally representative sample of individual tax returns from the IRS. [20]

Interest rate elasticity of household saving. Estimates of the elasticity of household saving with respect to the interest rate generally fall into one of two categories. The first is structural. Instead of estimating the interest rate elasticity of savings, researchers estimate – or use existing estimates of – the elasticity of inter-temporal substitution (EIS), on which the interest rate elasticity of household savings closely depends.

These studies then use a structural model to show how estimates of the EIS translate into savings elasticities depending on the values chosen for the other parameters. Examples of this approach include Attanasio and Wakefield (2010) who consider estimates of the EIS from .25 to 1. Using their preferred set of assumptions in a detailed life cycle model, they find that a half percentage point increase in r results in a 10 percent increase in new savings, or an implied A_r of .2. Using a similar procedure, Evans (1983) finds a range of estimates of A_r between .1 and 3.55, with most estimates falling between .1 and 1 depending on assumptions made about household discount rates, growth rates, and the EIS itself.

The second category includes reduced form estimates. Jappelli and Pistaferri (2007) find that changes in after-tax interest rates had no effect on demand for mortgage debt, implying a near-zero elasticity. DeFusco and Paciorek (2017) find that a 1 percentage point increase in the interest rate reduced mortgage debt by between 1.5 and 2 percent. On the larger end of the range, Best et al. (2020) estimate a reduced borrowing elasticity of .5, while Dunsky and Follain (2000) estimate an elasticity of 1. Given the wide range of estimates from the empirical literature, I consider values between .1 and 1.

Marginal propensities to save. I use the consumption elasticities estimated in Straub (2019) to construct the distribution of MPS out of permanent income. In that paper, he uses PSID data to estimate the elasticity of household consumption

[20] For details on the calculation of the income distribution see Appendix A.8 and https://www.cbo.gov/publication/58781.

Moment	Description	Value
ϵ_ℓ	Elas. income to $\ell \cdot \epsilon_\ell, \tau$	[-.1,-.25]
τ_ℓ	Average labor tax rate, US	0.35
$w\ell/k$	labor share divided by capital-output ratio	0.65/2
k_r	Interest rate elasticity of k	[-0.05,-0.65]
A_r	Average interest rate elasticity of savings	[0.10,1.00]
A_w	Average wage rate elasticity of savings	1.30
r	Real net interest rate	0.05
w_k	Elasticity of wage with respect to k	0.35
$\{\pi\theta_i\ell_i/\ell\}_{i\in I}$	Type-i labor income share	$\{0.05, 0.15, 0.15, 0.15, 0.50\}$
$\{\pi a_i/k\}_{i\in I}$	Type-i capital income share	$\{.05, .08, .08, .08, .71\}$
$\{MPS_i\}_{i\in I}$	Distribution of MPS	{.17,.23,.34,.43,.57}
σ_y	Inverse IES	2

Table 1: Values used in numerical exercise

to permanent income. His preferred estimate is .7, implying an approximate savings elasticity of around 1.3. Indeed, I use his implied value for the permanent income elasticity of savings for A_w. Because $\frac{\partial \log a_i}{\partial \log PI_i} = \frac{\partial a_i}{\partial PI_i}\frac{PI_i}{a_i}$, the savings elasticities implied by Straub (2019) can be combined with estimates of savings *rates* across different PI groups to recover the MPS out of PI across the income distribution. To generate estimates of savings rates by permanent income quintile, I rely on the methodology presented in Dynan et al. (2004) (DSZ), and re-estimate their specification on updated data from the Panel Study of Income Dynamics (PSID). Details can be found in Appendix A.9.

3.3.2 Comparative statics.

Using the values in Table 1, I begin by calculating the *direct* effect of the labor income redistribution on next period's capital. Table 2 reports the percent decline in next period's capital associated with the *direct* redistributive effect of a one percentage point increase in τ_ℓ. In other words, the table reports $\mathcal{K}_D$. For larger values of k_r, firm capital investment is more sensitive to borrowing costs, amplifying the effect of redistribution on capital accumulation. At the same time, as A_r increases, savings are more responsive to rising interest rates, dampening the effect.

I use these estimates to compare the relative size of the PI redistribution channel and the labor distortion channel for various values of welfare weights and moments. In particular, I am interested in how sensitive the relative strength of the PI redistri-

k_r	$A_r = 0.1$	$A_r = 0.5$	$A_r = 1.0$
$k_r = -0.050$	-0.06	-0.01	-0.01
$k_r = -0.325$	-0.18	-0.08	-0.04
$k_r = -0.650$	-0.22	-0.12	-0.08

Table 2: Direct Effect of PI Redistribution on Investment ($\mathcal{K}_D$).

bution channel is to (i) changes in the weight on future generations, γ; (ii) variance in the social welfare weights of each type, $\kappa_i \equiv \lambda_i(c_i^y)^{-\sigma_y}$, and therefore in the strength of the redistribution motive; and (iii) changes in key moments/elasticities whose values are not settled in the empirical literature.[21]

I consider both 'laissez-faire' weights in which $\kappa_i \equiv \kappa_i^L = 1/5$ for all types, and 'egalitarian' social welfare weights in which $\kappa_i \equiv \kappa_i^E = (y_i)^{-\sigma_y}/\bar{\lambda}$, where y_i correspond to each quintile's share of total income and $\bar{\lambda}$ is a normalizing constant set to ensure that the weights sum to 1.[22] I use the estimated distribution of MPS out of permanent income. I assume $A_r = .1$ and $k_r = -.65$.

Figure 3 plots the ratio of the welfare cost of the PI redistribution channel to the welfare cost of the direct labor distortion channel, for various values of γ and welfare weights. The ratio is straightforwardly increasing in γ, as the more weight the planner puts on future generations, the greater the relative impact of the PI redistribution channel. This ratio is also higher when the planner uses egalitarian welfare weights that reflect the unequal distribution of resources, rather than laissez-faire weights that rationalize the existing distribution. Intuitively, a higher savings rate and capital stock in future periods pushes up wages and pushes down rates of return, disproportionately benefiting households who earn a greater share of their income from wages rather than capital — in the data, lower income households.

From the figure, the PI redistribution channel is about 26 percent of the labor earnings distortion channel when welfare weights are egalitarian and the labor earnings elasticity is towards the bottom of the range estimated by Saez et al. (2012), while the labor distortion channel dominates when the elasticity is largest in magnitude. This simple back-of-the-envelope exercise shows that this channel may be meaningful enough to matter when determining optimal labor income tax formulas. In the next section, I solve a richer life-cycle model, and perform an analogous decomposition for

[21]Note that, as I vary welfare weights, I normalize so that they continue to sum to 1 to avoid mechanical changes in the size of each channel.

[22]The latter correspond roughly to the case in which $\lambda_i = 1$ for all types.

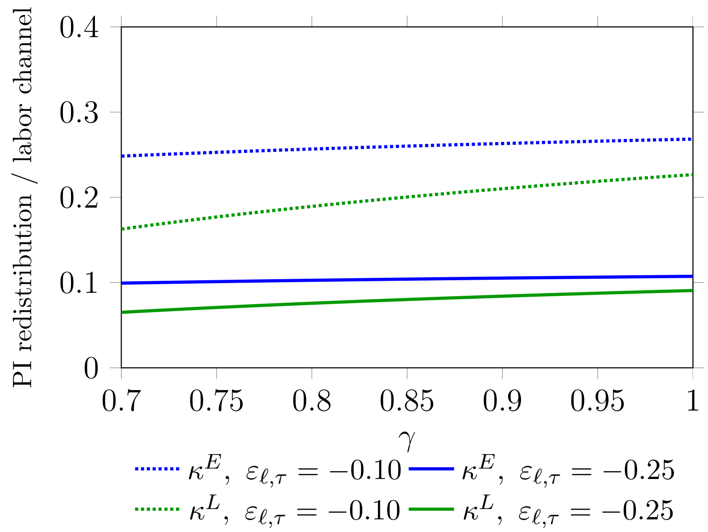

Figure 3: Ratio of the PI redistribution channel to the direct labor channel.

both small and large labor income tax reforms.

4 Quantitative Model.

In this section, I use the distribution of MPS out of permanent income derived in Appendix A.9 to calibrate a richer quantitative OLG model with a more realistic earnings life-cycle and uninsurable idiosyncratic labor productivity shocks. As in the simple model, to match the derived MPS out of permanent income I use a combination of non-homothetic preferences over lifetime consumption and type-dependent time preferences, but also allow for non-homothetic *warm glow* preferences over bequests as in De Nardi (2004) and Straub (2019).

I use this model to quantify the size of the trade-off between permanent income redistribution and investment in the long-run steady state. In particular, I study the same simple labor income redistribution scheme as in Section 3, a linear labor income tax that funds a lump-sum transfer, which is layered on top of a more realistic set of existing fiscal policies. In addition to redistributing permanent income from high to low productivity types and distorting labor supply, this tax impacts welfare by (i) mitigating risk exposure by providing insurance through redistributing resources from households with higher realized productivity to households with lower realizations, (ii) lowering precautionary savings and therefore aggregate capital, and (iii) redistributing resources from *older* higher-earning households to *younger* households, potentially impacting the aggregate savings rate and capital stock.[23]

I first decompose the direct and indirect effects of a small increase in the linear labor income tax on steady state capital. This decomposition is analogous to the one given in Lemma 4, where the direct effect captures only the effect of the change in the distribution of (in this case *expected*) permanent income on savings, holding labor supply constant at the previous steady state level. In the baseline model, I report the elasticity of total capital, total labor, and total consumption with respect to the labor income tax, as well as the share of the decline in capital, output, and consumption attributable only to the direct effect of the change in permanent income. I then solve for these elasticities in a model in which marginal propensities to save are targeted using only non-homothetic preferences (uniform discount rates), as well as a version of the model with capital-skill complementarity.

I then turn to the welfare effects of the redistributive tax for a range of values for the Pareto weights, analogous to the exercise in Figure 3. In particular, for both egalitarian and laissez-faire Pareto weights, I plot the decline in welfare attributable to the permanent income redistribution channel alongside the decline attributable to the labor distortion channel for a range of values for γ, the rate at which future generations are discounted.

4.1 Environment.

Households. There is a unit mass of households who each live for J periods. Households supply labor to firms in all but the final retirement period, in which they receive type-specific social security income SS_i. There is a constant fraction π_{ij} of age-j households of each permanent type $i \in I$. The permanent type determines households' expected labor productivity, as well as preference parameters like their discount rate, their dis-utility of labor, and their warm-glow bequest motive. A household born in year t who is age-j and type-i has labor productivity, $\theta_{ij}\epsilon_{i,t+j}$ where θ_{ij} is the permanent component of labor income for type-i age-j households and $\epsilon_{i,t+j}$ is a type-specific idiosyncratic labor productivity shock.[24]

[23]This channel is analogous to the change in the savings rate induced by other forms of intergenerational redistribution like pay-as-you-go social security or government debt, that redistribute resources between current generations with different savings rates.

[24]Note that the income shock process is itself type-specific.

Households receive $(1-\tau_{t+j}(w_t\ell_{ij,t+j}\theta_{ij}\epsilon_{i,t+j}))w_t\ell_{ij,t+j}\theta_{ij}\epsilon_{i,t+j}$ in after-tax labor income each period. The function $\tau_{t+j}(.)$ is a progressive labor income tax following Bénabou (2002) and Heathcote et al. (2017), given by equation (20). Here, $\bar{\tau}_\ell$ parameterizes the average level of the labor income tax, while γ_ℓ determines the degree of progressivity. I denote $\tau_{ij,t+j}$ as a type-i age-j household's progressive tax rate.

$$\tau_{ij,t+j} = 1 - (w_{t+j}\ell_{ij,t+j}\theta_{ij}\epsilon_{i,t+j})^{-\gamma_\ell}\bar{\tau}_\ell \tag{20}$$

Households can save or borrow in a one-period bond or capital at gross after-tax interest rate, $1+(1-\tau_K)r_{t+1}$. Households face a borrowing constraint $\underline{a}$ such that $a_{ij,t+j} \geq \underline{a}$. Type-i households receive share σ^i_π of aggregate profit flows each period, Π_t starting in middle age, as well as $a_{i0,t}(1+(1-\tau_K)r_t)(1-\tau_b)$ in after-tax bequest income when they are born. Note that all type-i households receive the same bequest transfer equal to the average type-i bequest the year before they are born, $a_{i0,t} = \bar{a}_{iJ,t-1}$.[25] Finally, households may receive a lump-sum transfer, T_t. Lifetime utility for a household born at time t is given by (21).

$$u(c_{ij,t+j},\ell_{ij,t+j},a_{iJ,t+J}) = \sum_{j=1}^{J}\beta_i^{j-1}\left(\frac{(c_{ij,t+j})^{1-\sigma_j}}{1-\sigma_j} - \psi^j_{\ell i}\frac{(\ell_{ij,t+j})^{1+\nu}}{1+\nu}\right) + \beta^J\psi_{ia}\frac{(a_{iJ,t+J}+\bar{a})^{1-\eta}}{1-\eta} \tag{21}$$

This model nests the same two micro-foundations for increasing MPS as in the simple model presented in Section 2, but adds in a non-homothetic warm-glow bequest motive. Whenever $\sigma_j > \sigma_{j+1}$ at all ages and $\sigma_J > \eta$, both consumption later in life and bequests are a luxury good. I follow Straub (2019) and include the term $\bar{a}$ in households' bequest motive, while also allowing ψ_{ia} to be type-dependent, in order to generate a mass of households who leave no bequests. Let $R^j_{t+j} = \prod_{\tau=0}^{j}(1+(1-\tau_k)r_{t+\tau})$. A type-i household born at time t has the following lifetime budget

[25] I make this simplifying assumption to avoid having to keep track of the history of idiosyncratic shocks across generations. This implies that labor productivity is the same for all members within a dynasty.

constraint, given by equation (22).

$$\frac{a_{iJ,t+J}}{R^J_{t+J}} + \sum_{j=0}^{J} \frac{c_{ij,t+j}}{R^j_{t+j}} = R_t a_{i0,t}(1-\tau_b) + \sum_{j=0}^{J-1} \frac{(1-\tau_{i,t+j}) w_{t+j} \ell_{ij,t+j} \theta_{ij} \epsilon_{i,t+j} + T_{t+j}}{R^j_{t+j}} + \sum_{j=1}^{J} \frac{\Pi_{t+j}\sigma^\pi}{R^j_{t+j}} + \frac{SS_i}{R^J_{t+J}} \tag{22}$$

Firms. There is a unit mass of monopolistically competitive intermediate goods firms indexed by $m \in [0,1]$ who rent labor and capital from households and produce a differentiated intermediate good, y^m_t according to a CES production function (23). A competitive final

$$y^m_t = Z_t \bigg(\alpha (k^m_t)^\gamma + (1-\alpha)(\ell^m_t)^\gamma \bigg)^{1/\gamma} \tag{23}$$

goods firm aggregates the intermediate goods using a Dixit-Stiglitz CES aggregator (24). This specification generates a standard expression for demand for each intermediate good (25). Because there are no nominal rigidities, intermediate goods firms all produce the same

$$Y_t = \left(\int_0^1 (y^m_t)^{\frac{\varepsilon-1}{\varepsilon}} dm \right)^{\frac{\varepsilon}{\varepsilon-1}} \tag{24}$$

$$y^m_t = Y_t \left(\frac{p^m_t}{P_t} \right)^{-\varepsilon} \tag{25}$$

level of output, employ the same labor and capital, and charge the same markup $\mu = \frac{\varepsilon}{\varepsilon-1}$ over marginal cost.

Government. The government taxes returns on capital, bequests, and labor income, issues debt B_t, pays social security benefits, and distributes uniform lump-sum transfers T_t. Government spending G_t is exogenous. The government's per-period budget constraint is given by the following expression.

$$B_t + \sum_I \sum_{j=0}^{J-1} w_t \int_0^1 \pi_{ij} \tau_{i,t} \theta_{ij} \epsilon_{i,t+j} \ell_{ij,t+j} + A_t r_t \tau_K + \sum_I \pi_{iJ} a_{iJt-1} \tau_b = T_t + r_t B_{t-1} + \sum_I \pi_{Ji} SS_i$$

Equilibrium. An equilibrium is defined as a sequence of prices, $\{r_t, w_t\}_{t\geq 0}$, individual and aggregate financial positions, $\{\{a_{ij,t+j}\}_{i\in I, j\in J}, A_t\}_{t\geq 0}$, policies, $\{\tau_t, \tau_{bt}, \tau_{kt}, B_t, G_t, T_t\}$, individual household and firm allocations, $\{\{c_{ij,t+j}, \ell_{ij,t+j}\}_{i\in I, j\in J}, \{y_t^m, n_t^m, k_t^m\}_{m\in[0,1]}\}_{t\geq 0}$, and aggregate allocation, $\{k_t, \ell_t, C_t\}_{t\geq 0}$ such that the following conditions hold: (i) Households' first order conditions and budget constraints hold for each generation and productivity type, at each age and history of shocks, (ii) the intermediate goods firms' first order conditions, production function and demand hold:

$$w_t \mu = \left(\frac{y_t^m}{\ell_t^m}\right)^{\frac{1}{\rho}} (1-\alpha) Z_t \tag{26}$$

$$(r_t + \delta)\mu = \left(\frac{y_t^m}{k_t^m}\right)^{\frac{1}{\rho}} \alpha Z_t \tag{27}$$

Finally, the government budget constraint, labor market clearing, asset market clearing, and resource constraint hold.

4.2 Calibration

One period in the model corresponds to 15 years. All flow variables (for example, rates of return, discount factors, depreciation rates, wage flows) are reported as annual values in the text. Annual output is normalized to 1. Net foreign assets are determined residually from asset market clearing, and in the calibration take a value of $NFA/K \approx -0.43$, reflecting the fact that calibrated household savings fall short of the sum of the domestic capital stock and government debt. The implied *annual* real rate of return is $r = 3.6\%$.

Firms. In the baseline model I assume $\gamma = 0$ so that the production function is Cobb-Douglas. I normalize aggregate output to 1 and set Z accordingly. The labor share of income, $1 - \alpha$, is set to 0.724 and the annual depreciation rate to 8.6%, consistent with a capital-to-annual-output ratio of 2.1 and an investment share of 18%. The markup is set to $\mu = 1.08$, implying a profit share of 7.5% and a labor share net of profits of 0.67.

Government Policy. I parameterize the labor income tax function following Heathcote et al. (2017) (HSV), setting the level parameter $\bar{\tau} = 0.80$ and the progressivity parameter $\gamma_\ell = 0.181$. I follow De Nardi (2004) and Straub (2019) and set the bequest tax to $\tau_b = 0.10$ and the capital income tax to $\tau_K = 0.15$. Social security payments are set by income quintile following Huggett and Ventura (2000) and Straub (2019). Lump-sum transfers are initially set to 0. Government debt relative to GDP is set to $B/Y = 1.05$, consistent with the ratio of publicly held federal debt to GDP in 2019. Government spending G is calibrated residually from the government budget constraint, resulting in $G/Y = 0.22$.

Household Income. Households live for 60 years (4 periods of 15 years each), from age 25 to 85, and retire in the final period. I set $I = 5$ equally-sized permanent income types and I estimate permanent labor productivity by quintile and age from the PSID following the procedure in Dynan et al. (2004) (see Appendix A.9 for details). Labor productivity parameters θ_{ij} are normalized so that $\sum_i \pi \frac{1}{4} \sum_j \theta_{ij} = 1$, with relative productivity matching relative permanent labor income by type and age in the data. Idiosyncratic productivity shocks follow type-specific 3-state Markov chains calibrated to match the mean and variance of *5-year* income growth rates from Guvenen et al. (2015). Profit shares σ^π are set to match the 2019 wealth Lorenz curve (Aladangady and Forde, 2021).[26]

Household Preferences. Each household type has a type-specific discount factor β_i, bequest motive strength ψ_{ia}, and age-dependent labor disutility $\psi^j_{\ell i}$. These preference parameters, along with the age-varying risk aversion $\{\sigma_j\}_{j=1}^4$ and bequest motive parameters η and $\bar{a}$, are calibrated to match the distribution of MPS out of permanent income derived in Appendix A.9, and savings rates by age and income quintile from the PSID. With five β_i, four σ_j, and η, the model is over-identified relative to the five MPS targets alone—as discussed in Section 2, both type-dependent discount factors and non-homothetic preferences can generate increasing MPS.

I use a two-stage approach to resolve this. In the first stage, for a given degree of non-homotheticity, $\{\sigma_j\}$, η, and $\bar{a}$, the discount factors β_i are chosen to exactly match MPS by quintile. The labor dis-utility parameters $\psi^j_{\ell i}$ are set to target unit labor supply in expectation for each type-age cell, and the bequest parameters ψ_{ia} target bequests at approximately 1% of GDP. Bequest levels are calibrated to ensure that $a_{i0} = a_{iJ}$ in the steady state. Then, in the second stage, I use savings rates by age and quintile from the PSID to discipline the degree of non-homotheticity, which governs how much of the MPS gradient is attributable to scale dependence versus type dependence. A full table of calibrated parameter values is provided in Appendix Table 8.

[26] See their Figure 2.

Table 3: Savings Rates by Age and Permanent Income Quintile: Model vs. Data

	Data (PSID 1999)			Baseline Model		
	Ages 20–35	Ages 35–50	Ages 50–65	Ages 20–35	Ages 35–50	Ages 50–65
Quintile 1	0.07	0.12	0.19	0.16	0.26	0.26
Quintile 2	0.12	0.17	0.22	0.06	0.16	0.17
Quintile 3	0.18	0.26	0.34	0.16	0.29	0.29
Quintile 4	0.24	0.33	0.39	0.16	0.29	0.31
Quintile 5	0.39	0.44	0.47	0.33	0.39	0.41

Note. Data savings rates are from the PSID using the 1999 consumption measures. Model savings rates are computed as $a_{ij}/(c_{ij} + a_{ij})$ where a_{ij} and c_{ij} are expected assets and consumption for type i at age j in the calibrated steady state.

I also calibrate a version of the model with a uniform β, so that the MPS gradient is generated entirely by non-homothetic preferences as in Straub (2019) and Mian et al. (2021). Non-homotheticity alone generates a substantial MPS gradient but cannot exactly replicate the empirical distribution (see Figure 4). I use this purely scale-dependent calibration in the policy experiments below to assess how much of the redistribution-investment trade-off survives absent type dependence.

Calibration Results. Figure 4 plots the MPS out of permanent income by quintile in the data (with standard errors from the preferred specification), the baseline model with type-specific β_i, and an alternative calibration with a uniform discount factor. The baseline model matches the empirical MPS closely across all five quintiles. The uniform-β model, which relies solely on non-homothetic preferences to generate the MPS gradient, produces a substantial spread — MPS ranging from approximately 0.14 for Q1 to 0.57 for Q5 — but cannot match the full empirical pattern. In particular, the uniform-β model underpredicts MPS for Q1 (0.14 vs. 0.17) and Q3 (0.29 vs. 0.34), while overpredicting Q2 (0.28 vs. 0.23), where borrowing constraints and low income make MPS highly sensitive to patience.

Table 3 reports savings rates by age and quintile in the model and the data. The model matches the broad pattern of savings rates increasing with both age and permanent income. Savings rates for Q4 and Q5 are well-matched at all ages, though the model underpredicts savings rates in the first period for Q1–Q3 due to the borrowing constraint binding for low-income young households.

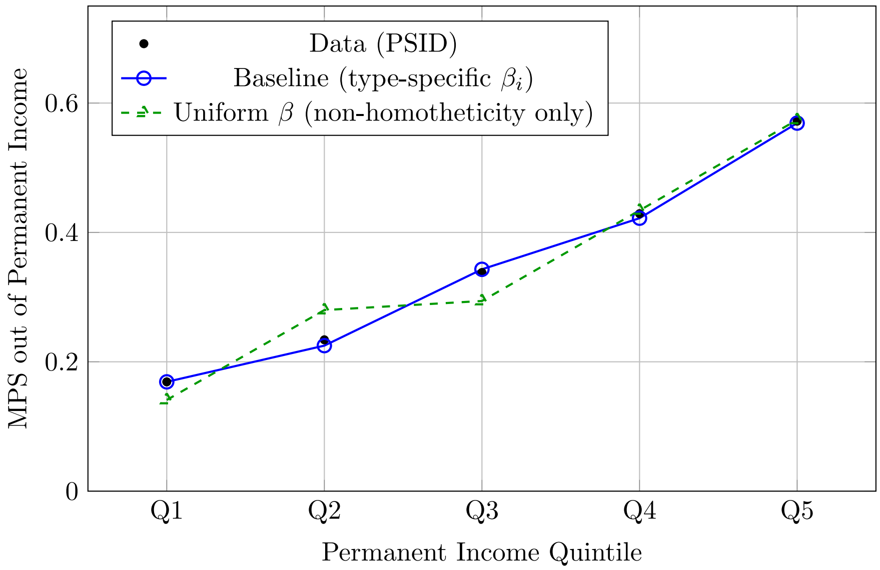

Figure 4: MPS out of Permanent Income: Model vs. Data.

Data points are MPS estimates constructed by combining PSID savings rates by permanent income quintile (1999 consumption measures, with household controls) with the permanent income elasticity of savings implied by Straub (2019) (1.3). The baseline model matches MPS targets using type-specific discount factors and non-homothetic preferences.

4.3 Alternative Calibrations.

I consider two alternative calibrations of the quantitative model, each designed to isolate a different mechanism through which redistribution affects capital accumulation.

Uniform β (non-homothetic preferences only). In the first alternative, I restrict the discount factor to be common across all types and rely solely on the non-homotheticity of preferences to generate the MPS gradient. I optimize over $(\beta, \sigma_1, \sigma_2, \sigma_3, \sigma_4)$ to minimize the sum of squared deviations between model and empirical MPS, recalibrating $\psi^j_{\ell i}$ at each candidate to ensure labor targets are met. The best-fit uniform discount factor is $\beta = 0.81$ with $\sigma_j = \{2.67, 2.38, 2.30, 2.05\}$. As shown in Figure 4, non-homotheticity alone generates a substantial MPS gradient but cannot fully match the empirical distribution. This calibration is useful for decomposing how much of the redistribution-investment trade-off arises from the scale-dependent (non-homothetic) channel versus the type-dependent (β_i) channel.

Capital-skill complementarity. In the second alternative, I replace the Cobb-Douglas production function with a nested CES specification featuring capital-skill complementarity following Krusell et al. (2000):

$$Y = Z\Big[\alpha_s Q^{\zeta_s} + (1-\alpha_s)(\ell^s)^{\zeta_s}\Big]^{1/\zeta_s}, \quad Q = \Big[\alpha_c K^{\zeta_c} + (1-\alpha_c)(\ell^c)^{\zeta_c}\Big]^{1/\zeta_c} \tag{28}$$

where ℓ^c and ℓ^s denote complementary and substitutable labor inputs, respectively. Following their estimates, I set $\zeta_s = 0.401$ (elasticity of substitution 1.67 between Q and ℓ^s) and $\zeta_c = -0.495$ (elasticity 0.67 between K and ℓ^c). Each household type i allocates a fraction ρ_i of its effective labor to the complementary sector, with ρ_i increasing in permanent income to proxy for college share. Type-specific wages are $w_i = \rho_i w^c + (1-\rho_i)w^s$. To ensure that household problems are identical across the baseline and capital-skill models (where superscript *bl* denotes baseline), I set $\theta_{ij}^{cs} = (w^{bl}/w_i)\theta_{ij}^{bl}$ so that $w_i\theta_{ij}^{cs} = w^{bl}\theta_{ij}^{bl}$. The firm parameters (α_c, α_s, Z) and labor composition (ℓ^c, ℓ^s) are determined jointly in a fixed-point loop, targeting the same labor share and output as the baseline.

4.4 Permanent Income Redistribution in the Steady State.

Here I consider a policy experiment in which – on top of existing fiscal policy – the government adds an additional linear labor income tax, τ, that funds a uniform lump-sum transfer. This simple policy experiment allows me to cleanly isolate the impact of the change in the permanent income distribution on capital and welfare, and compare the size of this channel to the labor supply distortion channel. Doing so requires decomposing the total steady state decline in capital investment generated by the policy into the effect of the permanent income redistribution channel and all other channels as in Proposition 3. From Lemma 4 we saw that, in the simple model in Section 3, aggregate household savings changed in response to the labor income tax both due to the redistribution of permanent income and in response to the general equilibrium changes in labor income.

Here, in addition to general equilibrium changes in labor supply, the policy also changes the savings supply by reducing idiosyncratic risk and by shifting after-tax income over the life cycle.[27] Before any general equilibrium adjustment in prices or labor supply from their steady state levels, the local effect of a small increase in τ on a household's expected permanent income is given by equation (29), where $\bar{\ell}$ is

[27]The latter two channels were obscured in the simple model due to the absence of idiosyncratic risk and the single period of labor earnings.

aggregate steady state labor supply per worker.[28] This is the quantitative model's analog of the numerator of equation (16). Intuitively, the proceeds of the labor income tax fund a new uniform transfer equal to $\tau w \ell$, and therefore $\tau w \bar{\ell}$ per working-age household. Type-i age-j households earn $w\theta_{ij}\ell_{ij}\epsilon$, and this labor income is now taxed at rate τ in addition to the progressive labor income tax. Given that the model is calibrated to ensure that $\mathbb{E}[\ell_{ij}\epsilon] = 1$, households' expected new tax liability is simply $\tau w \theta_{ij}$. Discounting each term by the after-tax steady state interest rate gives the following expression for each household's partial-equilibrium (direct) change in permanent income as a result of the new tax.

$$\Delta E[PI_i] = \tau w \sum_J \frac{(\bar{\ell} - \theta_{ij})}{(1 + (1 - \tau_k) r)^j} \tag{29}$$

In order to isolate the impact of this change in the permanent income distribution alone, I construct a set of hypothetical lump-sum taxes and transfers, $\{T_{ij}\}$, that equal the partial equilibrium change in a household's expected permanent income from the labor income tax. These hypothetical taxes are constructed to satisfy the following two conditions.

$$\sum_J \frac{T_{ij}}{(1+r)^j} = \Delta E[PI_i]$$
$$T_i = \frac{T_{ij}}{(1+r)^j} \text{ for all } j \in J$$

By constructing the lump-sum taxes in this way, I ensure that the hypothetical taxes reflect the change in a type-i household's permanent income without directly amplifying or dampening idiosyncratic risk, without distorting labor supply, and without redistributing resources over the life cycle.

4.4.1 Positive effects of redistribution

Table 4 reports the percent change in aggregate capital, labor, output, and consumption following a 1 percentage point increase in τ (column 1), as well as the direct (partial equilibrium) effect of the permanent income redistribution channel on these quantities, holding aggregate labor constant at its steady-state level (column 2). From

[28] See Appendix A.6.4 for the analogous expression in the case with capital-skill complementarity.

the table, a 1 percentage point increase in the new linear labor income tax decreases steady state capital by 1.01 percent, roughly a third of which can be attributed to the permanent income redistribution channel alone. Interestingly, the direct PI channel semi-elasticity of -0.33 is a close match to the back-of-envelope estimate of -0.27 from the simple model.

Table 4: Effect of a 1pp Increase in τ on Aggregate Quantities (% Change)

	Full Effect	PI Channel	PI Share (%)
ΔK	-1.01	-0.34	33.9
ΔL	-0.34	0	—
ΔY	-0.52	-0.09	18.1
ΔC	-0.7	-0.11	15.2

Note. Column 1 reports the percent change in each aggregate quantity in the new steady state following a 1 percentage point increase in the linear labor income tax τ, with the additional revenue rebated as a uniform lump-sum transfer. Column 2 reports the direct effect of the permanent income redistribution channel, holding aggregate labor fixed at its steady-state level. Column 3 reports the PI Channel as a share of the Full Effect.

I then examine the size of the permanent income redistribution channel for larger policy changes, as well as in the alternative model calibrations with only non-homothetic preferences (a uniform β), and with capital-skill complementarity. These results are presented in Table 5. The effect of a small 1 percentage point increase in τ is larger in the uniform β model than in the baseline model, as this model generates a slightly steeper gradient in the MPS distribution than the baseline model, which is able to exactly match the empirical distribution. In particular, the first and third quintiles have lower MPS out of PI than in the baseline, while the fourth and fifth quintile have slightly higher MPS. This implies a greater direct effect of PI redistribution on savings. However, there is a clear difference in how sensitive the elasticity of capital to permanent income redistribution is to the *size* of the redistribution between the two models. The capital elasticity is relatively stable in the baseline model, but drops more dramatically in the uniform-β model as the size of the redistribution increases.

This reflects the role of type vs. scale-dependence highlighted in Figure 2 in the simple model. Intuitively, if the only reason the lower-productivity households save less out of a marginal dollar is *because they are lower income*, then as we compress the income distribution we also compress the MPS distribution, dampening the subsequent impact of additional redistribution on savings. Therefore, when heterogeneous MPS out of PI are generated entirely with non-homothetic preferences rather than permanent types, every incremental increase in the degree of redistribution erodes the size of the PI redistribution channel.

The results for the model with capital-skill complementarity are more surprising, given the results from Lemma 5. The Lemma argued that – under certain restrictive assumptions – moving toward a model in which lower-income types are more substitutable with capital would amplify the effect of PI redistribution on capital, as the wages (and therefore the PI) of high-MPS households would be especially sensitive to a decline in capital. In the quantitative model, however, this force is dominated by the lower interest rate elasticity of capital (equivalently the higher elasticity of the interest rate to the capital stock) in steady state. At the given calibration, equilibrium interest rates are nearly 50 percent more elastic to capital relative to the baseline model. This implies that, as PI redistribution lowers the supply of savings and capital, the decline in capital spurs a larger increase in r, drawing in more savings via the substitution effect and ultimately dampening the channel. This is the key general equilibrium feedback effect discussed in Section 2.5.2.

Table 5: PI Channel: % Change in Aggregates per 1pp Increase in τ

	$\tau = 0.01$	$\tau = 0.05$	$\tau = 0.10$
$\%\Delta K$ *per 1pp*			
Baseline	-0.34	-0.34	-0.32
Uniform β	-0.37	-0.34	-0.33
Capital-skill	-0.23	-0.25	-0.24
$\%\Delta Y$ *per 1pp*			
Baseline	-0.09	-0.09	-0.09
Uniform β	-0.1	-0.09	-0.09
Capital-skill	-0.06	-0.07	-0.07
$\%\Delta C$ *per 1pp*			
Baseline	-0.11	-0.11	-0.1
Uniform β	-0.12	-0.11	-0.1
Capital-skill	-0.08	-0.09	-0.09

Note. Each entry reports the percent change in the aggregate quantity attributable to the permanent income redistribution channel, per 1 percentage point increase in the linear labor tax τ, holding labor supply fixed. Baseline uses type-specific β_i; Uniform β uses a common discount factor with non-homothetic preferences only; Capital-skill adds capital-skill complementarity in production.

4.4.2 Welfare decomposition

Following the same procedure as in Proposition 3, I solve for the marginal change in steady state social welfare induced by τ. In the simple model, additional labor income taxes impacted welfare through direct redistribution, through direct changes in after-tax labor income, and through changes in factor prices induced by the decreases in aggregate labor and capital. Relative to the simple model, redistribution now also impacts welfare through several additional channels: (i) a reduction in exposure to idiosyncratic labor income risk, (ii) lowering bequests for wealthy households as well as bequest tax revenue, and (iii) a decline in capital tax revenue.

Proposition 4 *Define the social welfare weights, $\kappa_{ij} \equiv \frac{\gamma^{J-j}}{(1+\bar{r})^j}\lambda_i u'(c_{i0})$. Let $\bar{r} = (1-\tau_k)r$ be the after-tax steady state interest rate and $\bar{\tau}_{ij}$ be the average progressive tax rate for type-i age-j households over idiosyncratic states. Finally, denote the general-equilibrium elasticities of labor and capital with respect to τ as $\epsilon_{\ell,\tau}$ and $\epsilon_{k,\tau}$ respectively. Then the total change in social welfare induced by τ is given by:*

$$\frac{d\mathbb{SW}}{d\tau} = \sum_I \sum_{j=0}^{J-1} \kappa_{ij} \left(\underbrace{w\ell\left(1 - \frac{\pi\theta_{ij}}{\ell}\right)}_{\text{Direct redistri.}} + \underbrace{\tau_b \sum_I \pi\, db_{i0}}_{\text{bequest tax revenue}} \right) + \underbrace{\mathcal{W}_R}_{\text{Risk}} + \underbrace{\sum_I \pi\kappa_{i0} db_{i0}(1-\tau_b)r}_{\text{direct effect of lower bequests}}$$

$$\underbrace{w\ell \sum_I \sum_{j=0}^{J-1} \kappa_{ij} \left(\frac{\pi\theta_i}{\ell} d\bar{\tau}_{ij} + \mathbb{E}\left[\frac{\pi\theta_{ij}\ell_{ij}\epsilon_i\tau_{ij}}{\ell} \left(\frac{d\ell_{ij}}{\ell_{ij}} + \frac{d\tau_{ij}}{\tau_{ij}} \right) \right] \right)}_{\text{shift in expected progr. labor tax incidence}} \Bigg) + \underbrace{\Delta\epsilon_{\ell,\tau}}_{\text{impact } d\ell \text{ on prices}}$$

$$\underbrace{w\ell \sum_I \pi \sum_{j=0}^{J-1} \kappa_{ij} \left(\tau_k \frac{\alpha}{1-\alpha}\epsilon_{k,\tau} \right)}_{\downarrow \tau_k \text{ revenue}} + \underbrace{\Delta\epsilon_{k,\tau}}_{\text{impact } dk \text{ on prices}}$$

Where Δ is defined as:

$$\Delta \equiv w\ell \sum_I \left(\sum_{j=0}^{J-1} \pi\kappa_{ij} \left(\frac{\theta_{ij}}{\ell}((1-\bar{\tau}_{ij}) - \frac{a_{ij}}{k}(1-\tau_k^j) - \tau_k \frac{A}{k} + \frac{B}{k} + \mathbb{E}\left[\frac{\theta_{ij}}{\ell}\epsilon_i\tau_{ij}\ell_{ij} \right] \right) - \pi\kappa_{iJ}\frac{a_{iJ}(1-\tau_k)}{k} \right)$$

For the complete expression for $\mathcal{W}_R$ and full derivation, see Appendix A.10.

As in the simple model, the labor income tax redistributes income directly from working households with above-average labor earnings to households with below-average labor earnings. If younger, lower productivity households are weighted more heavily, the direct redistribution effect will improve social welfare.

The labor supply distortion shifts the expected incidence and tax revenue generated by the existing progressive labor income tax, the analog of the direct labor distortion effect in Proposition 3. Furthermore, as in the simple model, the general equilibrium declines in capital and labor impact social welfare by changing factor prices, redistributing income from capital income earners to wage earners. Now, the decline in capital and labor also decreases capital tax revenue through lower rates of return, and raises average labor tax revenue through higher wages. As in the simple model, I denote the sum of these effects by Δ.

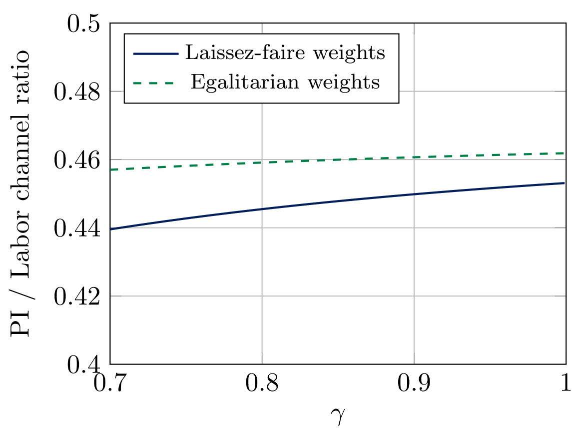

Figure 5: Ratio of PI to Labor Welfare Channels.

Note. The figure plots the ratio of the PI redistribution channel to the labor distortion channel in the marginal welfare effect of increasing τ, for laissez-faire and egalitarian welfare weights. A ratio of 0.5 means the PI channel is half as large as the labor channel.

Unlike in the simple model, a primary channel through which redistribution improves welfare is by mitigating exposure to idiosyncratic labor income risk. The linear tax and transfer redistributes income from households with higher-than-average realizations of labor productivity to those with lower-than-average, partially filling in for missing insurance markets. This channel is captured by the $\mathcal{W}_R$ term – defined in Appendix equation (A.30) – which is equal to the impact of τ on the expected product of a household's marginal utility and the deviation of their realized after-tax labor income from the average after-tax income of type-i age-j households. The insurance value of redistribution is a central channel that has been studied in the literature (Heathcote et al. (2017)), and is distinct from the impact of the tax on *average* after-tax labor income.

Additionally, because both capital income and bequests are taxed, the decline in bequests and capital induced by the tax lowers tax revenue available to fund the lump-sum transfer, while also directly impacting the after-tax bequest income of young households who would have received inheritances.

To facilitate the closest comparison between the decomposition and numerical exercise performed in Section 3, I compare the size of the 'expected' labor distortion channel — defined as the change in expected after-tax labor income induced by the change in labor — to the permanent income redistribution channel. The former is equal to the sum of the effect of distorted labor on expected progressive tax incidence and prices, while the latter is the sum of the effect of capital on capital tax revenue and prices, multiplied by the elasticity of capital associated with the PI Channel presented in Table 4. Figure 5 plots the ratio of the PI redistribution channel to the labor distortion channel. Over a range of Pareto weights, the PI redistribution channel is between 44 and 47 percent of the size of the labor distortion channel.

5 Conclusion.

Redistribution involves trade-offs. The classical one — equity against distorted labor income — has been the primary framework of the optimal labor income taxation literature. This paper argues that a second trade-off, largely overlooked, operates alongside it: when marginal propensities to save increase with permanent income, redistribution reduces aggregate savings and the long-run capital stock, creating a tension between intra- and inter-generational equality. A quantitative simulation suggests that this channel is nearly half the size of the labor income distortion channel.

These findings have implications for the design of redistributive tax policy. Taking existing debt and inter-generational transfer programs as fixed, the results suggest that the optimal labor income tax rate may be substantially lower than existing formulas imply. If however, the government is able to boost savings rates in using other reforms, my analysis suggests that more generous redistribution schemes should be paired with policies that boost the savings rate.

References

Aditya Aladangady and Akila Forde. Wealth Inequality and the Racial Wealth Gap, 2021.

Anthony Atkinson and Agnar Sandmo. Welfare implications of the taxation of savings. *Economic Journal*, 90(359):529–49, 1980.

Orazio P. Attanasio and Matthew Wakefield. The effects on consumption and saving of taxing asset returns. In *Dimensions of Tax Design: The Mirrlees Review*, pages 675–736. Oxford University Press, Oxford and New York, 2010.

Roland Bénabou. Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency? *Econometrica*, 70(2):481–517, 2002. ISSN 0012-9682.

Michael Carlos Best, James S Cloyne, Ethan Ilzetzki, and Henrik J Kleven. Estimating the elasticity of intertemporal substitution using mortgage notches. *The Review of Economic Studies*, 87(2):656–690, 2020.

Cesar Blanco and Sebastian Diz. Optimal monetary policy with non-homothetic preferences. MPRA Paper 107427, University Library of Munich, Germany, 2021.

Mariacristina De Nardi. Wealth inequality and intergenerational links. *Review of Economic Studies*, 71(3), 2004.

Mariacristina De Nardi and Giulio Fella. Saving and Wealth Inequality. *Review of Economic Dynamics*, 26:280–300, October 2017.

Anthony A. DeFusco and Andrew Paciorek. The interest rate elasticity of mortgage demand: Evidence from bunching at the conforming loan limit. *American Economic Journal: Economic Policy*, 9(1):210–40, February 2017.

Peter A. Diamond. National debt in a neoclassical growth model. *The American Economic Review*, 55(5):1126–1150, 1965.

Sebastian Doerr, Thomas Drechsel, and Donggyu Lee. Income Inequality and Job Creation, 2023.

Robert M. Dunsky and James R. Follain. Tax-induced portfolio reshuffling: The case of the mortgage interest deduction. *Real Estate Economics*, 28(4):683–718, 2000.

Karen E. Dynan, Jonathan Skinner, and Stephen P. Zeldes. Do the rich save more? *Journal of Political Economy*, 112(2):397–444, 2004.

Julio Dávila, Jay H. Hong, Per Krusell, and José-Víctor Ríos-Rull. Constrained efficiency in the neoclassical growth model with uninsurable idiosyncratic shocks. *Econometrica*, 80(6):2431–2467, 2012. ISSN 00129682, 14680262.

Owen J. Evans. Tax policy, the interest elasticity of saving, and capital accumulation: Numerical analysis of theoretical models. *The American Economic Review*, 73(3): 398–410, 1983.

Andreas Fagereng, Martin Blomhoff Holm, Benjamin Moll, and Gisle Natvik. Saving behavior across the wealth distribution: The importance of capital gains. *NBER Working Paper*, (26588), 2019. Revise and Resubmit, *Review of Economic Studies.*

Andreas Fagereng, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri. Heterogeneity and persistence in returns to wealth. *Econometrica*, 88(1):115–170, 2020. doi: 10.3982/ECTA14835.

Aart Gerritsen, Bas Jacobs, Alexandra Rusu, and Kevin Spiritus. Optimal Taxation of Capital Income with Heterogeneous Rates of Return. June 2020.

Mikhail Golosov, Maxim Troshkin, Aleh Tsyvinski, and Matthew Weinzierl. Preference heterogeneity and optimal capital income taxation. *Journal of Public Economics*, 97:160–175, 2013.

Mikhail Golosov, Maxim Troshkin, and Aleh Tsyvinski. Redistribution and social insurance. *American Economic Review*, 106(2):359–86, February 2016.

Fatih Guvenen, Fatih Karahan, Serdar Ozkan, and Jae Song. What do data on millions of u.s. workers reveal about life-cycle earnings risk? Working Paper 20913, National Bureau of Economic Research, January 2015.

Koichi Hamada. Lifetime equity and dynamic efficiency on the balanced growth path. *Journal of Public Economics*, 1(3):379–396, 1972. ISSN 0047-2727.

Jonathan Heathcote, Kjetil Storesletten, and Giovanni L. Violante. Optimal Tax Progressivity: An Analytical Framework*. *The Quarterly Journal of Economics*, 132(4):1693–1754, 06 2017.

Joachim Hubmer, Elin Halvorsen, Sergio Salgado, and Serdar Ozkan. Why are the wealthiest so wealthy? New longitudinal empirical evidence and implications for theories of wealth inequality. *Econometrica*, 2026. Conditionally Accepted.

Mark Huggett and Gustavo Ventura. Understanding why high income households save more than low income households. *Journal of Monetary Economics*, 45(2): 361–397, 2000. ISSN 0304-3932.

Ayse Imrohoroglu, Cagri S. Kumru, and Arm Nakornthab. Revisiting Tax on Top Income. Anu working papers in economics and econometrics, Australian National University, College of Business and Economics, School of Economics, February 2018.

Tullio Jappelli and Luigi Pistaferri. Do people respond to tax incentives? an analysis of the italian reform of the deductibility of home mortgage interests. *European Economic Review*, 51(2):247–271, 2007.

Kenneth L. Judd. Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics*, 28(1):59–83, 1985. ISSN 0047-2727.

Per Krusell, Lee E. Ohanian, José-Víctor Ríos-Rull, and Giovanni L. Violante. Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica*, 68(5):1029–1053, 2000.

Atif Mian, Ludwig Straub, and Amir Sufi. Indebted Demand*. *The Quarterly Journal of Economics*, 136(4):2243–2307, 03 2021. ISSN 0033-5533.

J. A. Mirrlees. An exploration in the theory of optimum income taxation. *The Review of Economic Studies*, 38(2):175–208, 1971.

Nobuhiro Okuno and Akira Yakita. Public investment and income distribution: A note. *The Quarterly Journal of Economics*, 96(1):171–176, 1981.

Pierre Pestieau and Uri M. Possen. Optimal growth and distribution policies. *Journal of Public Economics*, 9(3):357–372, 1978. ISSN 0047-2727.

Thomas Piketty and Emmanuel Saez. Chapter 7 - optimal labor income taxation. In Alan J. Auerbach, Raj Chetty, Martin Feldstein, and Emmanuel Saez, editors, *handbook of public economics, vol. 5*, volume 5 of *Handbook of Public Economics*, pages 391–474. Elsevier, 2013a.

Thomas Piketty and Emmanuel Saez. A theory of optimal inheritance taxation. *Econometrica*, 81(5):1851–1886, 2013b.

Alessandra Pizzo. The welfare effects of tax progressivity with frictional labor markets. *Review of Economic Dynamics*, 49:123–146, July 2023.

Emmanuel Saez and Stefanie Stantcheva. A simpler theory of optimal capital taxation. *Journal of Public Economics [also: NBER Working Paper 22664]*, 162: 120–142, 2018.

Emmanuel Saez, Joel Slemrod, and Seth H. Giertz. The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature*, 50(1):3–50, March 2012. doi: 10.1257/jel.50.1.3.

Karl Schulz. Redistribution and Return Inequality. June 2021.

Eytan Sheshinski. The optimal linear income-tax. *The Review of Economic Studies*, 39(3):297–302, July 1972.

Eytan Sheshinski. Income taxation and capital accumulation. *The Quarterly Journal of Economics*, 90(1):138–149, 1976.

Ludwig Straub. Consumption, savings, and the distribution of permanent income. Revise and resubmit, *Econometrica*, 2019.

Iván Werning. Optimal Fiscal Policy with Redistribution. *The Quarterly Journal of Economics*, 122(3):925–967, 08 2007.

A Appendix.

A.1 Proof of Lemma 1

A.1.1 Marginal propensities to save out of permanent income.

The households' Euler equation is:

$$(c_{it}^{y})^{-\sigma_y} = \beta_i(1-\tau_{kt+1})(1+r_{t+1})(c_{it+1}^{o})^{-\sigma_o}$$

The household's lifetime budget constraint is:

$$w_t\theta_i + T_{it}^{y} + \frac{\hat{T}_{t+1}^{o}}{(1+r_{t+1})} = PI_{it} = c_{it}^{y} + \frac{c_{it+1}^{o}}{(1+r_{t+1})(1-\tau_{kt+1})}$$

Using the Euler equation, calculate the derivatives of c_{it+1}^{o} with respect to c_{it}^{y}.

$$\frac{\partial c_{it+1}^{o}}{\partial c_{it}^{y}} = \frac{\sigma_y (c_{it}^{y})^{-\sigma_y-1}}{\beta_i(1-\tau_{kt+1})(1+r_{t+1})(\sigma_o)(c_{it+1}^{o})^{-\sigma_o-1}}$$

Using the lifetime budget constraint, solve for the total derivative of c_{it}^{y} with respect to PI_{it}.

$$\frac{\partial c_{it}^{y}}{\partial PI_{it}} = \left(1 + \frac{1}{(1-\tau_{kt+1})(1+r_{t+1})}\left(\frac{\partial c_{it+1}^{o}}{\partial c_{it}^{y}}\right)\right)^{-1}$$

Case 1 (homothetic preferences and uniform β): Using the Euler equations, it is clear that when $\sigma_y = \sigma_o$, the ratio between c_{it}^{y}, c_{it+1}^{o} depends only on β_i and the after-tax rate of return:

$$\frac{c_{it}^{y}}{c_{it+1}^{o}} = (\beta_i(1-\tau_{kt+1})(1+r_{t+1}))^{-\frac{1}{\sigma_y}}$$

This implies uniform marginal propensities to consume out of permanent income as:

$$\frac{\partial c_{it+1}^{o}}{\partial c_{it}^{y}} = \frac{(\beta_i(1-\tau_{kt+1})(1+r_{t+1}))^{-\frac{1}{\sigma_y}(-\sigma_y-1)}}{\beta_i(1-\tau_{kt+1})(1+r_{t+1})}$$

When $\beta_i = \beta$ for all i, this expression is identical across types, and therefore marginal propensities to save out of additional permanent income are also uniform across households.

Case 2 (non-homothetic preferences or heterogeneous β): Using the Euler equations, it is clear that when $\sigma_y > \sigma_o$ the ratio between c^y_{it} and c^o_{it+1} is *decreasing* in c^o_{it+1} and therefore decreasing in PI_{it}:

$$\frac{c^y_{it}}{c^o_{it+1}} = (\beta_i(1-\tau_{kt+1})(1+r_{t+1}))^{-\frac{1}{\sigma_y}}(c^o_{it+1})^{\frac{\sigma_o}{\sigma_y}-1}$$

Since $\sigma_o < \sigma_y$, the exponent $\frac{\sigma_o}{\sigma_y} - 1 < 0$, so the ratio c^y/c^o is decreasing in c^o and hence in PI. Alternatively, if preferences are homothetic ($\sigma_y = \sigma_o$) but $\beta_H > \beta_L$, then the ratio $c^y_{it}/c^o_{it+1} = (\beta_i(1-\tau_{kt+1})(1+r_{t+1}))^{-1/\sigma_y}$ is decreasing in β_i, which is higher for high-productivity types with higher permanent income, so again c^y/c^o is decreasing in PI.
Plug these results into the derivative of c^o_{it+1} with respect to c^y_{it}:

$$\begin{aligned}\frac{\partial c^o_{it+1}}{\partial c^y_{it}} &= \frac{\sigma_y(c^y_{it}/c^o_{it+1})^{-\sigma_y-1}}{\beta_i(1-\tau_{kt+1})(1+r_{t+1})(\sigma_o)(c^o_{it+1})^{\sigma_y-\sigma_o}} \\ &= f(\beta_i,\tau_{kt+1},r_{t+1})(c^o_{it+1})^{1-\frac{\sigma_o}{\sigma_y}}\end{aligned}$$

Since $1-\sigma_o/\sigma_y > 0$, $\frac{\partial c^o_{it+1}}{\partial c^y_{it}}$ is increasing in c^o_{it+1}, and therefore increasing in PI_{it}.

Finally, since $\frac{\partial c^y_{it}}{\partial PI_{it}} = \left(1+\frac{1}{(1-\tau_{kt+1})(1+r_{t+1})}\frac{\partial c^o_{it+1}}{\partial c^y_{it}}\right)^{-1}$ is decreasing in $\frac{\partial c^o_{it+1}}{\partial c^y_{it}}$, and $\frac{\partial c^o_{it+1}}{\partial c^y_{it}}$ is increasing in PI_{it}, it follows that $\frac{\partial c^y_{it}}{\partial PI_{it}}$ is decreasing in PI_{it}. Therefore $MPS_i = 1 - \frac{\partial c^y_{it}}{\partial PI_{it}}$ is increasing in permanent income.

A.2 Proof of Lemma 2

A.2.1 Equilibrium capital.

Plugging the household's period budget constraints into their Euler equations gives you:

$$\begin{aligned}&(w_t\theta_i + T^y_{it} - a_{it})^{-\sigma_y} \\ &= \beta_i(1-\tau_{kt+1})(1+r_{t+1})(a_{it}(1+r_{t+1})(1-\tau_{kt+1}) + T^o_{t+1})^{-\sigma_o}\end{aligned} \tag{A.1}$$

Equations (A.1) implicitly define a_{it} as a function $a^E_{it}(.)$ of prices and the policies a household is subject to over their lifetime: τ_{kt+1}, T^y_{it}, and T^o_{t+1}. Substituting the labor market clearing conditions into the firms' optimality conditions, then substituting these conditions in for prices, the asset market clearing condition can now be written as:

$$k_{t+1} = \sum_{i \in I} \pi a^E_{it}(T_{it}, T^o_{t+1}, \tau_{k_{t+1}}, k_{t+1}, k_t) \tag{A.2}$$

Equation (A.2) defines k_{t+1} and an implicit function $k^E_{t+1}(T_{it}, T^o_{t+1}, \tau_{kt+1}, k_t)$.

A.2.2 Convergence to a unique steady state for $k_0 > 0$.

From equation (A.1), it is immediate that a_{it}, and therefore k_{t+1} is increasing in w_t for both types. Substituting in $w_t = F_\ell(k_t, 1)$, we can use (A.1) and (A.2) to solve for $\frac{dk^E_{t+1}(.)}{dk_t}$ at a given set of policies. By the implicit function theorem, we have that

$$\frac{dk^E_{t+1}(.)}{dk_t} = -\frac{\sum_I \frac{da_{it}}{dk_t}}{\sum_I \frac{da_{it}}{dk_{t+1}} - 1}$$

Where $\frac{da_{it}}{dk_{t+1}} = \frac{da_{it}}{dr_{t+1}} \frac{\partial r_{t+1}}{\partial k_{t+1}}$. Given the assumed production function, $\frac{\partial r_{t+1}}{\partial k_{t+1}} < 0$. To see that $\frac{da_{it}}{dr_{t+1}}$ is positive under the parameter assumptions, define the implicit function:

$$F(a_{it}, r_{t+1}) = (c^y_{it})^{-\sigma_y} - \beta_i(1 - \tau_{kt+1})(1 + r_{t+1})(c^o_{it+1})^{-\sigma_o} = 0$$

then we have that:

$$\frac{\partial a_{it}}{\partial r_{t+1}} = -\frac{\partial F / \partial r_{t+1}}{\partial F / \partial a_{it}}$$

Computing the partial derivatives:

$$\frac{\partial F}{\partial a_{it}} = \sigma_y (c^y_{it})^{-\sigma_y - 1} + \beta_i (1 - \tau_{kt+1})^2 (1 + r_{t+1})^2 \sigma_o (c^o_{it+1})^{-\sigma_o - 1} > 0$$
$$\frac{\partial F}{\partial r_{t+1}} = -\beta_i (1 - \tau_{kt+1})(c^o_{it+1})^{-\sigma_o} \left[1 - \sigma_o \frac{a_{it}(1 - \tau_{kt+1})(1 + r_{t+1})}{c^o_{it+1}} \right]$$

Define $s_{it} \equiv \frac{a_{it}(1-\tau_{kt+1})(1+r_{t+1})}{c^o_{it+1}} \in [0,1]$ as long as $T^o_t \geq 0$ and $\sigma_o \leq 1$. Then:

$$\frac{\partial F}{\partial r_{t+1}} = -\beta_i(1-\tau_{kt+1})(c^o_{it+1})^{-\sigma_o}[1-\sigma_o s_{it}]$$

If $s_{it} \in [0,1]$ and $\sigma_o > 0$, we have $1-\sigma_o s_{it} \geq 0$. Therefore:

$$\frac{\partial a_{it}}{\partial r_{t+1}} = -\frac{(-)}{(+)} > 0 \rightarrow \sum_I \frac{da_{it}}{dk_{t+1}} < 0$$

As long as $T^o_t \geq 0$ and $\sigma_o \leq 1$, then $1 - \sum_I \frac{da_{it}}{dk_{t+1}} \geq 1$.

Next, I establish that $\pi \sum_I \frac{da_{it}}{dk_t} \in (0,1)$. This follows simply because, with constant returns to scale production, $\frac{dy_t}{dk_t} < 1$. The increase in output is divided between the old through capital income, whose marginal propensity to save is 0, and the young through labor income whose MPS $\in (0,1)$.

Let $\bar{k}$ be defined by $f(\bar{k},1) = \delta\bar{k}$, so that net output is zero at $\bar{k}$; since capital cannot be negative, $k_t \in [0,\bar{k}]$ for all t, which is a compact set. Together with the contraction established above, the Banach fixed-point theorem guarantees convergence to a unique steady state.

Putting these results together, it is immediate that $k^E_{t+1}(k_t)$ **is a contraction**. Specifically,

$$\left|\frac{dk^E_{t+1}(.)}{dk_t}\right| = \left|\frac{\sum_I \frac{da_{i,t}}{dk_t}}{1-\sum_I \frac{da_{it}}{dk_{t+1}}}\right| < 1$$

Therefore, if policy is kept constant, the economy converges to a unique steady state.

A.2.3 Comparative statics.

The Lemma states that $k^E(.)$ is non-increasing in $T_{Lt} - T_{Ht}$ and non-increasing in $T^o_{t+1} - \sum_I T^y_{it}$. That is, equilibrium capital is non-increasing in both net transfers to the poor and net transfers to the old.

a) Define the following auxiliary policies, T_{Lt}, T_{Ht}, $\hat{T}^o_t$, T^y_t, and T_t such that the following hold.

$$T_{Ht} = -T_{Lt}\ ;\ \hat{T}^o_t = -T^y_t\ ;\ \tau_{kt}k_t(1+r_t) = T_t$$

Then the policy tools available to the government can be rewritten in the following way.

$$T_{it}^y = T_{it} + T_t^y + T_t$$
$$T_t^o = \hat{T}_t^o + T_t$$

That is, T_{it} captures net transfers from high to low skill types, $\hat{T}_t^o$ captures net transfers from the current young to the current old, and T_t captures the uniform lump-sum transfer from capital tax revenue.

The government's constraints can now be rewritten as:

$$\tau_{kt} k_t (1 + r_t) = T_t$$
$$\hat{T}_t^o \geq \underline{\mathrm{T}}$$
$$(w_t \theta_H - w_t \theta_L)/2 = \bar{T}_t^L \geq$$
$$T_{Lt} \geq -(w_t \theta_L - \hat{T}_t^o + T_t) = \underline{\mathrm{T}}_t^L$$

Intuitively, the uniform lump-sum transfer is funded by capital taxes, the net transfer to the old must exceed the political lower bound but not violate the non-negativity constraint of any household, and the net transfer to the low-skilled must be large enough to respect the non-negativity constraint of the low-skill types but no larger than the transfer needed to equalize income between the two types.

b) *Claim.* Holding all other policy constant, k_{t+1}^E is decreasing in T_{Lt} whenever $\beta_L < \beta_H$ or $\sigma_y > \sigma_o$. When $\beta_L = \beta_H$ and $\sigma_y = \sigma_o$, $dk_{t+1}^E / dT_{Lt} = 0$.

Proof. Using the results from Lemma 1, the marginal propensity to save out of additional permanent income is smaller for low-skill households. Therefore, the numerator of the following expression is negative. The denominator is positive, as established in Section A.2.2, where it is shown that $1 - \sum_{i \in I} \frac{da_{it}}{dk_{t+1}} \geq 1 > 0$.

The total derivative of k_{t+1}^E with respect to T_{Lt} is:

$$\frac{dk_{t+1}^E}{dT_{Lt}} = \pi \left(\frac{\partial a_{Lt}}{\partial T_{Lt}} - \frac{\partial a_{Ht}}{\partial T_{Ht}} \right) \left(1 - \sum_{i \in I} \frac{a_{it}}{k_{t+1}^E} \frac{\mathrm{dlog} a_{it}}{\mathrm{dlog} r_{t+1}} \frac{\mathrm{dlog} r_{t+1}}{\mathrm{dlog} k_{t+1}} \right)^{-1} < 0$$

When $\beta_L = \beta_H$ and $\sigma_y = \sigma_o$, Lemma 1 (see Section A.1) implies that MPS are equal across types, so $\frac{\partial a_{Lt}}{\partial T_{Lt}} = \frac{\partial a_{Ht}}{\partial T_{Ht}}$ and therefore $dk^E_{t+1}/dT_{Lt} = 0$. $\square$

c) Holding all other policy constant, k^E_{t+1} is decreasing in $\hat{T}^o_t$ and $\hat{T}^o_{t+1}$.

Proof. For $\hat{T}^o_{t+1}$: from equation (A.1) we can show the following.

$$\frac{\partial a_{it}}{\partial \hat{T}^o_{t+1}} = \frac{-1}{\frac{\sigma_y}{\sigma_o}\frac{c^o_{it+1}}{c^y_{it}} + (1 + r_{t+1})(1 - \tau_{kt+1})} < 0$$

For $\hat{T}^o_t$: from the decomposition in part (a), $T^y_t = -\hat{T}^o_t$, so an increase in $\hat{T}^o_t$ reduces the uniform transfer to the young by the same amount. Since $T^y_{it} = T_{it} + T^y_t + T_t$, we have $\frac{\partial T^y_{it}}{\partial \hat{T}^o_t} = -1$ for both types. Since savings are increasing in income, $\frac{\partial a_{it}}{\partial T^y_{it}} > 0$, it follows that:

$$\frac{\partial a_{it}}{\partial \hat{T}^o_t} = -\frac{\partial a_{it}}{\partial T^y_{it}} < 0$$

In both cases the denominator $1 - \sum_{i \in I} \frac{da_{it}}{dk_{t+1}} \geq 1 > 0$ by Section A.2.2, so $dk^E_{t+1}/d\hat{T}^o_t < 0$ and $dk^E_{t+1}/d\hat{T}^o_{t+1} < 0$. $\square$

A.3 Proof of Proposition 1

The unconstrained planner's problem is the following.

$$\max_{\{\{c^y_{it}, c^o_{it}\}_{i\in I}, k_{t+1}\}_{t\geq 0}} \sum_{t=0}^{\infty} \gamma^t \sum_{i\in I} \lambda_i \pi \left(\frac{(c^y_{it})^{1-\sigma_y}}{1-\sigma_y} + \beta_i \frac{(c^o_{it+1})^{1-\sigma_o}}{1-\sigma_o} \right) + \gamma^{-1} \sum_{i \in I} \beta_i \lambda_i \pi \left(\frac{(c^o_{i0})^{1-\sigma_o}}{1-\sigma_o} \right)$$

subject to:

$$k_{t+1} = k_t(1-\delta) + f(k_t, 1) - \pi \sum_I (c^y_{it} + c^o_{it}) \tag{A.3}$$

$$c^y_{it} \geq 0,\ c^o_{it} \geq 0,\ k_{t+1} \geq 0 \ \text{ for all } t \geq 0 \tag{A.4}$$

Let μ_t be the lagrange multiplier on the resource constraint (A.3) at time t. The unconstrained planner's first order conditions with respect to k_t for $t \geq 1$ are:

$$\mu_{t-1} = \mu_t(f_k(k_t, \ell_t) + 1 - \delta) \tag{A.5}$$

The planner's first order conditions with respect to c_{it}^y and c_{it}^o for $t \geq 0$ are:

$$\mu_t = \pi\gamma^t \lambda_i (c_{it}^y)^{-\sigma_y}$$
$$\mu_t = \pi\gamma^{t-1} \beta_i \lambda_i (c_{it}^o)^{-\sigma_o}$$

Within each period, this implies that consumption and asset allocations across types should be set so that the following conditions hold.

$$\left(\frac{c_{Ht}^y}{c_{Lt}^y}\right)^{\sigma_y} = \frac{\lambda_H}{\lambda_L} \tag{A.6}$$

For type-i households across generations, the following also must hold.

$$\frac{1}{\gamma}(c_{it-1}^y)^{-\sigma_y} = (c_{it}^y)^{-\sigma_y}(f_k(k_t, \ell_t) + (1-\delta)) \tag{A.7}$$

Intuitively, the costs of greater capital: less consumption for the previous generation, must equal the benefits: more consumption for the current generation.

The steady state version of this condition is:

$$f_k(k^{gr}, \ell) + 1 - \delta = \frac{1}{\gamma} \tag{A.8}$$

Here k^{gr}, implicitly defined by the above equation, is the modified Golden Rule capital stock.

A.4 Proof of Corollary 1

To see that the planner can implement the first best allocation with an unrestricted set of lump-sum transfers, it is sufficient to characterize the set of policies, $\{T_{Lt}^y, T_{Ht}^y, T_t^o\}_{t\geq 0}$ that implements $\mathcal{A}_u$ as an equilibrium.

First, set $w_t = f_\ell(\bar{k}_t, 1)$ and $r_t = f_k(\bar{k}_t, 1) - \delta$ for all t. Plug these factor prices into the household's lifetime budget constraint. Set $\tau_{kt} = 0$ in all periods. Use the households' savings functions and first-period budget constraint to write:

$$\frac{\bar{c}_{Ht}^y}{\bar{c}_{Lt}^y} = \frac{w_t\theta_H + T_{Ht}^y - a_{Ht}(.)}{w_t\theta_L + T_{Lt}^y - a_{Lt}(.)} = \left(\frac{\lambda_H}{\lambda_L}\right)^{\frac{1}{\sigma_y}} \tag{A.9}$$

Then, simply use equation (A.9), the asset market clearing condition, and the government's budget constraint to form a system of 3 equations in 3 unknowns: T^y_{Lt}, T^y_{Ht}, T^o_t in each period.

A.5 Proof of Proposition 2

A.5.1 Implementability

If an allocation satisfies the political constraint (4) and following conditions for all $t \geq 0$, then it can be implemented as a competitive equilibrium for some sequence of prices and policies.

$$k_{t+1} = k^E_{t+1}(T^y_{Lt}, T^y_{Ht}, \hat{T}^o_{t+1}, \tau_{k_{t+1}}, k_t) \tag{A.10}$$

$$a_{it} = a^E_{it}(T^y_{it}, \hat{T}^o_{t+1}, \tau_{k_{t+1}}, k_t) \tag{A.11}$$

$$c^y_{it} = f_\ell(k_t, 1)\theta_i + T^y_{it} - \hat{T}^o_t + T_t - a_{it} \;\; ; \; c^o_{it} = a_{it-1}(1 + f_k(k_t) - \delta)(1 - \tau_{kt}) + \hat{T}^o_t \tag{A.12}$$

$$T_{Ht} = -T_{Lt} \tag{A.13}$$

$$\hat{T}^o_t = -T^y_t \tag{A.14}$$

$$T_t = k_t(1 + r_t)\tau_{kt} \tag{A.15}$$

Necessity. The necessity of equations (A.10) and (A.11) follows directly from the construction of the functions $a^E_{it}(.)$ and $k^E_{t+1}(.)$ using the results from Lemma 2 derived in Appendix A.2.1. The following conditions are equilibrium conditions and therefore necessary.

Sufficiency. It is sufficient to show that if an allocation satisfies conditions (A.10)-(A.15) for a given set of policies, there exists a set of prices that implements this allocation as a competitive equilibrium. Suppose an allocation satisfied equations (A.10)-(A.15). Then simply set prices, $w_t = f_\ell(k_t, 1)$ and $r_t = f_k(k_t, 1) - \delta$. By construction, households' Euler equations are satisfied if equation (A.11) is satisfied. Because equation (A.10) is satisfied, asset market clearing conditions are satisfied.

The final conditions are the households' budget constraints and the government's budget constraints, ensuring the resource constraint is satisfied by Walras Law.

A.5.2 Constrained planner's first order conditions

Using the alternative policy representation from Appendix A.2.3, the first order conditions with respect to T_{Lt} is:

$$\gamma^t\left(\lambda_L(c^y_{Lt})^{-\sigma_y} - \lambda_H(c^y_{Ht})^{-\sigma_y}\right) + \sum_{j=t}^{\infty}\left(\gamma^j\left(\lambda_L(c^y_{Lj})^{-\sigma_y}\frac{a_{Lj}}{1+r_{j+1}} + \lambda_H(c^y_{Hj})^{-\sigma_y}\frac{a_{Hj}}{1+r_{j+1}}\right)\frac{\partial r_{j+1}}{\partial k_{j+1}}\right.$$
$$\left.\gamma^{j+1}\left(\lambda_L(c^y_{Lj+1})^{-\sigma_y}\theta_L + \lambda_H(c^y_{Hj+1})^{-\sigma_y}\theta_H\right)\frac{\partial w_{j+1}}{\partial k_{j+1}} + \tau_{kj+1}\gamma^t\Theta\left(1+r_{j+1}+k_{j+1}\frac{\partial r_{j+1}}{\partial k_{j+1}}\right)\right)\frac{dk_{j+1}}{dT_{Lt}} = 0$$

Here, the term Θ is defined as:

$$\Theta = \sum_I\left(\frac{\lambda_i(c^y_{it})^{-\sigma_y}}{(1+r_{t+1})(1-\tau_{kt+1})} + \gamma\lambda_i(c^y_{it+1})^{-\sigma_y}\right)$$

Note that here I am using that, $\frac{da^y(.)}{dT_L}(-(c^y_{it})^{-\sigma_y} + \beta_i(1+r_{t+1})(1-\tau_{kt+1})(c^o_{it+1})^{-\sigma_o} = 0$, an application of the envelope theorem.

Define Δ_j as the term inside the large parentheses divided by γ^{t+j}, so that the above can be written as:

$$\gamma^t\left(\lambda_L(c^y_{Lt})^{-\sigma_y} - \lambda_H(c^y_{Ht})^{-\sigma_y}\right) + \sum_{j=t}^{\infty}\gamma^j\Delta_j\frac{dk_{j+1}}{dT_{Lt}} = 0 \tag{A.16}$$

Next, use the firms' production function to get the partial derivatives of factor prices with respect to k_{t+1}.

$$\frac{\partial r_{t+1}}{\partial k_{t+1}} = \frac{\nu-1}{\nu}f_k\left(\frac{f_k}{y} - \frac{1}{k}\right) = \frac{\nu-1}{\nu}\frac{f_k}{ky}(f_k k - y) = \frac{\nu-1}{\nu}\frac{f_k}{ky}f_\ell$$
$$\frac{\partial w_{t+1}}{\partial k_{t+1}} = \frac{\nu-1}{\nu}f_k\frac{f_\ell}{y}$$

Therefore, we have that $\frac{\partial r_{t+1}}{\partial k_{t+1}}\frac{\partial k_{t+1}}{\partial w_{t+1}} = \frac{1}{k}$. Dividing through by γ^t, plugging this into the above expression, and collecting terms gives:

$$\Delta_t = \Bigg(\Bigg(-\Bigg(\lambda_L(c^y_{Lt})^{-\sigma_y}\frac{a_{Lt}}{k_{t+1}} + \lambda_H(c^y_{Ht})^{-\sigma_y}\frac{a_{Ht}}{k_{t+1}}\Bigg)(1+r_{t+1})^{-1} +$$

$$\gamma\Bigg(\lambda_L(c^y_{Lt+1})^{-\sigma_y}\theta_L + \lambda_H(c^y_{Ht+1})^{-\sigma_y}\theta_H\Bigg) + \tau_{kt+1}\Theta\Bigg)\frac{\partial w_{t+1}}{\partial k_{t+1}} + \tau_{kt+1}\Theta(1+r_{t+1})\Bigg)$$

The sign of Δ_t determines whether an additional unit of k_{t+1} is welfare improving.

The first order condition with respect to τ_{kt+1} is:

$$\pi\gamma^t\Bigg(\sum_I (c^o_{it+1})^{-\sigma_o}\beta_i(1+r_{t+1})(k_{t+1}-a_{it})\Bigg) +$$

$$\gamma^{t+1}\Bigg(\pi\sum_I \lambda_i(c^y_{it+1})^{-\sigma_y}\Bigg)(1+r_{t+1})k_{t+1} + \sum_{j=t}^{\infty}\gamma^{t+j}\Delta_j\frac{dk_{t+j}}{d\tau_{kt+1}} = 0$$

A.5.3 Proof of part (1)

Part 1 of the Proposition follows directly from combining the results in Lemma 1 and Lemma 2 with the above first order conditions. If β_i is uniform and preferences are homothetic, then by Lemma 1, marginal propensities to save are uniform. Then Lemma 2 shows that intra-generational transfers have no effect on capital.

Plugging $\frac{dk_{t+1}}{dT_{Lt}} = 0$ into equation (A.16) leaves only the expression:

$$\Bigg(\lambda_L(c^y_{Lt})^{-\sigma_y} - \lambda_H(c^y_{Ht})^{-\sigma_y}\Bigg) = 0$$

which is only satisfied at the 'ideal' level of inequality.

A.5.4 Proof of part (2)

To prove part (2), I proceed in four steps:

(i) Show that if the planner attempts to implement first-best equality *or greater* in each period, for a given $\hat{T}^o_t = \bar{T}$, k_t will converge to some $k^* \leq k_f < k^{gr}$. This implies that there exists some $\tau \geq 0$ such that $k_t < k^{gr}$ for all $t > \tau$.

(ii) If $k_t < k^{gr}$, then $\Delta_{t-1} > 0$. That is, if the capital stock is below the Golden Rule level, capital is unambiguously welfare improving in period t. Moreover, define $\bar{k}$ as the capital stock at which $\Delta_t = 0$. We show that $\bar{k} > k^{gr}$.

(iii) Establish that if the constrained optimal allocation features first-best equality or greater in every period, the planner's first order condition cannot hold. This directly contradicts optimality.

(iv) Use parts (ii) and (iii) to argue that if $k_0 < k^{gr}$, the constrained optimal solution features less than first-best equality in every period.

The claims in Part (2) follow directly from the above.

i) If the planner implements first-best equality *or greater*, $(c^y_{Ht}/c^y_{Lt})^{\sigma_y} \leq \frac{\lambda_H}{\lambda_L}$, in every period with $\hat{T}^o_t = \bar{T}$ and $\tau_{kt} = 0$, then k_t converges to some $k^* \leq k_f < k^{gr}$. Therefore $\exists$ a $\tau \geq 0$ such that $k_t < k^{gr}$ for all $t > \tau$.

Note that setting $\hat{T}^o_t = \bar{T}$ represents the most favorable case for the planner: intergenerational transfers are pushed to their political limit, giving the planner the maximum non-distortionary ability to boost the capital stock. For any $\hat{T}^o_t < \bar{T}$, the capital stock converges to a steady state strictly below k_f, making the trade-off strictly more severe.

To see this, first consider the case of exactly first-best equality. Set $\hat{T}^o_t = \bar{T}$ and $\tau_{kt} = 0$ in all periods. Define the function $T(k_t)$ as the level of transfers that implement first-best equality in period t for a given capital stock, k_t. Next, let $f(k_t, T_L) = k^E_{t+1}$ be the equilibrium capital stock defined in Appendix A.2. Finally define the composite function, $g(k) = f(k, T(k))$.

From Appendix A.2, we know that $k^E_{t+1}(.) = f(k_t, T_{Lt})$ is continuous and unique in k_t and T_{Lt}.

The full equality steady state k_f satisfies:

- $k_f = f(k_f, T_f)$ where $T_f = T(k_f)$
- $c^y_L(k_f, T_f) = c^y_H(k_f, -T_f)$

Therefore $g(k_f) = f(k_f, T(k_f)) = f(k_f, T_f) = k_f$.

Suppose $k_t = \tilde{k} > k_f$. Let $\tilde{T}_L$ be the level of transfers needed to maintain $\tilde{k}$ as a steady state such that $\tilde{k} = f(\tilde{k}, \tilde{T}_L)$. Because $\tilde{k} > k_f$, $T(\tilde{k}) > \tilde{T}_L$. Therefore, $g(\tilde{k}) < \tilde{k}$. A symmetric argument implies $g(\tilde{k}) > \tilde{k}$ if $\tilde{k} < k_f$. Then we have:

- If $k_0 < k_f$, then $k_1 = g(k_0) > k_0$, and by induction, $k_t < k_{t+1} < k_f$ for all t.
- If $k_0 > k_f$, then $k_1 = g(k_0) < k_0$, and by induction, $k_t > k_{t+1} > k_f$ for all t.

- If $k_0 = k_f$, then $k_t = k_f$ for all t.

If $k_0 < k_f$, then $\{k_t\}$ is increasing and bounded above by k_f. If $k_0 > k_f$, then $\{k_t\}$ is decreasing and bounded below by k_f. In both cases, $\{k_t\}$ is a bounded monotone sequence. By the Monotone Convergence Theorem, $\lim_{t\to\infty} k_t = k^*$ where $k^* = g(k^*)$ is a fixed point of g. By monotonicity of g, k_f is the unique fixed point, hence $k^* = k_f$.

Suppose the planner had been implementing first-best equality in every previous period. Then by convergence, for any $\varepsilon > 0$, $\exists$ some $\tau > t$ such that $|k_t - k_f| < \varepsilon$. By assumption, $k_f < k^{gr}$. Let $\varepsilon = (k^{gr} - k_f)/2$. Then there exists a period $\tau \geq 0$ such that $k_t < k^{gr}$.

Since k^E_{t+1} is non-increasing in the degree of redistribution (Lemma 2), implementing redistribution weakly greater than first-best in every period implies k_t converges to some $k^* \leq k_f < k^{gr}$. Therefore the same conclusion holds: $\exists\ \tau \geq 0$ such that $k_t < k^{gr}$ for all $t > \tau$.

ii) If $k_t < k^{gr}$ then $\Delta_{t-1} > 0$. Moreover, $\bar{k} > k^{gr}$.

If $k_t < k^{gr}$ this implies $\Delta_{t-1} > 0$. To see this, simply note that $k_t < k^{gr} \to r_t > 0$. Given the assumptions on the permissible degree of redistribution and the assumed restrictions on Pareto weights, low-productivity types will always have higher welfare weights. That is, $\lambda_L(c^y_{Lt})^{-\sigma_y} \geq \lambda_H(c^y_{Ht})^{-\sigma_y}$.

By assumption, the share of total labor income of low-types (with higher welfare weights), θ_L, is as high or higher than their share of total capital income, $\frac{a_{Lt}}{k_{t+1}}$, because high-types have higher savings rates. This is sufficient for capital to be unambiguously welfare improving — equivalently $\Delta_{t-1} > 0$.

Therefore, using the results from part (i), $\Delta_{t-1} > 0$ for all $t > \tau$.

Note ($\bar{k} > k^{gr}$). Define $\bar{k}$ as the capital stock at which $\Delta_t = 0$. We claim $\bar{k} > k^{gr}$. To see this, evaluate Δ_t at $k_t = k^{gr}$. The pure capital accumulation motive (the $r_t > 0$ term) is zero at k^{gr}, but Δ_t also contains a distributional component: higher capital raises wages and lowers returns, redistributing factor income toward labor. Since low-productivity households earn a weakly greater share of income from labor than from capital (by construction, as high-types save more), and since they carry weakly higher welfare weights (by the assumption on Pareto weights, which holds as long as we have less-than-first-best equality), this distributional component is strictly positive at k^{gr}. Therefore $\Delta_t > 0$ at $k_t = k^{gr}$, which implies $\bar{k} > k^{gr}$. $\square$

iii) If the constrained planner implements first-best equality or greater, $(c^y_{Ht}/c^y_{Lt})^{\sigma_y} \leq \frac{\lambda_H}{\lambda_L}$, in every period, the planner's first order condition cannot hold.

Consider a policy that implements $(c^y_{Ht}/c^y_{Lt})^{\sigma_y} \leq \frac{\lambda_H}{\lambda_L}$ for all $t \geq 0$. Equation (A.16) becomes:

$$\gamma \underbrace{\left(\lambda_L (c^y_{Lt})^{-\sigma_y} - \lambda_H (c^y_{Ht})^{-\sigma_y}\right)}_{\leq 0} + \sum_{j=t}^{\infty} \gamma^j \underbrace{\Delta_j}_{>0} \underbrace{\frac{dk_{j+1}}{dT_{Lt}}}_{<0} = 0$$

The first term is ≤ 0 — equal to zero at exactly first-best equality and strictly negative at greater-than-first-best equality. The second term is strictly negative since $\Delta_j > 0$ for all $j > \tau$ (from step (ii)) and $\frac{dk}{dT_L} < 0$. Therefore the sum of both terms is strictly negative for all $t > \tau$, contradicting the FOC. This rules out both first-best equality and greater-than-first-best equality in the constrained optimal allocation.

iv) If $k_0 < k^{gr}$ then $c^y_{Ht}/c^y_{Lt} > (\lambda_H/\lambda_L)^{1/\sigma_y}$ for all $t \geq 0$.

From step (iii), the constrained optimal allocation features less-than-first-best equality at some $\tau \geq 0$. It remains to show this persists for all $t > \tau$. Since $\bar{k} > k^{gr}$ (established in the note following step (ii)), and the planner has no incentive to push k_t above $\bar{k}$ — doing so would make $\Delta_t < 0$, meaning additional capital is welfare-reducing — the constrained optimal capital stock satisfies $k_t < \bar{k}$ for all t. Therefore $\Delta_t > 0$ for all t, and the FOC (A.16) requires:

$$\gamma \left(\lambda_L (c^y_{Lt})^{-\sigma_y} - \lambda_H (c^y_{Ht})^{-\sigma_y}\right) > 0$$

in every period, which is equivalent to less-than-first-best equality in every period. □

A.6 Comparative Statics.

A.6.1 Proof of Lemma 3.

From asset market clearing, equilibrium capital satisfies $k_{t+1} = \sum_I \pi a_{it}$. Taking the total derivative with respect to T_{Lt}:

$$\frac{dk_{t+1}}{dT_{Lt}} = \sum_I \pi \frac{da_{it}}{dT_{Lt}} = \sum_I \pi \left(\frac{\partial a_{it}}{\partial T_{it}} \frac{\partial T_{it}}{\partial T_{Lt}} + \frac{\partial a_{it}}{\partial r_{t+1}} \frac{dr_{t+1}}{dk_{t+1}} \frac{dk_{t+1}}{dT_{Lt}} \right)$$

Since $T_{Lt} = -T_{Ht}$ (a transfer from H to L), we have $\frac{\partial T_{Lt}}{\partial T_{Lt}} = 1$ and $\frac{\partial T_{Ht}}{\partial T_{Lt}} = -1$. By definition, $\frac{\partial a_{it}}{\partial T_{it}} = MPS_{it}$. Collecting terms:

$$\frac{dk_{t+1}}{dT_{Lt}} = \pi \left(MPS_{Lt} - MPS_{Ht}\right) + \frac{dk_{t+1}}{dT_{Lt}} \sum_I \pi \frac{\partial a_{it}}{\partial r_{t+1}} \frac{dr_{t+1}}{dk_{t+1}}$$

Rearranging and collecting $\frac{dk_{t+1}}{dT_{Lt}}$ on the left-hand side:

$$\frac{dk_{t+1}}{dT_{Lt}} \left(1 - \sum_I \pi \frac{\partial a_{it}}{\partial r_{t+1}} \frac{dr_{t+1}}{dk_{t+1}}\right) = \pi \left(MPS_{Lt} - MPS_{Ht}\right)$$

Multiplying and dividing the sum by $\frac{a_{it}}{k_{t+1}} \cdot \frac{r_{t+1}}{r_{t+1}}$ to express in elasticity form:

$$\sum_I \pi \frac{\partial a_{it}}{\partial r_{t+1}} \frac{dr_{t+1}}{dk_{t+1}} = \sum_I \frac{a_{it}}{k_{t+1}} \frac{d\log a_{it}}{d\log r_{t+1}} \frac{d\log r_{t+1}}{d\log k_{t+1}} = (k_{t+1}^{r_{t+1}})^{-1} \mathcal{A}_t^{r_{t+1}}$$

where $\mathcal{A}_t^{r_{t+1}} \equiv \sum_I \frac{a_{it}}{k_{t+1}} \frac{d\log a_{it}}{d\log r_{t+1}}$ is the asset-weighted interest rate elasticity of household savings and $k_{t+1}^{r_{t+1}} \equiv \left(\frac{d\log r_{t+1}}{d\log k_{t+1}}\right)^{-1}$ is the partial equilibrium elasticity of firms' capital demand with respect to the interest rate. Dividing through gives:

$$\frac{dk_{t+1}}{dT_{Lt}} = \frac{\pi MPS_{Lt} - \pi MPS_{Ht}}{1 - (k_{t+1}^{r_{t+1}})^{-1} \mathcal{A}_t^{r_{t+1}}}$$

The denominator is positive, as established in Section A.2.2. The numerator is negative whenever $MPS_{Lt} < MPS_{Ht}$, which holds by Lemma 1 whenever $\beta_L < \beta_H$ or $\sigma_y > \sigma_o$. Therefore $\frac{dk_{t+1}}{dT_{Lt}} < 0$ under these conditions, and equals zero when MPS are uniform. $\square$

A.6.2 Heterogeneous rates of return.

There are two types of firms, high and low productivity. Each firm produces output according to the following CES production function, where A_i denotes type-specific firm productivity.

$$y_i(k_{it}^d, \ell_{it}^d) = A_i f^i(k_{it}^d, \ell_{it}^d) = A_i \left(\alpha_k k_{it}^{d\zeta} + \alpha_\ell \ell_{it}^{d\zeta}\right)^{\frac{1}{\zeta}}$$

Because high(low)-labor-productivity households are restricted to lending capital to high(low)-productivity firms, each firm offers a type-dependent return, r_{it}, but both firms hire from a single labor market and face the same wage. Firms' first order conditions are:

$$
\begin{aligned}
w_t = A_L f_\ell^L(k_{Lt}, \ell_{Lt}^d) &= A_H f_\ell^H(k_{Ht}, \ell_{Ht}^d) \\
r_{Lt} &= A_L f_k^L(k_{Lt}^d, \ell_{Lt}^d) \\
r_{Ht} &= A_H f_k^H(k_{Ht}^d, \ell_{Ht}^d)
\end{aligned}
$$

Market clearing conditions are now given by the following:

$$
\begin{aligned}
\pi_{Lt} a_{Lt} &= k_{Lt+1}^d \\
\pi_{Ht} a_{Ht} &= k_{Ht+1}^d \\
\pi_{Lt}\ell_{Lt} + \pi_{Ht}\ell_{Ht} &= \ell_{Lt}^d + \ell_{Ht}^d
\end{aligned}
$$

As a result, when $A_L < A_H$, $r_{Lt} < r_{Ht}$.

A.6.3 Capital-skill complementarity.

Capital-skill complementarity. An important dimension of the redistribution-investment trade-off is that more capital boosts wages in future periods and improves welfare for future generations. If low-income households – presumably with higher welfare weights – tend to work in occupations that are substitutable with capital, this could lower the welfare benefit of future capital and dampen the trade-off. Consider a variant of the simple model with the following nested-CES production function (A.17).

$$
y_t = \left((\alpha_\ell (\ell_t^c)^{\zeta_c} + \alpha_k (k_t)^{\zeta_c})^{\frac{\zeta_s}{\zeta_c}} + (\ell_t^s)^{\zeta_s} \right)^{\frac{1}{\zeta_s}} \tag{A.17}
$$

Suppose that $\zeta_c < 0 < \zeta_s$, implying that capital and complementary labor are combined in an inner nest, which is then substitutable with ℓ_t^s. In this case, it is easy to show that capital increases wages for complementary workers more than for substitutable workers.[29] Suppose that all low-skill households supply a fraction ρ_L

[29] See Appendix A.6 or Krusell et al. (2000) for a discussion.

of their labor exogenously to complementary occupations and the remaining fraction to substitutable labor. Let ρ_H be the analogous fraction for high-skill workers. How would increasing the share of low-productivity workers in substitutable occupations change the redistribution-investment trade-off?

The answer depends on the relative strength of two competing forces. On the one hand, low-productivity households will have higher marginal utilities of consumption, and will therefore tend to have higher welfare weights. The greater the share of low-productivity households working in substitutable occupations, the less they benefit from additional capital relative to high-productivity workers. This dampens the welfare benefit of additional investment. At the same time however, a greater share of high-productivity households working in complementary occupations amplifies the impact of redistribution on future capital. Intuitively, when redistribution lowers the savings rate and capital stock, this lowers wages disproportionately for high-MPS workers, exacerbating the impact on capital in future periods.

Which force dominates depends on the relative strength of the planner's redistribution motive. In fact, it is possible to find λ_H/λ_L sufficiently large that an increase in ρ_H/ρ_L – the share of high-types in complementary occupations relative to low-types – actually *increases* the trade-off. This result is presented formally for a simple case in Lemma 5.

Lemma 5 *Suppose the economy has converged to a steady state and the initial complementary labor shares, ρ_H and ρ_L are set such that $w^s = w^c$. Consider an increase in ρ_H/ρ_L that keeps total complementary labor, ℓ^c and total substitutable labor, ℓ^s constant.*

There exists a threshold, $\bar{\lambda}$ such that if $\lambda_H/\lambda_L > \bar{\lambda}$, the redistribution trade-off is increasing in $\rho_H - \rho_L$. That is, at the constrained optimal allocation, $\mathcal{A}^*$, the optimal degree of inequality, c^{y*}_{Ht}/c^{y*}_{Lt} is increasing in $\rho_H - \rho_L$.

Proof. Lemma 5 says that when the planner's redistribution motive is sufficiently weak, the redistribution-investment trade-off is amplified by capital-skill complementarity. As more high-skill households work in complementary occupations and more low-skill workers work in substitutable occupations, additional capital is less beneficial to future generations, but the effect of redistribution on capital accumulation is more pronounced because the permanent income of high-MPS households is especially sensitive to the loss of capital. When the planner's redistributive motive is weak enough, the latter force dominates, and it is optimal for the planner to tolerate greater intra-generational inequality.

Production function:

$$y_t = \left[\left(\alpha k_t^{\zeta_c} + (1-\alpha)(\ell_t^c)^{\zeta_c}\right)^{\frac{\zeta_s}{\zeta_c}} + (\ell_t^s)^{\zeta_s}\right]^{\frac{1}{\zeta_s}}$$

Define the inner CES aggregate (capital and complementary labor):

$$Q_t = \left[\alpha k_t^{\zeta_c} + (1-\alpha)(\ell_t^c)^{\zeta_c}\right]^{\frac{1}{\zeta_c}}$$

So: $y_t = \left[(Q_t)^{\zeta_s} + (\ell_t^s)^{\zeta_s}\right]^{\frac{1}{\zeta_s}}$

Parameter interpretation: $\zeta_c < 0$ implies capital and complementary labor are complements (elasticity $\sigma_{kc} = \frac{1}{1-\zeta_c} < 1$), and $0 < \zeta_s < 1$ implies the (k, ℓ^c) bundle and substitutable labor are substitutes (elasticity $\sigma = \frac{1}{1-\zeta_s} > 1$).

i) Show $\frac{\partial w_t^c/\partial k_t}{\partial w_t^s/\partial k_t} > 1$ when $\zeta_c < 0$ and $\zeta_s \in (0,1)$

Marginal products:

$$w_t^c = (1-\alpha)\left(\frac{\ell_t^c}{Q_t}\right)^{\zeta_c - 1}\left(\frac{Q_t}{y_t}\right)^{\zeta_s - 1}$$
$$w_t^s = \left(\frac{\ell_t^s}{y_t}\right)^{\zeta_s - 1}$$
$$r_t + \delta = \alpha\left(\frac{k_t}{Q_t}\right)^{\zeta_c - 1}\left(\frac{Q_t}{y_t}\right)^{\zeta_s - 1}$$

Derivatives with respect to capital:

$$\frac{\partial w_t^c}{\partial k_t} = \alpha(1-\alpha)\left(\frac{\ell_t^c}{Q_t}\right)^{\zeta_c - 1}\left(\frac{k_t}{Q_t}\right)^{\zeta_c - 1}\frac{Q_t^{\zeta_s - 1}}{y_t^{\zeta_s}}\left[(\zeta_s - \zeta_c) + (1-\zeta_s)\phi_Q\right]$$
$$\frac{\partial w_t^s}{\partial k_t} = (1-\zeta_s)\alpha\left(\frac{\ell_t^s}{y_t}\right)^{\zeta_s - 1}\left(\frac{k_t}{Q_t}\right)^{\zeta_c - 1}\frac{Q_t^{\zeta_s - 1}}{y_t}$$

where $\phi_Q = \frac{Q_t^{\zeta_s}}{y_t^{\zeta_s}} \in (0,1)$ is the CES share of the (k, ℓ^c) bundle.

Signs: With $\zeta_c < 0$ and $0 < \zeta_s < 1$, both derivatives are positive since $(\zeta_s - \zeta_c) > 0$, $(1-\zeta_s) > 0$, and $\phi_Q > 0$.

Taking the ratio gives

$$\frac{\frac{\partial w_t^c}{\partial k_t}}{\frac{\partial w_t^s}{\partial k_t}} = \frac{(1-\alpha)}{(1-\zeta_s)} \left(\frac{\ell_t^c}{Q_t}\right)^{\zeta_c - 1} \left(\frac{y_t}{\ell_t^s}\right)^{\zeta_s - 1} y_t^{1-\zeta_s} \left[(\zeta_s - \zeta_c) + (1-\zeta_s)\phi_Q\right]$$

Assuming $w_t^c = w_t^s$ as in the text:

$$\left.\frac{\partial w^c/\partial k}{\partial w^s/\partial k}\right|_{w^c = w^s} = \frac{(\zeta_s - \zeta_c) + (1-\zeta_s)\phi_Q}{(1-\zeta_s)} \cdot \phi_Q^{(1-\zeta_s)/\zeta_s} = \frac{1}{\mathcal{C}} > 1$$

Result: $\zeta_s \in (0,1)$ and $\phi_Q \in (0,1)$, and $\zeta_c < 0$. Therefore $\mathcal{C} < 1$. Complementary labor benefits more from capital accumulation due to direct complementarity within the Q-bundle, while substitutable labor only benefits through general equilibrium effects.

Note that the derivative of rental rate with respect to capital:

$$\frac{\partial r_t}{\partial k_t} = \alpha^2 \left(\frac{k_t}{Q_t}\right)^{\zeta_c - 1} \frac{Q_t^{\zeta_s - 1}}{y_t^{\zeta_s}} \left[(\zeta_c - 1) + (\zeta_s - \zeta_c)\frac{k_t^{\zeta_c}}{Q_t^{\zeta_c}} + (1-\zeta_s)\frac{Q_t^{\zeta_s}}{y_t^{\zeta_s}} \frac{k_t^{\zeta_c}}{Q_t^{\zeta_c}}\right]$$

which implies that:

$$\frac{\partial r_t}{\partial k_t} = \frac{\alpha}{1-\alpha} \cdot \frac{\ell_t^c}{k_t} \cdot \frac{\partial w_t^c}{\partial k_t}$$

ii) Derive $\rho_L(\rho_H)$ to maintain aggregate quantities and income distributions.

As in the text, let ρ_L (or ρ_H) be the share of low (high)-skill labor in complementary occupations. Therefore, type-i households' labor income is given by:

$$\rho_i w_t^c \theta_i + (1-\rho_i) w_t^s \theta_i$$

Wages for each type are determined by their task allocation:

$$w_L = \rho_L w^c + (1-\rho_L) w^s$$
$$w_H = \rho_H w^c + (1-\rho_H) w^s$$

Total labor income for each type:

$$y_L = \theta_L w_L, \quad y_H = \theta_H w_H$$

We impose two aggregate labor market clearing conditions:

$$\begin{aligned} \pi\rho_L\theta_L + \pi\rho_H\theta_H &= \bar{\ell}^c \\ \pi(1-\rho_L)\theta_L + \pi(1-\rho_H)\theta_H &= \bar{\ell}^s \end{aligned}$$

Since $w^c = w^s$ by assumption, $w_i(\rho_i) = w^c$ regardless of ρ_i, so total labor income $\theta_i w_i = \theta_i w^c = \bar{y}_i$ is unchanged for any (ρ_L, ρ_H) — income is preserved by construction. The two aggregate labor market clearing conditions then determine ρ_L as a function of ρ_H.

$$\begin{aligned} \pi\frac{\rho_L\bar{y}_L}{w_L(\rho_L)} + \pi\frac{\rho_H\bar{y}_H}{w_H(\rho_H)} &= \bar{\ell}^c \\ \pi\frac{(1-\rho_L)\bar{y}_L}{w_L(\rho_L)} + \pi\frac{(1-\rho_H)\bar{y}_H}{w_H(\rho_H)} &= \bar{\ell}^s \end{aligned}$$

Define:

$$R_H(\rho_H) \equiv \pi\frac{\rho_H\bar{y}_H}{\rho_H w^c + (1-\rho_H)w^s}$$

Then ρ_L as a function of ρ_H is given by:

$$\rho_L(\rho_H) = \frac{(\bar{\ell}^c - R_H(\rho_H))w^s}{\pi\bar{y}_L - (\bar{\ell}^c - R_H(\rho_H))(w^c - w^s)}$$

Since $\frac{dR_H}{d\rho_H} = \pi\bar{y}_H\frac{w^s}{w_H(\rho_H)^2} > 0$, we have $\frac{d\rho_L}{d\rho_H} < 0$.
Therefore, it is possible to simultaneously increase ρ_H (high-skilled allocate more to complementary tasks) and decrease ρ_L (low-skilled allocate more to substitutable tasks) while holding aggregate labor supplies $(\bar{\ell}^c, \bar{\ell}^s)$ and total incomes $(\bar{y}_L, \bar{y}_H)$ constant. Because preferences and the distribution of income is the same, savings, capital, and prices will also be the same.

Plugging these results into the constrained planner's first order condition adapted from section A.5.2, we have:

$$\gamma^t \sum_I \lambda_i (c^y_{it})^{-\sigma_y} \frac{dT_{it}}{dT_{Lt}} + \sum_{j=t}^{\infty} \Bigg(\gamma^j \frac{\partial r_{j+1}}{\partial k_{j+1}} \sum_I \lambda_i (c^y_{ij})^{-\sigma_y} \frac{a_{ij}}{1+r_{j+1}}$$

$$\gamma^{j+1} \frac{\partial w^c_{j+1}}{\partial k_{j+1}} \sum_I \lambda_i (c^y_{ij+1})^{-\sigma_y} \theta_i (\rho_i + (1-\rho_i)\mathcal{C}) + \tau_{kj+1} \gamma^j \Theta_j \left(1 + r_{j+1} + k_{j+1} \frac{\partial r_{j+1}}{\partial k_{j+1}}\right) \Bigg) \frac{dk_{j+1}}{dT_{Lj}} = 0$$

The term Θ is defined as before:

$$\Theta_j = \sum_I \left(\frac{\lambda_i (c^y_{ij})^{-\sigma_y}}{(1+r_{j+1})(1-\tau_{kj+1})} + \gamma \lambda_i (c^y_{ij+1})^{-\sigma_y} \right)$$

Because by assumption the economy has converged to a steady state, and using the definition of Δ from equation (A.16) in Appendix A.5.2, the above can be written:

$$\sum_I \lambda_i (c^y_i)^{-\sigma_y} \frac{dT_i}{dT_L} + \frac{\Delta}{1 - \gamma \frac{dk^E(.)}{dk}} \frac{dk}{dT_L} = 0$$

Set ρ_L and ρ_H as above to match the same aggregate moments, implying the same $\mathcal{C}$, w^c_t, w^s_t, k_t etc. for all $t \geq 0$. Because preferences and incomes are the same, c^y_{it} is also the same.

This implies that the planner's first order condition is identical between the two models except for:

1. $\sum_I \lambda_i (c^y_i)^{-\sigma_y} \theta_i (\rho_i + (1-\rho_i)\mathcal{C})$
2. $\frac{dk^E(.)}{dk}$

iii) Show that the welfare benefits of additional capital are muted as $\rho_H - \rho_L$ increases.

The derivative of the first term with respect to ρ_H is:

$$\lambda_L (c^y_L)^{-\sigma_y} \theta_L \left(\frac{d\rho_L}{d\rho_H} (1-\mathcal{C}) \right) + \lambda_H (c^y_H)^{-\sigma_y} \theta_H (1-\mathcal{C})$$

With egalitarian Pareto weights, $\lambda_i = 1$ for both types, because $\mathcal{C} < 1$, the first term is unambiguously decreasing in $\rho_H - \rho_L$ for sufficiently low $\mathcal{C}$. Intuitively, the more low-income types work in substitutable occupations, the less they benefit from additional capital. This makes the welfare costs of losing capital less severe as capital primarily benefits the high-income types in future generations.

However, as $\lambda_H/\lambda_L \to \infty$ the derivative of the first term vanishes. Therefore, there exists a threshold in which if $\lambda_H/\lambda_L > \bar{\lambda}$, the above derivative is 0.

iv) Show that $\frac{dk^E(.)}{dk}$ increases with $\rho_H - \rho_L$.

Because k_t is predetermined, $\frac{dk_{t+1}}{dT_{Lt}}$ is unaffected by wages in period-t and therefore unaffected by $\rho_H - \rho_L$.

However using the results from Section A.2.2, $\frac{dk^E(.)}{dk}$ is given by:

$$\frac{dk^E(.)}{dk} = \frac{\pi \sum_I \frac{da_{it}}{dk_t}}{1 - \pi \sum_I \frac{da_{it}}{dk_{t+1}}}$$

Then simply note that $\pi \sum_I \frac{da_{it}}{dk_t}$ is equal to $\pi \sum_I MPS_i \theta_i (\rho_i + (1-\rho_i)\mathcal{C}) \frac{dw^c}{dk}$. By assumption, $MPS_H > MPS_L$. Therefore $\frac{dk^E(.)}{dk}$ is increasing in $\rho_H - \rho_L$.

Conclusion. Combining parts (iii) and (iv): as $\rho_H - \rho_L$ increases, the welfare benefit of additional capital (the first term in the steady-state FOC) falls, while $\frac{dk^E(.)}{dk}$ rises, making $1 - \gamma \frac{dk^E(.)}{dk}$ smaller and amplifying the second term. When $\lambda_H/\lambda_L > \bar{\lambda}$, the first effect vanishes and the second dominates — the planner must accept greater intra-generational inequality to satisfy the FOC, so the constrained optimal degree of inequality c_H^{y*}/c_L^{y*} is increasing in $\rho_H - \rho_L$. □

A.6.4 Direct effect of linear labor income tax with capital-skill complementarity.

In this case, the uniform transfer is equal to $\tau\left(w^c \ell^c + w^s \ell^s\right)$ =, while each type-i household's tax burden at age-j is equal to $\tau \ell_{ij} \theta_{ij} w_i \epsilon$, where $w_i = \rho_i w^c + (1-\rho_i) w^s$ and ϵ is their realization of the idiosyncratic shock. The change in the flow of after-ax labor income at age-j is therefore:

$$\tau\left(w^c(\bar{\ell}^c - \theta_{ij}\rho_i \ell_{ij}\epsilon) + w^s(\bar{\ell}^s - \theta_{ij}(1-\rho_i)\ell_{ij}\epsilon)\right)$$

We take the expectation of the present value of these flows, noting again that we

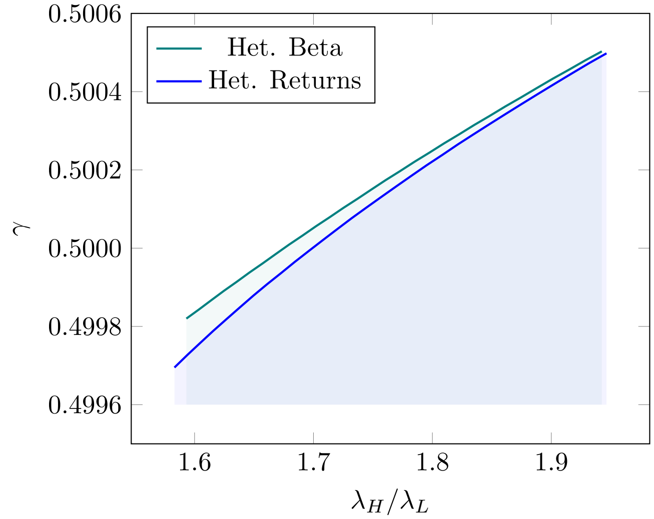

Figure 6: Heterogeneous Beta vs. Heterogeneous Returns

have calibrated the model so that $\mathbb{E}[\ell_{ij}\epsilon] = 1$. This gives:

$$\Delta\mathbb{E}[PI_i] = \tau \sum_j \left(w^c(\bar{\ell}^c - \theta_{ij}\rho_i) + w^s(\bar{\ell}^s - \theta_{ij}(1-\rho_i)) \right) (R^{j-1})^{-1}$$

A.6.5 Heterogeneous rates of return.

If instead of type-dependent discount factors, households had type-dependent rates of return, the trade-off between redistribution and investment would be amplified.[30] Consider a variant of the simple model in which households with high labor productivity had exclusive access to lend capital to higher productivity firms, while low-productivity households only have access to low productivity capital. Both high and low-productivity firms hire from the same labor market and are subject to the same wage rate. As a result, high-productivity firms offer high-labor-productivity type households a higher return, $r_{Ht} > r_{Lt}$.[31]

Figure 6 again plots the portion of the Pareto Frontier that is implementable with fiscal policy in a model with type-dependent discount factors (green) and type dependent rates-of-return (blue). The two examples are calibrated to exactly match aggregate output and the labor share, and closely match the marginal propensities

[30] Type dependent rates of return have been documented empirically. See Fagereng et al. (2020).
[31] For more details on this model variant, see Appendix A.6.

to save for both types. From the figure, we can see that at the degree of redistribution increases, the trade-off in the type-based returns case is more severe than the heterogeneous β despite being calibrated to match the same initial distribution of MPS. Intuitively, when high-income households have access to more productive capital, redistribution moves resources away from those who would not only save *more* but would also invest their capital more productively.

A.7 Proof of Proposition 3

A.7.1 Proof of Lemma 4.

The equilibrium conditions are the following:

$$(c_{it}^{y})^{-\sigma_y} = \beta_i(1+r_{t+1})(c_{it+1}^{o})^{-\sigma_o} \tag{A.18}$$

$$c_{it}^{y} = w_t\theta_i(1-\tau_{\ell t})\ell_{it} + T_t - a_{it}\,; \quad c_{it}^{o} = a_{it-1}(1+r_t) \tag{A.19}$$

$$\ell_{it} = ((1-\tau_{\ell t})w_t\theta_i)^{\frac{1}{\nu_i}}(c_{it}^{y})^{-\frac{\sigma_y}{\nu_i}} \tag{A.20}$$

$$w_t = f_\ell(k_t,\ell_t)\,; \quad r_t = f_k(k_t,\ell_t) - \delta \tag{A.21}$$

$$T_t = \tau_{\ell t}\ell_t w_t \tag{A.22}$$

Combine the households' first and second period budget constraints, (A.19) into a lifetime budget constraint. Substitute in households labor supply condition (A.20) in to get $\ell_{it}(c_{it}^{y})$ and Euler equation (A.18) to get $c_{it}^{o}(r_{t+1}, c_{it}^{y})$. The lifetime budget constraint then gives a single implicit function for c_{it}^{y} as a function of r_{t+1} and PI_{it}.

$$c_{it}^{y} + \frac{c_{it}^{y}(c_{it+1}^{o}, r_{t+1})}{1+r_{t+1}} = (1-\tau_{\ell t})w_t\theta_i\ell_{it}(c_{it}^{y}, (1-\tau_{\ell t})w_t\theta_i) + T_t = PI_{it}$$

The total general equilibrium change in c_{it}^{y} as a result of the policy can be written as:

$$\frac{dc_{it}^{y}}{d\tau_{\ell t}} = \frac{\partial c_{it}^{y}}{\partial r_{t+1}}\frac{dr_{t+1}}{d\tau_{\ell t}} + \frac{\partial c_{it}^{y}}{\partial PI_{it}}\frac{dPI_{it}}{d\tau_{\ell t}}$$

Here, $\frac{dr_{t+1}}{d\tau_{\ell t}}$ and $\frac{dPI_{it}}{d\tau_{\ell t}}$ are the total general equilibrium changes resulting from the tax.

The latter can be decomposed in the following way:

$$\frac{dPI_{it}}{d\tau_{\ell t}} = \underbrace{\theta_i(1-\tau_{\ell t})\left(\frac{dw_t}{d\tau_{\ell t}}\ell_{it} + w_t\frac{d\ell_{it}}{d\tau_{\ell t}}\right)}_{\text{Individual Earnings}} + \underbrace{\tau_{\ell t}\left(\frac{dw_t}{d\tau_{\ell t}}\ell_t + \frac{d\ell_t w_t}{d\tau_{\ell_t}}\right)}_{\text{Lump-Sum Transfer}} + \underbrace{w_t(\ell_t - \theta_i\ell_{it})}_{\text{Direct Change in PI}}$$

Using the first period budget constraint, the change in steady state savings for type-i households as a result of a permanent change policy can then be written in the following way:

$$\frac{da_i}{d\tau_\ell} = \frac{dPI_i}{d\tau_\ell} - \frac{dc_i^y}{d\tau_\ell} = \left(1 - \frac{\partial c_i^y}{\partial PI_i}\right)\frac{dPI_i}{d\tau_\ell} - \frac{\partial c_i^y}{\partial r}\frac{dr}{d\tau_\ell}$$

Note that $\frac{dPI_i}{d\tau_\ell}$ and $\frac{dr}{d\tau_\ell}$ are the general equilibrium changes as a result of the permanent increase in the tax. Define a household's marginal propensity to save out of permanent income, $MPS_i = \left(1 - \frac{\partial c_i^y}{\partial PI_i}\right)$ which we have established is only unique to a type-i household if preferences are non-homothetic or βs are heterogeneous. Let $M\bar{P}S$ be the average MPS_i.

The total change in steady state capital following a permanent increase in the tax as:

$$\frac{dk}{d\tau_\ell} = \sum_I \pi \frac{da_i}{d\tau_\ell}$$

This can further be decomposed into changes associated with differing marginal propensities to save, and changes associated with the aggregate labor distortion.

$$\frac{dk}{d\tau_\ell} = \underbrace{M\bar{P}S\left(\frac{dw}{d\tau_\ell}\ell + \frac{d\ell}{d\tau_\ell}w\right)}_{\text{Average change in labor income}} + \underbrace{\sum_I \pi MPS_i\Big(w(\ell - \theta_i\ell_i)\Big)}_{\text{Direct effect of PI redistribution}} +$$
$$\underbrace{\sum_I \pi\Big(MPS_i - M\bar{P}S\Big)\left(\theta_i(1-\tau_\ell)\left(\ell_i\frac{dw}{d\tau_\ell} + w\frac{d\ell_i}{d\tau_\ell}\right)\right)}_{\text{Covariance between labor earnings distortion and MPS}} + \underbrace{\sum_I \pi\frac{da_i}{dr}\frac{dr}{d\tau_\ell}}_{\text{Interest Rate Effect}}$$

Define the elasticity of aggregate savings to the interest rate

$$A_r \equiv \sum_I \pi \frac{a_i}{k}\frac{da_i}{dr}\frac{r}{a_i} \tag{A.23}$$

Then, using the result that for CRS production functions, $\frac{\partial r}{\partial \ell} = -\frac{\partial w}{\partial \ell}\frac{\ell}{k}$ the interest rate effect can be written as:

$$A_r\left(\frac{dr}{dk}\frac{k}{r}\frac{dk}{kd\tau_\ell} + \frac{dr}{d\ell}\frac{\ell}{r}\frac{d\ell}{\ell d\tau_\ell}\right) = A_r\left(r_k\frac{dk}{kd\tau_\ell} - \frac{dw}{d\ell}\frac{\ell}{k}\frac{w}{w}\right)$$

Above I use the fact that $A = k$. Similarly, I define the elasticity of aggregate savings to the wage

$$A_w \equiv \sum M\bar{P}S + Cov\left(MPS_i, \theta_i\frac{\ell_i}{\ell}\right)(1-\tau_\ell) * \frac{w}{k} \tag{A.24}$$

Denote $k_r \equiv \frac{dr}{dk}\frac{k}{r}$. Multiplying the entire term by $\frac{1}{k}$ and the first term $\frac{\ell}{\ell}\frac{w}{w}$ gives the following expressions for the semi-elasticity of steady state capital to τ_ℓ.

$$\frac{dk}{d\tau_\ell}\frac{1}{k} = \frac{\frac{w\ell}{k}\left(M\bar{P}S\cdot\epsilon_{\tau_\ell} + \sum_I MPS_i\left(1-\frac{\theta_i\ell_i}{\ell}\right) + Cov\left(MPS_i, \epsilon_{i\tau_\ell}\right) - A_r w_\ell \epsilon_{\tau_\ell}\right)}{\left(1 - A_r k_r^{-1} - A_w w_k\right)}$$

Where, as in the text, I define the (*total, general equilibrium*) semi-elasticity of aggregate labor income with respect to the policy, $\epsilon_{\tau_\ell} \equiv \frac{d(w\ell)}{d\tau_\ell(w\ell)}$ and the individual semi-elasticity of labor earnings with respect to the tax for type-i, $\epsilon_{i\tau_\ell}$ analogously.

This change in capital can be decomposed into the direct effect, $\mathcal{K}_D$ and the indirect effect, $\mathcal{K}_I$:

$$\frac{dk}{kd\tau_\ell} = \underbrace{\frac{\frac{w\ell}{k}\left(\sum_I MPS_i\left(1-\frac{\theta_i\ell_i}{\ell}\right)\right)}{\left(1 - A_r k_r^{-1} - A_w w_k\right)}}_{\mathcal{K}_D} + \underbrace{\frac{\frac{w\ell}{k}\left(M\bar{P}S\cdot\epsilon_{\tau_\ell} + Cov\left(MPS_i, \epsilon_{i\tau_\ell}\right) - A_r w_\ell \epsilon_{\tau_\ell}\right)}{\left(1 - A_r k_r^{-1} - A_w w_k\right)}}_{\mathcal{K}_I}$$

A.7.2 Marginal impact of τ_ℓ on social welfare.

Steady state social welfare is defined as in the text.

$$SW \equiv \sum_{i\in I} \lambda_i \pi \left(\gamma \left(\frac{(c_i^y)^{1-\sigma_y}}{1-\sigma_y} - \psi_i \frac{\ell_i^{1+\nu_i}}{1+\nu_i} \right) + \beta_i \frac{(c_i^o)^{1-\sigma_o}}{1-\sigma_o} \right)$$

The first order conditions with respect to τ_ℓ are:[32]

$$\frac{dSW}{d\tau_\ell} = \sum_I \lambda_i \pi \left(\gamma (c_i^y)^{-\sigma_y} \left(\frac{dT}{d\tau_\ell} - w\ell_i\theta_i + \theta_i(1-\tau_\ell)\ell_i \frac{dw}{d\tau_\ell} \right) + \beta_i (c_i^o)^{-\sigma_o} a_i \frac{dr}{d\tau_\ell} \right)$$

The tax changes steady state social welfare, first by directly changing the after-tax labor income of the young, and second, by changing factor prices. Plugging in the government's budget constraint for $\frac{dT}{d\tau_\ell}$, and using the fact that in the steady state equilibrium, $\beta_i(c_i^o)^{-\sigma_o} = \frac{(c_i^y)^{-\sigma_y}}{1+r}$, the above can be re-written as:

$$\frac{dSW}{d\tau_\ell} = \sum_I \lambda_i \pi (c_i^y)^{-\sigma_y} \left(\gamma \left((w\ell - w\ell_i\theta_i) + \tau_\ell \frac{d(w\ell)}{d\tau_\ell} + \theta_i(1-\tau_\ell)\ell_i \frac{dw}{d\tau_\ell} \right) + \frac{a_i}{1+r} \frac{dr}{d\tau_\ell} \right)$$

Using the firm's optimality conditions, I can express the total change in w and r as functions of the total general equilibrium change in ℓ and k. Further, I can use the firm's first order conditions to write the ratio between $\frac{\partial r_t}{\partial k_t}$ and $\frac{\partial w_t}{\partial k_t}$ as well as $\frac{\partial r_t}{\partial \ell_t}$ and $\frac{\partial w_t}{\partial \ell_t}$:

$$\begin{aligned} \frac{\partial r}{\partial k} &= (1-\zeta)\frac{f_k}{ky}(f_k k - y) = -(1-\zeta)\frac{f_k f_\ell \ell}{ky} \\ \frac{\partial w}{\partial k} &= (1-\zeta)\frac{f_k f_\ell}{y} \qquad \Rightarrow \qquad \frac{\partial w/\partial k}{\partial r/\partial k} = -\frac{k}{\ell} \\ \frac{\partial w}{\partial \ell} &= (1-\zeta)\frac{f_\ell}{\ell y}(f_\ell \ell - y) = -(1-\zeta)\frac{f_k f_\ell k}{\ell y} \\ \frac{\partial r}{\partial \ell} &= (1-\zeta)\frac{f_k f_\ell}{y} \qquad \Rightarrow \qquad \frac{\partial r/\partial \ell}{\partial w/\partial \ell} = -\frac{\ell}{k} \end{aligned}$$

[32] Note that the Envelope Theorem applies here – households are on their inter and intra-temporal optimality conditions and therefore, there are no direct effects of changing savings levels or labor supply on welfare.

Using this result gives:

$$\frac{dSW}{d\tau_\ell} = w\ell \sum_I \lambda_i \pi (c_i^y)^{-\sigma_y} \Bigg(\gamma \Bigg((1 - \frac{\ell_i}{\ell}\theta_i) + \frac{\tau_\ell}{w\ell} \Bigg(\frac{d(w\ell)}{d\ell}\frac{d\ell}{d\tau_\ell} + \frac{\ell dw}{dk}\frac{dk}{d\tau_\ell} \Bigg) \Bigg) + \frac{1}{w} \Bigg(\frac{dw}{d\ell}\frac{d\ell}{d\tau_\ell} + \frac{dw}{dk}\frac{dk}{d\tau_\ell} \Bigg) \Bigg(\gamma\theta_i (1-\tau_\ell)\frac{\ell_i}{\ell} + \frac{a_i}{k(1+r)} \Bigg) \Bigg)$$

As in the text, denote the steady state social welfare weight of type-i households, $\kappa_i \equiv \lambda_i (c_i^y)^{-\sigma_y}$. Without loss of generality, I assume that λ_i are set such that $\sum_I \kappa_i = 1$. I define $\epsilon_{x,y}$ as the elasticity of variable x with respect to variable y. Collecting terms attributable to the impact of the labor supply distortion on labor income:

$$\epsilon_\ell^D \equiv \sum_I \kappa_i \Bigg(\theta_i (1-\tau_\ell) \frac{\pi \ell_i}{\ell} \epsilon_{w,\ell} + \tau_\ell \epsilon_{w\ell,\ell} \Bigg) \tag{A.25}$$

This term is equal to the total (general-equilibrium) elasticity of after-tax labor income with respect to aggregate labor supply. When labor supply is distorted, it changes the wage, which in turn changes after-tax labor incomes for type-i households (first term of A.25). Distorted labor supply also changes the size of the lump-sum transfer (second term of A.25). When multiplied by the elasticity of aggregate labor supply to the tax, $\epsilon_{\ell,\tau}$, the direct effects captured in ϵ_ℓ^D are the closest analog to the general equilibrium elasticity of after-tax income to the tax used in Piketty and Saez (2013a) in a static model with no capital.

In a dynamic model with capital, the change in the labor supply also impacts rates of return, which are weighted by the product of each types' social welfare weight and their asset income share, $\frac{\pi a_i}{k}$, and discounted by (1+r). Distorted labor also impacts household savings, and therefore the capital stock. Together, these additional indirect effects of distorted labor are given by the following expression:

$$\epsilon_\ell^I \equiv \Bigg(\epsilon_{w,\ell} \sum_I \kappa_i \frac{\pi a_i}{k(1+r)} + \epsilon_{w,k} \Delta \mathcal{K}_I \Bigg) \tag{A.26}$$

Where Δ, defined as in the text, captures the welfare impact of additional capital, and $\mathcal{K}_I$ are the indirect effects of labor on capital accumulation defined in Lemma 4.

I denote the elasticity of total income with respect to the labor supply as the sum of the direct effects through labor income, and the indirect effects through changes in capital and rates of return:

$$\epsilon_\ell = \left(\epsilon_\ell^D + \epsilon_\ell^I\right)\epsilon_{\ell,\tau} \tag{A.27}$$

Next, I collect terms attributable to the steady state change in capital induced by the redistribution of permanent income *alone*, holding labor supply constant:

$$C \equiv \gamma \sum_I \kappa_i \left(\tau_\ell + (1-\tau_\ell)\frac{\theta_i \ell_i}{\ell} - \frac{a_i}{k\gamma(1+r)}\right)\epsilon_{w,k}\mathcal{K}_D \equiv \gamma \Delta \epsilon_{w,k}\mathcal{K}_D$$

Here, I define the terms inside the parenthesis as Δ, as in the text. Putting these terms together and dividing by γ, the total change in steady state social welfare can be written as:

$$\frac{dSW}{d\tau_\ell} \propto \sum_I \kappa_i \left(1 - \frac{\pi\theta_i\ell_i}{\ell}\right) + \epsilon_\ell + \Delta\epsilon_{w,k}\mathcal{K}_D \tag{A.28}$$

A.8 Estimating income distributions.

Using estimates from the Congressional Budget Office (CBO) report on the Distribution of Household Income in 2019 (https://www.cbo.gov/publication/58781), on both the shares and composition of before tax and transfer income by quintile, we can construct the income distribution measure used in the text. For type-i quintiles, their share of total income is:

$$s_i = \frac{\pi(w_t\theta_i\ell_{it} + a_{it}r_t)}{w_t\ell_t + k_t r_t}$$

One can use the ratio of labor income to capital income, rat_i to write $a_{it}r_t = s_i(1+rat_i)^{-1}$. and $w_t\theta_i\ell_{it} = s_i(1+1/rat_i)^{-1}$. Then the type-i share of labor income, $\frac{\pi\theta_i\ell_{it}}{\ell_t} = s_i(1+1/rat_i)^{-1}/(1-\alpha)$ and share of capital income, $\frac{\pi a_{it}}{k_t} = s_i(1+rat_i)^{-1}/\alpha$ where $(1-\alpha)$ is the aggregate labor share.

Quintile	1	2	3	4	5
Share total in 2019	.05	.14	.14	.14	53
Labor/capital income in 2019	61/10	69/6	69/6	69/6	70/15

Table 6: Estimates from the CBO

Note: I use the sum business income, capital income, and capital gains for capital income.

A.9 Details on estimation of MPS.

Dynan et al. (2004) (DSZ) separate households into permanent income quintiles and estimate 'active' savings rates by quintile.[33] DSZ use multiple approaches to generate a measure of permanent income, including using both past and future household income as an instrument. They show that savings rates increase monotonically with permanent income. However, because consumption data had not yet been added to the PSID, their measure of active savings had to be imputed. Starting in 1999, the PSID added questions about household consumption to the survey, followed by another wave of additional consumption variables in 2005. Here, I replicate their estimation procedure with the updated data which allows me to measure active savings directly as after-tax income less consumption. Appendix Table 7 reports savings rates by permanent income quintile using both the 1999 consumption measures and the 2005 consumption measures. I combine these updated estimates of savings rates by permanent income quintile with a savings elasticity implied by Straub (2019) (1.3) to get my measure of MPS out of PI.

A.10 Welfare decomposition in the quantitative model.

Define steady-state social welfare as in the text as the sum over all types and currently living generations of expected utility, weighted by type-specific Pareto weights, λ_i:

$$\mathbb{SW} \equiv \sum_I \pi\lambda_i \left(\sum_J (\beta_i)^{j-1}\gamma^{J-j}\mathbb{E}\left[\frac{c_{ij}^{1-\sigma_j}}{1-\sigma_j} - \psi_{\ell,ij}\frac{\ell_{ij}^{1+\nu}}{1+\nu}\right] + \beta^J\psi_{ia}\mathbb{E}\left[\frac{a_{iJ}^{1-\eta}}{1-\eta}\right]\right)$$

Here as in the simple model, generations are discounted at rate γ. The total derivative of steady state social welfare with respect to the tax is given by the following

[33] Here, active savings refers to a measure of income - consumption, as opposed to a change in wealth which could also capture increasing asset prices.

Table 7: PSID Summary Statistics

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
			All Ages		
Median Income	28,046	48,301	72,130	103,554	172,275
College Share	0.39	0.49	0.59	0.69	0.86
Saving Rate ('99)	0.13	0.18	0.26	0.33	0.44
Saving Rate ('05)	-0.00	0.04	0.13	0.19	0.31
			Ages 20–35		
Median Income	22,925	39,295	57,067	80,113	125,130
Saving Rate ('99)	0.07	0.12	0.18	0.24	0.39
Saving Rate ('05)	-0.11	-0.03	0.02	0.11	0.24
			Ages 35–50		
Median Income	30,113	52,179	76,887	109,110	177,869
Saving Rate ('99)	0.12	0.17	0.26	0.33	0.44
Saving Rate ('05)	-0.02	0.04	0.10	0.20	0.31
			Ages 50–65		
Median Income	30,405	51,089	76,601	111,123	190,758
Saving Rate ('99)	0.19	0.22	0.34	0.39	0.47
Saving Rate ('05)	0.07	0.09	0.22	0.25	0.35
Observations ('99)	9,500	9,085	8,661	8,354	8,429
Observations ('05)	7,870	7,370	6,903	6,692	6,938

This table reports summary statistics for the PSID data by age group and permanent income quintile. Real median income is reported in 2019 dollars. Saving is calculated as annual total post-tax income less consumption. The savings rate is equal to savings over current total income. Saving Rate ('99) is the average savings rate using only the 1999 consumption measures. Saving Rate ('05) uses the 2005 consumption measures. College share is the fraction of household heads with at least a college degree.

expression:

$$\frac{d\mathbb{SW}}{d\tau} = \sum_I \pi\lambda_i\Bigg(\gamma^J u'(c_{i0})\Big(\theta_{i0}\ell_{i0}((1-\tau_{i0}-\tau)dw + w(d\tau_{i0}-1)) + dT + db_{i0}(1-\tau_b)\Big)+$$
$$\sum_{j=1}^{J}\gamma^{J-j}\mathbb{E}\Big[\beta_i^j u'(c_{ij})\Big(\epsilon\theta_{ij}\ell_{ij}((1-\tau_{ij}-\tau)dw + w(d\tau_{ij}-1)) + dT + a_{ij}(1-\tau_k)dr\Big)\Big]\Bigg)$$

Note that here we apply the envelope theorem as before, households are optimizing with respect to their labor, savings, and bequest choices, and therefore changes in ℓ_{ij} and a_{ij} have no direct effect on welfare. Using the government's budget constraint, I can write the total general equilibrium change in the lump-sum tax, dT as:

$$\begin{aligned} dT = d\tau\ell w + dRev \quad &\text{where } \tau_{ij} \text{ is the progressive tax and} \\ dRev \equiv w\ell\mathbb{E}\left[\frac{\theta_i}{\ell}\pi\epsilon_i(\tau_{ij}\ell_{ij}\frac{dw}{w} + \tau_{ij}d\ell_{ij} + \ell_{ij}d\tau_{ij})\right] &+ \tau(\ell dw + wd\ell) + \\ \tau_k(rdA + Adr) &+ \sum_I \pi_{i0}db_{i0}\tau_b + Bdr \end{aligned} \tag{A.29}$$

The first term in $dRev$ is the total change in expected revenue from the progressive tax. Define $\mathcal{W}_P$ as the ratio of this expected change over total labor income, $w\ell$. Again, using the result that $r = \frac{\alpha}{1-\alpha}\frac{w\ell}{k}$ and that $dr = -dw\frac{\ell}{k}$, we can write dT as:

$$dT = w\ell\left(1 + \mathcal{W}_P + \tau_k\left(\frac{\alpha}{1-\alpha}\frac{dk}{kd\tau} - \frac{A}{k}\frac{dw}{w}\right) - \frac{B}{k}\frac{dw}{w}\right) + \tau_b\sum_I \pi_{i0}db_{i0}$$

I denote $\bar{r} \equiv r(1-\tau_k)$. Note that τ_k^j is a shorthand where $\tau_k^0 = \tau_b$ (for the youngest households, whose initial wealth is a bequest subject to the bequest tax) and $\tau_k^j = \tau_k$ for $j \geq 1$ (for older households whose assets earn the after-capital-tax return). Using the fact that we are starting in steady state, that the initial $\tau = 0$, that $\mathbb{E}[\epsilon_i\ell_{ij}] = 1$, and that $u'(c_{ij}) = (1+(1-\tau_k)r)\beta_i\mathbb{E}[u'(c_{i,j+1})]$ we have:

$$\begin{aligned} = \sum_I \pi\lambda_i\Bigg(\sum_{j=0}^{J-1}\gamma^{J-j}\frac{u'(c_{i0})}{(1+\bar{r})^j}\Big(\theta_{ij}((1-\bar{\tau}_{ij})dw + w(d\bar{\tau}_{ij}-1)) + dT + a_{ij}(1-\tau_k^j)dr\Big) \\ + \underbrace{\gamma^0\frac{u'(c_{i0})}{(1+\bar{r})^J}a_{iJ}(1-\tau_k)dr}_{\text{retirement: no labor income, no } dT} \\ + \sum_{j=1}^{J}\gamma^{J-j}\mathbb{E}\Big[u'(c_{ij})\theta_{ij}\Big(\ell_{ij}\epsilon_i((1-\tau_{ij})dw + w(d\tau_{ij}-1) - (1-\bar{\tau}_{ij})dw + w(d\bar{\tau}_{ij}-1))\Big)\Big] \\ + \gamma^J u'(c_{i0})db_{i0}(1-\tau_b)r \end{aligned}$$

Here I have added an 'intelligent 0' in order to decompose the change in welfare into the expected change (first term) and the change associated with the mitigation of risk (second term), as well as changes related to the stock of bequests. Here $\bar{\tau}_{ij}$ is the *average* / expected progressive tax rate for type-i age-j households over all states. The risk term grows as the deviations of the *actual realized* change in after-tax labor income deviates from $\theta_{ij}((1-\bar{\tau}_{ij})dw + wd\bar{\tau}_{ij})$. For simplicity, I denote the welfare effect associated with changing idiosyncratic labor income risk exposure, $\mathcal{W}_R$. Finally, I decompose the change in factor prices and the shift in the incidence of the progressive tax into the contributions of the general equilibrium changes in capital and labor.

$$= w\ell \sum_I \pi\kappa_i \sum_{j=0}^{J-1} \frac{\gamma^{J-j}}{(1+\bar{r})^j} \Bigg(\underbrace{\left(1 - \frac{\theta_{ij}}{\ell}\right)}_{\text{Direct redist.}} + \underbrace{\frac{\theta_i}{\ell} d\bar{\tau}_{ij} + \mathbb{E}\left[\frac{\theta_{ij}\ell_{ij}\epsilon_i\tau_{ij}}{\ell}\left(\frac{d\ell_{ij}}{\ell_{ij}} + \frac{d\tau_{ij}}{\tau_{ij}}\right)\right]}_{\text{shift progr. tax incidence}}$$

$$\left(w_k \frac{dk}{k\tau} + w_\ell \frac{d\ell}{d\ell\tau}\right) \underbrace{\left(\frac{\theta_{ij}}{\ell}((1-\bar{\tau}_{ij}) - \frac{a_{ij}}{k}(1-\tau_k^j) - \tau_k \frac{A}{k} + \frac{B}{k} + \mathbb{E}\left[\frac{\theta_{ij}}{\ell}\epsilon_i \tau_{ij}\ell_{ij}\right]\right)}_{\text{impact of changing factor prices on working-age households}} +$$

$$\underbrace{\tau_k \frac{\alpha}{1-\alpha}\frac{dk}{d\tau k}}_{\downarrow \tau_k \text{ revenue}} \Bigg) - \left(w_k \frac{dk}{k\tau} + w_\ell \frac{d\ell}{d\ell\tau}\right) \underbrace{w\ell \sum_I \pi\kappa_i \frac{1}{(1+\bar{r})^J}\frac{a_{iJ}}{k}(1-\tau_k)}_{\text{Impact of dr on retirees}}$$

$$\underbrace{\mathcal{W}_R}_{\text{risk}} + \underbrace{\sum_I \sum_{j=0}^{J-1} \pi\kappa_i \frac{\gamma^{J-j}}{(1+\bar{r})^j}\left(\tau_b \sum_I \pi_{i0} db_{i0}\right) + \gamma^J \pi\kappa_i db_{i0}(1-\tau_b)r}_{\text{Welfare impact of lower bequests}}$$

Here, I am defining $\mathcal{W}_R$ as in equation (A.30).

$$\mathcal{W}_R \equiv \sum_{j=1}^{J} \gamma^{J-j} \mathbb{E}\Bigg[u'(c_{ij})\theta_{ij}\Bigg(\ell_{ij}\epsilon_i((1-\tau_{ij})dw + w(d\tau_{ij}-1) - (1-\bar{\tau}_{ij})dw + w(d\bar{\tau}_{ij}-1))\Bigg)\Bigg] \tag{A.30}$$

Finally, I can further group the terms above into the contributions of the general equilibrium changes in labor and capital.

$$\frac{d\mathbb{SW}}{d\tau} = \sum_I \pi_i \kappa_i \sum_{j=0}^{J-1} \frac{\gamma^{J-j}}{(1+\bar{r})^j} \Bigg(\underbrace{w\ell\left(1-\frac{\theta_{ij}}{\ell}\right)}_{\text{Direct redistri.}} + \underbrace{\tau_b \sum_I \pi_{i0} db_{i0}}_{\text{bequest tax revenue}} \Bigg) + \underbrace{\mathcal{W}_R}_{\text{Risk}} + \underbrace{\sum_I \pi\kappa_i \gamma^J db_{i0}(1-\tau_b)r}_{\text{direct effect of lower bequests}}$$

$$\underbrace{w\ell \sum_I \pi\kappa_i \sum_{j=0}^{J-1} \frac{\gamma^{J-j}}{(1+\bar{r})^j} \left(\frac{\theta_i}{\ell} d\bar{\tau}_{ij} + \mathbb{E}\left[\frac{\theta_{ij}\ell_{ij}\epsilon_i\tau_{ij}}{\ell} \left(\frac{d\ell_{ij}}{\ell_{ij}} + \frac{d\tau_{ij}}{\tau_{ij}} \right) \right] \right) \Bigg)}_{\text{shift progr. labor tax incidence}} + \underbrace{\Delta\epsilon_{\ell,\tau}}_{\text{impact } d\ell \text{ on prices}}$$

$$\underbrace{w\ell \sum_I \pi\kappa_i \sum_{j=0}^{J-1} \frac{\gamma^{J-j}}{(1+\bar{r})^j} \left(\tau_k \frac{\alpha}{1-\alpha} \epsilon_{k,\tau} \right)}_{\downarrow \tau_k \text{ revenue}} + \underbrace{\Delta\epsilon_{k,\tau}}_{\text{impact } dk \text{ on prices}}$$

Here the term Δ captures the welfare impact of changing factor prices and is defined as in the text:

$$\Delta \equiv w\ell \sum_I \pi\kappa_i \left(\sum_{j=0}^{J-1} \left(\frac{\theta_{ij}}{\ell}((1-\bar{\tau}_{ij}) - \frac{a_{ij}}{k}(1-\tau_k^j) - \tau_k \frac{A}{k} + \frac{B}{k} + \mathbb{E}\left[\frac{\theta_{ij}}{\ell} \epsilon_i \tau_{ij} \ell_{ij} \right] \right) - \frac{a_{iJ}(1-\tau_k)}{k(1+\bar{r})} \right)$$

Table 8: Calibrated Parameters: Baseline Model

Parameter	Value	Description
Firm		
α	0.276	Capital share = 0.255
δ (annual)	0.086	Investment/output = 0.18
Z	0.815	$Y = 1$
μ	1.081	Profit share = 7.5%
Government		
$\bar{\tau}$	0.80	HSV tax function level
γ_ℓ	0.181	HSV progressivity (Heathcote et al., 2017)
τ_K	0.15	Capital income tax
τ_b	0.10	Bequest tax (De Nardi, 2004)
B/Y	1.05	Federal debt/GDP (2019)
T	0.00	No lump-sum transfers at baseline
Preferences (common across types)		
σ_1	2.770	Age profile of savings rates (PSID 1999)
σ_2	2.459	
σ_3	2.266	
σ_4	2.000	
η	1.44	Bequest curvature
$\bar{a}$	2.00	Bequest shifter
ν	2.00	Frisch elasticity = 1
$\underline{a}$	0.00	Zero borrowing constraint
Preferences (type-specific)		
$\beta_1, \ldots, \beta_5$	0.847, 0.621, 0.778, 0.747, 0.679	MPS by quintile (PSID 1999)
$\psi_{a,4}$	0.009	Bequests/GDP $\approx$ 1%
$\psi_{a,5}$	0.017	
$\psi^j_{\ell i}$	(15 values, see Table 10)	$E[\ell^j_i] = 1$ for all (i, j)
Income Process		
θ_{ij}	Table 9	PSID income by quintile $\times$ age
σ^π	(0.00, 0.00, 0.04, 0.10, 0.86)	Profit shares (SCF 2019)
Shock process	Type-specific 3-state Markov	Income growth moments (Guvenen et al., 2015)
Equilibrium Aggregates		
K/Y	2.10	Target
r (annual)	3.6%	Implied by K/Y
w (annual)	0.67	Implied by labor share
NFA/K	−0.43	Residual
G/Y	0.22	Residual

Table 9: Productivity Parameters θ_{ij}

	Ages 20–35	Ages 35–50	Ages 50–65
Quintile 1	0.373	0.489	0.494
Quintile 2	0.639	0.848	0.830
Quintile 3	0.927	1.250	1.245
Quintile 4	1.302	1.773	1.806
Quintile 5	2.034	2.891	3.100

Note. θ_{ij} is the deterministic productivity component for type i at age j, normalized so that $\sum_i \pi \sum_j \frac{1}{4}\theta_{ij} = 1$. Source: PSID median labor income by permanent income quintile and age group.

Table 10: Labor Disutility Parameters $\psi^j_{\ell i}$

	Ages 20–35	Ages 35–50	Ages 50–65
Quintile 1	0.223	0.180	0.131
Quintile 2	0.096	0.110	0.107
Quintile 3	0.077	0.079	0.064
Quintile 4	0.049	0.053	0.045
Quintile 5	0.017	0.020	0.020

Note. $\psi^j_{\ell i}$ is calibrated so that $E[\theta_{ij} \cdot e \cdot \ell^j_i] = \theta_{ij}$ for each type-age cell, ensuring that each household supplies one unit of labor in expectation.