Redistribution and Investment

Wendy Morrison[*][†]

November 27th, 2023
Click here for current version

Abstract

This paper studies the trade-offs associated with income redistribution in an overlapping generations model in which savings rates increase with permanent income. By transferring resources from high savers to low savers, redistribution lowers aggregate savings, and depresses investment. I derive sufficient conditions under which this savings behavior generates a welfare trade-off between permanent income redistribution and capital accumulation in the short and long run. I quantify the size of this trade-off in two ways. First, I derive a sufficient statistic formula for the impact of this channel on welfare, and estimate the formula using U.S. household panel data. When redistribution is done with a labor income tax, the welfare costs associated with my channel are around 1/3 the size of those associated with labor supply distortions. Second, I solve a quantitative overlapping generations model with un-insurable idiosyncratic earnings risk in which savings rates increase with permanent income calibrated to the U.S. in 2019. In this setting, I find that around 17 percent of the trade-off between permanent income redistribution and average consumption can be attributed to my channel.

Keywords: Redistribution, Non-Homothetic Preferences, Optimal Capital Accumulation, Sufficient Statistics
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, and seminar participants at Columbia University for many helpful conversations.

[†]Columbia University, New York, NY 10027. Email: wam2133@columbia.edu, Phone: 610-283-6532

1 Introduction

What is the optimal amount of income redistribution? The existing literature has answered this question primarily by focusing on trade-offs between greater equity and inefficiencies introduced by distortionary taxation (Mirrlees 1971, Piketty and Saez 2013a, Werning, 2007). At the same time, empirical evidence suggests that marginal savings rates are increasing in permanent income (Dynan et al., 2004, De Nardi and Fella, 2017, Straub, 2019), which implies that redistribution may have additional effects on welfare by changing the permanent income distribution and lowering aggregate savings. In this paper, I explore the consequences of this savings behavior for the trade-offs associated with income redistribution in overlapping generations (OLG) models. In particular, I show that when marginal propensities to save out of permanent income increase over the income distribution, this implies an additional trade-off between redistribution and investment.

Intuitively, if lifetime savings increases with permanent income, all permanent redistributive policies — including non-distortionary lump-sum redistribution — will result in a transfer from households with a high marginal propensity to save (MPS) to households with a lower marginal propensity, lowering aggregate savings and putting upward pressure on interest rates in a closed or large open economy (Straub, 2019, Mian et al., 2020).[1] This increase in borrowing costs will curb firms' capital investment, reducing the long-run productive capacity of the economy. It is this potential trade-off between permanent income redistribution and optimal capital accumulation that will be the focus of this paper.[2]

I make several contributions towards better understanding this trade-off. The first is theoretical. In a simple OLG model, I present sufficient conditions for a welfare trade-off between non-distortionary (lump-sum) permanent income redistribution and capital accumulation in both the short and long run. I show that whether such a trade-off exists depends both on whether high-income households have higher MPS, and on the *desirability* of additional investment.[3] Intuitively, for there to be a trade-off, it must both be the case that redistribution dampens investment *and that boosting investment is welfare improving.* I explore how these conditions change when the government is given the ability to tax/subsidize capital, transfer from the young to the old, and issue debt, policies which are well known to alter the savings level.

[1] As long as the economy is not a small open economy and the domestic savings supply has some impact on interest rates. In Straub (2019) and Mian et al. (2021), an increase in inequality lowers interest rates when savings behavior is non-homothetic.

[2] As noted in Piketty and Saez (2013b) and Atkinson and Sandmo (1980), this trade-off is conceptually orthogonal to inefficiency concerns.

[3] Importantly, the presence of a trade-off does not depend on *why* high income households have greater marginal propensities to save.

I then show that the size of the welfare trade-off between permanent income redistribution and investment is large relative to other channels. Using a redistributive labor income tax as an illustrative case, I decompose the steady state (long run) welfare impact of a small increase in the tax into the benefits of greater equality, the standard efficiency costs of distorted labor supply, and the costs associated with my new channel. This decomposition facilitates a back-of-the-envelope comparison between the welfare impact of labor supply distortions, which depend on the long run aggregate labor supply elasticity, and my channel, which depends on the elasticity of capital to the permanent income distribution. I present a sufficient statistic formula for this elasticity, and estimate its terms using U.S. household panel data. I show that the size of my channel is between 1/5 and 1/2 that of labor supply distortions, depending on the estimated distribution of MPS, as well as the relative elasticity of firm investment and household savings to increased borrowing costs. These results suggest that this channel may be large enough to matter when determining optimal policy.

While illustrative of the relative importance of my savings channel for small policy changes in the long run, the sufficient statistic exercise cannot speak to the channel's importance for large policy changes in the short run. Higher MPS may result from *type specific* characteristics or *scale specific* characteristics that grow with income.[4] If MPS are scale-dependent, then more redistribution will compress the distribution of MPS, making my channel smaller.[5] Similarly, the importance of my channel depends on the interest rate elasticity of capital, which is likely much smaller in the short run than in the long run.

To answer these questions, I solve a quantitative OLG model with un-insurable idiosyncratic labor income risk, heterogeneous time preferences, and non-homothetic preferences over savings and bequests, similar in form to the model in De Nardi and Yang (2014). I calibrate the model to the United States in 2019 and consider the effect of the same redistributive labor income tax as in the previous section. In particular, I solve for the trade-off between greater redistribution and average consumption. I isolate the effects of my channel, and find that it can account for around 17 percent of the overall trade-off in the long run for large policy changes. I am currently in the process of quantifying the role my channel plays in the short run by examining the trade-off between redistribution and consumption along the transition path.

Framework and Methodology. I begin by studying a simple closed-economy OLG model with a possible motive for bequests and two labor productivity types. The model nests several

[4]For example, savings preferences or rates of return may be type specific, or may be scale dependent if preferences are non-homothetic or rates of return increase with income.

[5]Gaillard et al. (2023) make a similar point. They show that the optimal capital tax depends on the underlying mechanisms driving multiple dimensions of inequality.

prominent micro-foundations for marginal propensities to save that increase with permanent income. In particular both bequests and consumption later in life can be considered luxury goods (De Nardi, 2004, Mian et al., 2021, Straub, 2019), and high-productivity households may discount the future less than low-productivity households (De Nardi and Fella, 2017).[6] I first consider the impact of lump-sum permanent income redistribution from the high-productivity households to the low-productivity households on steady state welfare, defined as the Pareto weighted sum of each type's lifetime utility. An *unconstrained* planner who could choose any feasible allocation would redistribute resources until the Pareto-weighted marginal utility of consumption was equal across households — the *first best* level of inequality. A welfare trade-off exists whenever the optimal redistribution policy for a fiscal policy maker *constrained* to using the lump-sum tax results in more inequality than first best.

I find that a long-run trade-off exists whenever MPS increase with permanent income *and* when the steady state with the first-best level of inequality is dynamically efficient. Intuitively, suppose the fiscal authority sets redistribution policy to implement the optimal level of inequality. If MPS increase with permanent income and this steady state is dynamically efficient, reducing the amount of redistribution slightly will improve welfare by boosting savings, investment, and ultimately aggregate consumption.

I then consider the entire transition path, such that welfare is defined as the Pareto weighted *discounted* sum of the lifetime utility of all present and future generations. Now, the planner cares about both the short and long-run. In this case, the presence of a trade-off depends not only on the sufficient conditions for a long-run (steady state) trade-off, but also on the rate at which the planner discounts future generations. For a welfare trade-off to exist, the planner must put sufficiently high weight on future generations for the benefit of greater future capital to outweigh the costs of more inequality *and* less consumption today. Finally, I show that these sufficient conditions extend to a setting in which the the government has access to a broader set of fiscal policy tools, including debt, inter-generational transfers, and capital subsidies.[7]

Assuming the sufficient conditions derived above are satisfied, a natural question is whether the redistribution-investment trade-off is large relative to the standard equity-efficiency trade-off. To answer this question, I consider a simple redistribution scheme in which a lump-sum transfer is funded through a distortionary linear tax on labor income. I decompose the effect of a small increase in redistribution on steady state welfare into the effect of greater equality, the costs associated with distorted labor supply, and the costs

[6]The simple model does not take into account earnings risk or heterogeneous rates of return. I consider earnings risk in the quantitative model.

[7]These tools can be used to alter the savings supply in life cycle models, and could possibly be employed to offset the effect of changing the permanent income distribution (Diamond, 1965).

associated with my channel. The costs of distortions depend on the long run (steady state) elasticity of labor supply with respect to the tax rate, and I rely on a meta-analysis in Chetty et al. (2011) for an estimate of this term.[8] The redistribution-investment trade-off meanwhile, depends on the elasticity of the capital stock to changes in the permanent income distribution. I derive and estimate a sufficient statistic formula for this elasticity.

My formula shows that the size of this elasticity depends first on how marginal propensities to save (MPS) out of permanent income change over the income distribution. Intuitively, the greater the difference between the MPS of high and low-income households, the larger the impact of redistribution on aggregate savings. I estimate the MPS over the permanent income distribution using longitudinal data from the Panel Study of Income Dynamics (PSID). The sufficient statistic formula also depends on the interest rate elasticity of firm investment relative to the interest rate elasticity of household savings. Intuitively, whether a decline in savings increases savings (reduces household debt) or reduces investment more depends on whether households or firms are more sensitive to increased borrowing costs. I draw on a range of estimates from the literature of these elasticities. I then consider a set of extensions to the baseline formula. I find that the costs associated with my channel are between 1/5 and 1/2 of the size of those associated with distorted labor supply.

Finally, I solve a richer version of the simple life-cycle model with idiosyncratic income risk calibrated to United States economy in 2019. To calibrate the parameters governing savings behavior, I target my estimates of savings rates by permanent income quintile and age in the PSID. I calculate the impact on average steady state consumption of a lump-sum transfer funded by an increase in average labor income taxes. To isolate the effect of my channel, I solve for the direct effect of the redistribution on the permanent income distribution, and calculate the impact of just these changes on average consumption holding household labor supply constant at the original steady state level. In this case, the tax acts like a lump-sum permanent income transfer, and any change in capital – and consumption – can be attributed entirely to the direct effect of my channel.[9] I find that my channel can account for around 17 percent of the overall trade-off in the steady state. My next step will be to consider the relative importance of my channel over the transition path.

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)

[8] Chetty et al. (2011) show that while considerable disagreement exists regarding the Frisch (intertemporal) labor supply elasticity, more consensus exists between micro and macro economists regarding the steady state elasticity.

[9] In particular, I isolate the effects of permanent income changes from the effects of insurance, intergenerational transfers, and labor supply distortions.

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.[10] 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 of 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 of 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 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 impacts welfare in life-cycle models.

This paper also contributes the the empirical literature studying the relationship between permanent household income and savings. In order to estimate my sufficient statistic formula, I use the Panel Study of Income Dynamics (PSID) to estimate marginal propensities to save (MPS) by permanent income type, and find significantly higher MPS for high income households. These findings echo those of Dynan et al. (2004), who use the PSID in combination with several other data sets to estimate both average and marginal savings rates by permanent income group. Relative to this study, I take advantage of the fact that the

[10]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.

PSID added consumption data in 1999 in order to generate a more straightforward measure of 'active' savings. Straub (2019) uses the same data set to estimate a related statistic: the *elasticity of consumption* with respect to permanent income. Using the elasticity of savings implied by his findings in conjunction with estimates of savings *rates* by income group, I can generate additional estimates of the MPS and find that they are very similar to my direct estimates.

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, which increases relative borrowing costs for bank-dependent firms, lowering their employment share.

Layout. The rest of the paper proceeds as follows. In Section 2, I lay out the baseline overlapping generations model and establish its key properties. I derive sufficient conditions under which non-homothetic savings behavior generates a welfare trade-off between redistribution and capital accumulation. In Section 3, I present and calibrate the sufficient statistics formula to estimate the size of my channel relative to the size of labor supply distortions. In Section 4, I present the quantitative model and results. Section 5 concludes.

2 The Redistribution-Investment Trade-off

In this section I derive sufficient conditions for the existence of a welfare trade-off between non-distortionary redistribution and capital accumulation in a simple overlapping generations model. The model nests several leading sources of non-homothetic savings behavior. To derive the conditions, I consider the problem of a planner who aims to maximize social welfare, defined as the Pareto weighted sum of household utility. I begin by considering welfare in the long-run steady state. A planner free to choose any feasible steady state allocation would allocate resources between households to generate an *ideal* (first best) level of equality.

On the other hand, a constrained fiscal planner faces a *trade-off* between lump-sum redistribution and capital accumulation whenever it is optimal for fiscal policy to tolerate more inequality than the ideal level. I show how the conditions for such a trade-off change when social welfare is defined as the discounted sum of household utility along the transition path. In this case, the planner must weigh the benefits of greater equality and greater consumption now against the cost of less capital for future generations, making the rate at which the planner discounts future generations a key determinant of ideal policy. Finally, I consider how the sufficient conditions change when the government can alter the savings supply using a wider set of policy tools.

2.1 Environment

I begin with a variant of the canonical 2-generation overlapping generations closed-economy model with fixed exogenous labor supply. Time is discrete. Agents have perfect foresight over future variables and there is no uncertainty.

Households. There is a unit mass of households who each live for 2 periods, $h \in \{y, o\}$ and have heterogeneous labor productivity types, θ_i for $i \in \{L, H\}$ where $\theta_L < \theta_H$. There is a constant fraction, π_i of each productivity type with an equal share of each generation (no population growth). 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 gross rate of return R_t and may leave bequests $a^o_{i,t}$ to households with the same productivity type in the next period. Households born at time t receive $R_t a^o_{i,t-1}$ in inheritance when they are young. Capital depreciates at rate δ. Households also receive a type-specific lump-sum tax(transfer) T_{it}. A type-i household born in year t has lifetime utility given by equation (1).

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

Note that the discount factor, β_i may be type-specific and that the parameters σ_y, σ_o, and η, which govern the elasticity of inter-temporal substitution and bequests, may differ from one another. Households choose consumption when young and old and bequests to maximize (1) subject to their lifetime budget constraint:

$$c^y_{it} + \frac{c^o_{i,t+1} + a^o_{it+1}}{R_{t+1}} = R_t a^o_{it-1} + w_t\theta_i + T_i = PI_{it} \tag{2}$$

I define the right hand side of equation (2) as the household's permanent income, PI_i. Let a household's change in assets, $a^h_i - a^{h-1}_{it-1} = s^h_t$, their savings at age h. Note that when $\psi_a = 0$ households do not leave bequests and $s^y_t = a^y_t$.

Firms. There is a continuum of perfectly competitive firms who rent capital and labor from households and produce output subject to a Cobb-Douglas production function, $F(K_t, L_t) = K_t^\alpha L_t^{1-\alpha}$. The firm's first order conditions are standard and are given by

$$R_t = F_K(K_t, L_t) + 1 - \delta \quad \text{and} \quad w_t = F_L(K_t, L_t) \tag{3}$$

Government. The government runs a balanced budget each period. The transfer T_{it} is defined in terms of each generation's lifetime income. The government cannot make net transfers between living generations, and can only transfer resources between household types in the same generation. I consider the case of a government with access to inter-generational transfers and debt policy in the next section. The government budget constraint is given by the following expression.

$$\sum_{i \in I} \pi_i T_{it} = 0$$

Equilibrium. I define an allocation $x \equiv \{\{c_{it}^y, c_{it}^o\}_{i\in I}, K_t, L_t\}_{t\geq 0}$. An equilibrium is an allocation, a sequences of financial positions, $\{a_{it}^o, a_{it}^y\}_{i\in I, t\geq 0}$ a sequence of prices, $\{R_t, w_t\}_{t\geq 0}$, and policies $T \equiv \{T_{Lt}, T_{Ht}\}_{t\geq 0}$ 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 ($L_t = 1$), and the resource constraint (4) and asset market clearing condition (5) are satisfied.

$$\frac{1}{2}\sum_I \pi_i(c_{it}^y + c_{it}^o)K_{t+1} = F(K_t, 1) + (1-\delta)K_t \tag{4}$$

$$K_{t+1} = \frac{1}{2}\sum_I \pi_i(a_{it}^y + a_{it}^o) \tag{5}$$

I define the set of all *feasible* allocations, $\mathcal{X}$ as the set of allocations that satisfy the resource constraint. I define the set of all *implementable* allocations, $\mathcal{X}^I$ as the set of allocations for which prices and policies exist that implement all $x \in \mathcal{X}^I$ as an equilibrium. Note that when policy is held constant, the economy converges monotonically to the unique steady state (see Appendix 1 for a proof). Let χ_s denote the set of all feasible steady state allocations and χ_s^I be the set of all implementable steady state allocations.

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

Assumption. I assume that $\sigma_y \geq \sigma_o \geq \eta$ and that $\beta_H \geq \beta_L$.

The above assumption allows for the possibility that households with higher incomes have a higher propensity to save out of their lifetime income.[11] When any of the above inequalities are strict, the marginal propensity to save out of permanent income when young for high-productivity type households, $\frac{\partial s^y_{Ht}}{\partial PI_H}$ is greater than for low-productivity households. The same is true for the derivative of bequests with respect to permanent income when $\sigma_o > \eta$. 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 has no affect on aggregate savings or the interest rate. Therefore, the steady state capital stock is unaffected by fiscal policy. When the marginal propensity to save is higher for high-permanent-income households, a greater lump-sum transfer to the low types, T_L reduces aggregate savings. This puts upward pressure on the steady state interest rate R and reduces steady state capital, K. These results are summarized in the following Lemma.

Lemma 1 *(The Distribution of MPS and Steady State Capital) Let K be the steady state level of capital.*

Case 1: ($\psi_a = 0$). When either $\sigma_y > \sigma_o$ or $\beta_H > \beta_L$, the marginal propensity to save out of permanent income is higher for high-productivity types, $\frac{\partial s^y_{Ht}}{\partial PI_{Ht}} > \frac{\partial s^y_{Lt}}{\partial PI_{Lt}}$ and $\frac{\partial K}{\partial T_L} < 0$.

Case 2: ($\psi_a > 0$). When $\sigma_y > \sigma_o$ or $\beta_H > \beta_L$ or $\sigma_o > \eta$, the marginal propensity to save out of permanent income when young is higher for high-productivity types $\frac{\partial s^y_{Ht}}{\partial PI_{Ht}} > \frac{\partial s^y_{Lt}}{\partial PI_{Lt}}$, the sensitivity of bequests to permanent income is higher for high income types, $\frac{\partial a^o_{Ht+1}}{\partial PI_{Ht}} > \frac{\partial a^o_{Lt+1}}{\partial PI_{Lt}}$, and $\frac{\partial K}{\partial T_L} < 0$.

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).[12] Whenever $\sigma_o > \eta$, leaving bequests is a luxury good, and

[11] In the Appendix, I consider an alternative model in which the high-productivity types have higher rates of return than the lower productivity types.

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

households leave larger bequests as their lifetime income increases (De Nardi, 2004, Straub, 2019, Mian et al., 2021). Finally, I 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).

Lemma 1 states that steady state capital, K is unaffected by lump-sum permanent income redistribution when MPS out of permanenent income are uniform, but is decreasing in the degree of redistribution when the high-productivity types have higher MPS. Intuitively, in the latter case, this implies that redistribution takes permanent income from households with high MPS and gives it to households with lower MPS. This reduces aggregate savings, lowering the supply of loanable funds, pushing up borrowing costs, and ultimately decreasing the capital stock.

2.2 Redistribution and Steady State Welfare

I begin by considering the effect of an incremental change in steady state transfers to the low-productivity households, T_L on steady state social welfare. Consider a social planner with Pareto weights λ_i for each household type. Define steady state social welfare as in equation (6).

$$SW_s = \sum_I \lambda_i \pi_i \left(\frac{(c_i^y)^{1-\sigma_y}}{1-\sigma_y} + \beta_i \frac{(c_i^o)^{1-\sigma_o}}{1-\sigma_o} + \beta_i \psi_a \frac{(a_i^o)^{1-\eta}}{1-\eta} \right) \tag{6}$$

I define the steady state welfare weight of type-i households, $\omega_i \equiv \lambda_i (c_i^y)^{-\sigma_y}$. These weights are the product of the value of type-i utility to the planner and type-i households' marginal utility of consumption. Therefore, they reflect the marginal value from the planner's perspective of giving additional resources to a type-i household.[13] For a given set of Pareto weights, as the consumption of type-i households falls, their marginal utility of consumption, and therefore their welfare weight increases. For a given allocation therefore, the ratio between the welfare weights of the two household types characterizes the degree of inequality. I make the following assumption about the Pareto weights.

Assumption. I assume that $\lambda_H \geq \lambda_L$.

By assuming that the Pareto weights assigned to the high-productivity households are always higher, I ensure that the planner will never prefer allocations in which the low-

[13]Note that in equilibrium, the households' inter-temporal conditions ensure that these social welfare weights are proportional to the change in welfare of type-i households receiving additional consumption when old or being able to leave greater bequests.

productivity types consume *more* than the high-productivity types. I define the first-best steady state allocation, x_s^* and the optimal constrained steady state allocation, x_s^c, as well as their corresponding welfare weights, in the following way.

Definition. Define the optimal unconstrained steady state allocation, $x_s^* \equiv \underset{x \in \chi_s}{\text{argmax}}\ SW_s$. Let $\omega_i^* \equiv \lambda_i (c_i^{y*})^{-\sigma_y}$ for $i \in I$ be the welfare weights corresponding to this allocation. Define the optimal constrained allocation, $x_s^c \equiv \underset{x \in \chi_s^I}{\text{argmax}}\ SW_s$. Let $\omega_i^c \equiv \lambda_i (c_i^{yc})^{-\sigma_y}$ for $i \in I$ be the corresponding welfare weights.

The optimal unconstrained allocation is the allocation that maximizes steady state social welfare subject only to feasibility. The ratio of the welfare weights characterizes the first-best *ideal* level of inequality. If the ratio of welfare weights for a given allocation is *higher* than this ratio, the marginal utility of the low types is higher, implying a greater level of inequality. At the unconstrained optimum allocation, the welfare weights are equal (see Appendix A.2 for a proof). Intuitively, suppose $\omega_i > \omega_j$. Then the cost to the planner of redistributing resources away from type-j households would be outweighed by the benefit of redistributing towards type-i households, implying this allocation is not optimal.

How does the constrained optimum differ from this allocation? To build intuition, it is helpful to first examine how a small increase in the lump-sum transfer from high-types to low types affect steady state welfare. First, the transfer affects welfare directly by shifting resources between households with different welfare weights. If $\omega_L > \omega_H$, this direct effect of the policy will be positive. Second, if $\psi_a > 0$ and households have a bequest motive, the redistribution will equalize the bequest distribution as well, further increasing the lifetime resources of the low-productivity households. Finally, the policy may change the steady state level of aggregate capital, which in turn would affect welfare by increasing the total level of bequests, and through general equilibrium effects on household income. These results are summarized in Lemma 2.

Lemma 2 *(Welfare impact of redistribution) Denote K_{PI} as the semi-elasticity of the steady state capital stock to the amount of lump-sum redistribution, $\frac{dK}{dT_L}\frac{1}{K}$.*

The steady state change in social welfare, dSW_s from a small increase in T_L is:

$$dSW_s = \underbrace{\sum_I \pi_i \omega_i dT_i}_{\textit{Direct Effect}} + \underbrace{RK\frac{1}{2}\sum_I \pi_i \omega_i d\Gamma_i^b}_{\textit{Bequest Distribution}} + \underbrace{\left(\frac{1}{2}\sum_I \pi_i \omega_i a_i^0 R + wL\Theta w_K\right) K_{PI}}_{\textit{Change in Capital}}$$

Here, $\Theta \equiv \sum_I \omega_i \pi_i \left(\frac{\pi_y \theta_i}{L} - \left(\frac{a_i^y}{KR} + \frac{a_i^o}{K} \right) \right)$, *and* Γ_i^b *denotes type-i household bequest's share of total capital.*

For a proof, see Appendix A.3.

Lemma 2 says that the total effect of the transfer can be decomposed into the direct effect, the effect on the distribution of bequests, and the effect of changing the aggregate level of capital.[14] The change in steady state capital affects welfare in two ways. First, through the effect of a change in capital on factor prices, and second, whenever $\psi_a > 0$, through changes in aggregate bequests left. The welfare impact of the change in factor prices is summarized by the term Θ. What the total effect of these changes are on aggregate welfare depends on whether Θ is positive, which in turn depends both on the the rate of return, R at the current steady state and on the current steady state distribution of capital and labor income.

When the steady state gross rate of return, $R > 1$, the steady state is dynamically efficient, and more capital increases average consumption. Furthermore, when savings rates are increasing in permanent income, high-productivity households have a greater share of aggregate capital income than aggregate labor income. Therefore, if the welfare weight of the low-productivity households is higher than that of the high productivity households, when $R > 1$ and savings rates increase with permanent income, an increase in capital improves welfare ($\Theta > 0$) by both increasing average consumption and by increasing wage income relative to capital income, disproportionately benefiting low-income households.

Whether the planner faces a redistribution-investment trade-off depends on the welfare impact of additional capital at the steady state associated with the first best level of inequality. At this steady state, the direct benefit of redistributing resources from the high to the low types has been exhausted. That is, the economy is at the ideal level of equality. A small reduction in the amount of redistribution would therefore have no *direct* effect on steady state social welfare. If MPS out of permanent income are higher for high-types *and* additional capital at this steady state would increase welfare ($\Theta > 0$), the planner could improve welfare by reducing the degree of redistribution and tolerating a slightly higher level of inequality. Proposition 1 summarizes this result.

Proposition 1 *(Redistribution-Investment Trade-off) Let* $\bar{K}$ *be the steady state level of capital associated with the first-best level of inequality such that* $\omega_L = \omega_H$

[14] Note that this result relies on a standard application of the envelope theorem. Households are already optimizing with respect to bequests and therefore the policy has no first order effect on utility associated with bequests.

(1) If $\frac{\partial a_H^h}{\partial PI_H} = \frac{\partial a_L^h}{\partial PI_L}$ *for* $h \in \{y, o\}$*, then at the constrained optimal steady state,* $\omega_L^c = \omega_H^c$*.*

(2) If $\frac{\partial a_H^h}{\partial PI_H} > \frac{\partial a_L^h}{\partial PI_L}$ *for* $h \in \{y, o\}$ *and* $F_K(\bar{K}) > \delta$*, then at the constrained optimal steady state* $\omega_L^c > \omega_H^c$*.*

For a proof, see Appendix A.4.

Proposition 1 states that when MPS are uniform, the degree of inequality in the constrained optimal allocation is identical to first-best. Intuitively, redistribution has no effect on aggregate capital in this case, and the planner faces no trade-off between additional redistribution and capital accumulation. However, when the high types have higher MPS and the steady state corresponding to the first best level of inequality is dynamically efficient, the constrained optimal level of inequality is greater than first best.

To see why, suppose that the government were to use T_L to implement the first-best level of inequality. Because the choice of redistribution policy pins down the level of steady state capital, it may be the case that the marginal product of this level of capital, $F_K(\bar{K})$ is higher than the depreciation rate.[15] In this case, reducing T_L would boost the capital stock and increase average consumption. At the same time, the resulting small increase in inequality would have no *direct* impact on welfare, as the economy is currently optimizing with respect to the inequality level. Therefore, implementing the first-best level of inequality is not optimal, and the planner should reduce T_L until the costs of greater inequality equal the benefit of additional capital.

2.3 Redistribution and Welfare Along the Transition

To see how taking into account short run welfare affects the trade-off, I consider the problem of a social planner who weights each household type with Pareto weights, λ_i and discounts generations at constant rate, γ. Here, social welfare is defined as the infinite Pareto-weighted discounted sum of the lifetime utility of all households as in equation (7).

$$SW(x) = \sum_I \pi_i \lambda_i \sum_{t=0}^{\infty} \gamma^t \left(\frac{(c_{it}^y)^{1-\sigma_y}}{1-\sigma_y} + \gamma^{-1} \beta_i \frac{(c_{it}^o)^{1-\sigma_o}}{1-\sigma_o} + \beta_i \gamma^{-1} \psi_a \frac{(a_{it}^o)^{1-\eta}}{1-\eta} \right) \tag{7}$$

I characterize a second set of sufficient conditions for the existence of a redistribution-investment trade-off analogous to the one presented the previous section. That is, I outline conditions under which the optimal redistribution policy results in a level of intra-generational inequality that is greater than first best. As in the previous section, whether it

[15]Note that unconstrained the first-best level is the Golden Rule capital stock, in which $F_K(K) = \delta$.

is optimal for fiscal policy to implement the first best level of intra-generational inequality depends on the distribution of MPS out of permanent income, and on whether the steady state associated with the first-best level of inequality is dynamically efficient. However, now the existence of a trade-off also depends on the rate at which the planner discounts the future. These results are summarized in Proposition 2.

Proposition 2 *Again, let $\bar{K}$ be the steady state level of capital associated with the first-best level of inequality such that $\omega_L = \omega_H$.*

(1) When $\frac{\partial a^h_{Lt}}{\partial PI_{Lt}} = \frac{\partial a^h_{Ht}}{\partial PI_{Ht}}$ for $h \in \{y, o\}$, then in the constrained optimal allocation, $\omega^c_{Lt} = \omega^c_{Ht}$= the first-best level of inequality for all $t \geq 0$.

(2) When $\frac{\partial a^h_{Ht}}{\partial PI_{Ht}} > \frac{\partial a^h_{Lt}}{\partial PI_{Lt}}$ for $h \in \{y, o\}$ and $F_K(\bar{K}) > \delta$, then there exists a $\hat{\gamma} \in (0, 1)$ and $\tau > 0$ such that if $\gamma > \hat{\gamma}$, $\omega^c_{Lt} > \omega^c_{Ht}$ for all $t \geq \tau$.

For a proof, see Appendix A.5

Proposition 2 states that when MPS are uniform, the optimal allocation features the first-best level of inequality at every time horizon. Intuitively, because redistribution has no effect on capital accumulation, the planner faces no trade-off between redistribution and investment, and therefore it is optimal to simply use redistribution to achieve the first-best level of equality. However, when the high types have higher MPS and the steady state corresponding to the first best level of inequality is dynamically efficient, then as long as the planner puts *sufficiently high weight on future generations*, a trade-off exists between permanent income redistribution and investment, and the first best level of inequality is not optimal.

To see why, again consider a government who sets T_{Lt} every period in order to achieve the first-best path of inequality such that $\omega_{Lt} = \omega_{Ht}$ for all $t \geq 0$. Such a policy pins down a particular path for capital, $\{K_t\}_{t\geq 1}$. No matter the level of initial capital, K_0, by setting such a policy, the planner ensures that K_t will eventually converge to $\bar{K}$, the steady state level of capital associated with the first best level of equality. If this level of capital is dynamically efficient, then as K_t approaches $\bar{K}$, K_t will eventually become dynamically efficient for all $t \geq \tau$. T_{Lt} has been set to achieve optimal equality at time t, but because the high types have higher MPS, reducing T_{Lt} today will increase the entire future path of capital, and increase average consumption, C_t for all $t \geq \tau$. So long as the planner *does not discount the future too heavily*, the marginal benefit of additional future capital will outweigh the costs of higher inequality and lower aggregate consumption today, and the first-best level of inequality is not optimal at any horizon.

Intuitively, the planner can use redistribution policy in order to target a future path for capital. When deciding whether to increase savings to boost the future capital stock, they must weigh the benefits against the costs *both* of less average consumption today and greater inequality today. As long as capital is guaranteed to produce greater aggregate consumption in the future and the weight put on future generations is sufficiently large, a trade-off emerges and optimal policy will implement higher-than-first best levels of inequality in order to boost the future capital stock.

2.4 Redistribution with More Fiscal Policy Tools

In the previous 2 subsections, when high-income households had higher MPS out of permanent income, a trade-off emerged between redistribution and capital accumulation because the income distribution was the single tool available that allowed policy makers to alter the savings level. In reality, governments have many fiscal tools available that allow them to alter the level of savings, including debt management, inter-generational transfers, and capital taxes/subsidies. In this section, I show how the sufficient conditions needed for an equality-investment trade-off along the transition change when the government has access to a larger set of policy tools.

Government. Suppose the government can now issue age-specific lump-sum taxes(transfers), T_{ht} in addition to type-specific lump-sum taxes(transfers), T_{it}. The government can also tax or subsidize capital directly, τ_{Kt} and borrow at the prevailing interest rate. The government's per-period budget constraint is given by Equation (8).

$$\sum_I \pi_i T_{it} + \sum_A \pi_h T_{ht} + \tau_{Kt-1} R_t K_t = R_t B_{t-1} - B_t \tag{8}$$

Crucially, the government also faces a set of *political constraints*. Equation (9) implies that the government cannot redistribute lump-sum from the current old to the current young. Equation (10) says that the government can issue debt but cannot invest directly.

$$T_{yt} \geq 0 \geq T_{ot} \tag{9}$$

$$0 \geq B_t \tag{10}$$

I adopt these assumptions for their realism. To my knowledge no scheme to redistribute from the current old to the current young exists. The few permanent government surpluses we observe in the data tend to be the result of state-owned natural resources rather than fiscal policy. If these restrictions were relaxed and the government had a complete set of policy tools, they could implement the first best allocation and there would be no trade-off.

Households. Households are identical to those in the previous section. For simplicity, we begin by considering the case without bequests in which $\psi_a = 0$.[16] Household utility is still given by equation (1), however given the new fiscal policy, the type-i households' lifetime budget constraint is now given by the following.

$$c_{it}^y + \frac{c_{i,t+1}^o}{R_{t+1}(1-\tau_{Kt})} = w_t\theta_i + \frac{T_{ot+1}}{R_{t+1}(1-\tau_{Kt})} + T_{it} + T_{yt}$$

Equilibrium. An equilibrium is again defined as an allocation, a sequence of prices, $\{R_t, w_t\}_{t\geq 0}$, and policies $\kappa \equiv \{T_{Lt}, T_{Ht}, T_{yt}, T_{ot}, B_t\}_{t\geq 0}$ such that the household first order conditions, the firms' first order conditions, and the government's budget constraint (8) are satisfied, the labor market clears ($L_t = 1$), and the resource constraint (4) and asset market clearing condition (5) are satisfied. I again define the set of all *feasible* allocations, $\mathcal{X}$ and the set of all *implementable* allocations, $\mathcal{X}^I$ as before.

As in the previous section, when MPS grow over the income distribution, redistribution decreases steady state capital, as resources are transferred from those with a high propensity to save to those with a lower propensity. However, now the fiscal planner has additional tools to influence the savings rate. A well known feature of over-lapping generations models is the ability of fiscal policy that transfers resources from the current young to the current old to change the savings supply (Diamond 1965; Samuelson 1975). When $\psi_a = 0$ and households do not leave bequests, both social security schemes and debt transfer resources from the saving young to the non-saving old, resulting in a lower capital stock. These results are presented in Lemma 3.

Lemma 3 *(Debt and Social Security Lower Savings)*
When $\psi_a = 0$, for a given set of policies, $T_{Ht}, T_{Lt}, \tau_{Kt}$, steady state capital, K is decreasing in steady state inter-generational transfers, $T_{yt} - T_{ot}$ and steady state debt, B_t.

A proof of 3 can be found in Appendix A.6.1.

Lemma 3 implies that once the fiscal planner's political constraints bind, they can no longer rely on debt management or inter-generational transfers to increase the capital stock, and must trade-off the benefits of equality against the cost of lower future capital accumulation

[16]The case with bequests will be considered in an extension in the Appendix.

and distortions associated with a capital tax/subsidy. However, when the political constraints do not bind the planner has all the tools needed to achieve the optimal amount of capital accumulation while also achieving the first best level of inequality. Therefore, the presence of a redistribution-capital accumulation trade-off now depends on whether savings behavior is non-homothetic and on whether the steady state with the first best level of inequality, no capital tax, and *binding political constraints* is dynamically efficient. If so, then for sufficiently high γ, the optimal redistribution policy will result in a higher than first-best level of inequality. These results are summarized in Proposition 3.

Proposition 3 *Assume $\psi_a = 0$ and let $\bar{K}$ be the level of capital associated with the first-best level of inequality, $\tau_{Kt} = 0$, and binding political constraints ($T_{ot} = T_{yt} = 0$, and $B_t = \bar{B}$ for all t).*

(1) When $\frac{\partial a^h_{Lt}}{\partial PI_{Lt}} = \frac{\partial a^h_{Ht}}{\partial PI_{Ht}}$ for $h \in \{y, o\}$, then in the constrained optimal allocation, $\omega^c_{Lt} = \omega^c_{Ht}$= the first-best level of inequality for all $t \geq 0$.

(2) When $\frac{\partial a^h_{Ht}}{\partial PI_{Ht}} > \frac{\partial a^h_{Lt}}{\partial PI_{Lt}}$ for $h \in \{y, o\}$ and $F_K(\bar{K}) > \delta$, then there exists a $\hat{\gamma} \in (0, 1)$ and a $\tau > 0$ such that if $\gamma > \hat{\gamma}$, $\omega^c_{Lt} > \omega^c_{Ht}$ for all $t \geq \tau$.

For a proof, see Appendix A.7.

To see why, consider a hypothetical steady state in which $\tau_{Kt} = 0$, T_{Lt} and T_{Ht} are set to implement the first-best level of intra-generational equality, and both political constraints bind. That is, $T_{yt} = T_{ot} = 0$ and $B_t = \bar{B}$ for all $t \geq 0$. This set of policies is associated with a unique steady state level of capital, $\bar{K}$. If $\bar{K}$ is greater than or equal to the modified golden rule (first-best) level of capital, then the planner can use debt or transfers from the young to the old to lower the capital stock, while using T_{Lt} and T_{Ht} to achieve the first best level of inequality.

If instead $F_K(\bar{K}) > \delta$ and the steady state is dynamically efficient, then if debt, inter-generational transfers, and τ_{Kt} remain unchanged, K_t will eventually converge to $\bar{K}$ and the economy will eventually become dynamically efficient. That is, there exists some future period τ such that for all $t \geq \tau$, increasing capital increases aggregate consumption. If future generations are given sufficient weight – as $\gamma \to 1$, the welfare impact of this additional capital outweighs the costs of reduced consumption and greater inequality today. In this case, the planner will choose to keep inter-generational transfers and debt at their constrained level, so as to not further reduce the investment rate. They will set intra-generational redistribution policy and capital subsidies so that the benefits of capital to future generations equals the cost of deviating from the first best level of equality and the distortions created by the capital tax.

3 Is the Redistribution-Investment Trade-off Large?

The previous section presented sufficient conditions for a welfare trade-off between capital accumulation and the redistribution of permanent income. In this section, I explore whether this trade-off is relevant for real-world policy makers by asking how its size compares to other channels. While the derivations in Section 2 relied on type-specific lump-sum redistribution, I show that the welfare impact of my channel is large relative to other channels in the context of a more realistic redistribution policy. To do this, I present a slightly modified version of the previous section's model and consider the welfare effect of a uniform lump-sum transfer funded by a simple proportional labor income tax as in Werning (2007). In particular, I allow household labor supply to be endogenously determined, allowing for a direct comparison of the size of the welfare impact of my channel relative to the effect of labor supply distortions.

To facilitate this comparison, I derive a formula for the size of my channel in terms of estimable sufficient statistics. My formula shows that the effect of redistribution on capital accumulation depends not only on the degree to which MPS differ over the income distribution, but also on the relative interest rate elasticities of investment and household savings. I use PSID data to estimate households' marginal propensities to save out of permanent income, relying on estimates in the literature for the formula's other terms. I present a range of values for the size of my channel and show that even the lower-end estimates imply that the channel is large relative to labor supply distortions.

Households. Consider a variant of the overlapping generations economy presented in Section 2. Households' labor supply is now endogenous and supplied only in the first period of life. Type-i households choose ℓ_{it} and receive $(1 - \tau_\ell) w_t \ell_{it} \theta_i$ in after-tax labor income while young and are retired when old. As in the previous section, households can borrow or save each period with gross rate of return R_{t+1}. For simplicity, I consider the version of the model without bequests ($\psi_a = 0$), however none of the results presented below depend on this assumption. Household lifetime utility is now given by equation (11).

$$u(c^y_{it}, c^o_{it}, \ell_{it}) = \frac{(c^y_{it})^{1-\sigma_y}}{1-\sigma_y} + \beta_i \frac{(c^o_{it})^{1-\sigma_o}}{1-\sigma_o} - \frac{(\ell^h_{it})^{1+\gamma}}{1+\gamma} \tag{11}$$

In addition to paying labor income taxes, households receive a uniform lump-sum transfer,

T. The type-i households' lifetime budget constraint is given equation (12).

$$c_{it}^{y} + \frac{c_{i,t+1}^{o}}{R_{t+1}} = (1-\tau_{\ell t})w_t\theta_i\ell_{it} + T = PI_{it} \tag{12}$$

Firms. There are a continuum of perfectly competitive firms who produce output according to the constant-returns-to-scale production function (13).

$$F(K_t, L_t) = \left(\alpha K_t^{\frac{\rho-1}{\rho}} + (1-\alpha)L_t^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}} \tag{13}$$

Government. The government runs a balanced budget each period and can fund lump-sum transfers using linear taxes on labor income, $\tau_{\ell i}$.

$$\sum_{i\in I}\pi_i T_i = \sum_{i\in I}\pi_i\Big(w_t\ell_{it}\theta_i\tau_\ell\Big)$$

Equilibrium. An equilibrium is a sequence of quantities, $\{\{c_{it}^h, a_{it}^h\}_{i\in I, h\in H}, K_t, L_t\}_{t\geq 0}$, prices, $\{R_t, w_t\}$, and policies $\{T, \tau_\ell\}$ such that the household first order conditions, the firms' first order conditions and the government's budget constraint are satisfied, the labor market clears, and the goods and asset markets clear.

Again defining steady state social welfare as the Pareto weighted sum of household utility, I consider the incremental impact on steady state welfare of a small budget-balancing increasing in the uniform lump-sum transfer, dT, funded by a small increase in the labor income tax, $d\tau_\ell$. I show that the welfare effect this change in fiscal policy can be decomposed into the direct effects of redistributing labor income from those with higher-than-average labor income to those with lower-than average, the effects of distorted labor supply, and the non-homothetic savings channel. This decomposition is presented in Proposition 4.

Proposition 4 *(Effect of labor income redistribution on steady stare welfare) Define* Θ*,* w_K *as in Lemma 3, and let* w_L *be the labor elasticity of the wage. Define* K_{PI} *as the semi-elasticity of capital to the direct effects of the tax and* L_{τ_ℓ} *is the labor supply semi-elasticity* *with respect to* τ_ℓ. *Then the impact of an incremental increase in* τ_ℓ *on social welfare is:*

$$dSW = \underbrace{\sum_I \omega_i \pi_i (wL - w\theta_i \ell_i) d\tau_\ell}_{\textit{Direct Effect of Redistribution}} + \underbrace{wL(\Theta + \tau_\ell) w_K K_{PI}}_{\textit{Direct Effect of NH Savings}}$$
$$\underbrace{wL(\Theta + \tau_\ell)\left(w_L + \frac{\tau_\ell}{\Theta + \tau_\ell}\right) L_{\tau_\ell}}_{\textit{Direct Effect of Labor Distortion}} + \underbrace{wL(\Theta + \tau_\ell)(\mathcal{L} + \mathcal{K})}_{\textit{Feedback Effects}}$$

Where $\mathcal{L}$ *and* $\mathcal{K}$ *are defined as in equation (A.27) and (A.28).*

For a proof, see Appendix A.8.

Like lump-sum redistribution, the labor income tax affects social welfare directly by transferring lifetime income between households with potentially different social welfare weights, ω_i. Now however, because labor is endogenous, the redistribution policy also distorts the aggregate labor supply, which lowers taxable labor income directly and indirectly through changes in the equilibrium wage. When savings is non-homothetic and $\frac{\partial a^h_{it}}{\partial PI_i}$ co-varies with labor productivity, an additional welfare channel emerges that is proportional to K_{PI}, the semi-elasticity of steady state capital to the direct effect of the redistribution on the permanent income distribution. Intuitively, when savings behavior is non-homothetic, the policy lowers aggregate savings and capital by redistributing from high labor-income households with a high propensity to save to low labor income households.

Finally, the policy impacts welfare through interaction between the latter two channels. The distortion of labor supply caused by τ_ℓ impacts firms' incentives to invest in capital and households' incentive and ability to save. These effects are captured in the term $\mathcal{K}$. At the same time, the decline in capital affects firms' labor demand and households' incentives to work. These effects are summarized in the term $\mathcal{L}$. In this section, I focus on comparing the *direct* effects of labor distortions and non-homothetic savings, and consider the impact of these feedback effects in the quantitative model.

Using the results from Proposition 4, it is straightforward to compare the size of non-homothetic savings channel relative to the direct effects of labor distortions. Doing so requires a suitable estimate of the semi-elasticity of labor supply with respect to taxes, L_{τ_ℓ}. Recall that this semi-elasticity is a *partial equilibrium* steady state elasticity. In a meta-analysis of the existing empirical literature, Chetty et al. (2011) report an average *total* steady state (Hicksian) elasticity of labor supply to labor income taxes of around .5. This estimate is the sum of both the extensive and intensive margins. Using an estimate of the average income tax rate of around .35 (Piketty and Saez, 2013a) would imply a semi-elasticity of around 1.4.

Because this is a total elasticity, it will be an over-estimate of L_{τ_ℓ} making any comparison between the two channels a conservative (lower bound) estimate of the importance of the non-homothetic savings channel.

Assuming a positive Θ, the ratio $\frac{\tau_\ell}{\Theta+\tau_\ell}$ is bounded between 0 and 1. Therefore, setting the ratio equal to 1 again generates a lower bound estimate of the importance of my channel. The elasticity of the wage with respect to capital and labor are directly determined by the substitution elasticity between labor and capital, ρ and the labor share, α_L.[17] Finally, comparing the two channels requires an estimate of the semi-elasticity of capital to the direct effects of the tax, K_{PI}. In the following section, I present a sufficient statistic formula for this term.

3.1 A Sufficient Statistic Formula.

The term K_{PI} can be written as function of sufficient statistics. K_{PI} is defined as in equation (14). See Appendix A.9 for a proof.

$$K_{PI} = \underbrace{\sum_I \pi_i \frac{\partial s_i}{\partial PI_i}\left(\frac{wL - \theta_i \ell_i w}{wL}\right) d\tau_\ell}_{\Delta_{NH}} \frac{K_R}{K_R(1 - A_{wL}w_K) - A_R} \frac{wL}{K} \tag{14}$$

As before, w_K is the wage elasticity with respect to capital, while A_{wL} is defined as the labor income elasticity of aggregate household savings, K_R is the interest rate elasticity of firm investment in capital, and A_R is the interest rate elasticity of household savings(debt). If the interest rate elasticity of investment increases, the size of the channel increases. As the the interest rate elasticity of household savings increases, the denominator becomes larger in absolute value and the size of the channel shrinks. Intuitively, if firm investment is very responsive to the interest rate, then as the supply of savings contracts and borrowing costs rise, the impact on capital will be substantial. If however households are very responsive to interest rates, then as the supply of savings contracts and the interest rate increases, households increase their savings supply (decrease the debt), providing more loanable funds to firms, and dampening the effect of the redistribution on capital.

How large of an effect the policy will have on aggregate savings depends on the degree of non-homothetic savings behavior. This term is summarized by Δ_{NH}. The term $\frac{\partial s_i}{\partial PI_i}$ is a type-i household's marginal propensity to save out of permanent income. the term $\frac{wL-\theta_i\ell_i w}{wL}$ is a type-i household's net tax burden relative to average labor income, as high labor income households are net payers of the tax and low labor income households are net recipients.

[17] See Section X for details.

Therefore, the term Δ_{NH} represents the sum of each household's change in savings as a result of their change in permanent income.

3.2 Estimating Marginal Propensities to Save

In this subsection, I estimate the components of the Δ_{NH} statistic directly using data from the Panel Survey of Income Dynamics (PSID). This term requires estimates of the lifetime average marginal propensity to save (MPS) out of annual permanent income flow for each labor-productivity type, as well as the difference between average annual labor income and annual labor income for each age-productivity group. Estimating the latter is straightforward using labor income data.

Both Straub (2019) and Dynan et al. (2004) (henceforth DSZ) use the PSID to explore the relationship between savings behavior and permanent income. Straub (2019) uses consumption data beginning in 1999 to estimate the elasticity of consumption to permanent income. These estimates can be used to generate an implied savings *elasticity*.[18] While this is not the statistic in my formula, in principle, these implied savings elasticities could be combined with savings *rates* out of permanent income by permanent income type to generate marginal propensities to save. I report the results of combining this implied savings elasticity with my own estimates of savings out of permanent income.

DSZ use the PSID to estimate the marginal propensity to save out of permanent income by permanent income type in 2 ways. First, they use variation in the cross section and simply divide the change in median savings rates between income quintiles by the change in median income to trace out a marginal savings schedule. Second, they use time-series variation and regress the change in average individual household savings between an earlier and later sample on the change in household income. Their cross sectional MPS estimates use a change-of-wealth savings measure, which includes capital gains and therefore may not accurately reflect the supply of loanable funds available to firms (Gale and Potter, 2002). They provide time-series estimates for both the change-of-wealth measure and an 'active' savings measure corresponding to the change in wealth minus capital gains, corrected for inflation and reporting error.

I follow DSZ and exploit cross-sectional differences in permanent income. However, my approach differs from theirs in that I use variation in permanent income *within* a permanent income quintile rather than *across* permanent income quintiles, as my formula calls for the within-group MPS. Furthermore, I use a more direct measure of active savings: income less consumption $(Y_{it} - C_{it})$. The reason they were unable to consider a more straightforward

[18]For example, a consumption elasticity of .7 implies an approximate savings elasticity of 1.3.

income less consumption measure of active savings is that consumption data did not appear in the PSID until 1999. Thankfully, the introduction of a set of consumption questions into the survey in 1999, and an additional set in 2005, means that it is possible to observe both many years of a household's active savings – measured simply as income less consumption – as well as many years of past and future income.

Empirical Strategy. To estimate $\frac{\partial s_i}{\partial PI_i}$ and $\frac{wL - w\theta_i^h \ell_i^h}{wL}$, I first define productivity types as permanent labor income quintiles. I do not observe permanent labor income in the data, and therefore must estimate it. I provide detail on this estimation procedure below. I measure the savings of household j at time t (who is type-i, age-h), S_{hijt} as current total income less consumption and taxes. This measures so called 'active' savings – as opposed to changes in total wealth – which is the closest analog to s_i in the model and captures the change in loanable funds available to firms.[19] With an estimate for permanent income, $\hat{P}I_{ij}$ in hand, I estimate a quintile's average MPS using the following equation for each quintile.

$$S_{hijt} = \beta_{0i} + \beta_{1i}\hat{P}I_{ij} + X_{hijt} + \epsilon_{hijt} \tag{15}$$

Here, X_{hijt} is a vector of controls, ϵ_{hijt} is an error term, and β_{0i} is a constant. The estimated coefficient $\hat{\beta}_{1i}$ can be used as the estimate of $\frac{\partial s_i}{\partial PI_i}$.

This strategy identifies MPS out of permanent income using within quintile cross-sectional variation. For this specification to be valid, it must be the case that there is no third factor driving both household permanent income and savings behavior. An obvious candidate of such a factor would be age, and I include the age of the household's reference person in the set of controls X_{hijt}. Similarly, general macroeconomic conditions over a household's sample would affect both my estimate of their permanent income as well as their savings behavior. To address this I also include average annual labor income at year t.

Other factors such as lifestyle or innate preferences could also impact both savings and permanent income. In a second set of regressions I add additional household level controls for the reference person's education, martial status, and family size to attempt to capture these factors.

In the data, I only observe a household's total annual labor income flow, and cannot directly observe what fraction of their current income reflects permanent rather than transitory income. Furthermore, because current income is used to construct my measure of current savings, any measurement error in current income will bias my estimate of the marginal

[19] Measuring savings as changes in total wealth include capital gains. If the value of a household's assets, in particular their house, appreciates, this would increase total wealth but would not reflect new resources available for investment.

propensity to save. To deal with these issues, I follow both Straub (2019) and DSZ and use various proxies for the permanent component of income. The first is a simple symmetric income average, $\overline{Y}_{hijt}$ defined in the following way.

$$\overline{Y}_{hijt} = \frac{1}{T} \sum_{k=-(T-1)/2}^{(T-1)/2} Y_{h+k,ij,t+k}$$

As noted in Straub (2019), as $T \to \infty$, the symmetric income average measures average annual permanent income without noise, as the effects of the transitory income process are averaged away.

I also follow DSZ and use lagged labor income as a proxy. Let Y_{pij} be the average labor income of household j for a sample of four years prior to the start of their sample.[20] For a sample sufficiently far in the past and for a sufficiently low persistence transitory income process, Y_{pij} should be correlated only with the permanent component of Y_{hijt}. I regress current labor income on lagged income and an age group dummy variable as in equation (16), and use the fitted values $\hat{\beta}_1$ and $\hat{\beta}_2$ to predict $\hat{P}I_{hijt}$.

$$\hat{Y}_{hijt} = \beta_0 + \beta_1 Y_{pij} + \beta_2 \mathbb{1}_{\text{Age Group}} + \epsilon_{hijt} \tag{16}$$

Because the dataset is not balanced across age, for both measures of permanent income I calculate the permanent income quintile for each age separately. That is, household j is put in quintile i if their estimated permanent income is in the ith quintile *for their age group.* All regressions control for age group and average annual labor income in year t. I run an additional set with controls for household characteristics including marital status, family size, and education of the response person. PSID sample weights are used in all regressions and summary statistics, and robust standard errors are used to correct for heteroskedasticity.

The Data. The PSID is a longitudinal household survey which began in 1968. The survey was conducted annually until 1997, at which point it became bi-annual. In 1999, the survey added a large group of questions about household consumption, covering about 70% of the categories in the Consumption Expenditure Survey (CEX). In 2005, an additional set of categories was included. Because the 'active' saving concept I use is income less consumption, I will only be able to measure savings starting in 1999. However, I use income data starting in 1995 to construct my lagged income proxy for permanent income.

To construct my consumption measures, I simply add up all consumption categories

[20] For example, if I observe a household starting in 2001, I begin that household's sample in 2005 and set Y_{pij} to be their average labor income in the years prior to 2005.

Table 1: PSID Summary Statistics

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
			All Ages		
Median Income	34,147	50,644	70,193	103,175	174,726
Saving Rate ('99)	0.12	0.16	0.19	0.25	0.36
Saving Rate ('05)	0.00	0.02	0.06	0.12	0.23
			Ages 20-35		
Median Income	22238	29,178	45,502	75,098	121,959
Saving Rate ('99)	0.05	0.07	0.12	0.18	0.31
Saving Rate ('05)	-0.14	-0.09	-0.04	0.06	0.18
			Ages 35-50		
Median Income	33923	55,640	80,221	111,767	181,068
Saving Rate ('99)	0.08	0.15	0.19	0.26	0.35
Saving Rate ('05)	-0.07	0.02	0.05	0.12	0.22
			Ages 50-65		
Median Income	36,973	62,369	87,852	121,391	200,100
Saving Rate ('99)	0.16	0.24	0.28	0.33	0.41
Saving Rate ('05)	0.05	0.12	0.18	0.18	0.28
Observations ('99)	7,966	8,943	9,249	8,021	6,789
Observations ('05)	5989	7,164	7,500	6,367	5,321

This table reports summary statistics for the PSID data by age group and permanent income quintile. 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 ('05) uses the 2005 consumption measures.

together using the 1999 and 2005 set of categories respectively. The PSID total family income measure includes all taxable income of both the respondent and their spouse, as well as all transfer income and social security income. Savings is then simply calculated as total family income less taxes and consumption. I exclude households with missing data or with unrealistically high levels of any individual consumption category.[21] All variables are then put in terms of 2019 dollars. I drop households with missing income data, households younger than 25, households with fewer than 5 years of responses, and households with less than $1,000 of total family income. I split the sample into 4 equally sized age groups, g.

The PSID reports only pre-tax income, however I need post-tax income in order to properly measure household active saving. To estimate each household's annual total tax payment, I use the NBER TAXSIM program.

Results. Table 2 reports the results from the estimation procedure outlined above. Columns labeled '1999' and '2005' report estimates using the 1999 and 2005 measures of consumption respectively. Columns with household controls control for education, marital status, and

[21]Specifically, any household that spends more than a million dollars on any category.

Table 2: Marginal Propensity to Save Out of Permanent Income

	Symmetric Average				Lagged Income				Implied by Straub (2019)
	1999	2005	1999	2005	1999	2005	1999	2005	
Quintile 1	0.25	0.17	0.23	0.16	0.20	0.16	0.10	0.09	0.16
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	
Quintile 2	0.29	0.18	0.28	0.19	0.40	0.26	0.23	0.12	0.21
	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.07)	(0.06)	(0.07)	
Quintile 3	0.41	0.36	0.43	0.39	0.33	0.12	0.26	0.06	0.25
	(0.05)	(0.06)	(0.05)	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)	
Quintile 4	0.41	0.35	0.40	0.36	0.49	0.39	0.38	0.31	0.33
	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)	
Quintile 5	0.54	0.50	0.54	0.50	0.59	0.56	0.59	0.55	0.47
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	
Household Controls	No	No	Yes	Yes	No	No	Yes	Yes	
Implied Δ_{NH}	-.85	-.82	-.86	-.83	-.96	-.92	-1.01	-.86	-.77

Note. This table reports estimated marginal propensities to consume out of permanent income by permanent income quintile using PSID data. All regressions control for average age group and average labor income for a given year. All regressions use PSID sample weights and heteroskedasticity robust standard errors. Columns marked 1999 or 2005 use the 1999 or 2005 consumption data respectively. Household controls include marital status, family size, and education. The final columns multiplies the savings elasticity implied by Straub (2019) by average savings rates. Implied Δ_{NH} multiplies each quintile's MPS by the difference between each households labor income and average labor income, normalized by average labor income.

family size in addition to age group and average annual labor income in year t. The left four columns report the estimated MPS for each income quintile when the symmetric income average is used to construct the proxy for permanent income. A clear pattern emerges. Household in higher income quintiles tend to have higher marginal propensities to save out of permanent income. The differences are especially pronounced between the top quintile and the 3rd and 4th quintile, and between the 3rd and 4th quintile, and the bottom 2 quintiles. The panels on the right report estimates using lagged labor income to construct the proxy for permanent income. The results are largely similar.

The final column takes average savings rates out of current income for each permanent income quintile and multiplies them by 1.3 – the permanent income elasticity of savings implied by Straub (2019). Multiplying a rate by an elasticity generates the derivative required by the formula. The estimated MPS implied by this procedure are very similar to the direct estimates.

The final row of the table reports the value of Δ_{NH} implied by the MPS estimates. In particular, I take the estimated MPS for each income quintile and multiply it by the difference between that quintile's average labor income and the average labor income for the whole sample. These values ultimately become the inputs into the sufficient statistics formula.

3.3 Estimates from the literature

The remaining terms in the formula, namely the interest rate elasticity of savings, the interest rate elasticity of investment, and the wage elasticity of capital, have each been studied in the existing literature. I briefly review the range of estimates for each of these terms.

Interest Rate Elasticity of Capital. The determinants of firm investment are the subject of a large empirical and theoretical literature, much of it exploring the short-run effect of either user-costs or Tobin's q on investment. My sufficient statistic formula characterizes a steady-state relationship, and therefore requires an estimate of the long-run effect of interest rates on firm investment.

Recall that I assumed that $F(K_t, L_t)$ is a constant elasticity of substitution production function with elasticity parameter ρ.

$$Y = \left(\alpha K^{\frac{\rho-1}{\rho}} + (1-\alpha)L^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}}$$

Assume that $\delta = 0$. The firm's first order condition is:

$$\alpha\left(\frac{K}{Y}\right)^{\frac{-1}{\rho}} = R - 1 + \delta = r$$

Therefore the interest rate elasticity of the *ratio* of capital to output, $\frac{\partial \log(K/Y)}{\partial r} = -\rho$. For a Cobb-Douglas production function ($\rho = 1$), the elasticity of the capital-output ratio is -1. Caballero et al. (1995) estimate the long run elasticity of capital to user costs and find values ranging from close to 0 to -2 percent depending on the industry. In their handbook chapter, Caballero (1999) estimate between -.4 and -1 percent. I will consider estimates for ρ between -.8 and 1.1.

The elasticity of substitution between capital and labor, ρ along with the labor share pins down K_R. (see Appendix A.10 for derivation). I use a labor share of .7 along with a range of estimates of ρ to generate values for K_R. For example, a Cobb-Douglas production function with $\rho = 1$ implies that a 1 percent increase in r generates around a .38 percent decrease in aggregate capital.[22]

[22] In forthcoming work, Gomez and Gouin-Bonenfant estimate an elasticity around -1.

Discussion of K_R: If capital is responsive to borrowing costs in the long run – if not in the short run – a natural question is why the United States investment share has stayed relatively constant over the last several decades despite falling real interest rates. Leading explanations include a rise in market power that coincided with the fall in rates (Gutiérrez and Philippon (2018); Farhi and Gourio (2018); Eggertsson et al. (2021)), as well as an under-counting of investment due to the rising importance of hard-to-measure intangibles (Crouzet and Eberly, 2019). If the latter is true, then the apparent tepid response of investment to falling interest rates can be explained by mismeasurement. If a rise in market power explains the response, what if anything can be inferred about the interest rate *elasticity* of firm investment?

Consider the firm's first order condition with respect to capital under standard Dixit-Stiglitz monopolistic competition. Market power in this case this introduces a profit wedge between r and firms' marginal product of capital (see Appendix X for derivation).

$$\alpha\left(\frac{K}{Y}\right)^{\frac{-1}{\rho}} = \mu r$$

If μ rose as r fell, this would explain the lack of a response on the part of capital. However, it is immediate that the wedge μ (if it is constant) does not affect the *elasticity* of capital with respect to r.[23] Intuitively, the last several decades may have featured a *shift back* in the demand for loanable funds from firms alongside a *shift out* in the supply, causing a fall in interest rates with no increase in investment. However, if supply were to *shift back*, as a result of redistribution, this may still generate a sizable decline in investment.

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

[23] To see this, simply take logs then take the derivative with respect to r.

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. DeFusco and Paciorek (2017) find that a 1 percentage point increase in the interest rate reduced mortage 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. For the numerical exercises, I consider values between .1 and 1.

Wage elasticity with respect to capital. For a CES production function with elasticity parameter ρ, the elasticity of the marginal product of labor with respect to capital is not constant, but is pinned down by ρ and the labor share, σ_L (see Appendix A.10 for details). In the numerical exercises, I use the value for w_K that corresponds with the the value chosen for K_R.

3.4 Estimates of K_{PI}

Given a range of estimates for each of the formula's terms along with an estimate of Δ_{NH}, it's now possible to solve for a range of possible estimates of K_{PI}. This range is reported in Table 3.

Table 3: Estimates of K_{PI}

	$\Delta_{NH} = -.93$			Δ_{NH}=-.86			Δ_{NH}=-.77		
A_R:	0.1	0.5	1	0.1	0.5	1	0.1	0.5	1
ρ =.80	-0.53	-0.43	-0.34	-0.49	-0.40	-0.31	-0.43	-0.35	-0.27
ρ =.9	-0.49	-0.42	-0.33	-0.46	-0.39	-0.31	-0.40	-0.34	-0.27
$\rho = 1$	-0.47	-0.41	-0.33	-0.43	-0.38	-0.31	-0.38	-0.33	-0.27
ρ =1.1	-0.45	-0.40	-0.33	-0.42	-0.37	-0.31	-0.37	-0.33	-0.27

Note. This table presents estimates of semi-elasticity of the steady state capital stock to the change in the permanent income distribution associated with a percentage point change in the labor income tax. The parameter ρ denotes the substitution elasticity between capital and labor while A_R is the elasticity of household savings to the interest rate. See text for description of Δ_{NH}.

The first 3 columns of Table 3 report the estimates of K_{PI} using values of Δ_{NH} within the range of those estimated in the previous section. As the elasticity of substitution between capital and labor grows, the interest rate elasticity of capital increases, while the elasticity of the wage with respect to capital decreases. These forces have opposite effects on the size of K_{PI}. Intuitively, a greater K_R means firms are more responsive to the increase in interest rates when savings decline, pushing the elasticity up. However, when w_K decreases, a lower capital stock has less of an effect on wages and income, and therefore a smaller feedback effect on savings. Quantitatively, the latter dominates the former channel, and an increase in ρ results in a decrease in K_{PI}.

Meanwhile, an increase in the interest rate elasticity of household savings dampens K_{PI}. Intuitively, if households are very responsive to interest rate changes, then as aggregate savings contracts, and interest rates rise, households' will response by increasing their savings, dampening the overall effect of the redistribution. Finally, the magnitude of K_{PI} is straightforwardly increasing in the absolute value of Δ_{NH}.

With a range of estimates of K_{PI} in hand, I can now compare the size of the redistribution-investment trade-off to the impact of labor supply distortions. The ratio of the two channels is reported in Table 4.

Table 4: Non-homothetic savings relative to labor distortions

	$\Delta_{NH} = -.93$			Δ_{NH}=-.86			Δ_{NH}=-.76		
A_R:	0.1	0.5	1	0.1	0.5	1	0.1	0.5	1
ρ =.80	0.26	0.22	0.17	0.24	0.20	0.15	0.21	0.18	0.14
ρ =.9	0.31	0.26	0.21	0.29	0.24	0.20	0.25	0.22	0.17
ρ =1	0.37	0.33	0.27	0.35	0.30	0.25	0.31	0.27	0.22
ρ =1.1	0.46	0.41	0.34	0.43	0.38	0.32	0.38	0.33	0.28

Note. This table presents estimates of the ratio between the welfare effects of the non-homothetic savings channel and the labor supply distortion channel. The parameter ρ denotes the substitution elasticity between capital and labor while A_R is the elasticity of household savings to the interest rate. See text for description of Δ_{NH}.

The estimated ratio ranges between .14 on the low end to .46 on the high end. Using the mid-range estimates of both the interest rate elasticity of household saving, A_R and of the degree of non-homotheticity, Δ_{NH}, as well as Cobb-Douglas production ($\rho = 1$), the non-homothetic savings channel is just under 1/3 of the size of labor supply distortions. These results suggest that while the inefficiencies may still generate the *majority* of the costs associated with redistribution, the effects of non-homothetic savings behavior on capital accumulation are likely large enough to warrant attention from researchers and policy makers.

3.5 Extensions

In order to derive the sufficient statistic formula for K_{PI} a number of simplifying assumptions were made including a closed economy, no life-cycle earnings dynamics, and no growth. In the following section, I relax each of these assumptions and see how doing so changes the sufficient statistic formula.

Large Open Economy. An important simplifying assumption employed in the derivation of K_{PI} was a closed economy. In this case, $K = A$ in the steady state. Consider instead the case of a large open economy in which households can lend to and borrow from abroad, but that the domestic savings supply is large enough in comparison with international capital flows that the domestic savings supply partially determines the domestic interest rate.[24] In this case, the steady state asset marker clearing condition is $K = A + NFA$. In this case, the sufficient statistic formula becomes the following. See Appendix A.11 for a derivation.

$$\hat{K}_{PI} = \Delta_{NH} K_R \left(K_R(1 - A_w w_R) - \frac{A}{K} A_R - \frac{NFA}{K} NFA_R \right)^{-1}$$

Note that now, the interest rate elasticity of net foreign assets appears in the denominator. Intuitively, if foreign savings grow substantially as the domestic interest rate increases in response to redistribution, the general equilibrium effect on the interest rate will be muted, and the domestic capital stock will be less effected. Note also that the importance of the domestic interest rate elasticity and the foreign interest rate elasticity depends on their respective share of total assets. If foreign savings are less responsive to the domestic interest rate than domestic savings, then the larger the net foreign asset share of total assets, the smaller the denominator and the large the overall effect.

Arbitrary life-cycle horizon, H. Consider a variant of the overlapping generations economy presented in Section 2, now with H over-lapping generations who each live for H periods. As in the baseline formula, there are I productivity types, and mass π_{ih} of each type and age. Households earn labor income in multiple periods, and have age-varying labor productivity θ_{ih}. In this case, the sufficient statistic formula becomes the following.

$$\hat{K}_{PI} = \frac{K_R}{K_R(1 - A_w w_K) - A_R} \sum_{I,H} \frac{\partial a_i^h}{\partial PI_i} \left(\sum_H (wL - \theta_{ih} \ell_i^h w) \frac{1}{R^{h-1}} d\tau_\ell \right)$$

See Appendix A.12 for a derivation. Now, the degree of non-homotheticity depends on how

[24] This is likely the case in the United States.

the labor income tax changes permanent income over the life-cycle. If households tend to earn more income later in life, the labor income tax will have a smaller effect on permanent income for all households, as it will disproportionately decrease labor earnings later in life. This will dampen the total effect of the policy on savings and aggregate capital.

Balanced Growth. Suppose that the production function was now given by the following function, where Z_t is the level of labor-augmenting productivity.

$$Y_t = \left(\alpha K_t^{\frac{\rho-1}{\rho}} + (1-\alpha)(L_t Z_t)^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}}$$

Suppose that Z_t grows at constant rate, g. To allow for the possibility of a balanced growth path, household preferences are augmented in the following way.

$$u(c_{it}^y, c_{it}^o, \ell_{it}) = \frac{(c_{it}^y/Z_t)^{1-\sigma_y}}{1-\sigma_y} + \beta_i \frac{(c_{it}^o/Z_{t+1})^{1-\sigma_o}}{1-\sigma_o} - \frac{(\ell_{it})^{1+\gamma}}{1+\gamma}$$

Note that now, households get utility from consumption, normalized by aggregate labor productivity. This ensures that as the economy grows and per-capita income increases, households do not continuously increase their savings levels. Intuitively, as average income grows, so too do households' desire for additional consumption. With this normalization, a balanced growth path is possible. See Appendix X for details. In this case, the sufficient statistic formula is identical to equation (14).

4 Quantitative Model

The previous section provided a back-of-the-envelope estimate of the impact of the redistribution-investment channel on welfare relative to the impact of labor supply distortions. While illustrative, this evidence can only speak to the importance of this new channel *relative* to labor supply distortions. Furthermore, by construction, the decomposition presented in Proposition 4 only applies to small increases in the redistributive tax. By examining small changes, I am able to estimate the relative importance of my channel without specifying *why* high-income households have higher MPS out of permanent income. Their MPS serve as sufficient statistics.

For large policy changes however, the micro-foundations matter. If differences in savings behavior are the result of type-specific *preferences*, then a large amount income redistribution should have little effect on these differences. High-productivity households remain more patient despite the redistribution. However, if non-homothetic preferences are responsible for the differences in MPS, then as the degree of redistribution increases, the after-tax permanent income distribution condenses, shrinking the differences in MPS. To assess the welfare costs of large policy changes therefore, a quantitative model is needed.

Furthermore, the results from the previous section concerned the relative size of the new channel in the long run, but cannot speak to the impact of this channel in the short run. The size of the channel over the transition path may be much smaller if the short-run user-cost elasticity of capital is lower than the long run elasticity.

In order to address these points, I solve a quantitative over-lapping generations model that incorporates all of the sources of non-homothetic savings behavior present in the simple model, while adding un-insurable idiosyncratic risk and a more realistic earnings life cycle. Adding un-insurable idiosyncratic risk captures the pre-cautionary savings motive and generates a redistributive role of redistribution. As in the simple model, households inherit both the bequests left and the labor market skills of the previous generation. The model is therefore very similar to the one presented in De Nardi and Yang (2014).

I calibrate the model to the US in 2019 and choose parameters governing savings behavior in order to most closely match the savings patterns I estimated in the previous section. In this setting, I calculate the trade-off between labor income redistribution and average consumption. In particular, I study the effects of increasing the average labor income tax to fund a budget-balancing lump-sum transfer.

Relative to the simple model in the previous section, additional labor income redistribution impacts average consumption through several additional channels. In addition to distorting labor supply, increasing the level of labor income tax to fund a uniform transfer plays an insurance role, as it increases net-of-tax income for households experiencing negative labor productivity shocks. In addition, the tax leaves younger households, who earn relatively less, with relatively more after-tax income than high-earning older households. This shift of resources boosts the savings supply and capital stock (Atkinson and Sandmo, 1980).

In order to isolate the effect of the pure change in permanent income, I solve a supplementary model, in which labor supply is fixed at the steady state level in the baseline model. I then solve for the direct effect of an increase in the labor income tax and uniform transfer on each household type's expected permanent income, and solve for the effect of this hypothetical tax change on average consumption in the model with fixed labor. This exercise isolates the effect of the change in the permanent income distribution on capital.

A potentially important model ingredient missing from the baseline quantitative model is heterogeneous rates of return. Empirical evidence shows that wealthier households receive higher returns to financial wealth (Fagereng et al. (2016), Bach et al. (2020)), which may be compensation for greater risk taking or evidence of greater investment skill. Idiosyncratic returns may help explain higher savings rates at the top of the income distribution (Benhabib et al. (2019), De Nardi and Fella (2017)) Whether higher returns are *type dependent* or *scale dependent* will impact ultimate size of my channel, as scale dependent returns would imply that as redistribution increased, the distribution of returns, and therefore of marginal propensities to save, would compress, lowering the impact of my channel.[25]

In an extension of the

4.1 Environment.

Households. There is a mass of households indexed by $j \in [0,1]$ who each live for H 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 are I permanent labor productivity types. A household j born in year t who is age h and is type-i has labor productivity, $\theta_i^h e_{ij,t+h}$ where θ_i^h is the permanent component of labor income for type-i age-h households and $e_{ji,t+h}$ is that household's idiosyncratic labor shock that evolves according to an AR(1) process with persistence ρ_e and standard deviation σ_e. Households receive $(1-\tau_{\ell ijht}(w_t \ell^h_{ijt+h}\theta_i^h e_{ijt+h}))w_t\ell^h_{ijt+h}\theta_i^h e_{ijt+h}$ in after-tax labor income each period. Here, $\tau_{\ell ijht}(w_{t+h}\ell^h_{ijt+h}\theta_i^h e_{ijt+h})$ is a progressive labor income tax following Bénabou (2002) and Heathcote et al. (2017), given by equation (17). Here, $\bar{\tau}_\ell$ parameterizes the average level of the labor income tax, while γ determines the degree of progressivity.

$$\tau_{\ell ijth}(w_{t+h}\ell^h_{ijt+h}\theta_i^h e_{ijt+h}) = 1 - \bar{\tau}_\ell(w_{t+h}\ell^h_{ijt+h}\theta_i^h e_{ijt+h})^{-\gamma_\ell} \tag{17}$$

Households can save or borrow in a one-period bond, a^h_{ijt+h}, buy government bonds, B_t or capital K_{t+1} at gross after-tax interest rate $R_{t+1} = 1 + (1-\tau_K)r_{t+1}$. Households face a borrowing constraint $\underline{a}$ such that $\underline{a} < a^h_{ijt+h}$. Type-i household receive share σ^i_π of profit flows each period, as well as $R_t a^0_{it}(1-\tau_b)$ in after-tax bequest income when the are born. Note that all type-i households recieve the same bequest transfer equal to the average type-i bequest the year before they are born, $a^0_{it} = a^H_{it-1}$.[26] Finally, households may receive a lump-sum

[25] Gaillard et al. (2023) make a similar point. They show that that different mechanisms used to generate heterogeneous returns determine the long run elasticity of capital and the optimal capital tax.

[26] I make this simplifying assumption to avoid having to keep track of the history of idiosyncratic shocks across generations.

transfer, T_t. Lifetime utility for a household born at time t is given by (18).

$$u(c^h_{ijt+h}, \ell^h_{ijt+h}, a^H_{ijt+H}) = \sum_{h=1}^{H} \beta_i^{h-1}\left(\frac{(c^h_{ijt+h})^{1-\sigma_h}}{1-\sigma_h} - \psi_\ell \frac{(\ell^h_{ijt+h})^{1+\gamma}}{1+\gamma}\right) + \beta^H \psi_a \frac{(a^h_{ijt+H}+\bar{a})^{1-\eta}}{1-\eta} \tag{18}$$

Note that this model nests the same three sources of non-homothetic savings behavior as the simple model presented in Section 2. I follow Straub (2019) and include the term $\bar{a}$ in households' bequest motive in order to generate a mass of households who give no bequests. Let $R^h_{t+h} = \prod_{k=0}^{h} R_{t+k}$. A type-i household born at time t has the following lifetime budget constraint, given by equation (19).

$$a^H_{ijt+H} + \sum_H \frac{c^h_{ijt+h}}{R^h_{t+h}} = R_t a^0_{it}(1-\tau_b) + \sum_H \frac{(1-\tau_{\ell ijht}) w_{t+h} \ell^h_{ijt+h} \theta^h_i e_{ijt+h} + T_{t+h}}{R^h_{t+h}} + \frac{SS_i}{R^H_{t+H}} \tag{19}$$

Firms. There is 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 (20). A competitive final

$$y^m_t = Z_t\left(\alpha {k^m_t}^{\frac{\rho-1}{\rho}} + (1-\alpha){\ell^m_t}^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}} \tag{20}$$

goods firm aggregates the intermediate goods using a Dixit-Stiglitz CES aggregator (21). This specification generates standard expression for demand for each intermediate good (22). Because there are no nominal rigidities, intermediate goods firms all produce the same

$$Y_t = \left(\int_0^1 {y^m_t}^{\frac{\epsilon-1}{\epsilon}} dm\right)^{\frac{\epsilon}{\epsilon-1}} \tag{21}$$

$$y^m_t = Y_t\left(\frac{p^m_t}{P_t}\right)^{-\epsilon} \tag{22}$$

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

Government. The government taxes returns on capital, bequests, and labor income, and issues debt, B_t, pays social security, issues uniform lump-sum transfers, T_t. The government's per-period budget constraint is given by (23).

$$B_t + w_t \int_J \tau_{\ell ijht} \theta_i^h e_{ijht} \ell_{ijht}^h dj + A_t r_t \tau_K + \sum_I \pi_{Hi} a_{it-1}^H \tau_b = T_t + R_t B_{t-1} + \sum_I \pi_{Hi} SS_i \quad (23)$$

Equilibrium. An equilibrium defined as a sequence of prices, $\{R_t, w_t\}_{t\geq 0}$, individual and aggregate financial positions, $\{\{a_{jit}^h\}_{j\in J, i\in I, h\in H}, A_t\}_{t\geq 0}$, policies, $\{\tau_{\ell t}, \tau_{bt}, \tau_{Kt}, B_t, G_t, T_t\}$, individual household and firm allocations, $\{\{c_{ijt}^h, \ell_{ijt}^h\}_{i\in I, j\in J, h\in H}, \{y_t^m, n_t^m, k_t^m\}_{m\in[0,1]}\}_{t\geq 0}$, and aggregate allocation, $\{K_t, L_t, C_t\}_{t\geq 0}$ such that the following conditions hold. Households' first order conditions and budget constraints hold for each productivity type, generation, and history of shocks. The intermediate goods firms' first order conditions (24) and (25), production function and demand hold. The competitive goods firm's technology constraint

$$w_t \mu = \left(\frac{y_t^m}{\ell_t^m}\right)^{\frac{1}{\rho}} (1-\alpha) Z_t \quad (24)$$

$$(r_t + \delta)\mu = \left(\frac{y_t^m}{k_t^m}\right)^{\frac{1}{\rho}} \alpha Z_t \quad (25)$$

holds. Finally, the government budget constraint, labor market clearing, asset market clearing, and resource constraint hold. See Appendix X for a complete description of the equilibrium conditions.

4.2 Calibration

Macroeconomic targets. I set the depreciation rate, δ to target an investment share of 18%. I set the markup, μ to target a 7.5% profit share, and α to target a labor share of .67. In the baseline model I assume $\rho = 1$ and therefore the production function in Cobb-Douglas. I normalize the supply of labor to 1, set K so that the capital-output share is 2.5, and set Z to normalize aggregate output to 1. Net foreign assets are set to clear asset markets, which in the calibration results in a value of $NFA/Y = 2.85$. The net rate of return is $r = .03$.

Government Policy. I set average labor income taxes, $\bar{\tau}_\ell$ to .35 and γ_ℓ to .15 following Heathcote et al. (2017). I follow De Nardi (2004) and Straub (2019) and set the bequest tax equal to .1. I set the capital tax to .4. I follow Huggett and Ventura (2000), De Nardi and Yang (2014), Straub (2019) in setting social security payments by income quintile (see Appendix X for details). Lump-sum transfers are initially set to 0. Government debt relative to GDP, B/Y are set to .7. G is set to satisfy the government's budget constraint, which in the calibration results in a value of government spending to output, $G/Y = .27$.

Household Income. Labor productivity by permanent income type and age, θ_i^h are set so that $\sum \pi_{ih}\theta_i^h = 1$ and so that relative labor productivity matches relative income by type and age in the data. The parameters ρ_e and σ_e are set as in De Nardi and Yang (2014).[27] The AR(1) process is discretized into a Markov process with 2 states. I set the equity shares to match the 2019 wealth Lorenz curve (Aladangady and Forde, 2021). In order to match my own estimates from the PSID in the previous section, I assume $I = 5$ and $H = 4$ (3 working years, plus retirement).

Household Preferences. I set the inverse Frisch elasticity, γ to 1/.82 following Chetty et al. (2011), and the inverse elasticity of inter-temporal substitution for the median age, $\bar{\sigma}$ to 2.5 following the literature. The weight on the disutility of labor, ψ_ℓ and the weight on the bequest motive, ψ_a are set to clear labor markets and target bequests as 5% of GDP respectively. The remaining parameters, $\{\beta_i\}_{i\in I}$, η, $\bar{a}$, and σ_{nh}, where $\sigma_h = \sigma_{nh}\sigma_{h+1}$, are all jointly calibrated so that the savings rates by age and labor productivity type in the model target their counterparts in the data. Table 5 reports the savings rates by age and permanent income type in the data and the model.

Table 5: Savings rates out of current income

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
	Age Group: Young				
Model	-0.17	-0.10	0.12	0.29	0.41
Data	0.05	0.07	0.12	0.18	0.31
	Age Group: Middle				
Model	0.05	0.14	0.28	0.38	0.47
Data	0.08	0.15	0.19	0.26	0.35
	Age Group: Old				
Model	0.48	0.50	0.50	0.52	0.54
Data	0.16	0.24	0.28	0.33	0.41

This table reports savings rates out of current income by permanent income quintile and age in the baseline calibrated model alongside their analogous estimates in the PSID. See Section 3 for details on the estimation procedure.

[27]These parameters are correspond to a 5-year labor earnings shock. Households in my model have the same labor earnings for 15 years.

Table 6: Calibration

Parameter	Description	Value	Source or Target
Distribution of Income			
J	Age Groups	4	
I	Income Groups	5	
$\{\theta_i^y\}_{i\in I}$	Labor productivity (young)	$\{.35, .46, .72, .1.18, 1.92\}$	PSID data
$\{\theta_i^m\}_{i\in I}$	Labor productivity (middle)	$\{.53, .88, 1.26, 1.76, 2.85\}$	PSID data
$\{\theta_i^o\}_{i\in I}$	Labor productivity (old)	$\{.58, .98, 1.38, 1.91, 3.15\}$	PSID data
ρ_e	Persistence e	.8	De Nardi and Yang (2014). See text.
σ_e	Standard deviation e	.3	De Nardi and Yang (2014). See text.
$\{\pi_i\}_{i\in I}$	Distribution of profit income	$\{0, .05, .1, .15, .7\}$	US Wealth Lorenz Curve 2019
Macro Parameters			
μ	Markup	1.08	Profit share of 7.5% (BEA)
α	Capital share after profits	.24	Labor share of .67
δ	Capital depreciation	.07	Investment share of .1
Z	Aggregate Productivity	.76	Set Y=1
NFA	Net Foreign Assets	2.85	See text
Endogenously Calibrated Micro parameters			
$\bar{a}$	Bequest utility when $a_{it}^J = 0$	.1	See text
ψ_a	Bequest utility parameter	1.26	Bequests/GDP of .05
ψ_ℓ	Labor disutility weight	1.92	L=1
$\{\sigma_h\}_{h\in H}$	1/EIS	$\{3.62, 3.01, 2.5, 2.075\}$	See text
$\{\beta\}_{i\in I}$	Discount factor	$\{.87, .90, .93, .95, .98\}$	See text
$\underline{a}$	Borrowing limit	-.2	10% constrained
η	1/elasticity of bequests	1.97	See text
Fiscal Policy			
τ_K	Capital tax	.1	
$\bar{\tau}_\ell$	Average Labor Tax	.35	PSZ. See text.
τ_b	Bequest tax	.1	De Nardi (2004)
B	Government debt	.76	Debt held by public/GDP in 2019
S	Social security transfers	{.06,.13,.21,.33,.41}	See text
G	Government Spending	.27	See text
γ	Labor tax progressivity	.4	CEX (see text)

Note. This table contains the model parameters, their values, and their source or target in the data. Details on how the endogenously calibrated were calculated can be found in the text.

I then solve a supplementary version of the baseline model with exogenous fixed labor supply. All parameters and macro moments are the same, and the labor supply of households at a given age, productivity type, and history of income shocks is set at the corresponding level in the baseline model.

4.3 Policy Experiment

Suppose the fiscal authority increased the average labor income tax rate, $\bar{\tau}_\ell$, keeping the degree of progressivity γ constant, in order to fund a budget-balancing uniform lump-sum transfer, T. How does this policy affect average consumption? To answer this question, I solve for the steady state of the baseline model for a range of policies, $\{\bar{\tau}_\ell, T\}$.

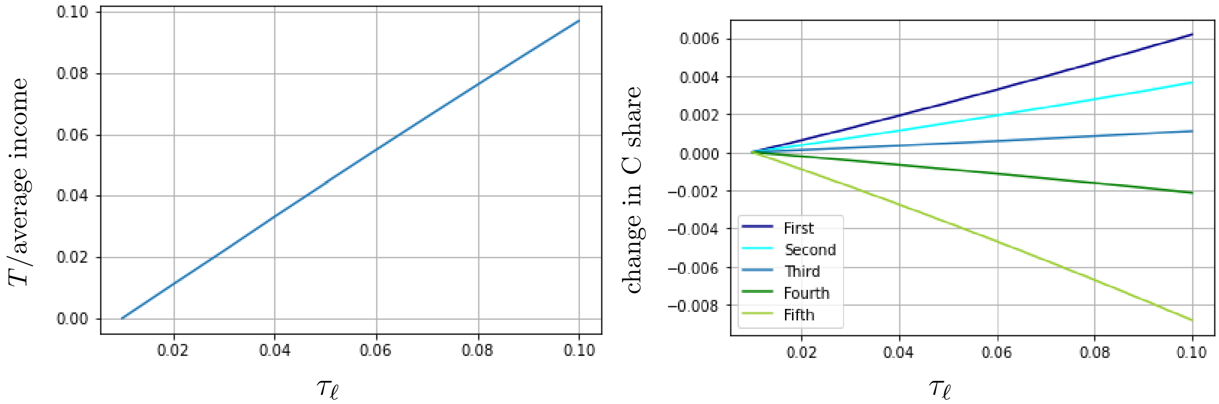

Figure 1: Effect of $\bar{\tau}_\ell$ on T and consumption distribution

Note. The left panel of this figure plots the set of budget balancing policies, $\{\bar{\tau}_\ell$,T$\}$. On the x-axis is the average level of labor income taxes, τ_ℓ while the y-axis plots the lump-sum transfer as a share of average income in the baseline calibration. The right panel plots the change in consumption share by income quintile as a result of the change in $\bar{\tau}_\ell$.

Figure 1 reports the effect of increasing $\bar{\tau}_\ell$ on the size of the lump sum transfer, T relative to the average income level and on the distribution of consumption. To fund a universal income transfer equal to 8 percent of average income, the average labor income tax would need to increase by 7 percentage points.

How much of this trade-off can be attributed to the *direct* effect of non-homothetic savings behavior? To isolate this channel, I solve for the direct effect of a percentage point increase in $\bar{\tau}_\ell$ – and the associated budget balancing transfer – on each household type's expected permanent income. For details on this calculation, see Appendix A.14. Given the change in permanent income for each type, ΔPI_i I define the age specific hypothetical lump-sum taxes, $\{T_{ih}\}_{h \in H}$ so that the following conditions hold.

$$\sum_H \frac{T_{ih}}{(1+r_0)^h} = \Delta PI_i \tag{26}$$

$$T_i = \frac{T_{ih}}{(1+r)^h} \text{ for all } h \in H \tag{27}$$

The first condition states that the present value of the lump-sum tax equals the present value of the change in permanent income from the direct effect of the labor income tax. The second condition states that the present value of the lump-sum tax at each age must all be equal, ensuring that the lump-sum tax does not alter the savings rate by shifting resources over the life-cycle.

I consider the effect of this hypothetical tax on average consumption, holding labor income constant at the baseline steady state level. Because I impose this tax in all idiosyncratic states, I remove the insurance role of the labor income redistribution. Therefore, the only way the hypothetical tax affects average consumption is through shifting permanent income from households with different MPS.

The left panel of Figure 2 plots the decline in steady state capital as the degree of redistribution increases. In the model with fixed labor, this decline in capital can be entirely attributed to the effect of non-homothetic savings behavior. From the figure, we can see that the direct effect of non-homothetic savings behavior can account for over half of the decline in steady state capital as the labor income redistribution increases. A 5 percentage point increase in average labor income taxes results in a 1 percent drop in steady state capital, .6 percent of which can be attributed to the direct effect of redistributing from households with a high MPS to households with a low MPS.

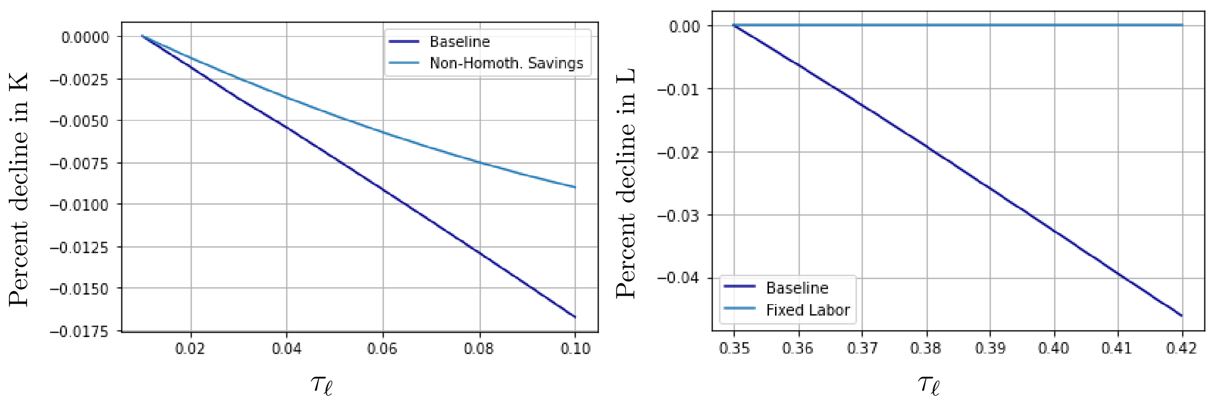

Figure 2: Effect of $\bar{\tau}_\ell$ on Investment and Labor Supply

Note. This left panel of this figure plots the percent change in the steady state capital stock as a function of the percentage point change in steady state average labor income taxes. The right panel plots the analogous decline in steady state labor supply. The dark blue line plots the change in the baseline model, while the light blue line plots the isolated affect of the non-homothetic savings channel.

The redistribution policy generates a sizeable labor supply distortion. In the baseline model, a 5 percentage point increase in the labor income tax generates over a 3 percent decline in the steady state labor supply. Because the decline in capital is around 75% larger in the baseline model, it is clear that the decline in labor supply amplifies the decline in aggregate capital. By construction, aggregate labor does not decline in the fixed labor model.

Next, I examine the impact of the redistribution policy on average consumption. Ultimately, policy makers care about the impact of the redistribution policy on welfare. As shown in the right hand side of Figure 1, increasing the degree of redistribution makes the after-tax income distribution — and therefore the distribution of consumption – more equal. All else equal, this effect will be welfare improving for a sufficiently egalitarian social welfare function. However, by decreasing the long-run labor supply and capital stock, the policy decreases the productive capacity of the economy, lowering average consumption. This effect is plotted in Figure 3. Because the decline in average consumption can be entirely attributed to the effect of non-homothetic savings behavior when labor supply is held fixed, the decline in average consumption can as well.

From Figure 3 we see that a 10 percentage point increase in the average labor income tax rate – which funds a budget balancing transfer – causes average steady state consumption to decline by 2.8 percent. From the figure, we can see that .5 percent or about 18 percent of the total decline in average consumption can be attributed to the direct effect on non-homothetic savings behavior alone. While this does not constitute the majority of the trade-off, these results suggest that the redistribution-investment trade-off may be large enough to influence optimal fiscal policy calculations.

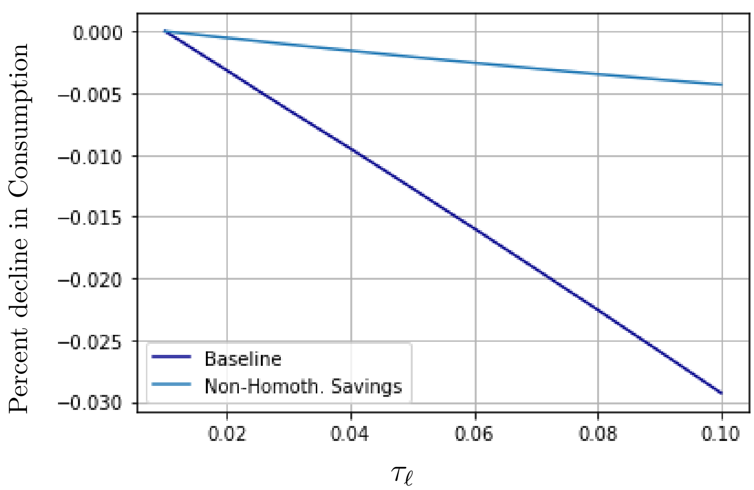

Figure 3: Effect of τ_ℓ on Consumption

Note. This figure plots the percent change in steady state average consumption as a function of steady state average labor income taxes, τ_ℓ. The dark blue line plots the change in the baseline model, while the light blue line plots the isolated affect of the non-homothetic savings channel.

4.4 Comparative Statics

The sufficient statistic formula pointed to several key determinants of the size of the redistribution-investment channel. In particular, the relative importance of this channel was shown to be increasing in the interest rate elasticity of firm investment and decreasing in both the interest rate elasticity of household savings and the capital elasticity of aggregate wages. How do the results in Figure 3 change when the primitives governing these parameters change?

Average Elasticity of Inter-temporal Substitution. The elasticity of inter-temporal substitution (EIS) governs households' interest rate elasticity of savings. In the baseline model, the average σ_h was set to 2.75, implying an average EIS of .36. How does the relative importance of non-homothetic savings change when the EIS, and therefore the interest rate elasticity of household savings decreases?

I consider the impact of the same hypothetical tax as in the previous section on aggregate consumption while varying the average EIS. I plot the results of this exercise in Figure 4.

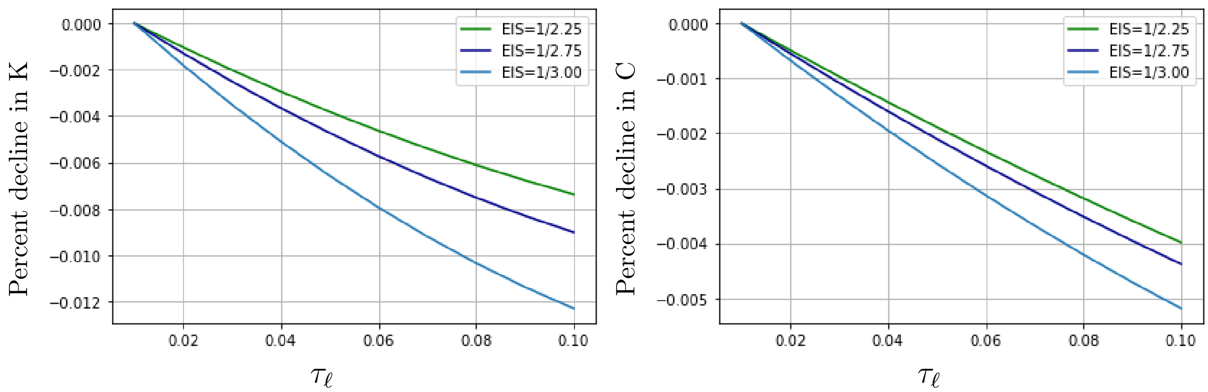

Figure 4: Effect of τ_ℓ on Consumption

Note. This figure shows how the average elasticity of inter-temporal substitution changes the size of the non-homothetic savings channel. The left panel plots the effect of the hypothetical lump-sum redistribution derived in Section 4.3 on aggregate capital. The right panel plots the effect on average consumption.

The left panel of Figure 4 plots the trade-off between redistribution and aggregate capital for various values of $\bar{\sigma}_h$. The right panel plots the analogous figure for average consumption. As the average EIS decreases, the impact of redistribution on capital increases. Intuitively, a smaller EIS implies a smaller interest rate elasticity of savings. Therefore, as redistribution pushes up interest rates, the savings response is more muted, and more capital is crowded out. As a result the impact on average consumption is larger.

4.5 Heterogeneous Rates of Return.

To be added.

4.6 The short run trade-off.

To be added.

5 Conclusion

The aim of this paper was to study the effect of non-homothetic savings behavior on the trade-offs associated with income redistribution in OLG models. When high permanent income households have larger marginal propensities to save out of permanent income than lower income households, all permanent redistribution policies transfer resources from high savers to low savers, lowering both the aggregate savings level and investment. Non-homothetic savings behavior generates a novel welfare trade-off between redistribution and capital accumulation that is present for all forms of redistribution policy.

I show that the existence of such a trade-off depends both on whether the rich do indeed have higher marginal propensities to save out of permanent income, and on whether achieving the first-best level of inequality would result in a savings level and aggregate capital stock below the golden-rule level. I present empirical evidence confirming the former condition. If one assumes that the United States is currently well above the ideal level of inequality, the latter condition is likely to hold as well.

Using labor income redistribution as an illustrative case, I show that the channel I highlight changes the impact of labor income redistribution policy on welfare in the long run. I derive a sufficient statistic formula for the size of my channel and show that it is substantial relative to the effect of labor supply distortions. This back-of-the-envelope analysis suggests that considering non-homothetic savings behavior is important for determining optimal redistribution policy.

The next step in this project is to use the quantitative model to study the importance of this channel in the short run. Redistribution affects long run capital by lowering the savings supply and pushing up interest rates. If firms are slow to respond to these increases – for example due to capital adjustment costs – the long run costs associated with my channel may take many years to materialize. In this case, the weight placed on future generations in the welfare calculus would become a more important ingredient in determining optimal redistribution.

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.

Laurent Bach, Laurent E. Calvet, and Paolo Sodini. Rich pickings? risk, return, and skill in household wealth. *American Economic Review*, 110(9):2703–47, September 2020.

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. doi: 10.1111/1468-0262.00293.

Jess Benhabib, Alberto Bisin, and Mi Luo. Wealth distribution and social mobility in the us: A quantitative approach. *American Economic Review*, 109(5):1623–47, May 2019.

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.

Ricardo J. Caballero. Chapter 12 aggregate investment. volume 1 of *Handbook of Macroeconomics*, pages 813–862. Elsevier, 1999.

Ricardo J. Caballero, Eduardo M. R. A. Engel, John C. Haltiwanger, Michael Woodford, and Robert E. Hall. Plant-level adjustment and aggregate investment dynamics. *Brookings Papers on Economic Activity*, 1995(2):1–54, 1995.

Raj Chetty, Adam Guren, Day Manoli, and Andrea Weber. Are micro and macro labor supply elasticities consistent? a review of evidence on the intensive and extensive margins. *American Economic Review*, 101(3):471–75, May 2011. doi: 10.1257/aer.101.3.471.

Nicolas Crouzet and Janice C Eberly. Understanding weak capital investment: the role of market concentration and intangibles. Working Paper 25869, National Bureau of Economic Research, May 2019.

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.

Mariacristina De Nardi and Fang Yang. Bequests and heterogeneity in retirement wealth. *European Economic Review*, 72:182–196, 2014. ISSN 0014-2921.

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 Dunsky and James Follain. Tax-induced portfolio reshuffling: The case of the mortgage interest deduction. *Real Estate Economics*, 28:683–718, 02 2000. doi: 10.1111/1540-6229.00816.

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

Gauti B. Eggertsson, Jacob A. Robbins, and Ella Getz Wold. Kaldor and piketty's facts: The rise of monopoly power in the united states. *Journal of Monetary Economics*, 124: S19–S38, 2021.

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, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri. Heterogeneity in returns to wealth and the measurement of wealth inequality. *American Economic Review*, 106(5):651–55, May 2016.

Emmanuel Farhi and François Gourio. Accounting for macro-finance trends: Market power, intangibles, and risk premia. Working Paper 25282, National Bureau of Economic Research, November 2018.

Alexandre Gaillard, Christian Hellwig, Philipp Wangner, and Nicolas Werquin. Inequality Accounting: A Tale of Tails, 2023.

William G. Gale and Samara R. Potter. An Economic Evaluation of the Economic Growth and Tax Relief Reconciliation Act of 2001. *National Tax Journal*, 55(1):133–186, March 2002. doi: 10.17310/ntj.2002.1.08.

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.

Germán Gutiérrez and Thomas Philippon. Ownership, concentration, and investment. *AEA Papers and Proceedings*, 108:432–37, May 2018. doi: 10.1257/pandp.20181010.

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.

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. doi: https://doi.org/10.1016/S0304-3932(99)00058-6.

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. doi: 10.1093/qje/qjab007.

Atif R Mian, Ludwig Straub, and Amir Sufi. The saving glut of the rich. Working Paper 26941, National Bureau of Economic Research, April 2020.

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. doi: 10.1016/j.red.2022.07.004.

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.

Paul A. Samuelson. Optimum social security in a life-cycle growth model. *International Economic Review*, 16(3):539–544, 1975.

Karl Schulz. Redistribution and Return Inequality. June 2021.

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, 2019. Revise and resubmit at Econometrica.

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

A Appendix 1

A.1 Proof of Lemma 1.

Non-homothetic savings behavior when $\sigma_y > \sigma_o$ or $\sigma_o > \eta$ or $\beta_H > \beta_L$.

The household's problem is the following:

$$\max_{\{c^y_{it}, c^o_{i,t+1}, a^o_{i,t+1}\}_{h\in\{0,1\}}} \frac{(c^y_{it})^{1-\sigma_y}}{1-\sigma_y} + \beta_i\left(\frac{(c^o_{it+1})^{1-\sigma_o}}{1-\sigma_o} + \psi_a\frac{(a^o_{it+1})^{1+\eta}}{1+\eta}\right)$$
$$\text{s.t. } c^y_{it} + \frac{c^o_{i,t+1} + a^o_{it+1}}{R_{i,t+1}} = R_{it}a^o_{it-1} + w_t\theta_i + T_{it}$$

If $\psi_a = 0$, the households' first order condition is:

$$(c^0_{it})^{-\sigma_y} = \beta_i R_{it}(c^1_{i,t+1})^{-\sigma_o}$$

If $\psi_a = 1$, the household's first order condition is:

$$(c^0_{it})^{-\sigma_y} = \beta_i R_{it}(c^1_{i,t+1})^{-\sigma_o}$$
$$(c^1_{i,t+1})^{-\sigma_o} = \psi_a(a^1_{i,t+1})^{-\eta}$$

Permanent income, PI_i is defined as:

$$PI_i = R_{it}a^o_{i,t-1} + w_t\theta_i + T_{it}$$

For each household type, the derivative of steady state savings to permanent income is given by the following expressions:

Case 1 ($\psi_a = 0$):

$$c^y_{it} + R_{i,t+1}^{\frac{1}{\sigma_o}-1}\beta_i^{\frac{1}{\sigma_o}}(c^y_{it})^{\frac{\sigma_y}{\sigma_o}} = PI_{it}$$
$$\frac{\partial c^y_{it}}{\partial PI_{it}} = \left((1 + \frac{\sigma_y}{\sigma_o}(c^y_{it})^{\frac{\sigma_y}{\sigma_o}-1}R_{i,t+1}^{\frac{1}{\sigma_o}-1}\beta_i^{\frac{1}{\sigma_o}}\right)^{-1} \tag{A.1}$$

This term is positive, meaning consumption increases with permanent income. When permanent income increases, c^y_{it} increases. From A.1 we can see then that $\frac{\partial c^y_{it}}{\partial PI_i}$ is straightforwardly decreasing in PI_i whenever $\sigma_y > \sigma_o$. This derivative is also decreasing in β_i. This implies that $\frac{\partial a^y_{it}}{\partial PI_i} = \frac{\partial s^h_t}{\partial PI_i}$ is *increasing* in β_i and increasing in PI_{it} whenever $\sigma_y > \sigma_o$.

Case 2 ($\psi_a > 0$)**:**

$$c^y_{it} + R_{it}^{\frac{1}{\sigma_o}-1}\beta_i^{\frac{1}{\sigma_o}}(c^y_{it})^{\frac{\sigma_y}{\sigma_o}} + R_{it}^{\frac{1}{\eta}-1}(\psi_a\beta_i)^{\frac{1}{\eta}}(c^y_{it})^{\frac{\sigma_y}{\eta}} = PI_{it}$$

$$\frac{\partial c^y_{it}}{\partial PI_{it}} = \left((1 + \frac{\sigma_y}{\sigma_o}(c^y_{it})^{\frac{\sigma_y}{\sigma_o}-1}R_{it+1}^{\frac{1}{\sigma_o}-1}\beta_i^{\frac{1}{\sigma_o}} + \frac{\sigma_y}{\eta}(c^y_{it})^{\frac{\sigma_y}{\eta}-1}R_{it+1}^{\frac{1}{\eta}-1}(\psi_a\beta_i)^{\frac{1}{\eta}}\right)^{-1} \tag{A.2}$$

Again, this term is positive, meaning the derivative is decreasing in permanent income whenever $\sigma_y > \sigma_o$ or $\sigma_o > \eta$. It is also straightforwardly decreasing in β_i. The derivative of bequests, a^o_{it+t} with respect to permanent income is:

$$\frac{\partial a^o_{it+1}}{\partial PI_{it}} = \left(\frac{\eta}{\sigma_y}(\psi_a\beta_i R_{it+1})^{-1/\sigma_y}(a^o_{it+1})^{\frac{\eta}{\sigma_y}-1} + \frac{\eta}{\sigma_o}(\psi_a^{-1/\sigma_o}R^{-1}_{it+1})(a^o_{it+1})^{\frac{\eta}{\sigma_o}-1} + \frac{1}{R}\right)^{-1} \tag{A.3}$$

Because this derivative is positive, bequests increase with permanent income. Therefore, whenever permanent income increases and $\eta < \sigma_y$ or $\eta < \sigma_o$, the denominator in A.3 decreases and $\frac{\partial a^0_{it+1}}{\partial PI_{it}}$ increases. Therefore, the derivative of bequests

Because the derivative of consumption with respect to PI is decreasing in PI, $\frac{\partial a^y_{it}}{\partial PI_{it}} = \frac{\partial s^y_{it}}{\partial PI_{it}}$ is increasing in PI_{it}.

Derivatives of K with respect to T. Lemma 1 stated that when $\frac{\partial a^y_{it}}{\partial T_{it}}$ and $\frac{\partial a^o_{it}}{\partial T_{it}}$ are constant over types, K_{t+1} is unaffected by fiscal policy and therefore is pinned down by K_t. Instead, when $\frac{\partial a^y_{it}}{\partial T_{it}}$ and $\frac{\partial a^o_{it}}{\partial T_{it}}$ are higher for high productivity types, K_{t+1} can be written as a function of K_t and $\{T_{it}\}_{i\in I}$.

Case 1 ($\psi_a = 0$)**:** Use the firms' FOC to write $R_{t+1} = (1 + F_K(K_{t+1}) - \delta)$ and $w_t(K_t) = F_L(K_t)$. Then use the households' Euler equations and budget constraints to write a^y_{it} as a function of $w_t(K_t)$, T_{it}, and $R_{t+1}(K_{t+1})$:

Use the government budget constraint to write $T_{Ht}(T_{Lt})$ and the asset market clearing condition to write the entire system in terms of K_{t+1}, K_t, and T_{it}.

$$K_{t+1} = \sum_I \pi_i a^y_{it}\Big(T_{it}(T_{Lt}), w_t(K_t), R_{t+1}(K_{t+1})\Big)$$

Take the total derivative with respect to T_{Lt} :

$$\frac{dK_{t+1}}{dT_{Lt}} = \left(\sum_I \pi_i \frac{\partial a^y_{it}}{\partial T_{it}}\right)\left(1 - \sum_I \pi_i \frac{\partial a^y_{it}}{\partial R_{t+1}}\frac{\partial R_{t+1}}{\partial K_{t+1}}\right)^{-1}$$

Note that because savings is increasing the interest rate, and the interest rate is decreasing in K_{t+1} by the firms' FOC, the entire denominator positive.

By the government's budget constraint, $\pi_H \frac{\partial a^y_{Ht}}{\partial T_{Lt}} = \pi_H \frac{\partial a^y_{Ht}}{\partial T_{Ht}}\frac{\partial T_{Ht}}{\partial T_{Lt}} = -\pi_L \frac{\partial a^y_{HT}}{\partial T_{Ht}}$. By assumption $\frac{\partial a^y_{Lt}}{\partial T_{Lt}} \leq \frac{\partial a^y_{Ht}}{\partial T_{Ht}}$, implying that the numerator is either equal to 0 when savings behavior is homothetic or negative.

Case 2 ($\psi_a > 0$): Again, use the firms' FOC to write $R_{t+1} = (1 + F_K(K_{t+1}) - \delta)$ and $w_t(K_t) = F_L(K_t)$. Then use the households' Euler equations and budget constraints to write a^y_{it} as a function of $w_t(K_t)$, $w_{t+1}(K_{t+1})$, T_{it}, $a^o_{i,t+1}$, $R_t(K_t)$, and $a^o_{i,t-1}$, and $R_{t+1}(K_{t+1})$. Now, use household's optimal bequest condition and second period budget constraint to write a^y_{it} as a function of $w_{t+1}(K_{t+1}), a^o_{i,t+1}, R_{t+1}(K_{t+1})$, and T_{it}.

These two equations jointly determine a^y_{it} and $a^o_{i,t+1}$ as a function of $w_t(K_t)$, $w_{t+1}(K_{t+1})$, $R_t(K_t)$, $R_{t+1}(K_{t+1})$, $a^o_{i,t-1}$, and T_{it}. The asset market clearing condition is:

$$2K_{t+1} = \sum_I \pi_i \Big(a^y_{it}((w_{t+j}(K_{t+j}), R_{t+j}(K_{t+j}))_{j\in(0,1)}, a^o_{i,t-1}, T_{it}) + a^o_{it}((w_{t+j}(K_{t+j}), R_{t+j}(K_{t+j}))_{j\in(-1,0)}, a^o_{t-2}, T_{i,t-1}) \Big)$$

For a given K_t, K_{t-1}, a^o_{t-1}, a^o_{t-2}, the total derivatives of K_{t+1} with respect to T_{it} and $T_{i,t-1}$ are:

$$\frac{dK_{t+1}}{dT_{it}} = \sum_I \pi_i \left(\frac{\partial a^y_{it}}{\partial T_{it}}\right)\left(2 - \sum_I \pi_i \frac{\partial a^y_{i}}{\partial R_{t+1}}\frac{\partial R_{t+1}}{\partial K_{t+1}}\right)^{-1}$$

$$\frac{dK_{t+1}}{dT_{i,t-1}} = \sum_I \frac{\pi_i}{2}\left(\frac{\partial a^o_{it}}{\partial T_{i,t-1}}\right)$$

Using the identical logic as in Case 1, $\frac{dK_{t+1}}{dT_{Lt}}$ is negative and $\frac{dK_{t+1}}{dT_{L,t-1}}$ is negative.

Proof of convergence when T_i is fixed. Finally, note that for a given $\{T_i\}$ policy, as long as $K_0 > 0$, K_{t+1} converges to a unique steady state associated with T_i, $K_{ss}(\{T_i\})$. To see this, note that the law of motion for capital is:

$$\frac{K_{t+1} - K_t}{K_t} = \frac{F(K_t, 1)}{K_t} - \delta - \frac{C_t(K_{t+1}, K_t, K_{t-1})}{K_t}$$

If I can show that $\frac{F(K_t,1)}{K_t} - \frac{C_t}{K_t}$ is monotonically decreasing, then that is sufficient to prove there is a unique non-zero steady-state level of capital $\bar{K}$ where this term δ. Then for $K_t > \bar{K}$, $K_{t+1} - K_t < 0$.

Note that $\frac{F(K_t,1)}{K_t} - \frac{C_t}{K_t} = \frac{K_{t+1}}{K_t} - 1 + \delta$, so it is sufficient to show that $\frac{K_{t+1}}{K_t}$ is decreasing in K_t. By the quotient rule, $\frac{\partial(K_{t+1}.K_t)}{\partial K_t} = \frac{K_t \partial K_{t+1}/\partial K_t - K_{t+1}}{K_t K_t} = \frac{\partial K_{t+1}/\partial K_t - K_{t+1}/K_t}{K_t}$. Therefore, as long as $\partial K_{t+1}/\partial K_t < K_{t+1}/K_t$ for all $K_t > 0$, this term is negative and the result is established.

The term $\frac{\partial K_{t+1}}{\partial K_t} - \frac{K_{t+1}}{K_t}$ is:

$$\frac{\sum_I \pi_i \frac{\partial a_{it}}{\partial w_t} \frac{\partial w_t}{\partial K_t}}{1 - \sum_I \pi_i \left(\frac{\partial a_{it}}{\partial R_{t+1}} \frac{\partial R_{t+1}}{\partial K_{t+1}} \right)} - \frac{K_{t+1}}{K_t}$$

Note that because $\frac{\partial R_{t+1}}{\partial K_{t+1}} < 0$, the denominator of the first term is positive, meaning $\frac{\partial K_{t+1}}{\partial K_t} > 0$. Multiplying the above by $\frac{K_t}{K_{t+1}}$:

$$\frac{\sum_I \pi_i \frac{\partial a_{it}}{\partial w_t} \frac{\partial w_t}{\partial K_t} K_t}{K_{t+1} - \sum_I \pi_i \left(\frac{\partial a_{it}}{\partial R_{t+1}} \frac{\partial R_{t+1}}{\partial K_{t+1}} K_{t+1} \right)} - 1$$

Multiplying the first term by $\frac{A_t}{A_t}$, and $\frac{w_t}{w_t}$ or $\frac{W_{t+1}}{w_{t+1}}$ in the numerator and denominator respectively.

$$\frac{w_K \sum_I \frac{\pi_i a_{it}}{A_t} \frac{\partial \log a_{it}}{\partial \log w_t}}{\frac{K_{t+1}}{A_t} - \sum_I \frac{\pi_i a_{it}}{A_t} \left(\frac{\partial \log a_{it}}{\partial \log R_{t+1}} R_K \right)} - 1$$

Using the fact that $K_{t+1} = A_t$, this can be written as,

$$\frac{w_K \sum_I \frac{\pi_i a_{it}}{A_t} \frac{\partial \log a_{it}}{\partial \log w_t}}{1 - R_K \sum_I \frac{\pi_i a_{it}}{A_t} \left(\frac{\partial \log a_{it}}{\partial \log R_{t+1}} \right)} - 1$$

Using the Cobb-Douglas assumption, $w_K = (1 - \alpha_L)$ and $R_K = -\alpha_L$ (see Appendix

A.10), where α_L is the labor share, wL/Y.

$$\frac{(1-\alpha_L)\sum_I \frac{\pi_i a_{it}}{A_t}\frac{\partial \log a_{it}}{\partial \log w_t}}{1+\alpha_L \sum_I \frac{\pi_i a_{it}}{A_t}\left(\frac{\partial \log a_{it}}{\partial \log R_{t+1}}\right)} - 1$$

To see that the denominator is greater than 1, note that a_{it} is increasing in R_{t+1} and therefore the term $\alpha_L \sum_I \frac{\pi_i a_{it}}{A_t}\left(\frac{\log a_{it}}{\partial \log R_{t+1}}\right) > 0$. To see that the numerator is less than 1, note that Cobb-Douglas production means that $(1-\alpha_L) = \frac{rK}{Y}$. Rearranging the numerator gives:

$$\frac{rK}{Y}\frac{w}{A}\sum_I \pi_i \frac{\partial a_{it}}{\partial w_t} = r * \frac{wL}{Y}\sum_I \pi_i \frac{\partial a_{it}}{\partial w_t}$$

Here the right hand side uses the assumption that I've assumed $L = 1$ and that $K = A$. Consumption smoothing ensures that $\sum_I \pi_i \frac{\partial a_{it}}{\partial w_t} < 0$, and because $r < 1$ and $wL/Y < 1$, this product is straightforwardly less than 1.

Therefore $\frac{F(K_t,1)}{K_t} - \frac{C_t}{K_t}$ is monotonically decreasing, implying a unique stable steady state.

Characterizing the Steady state. Assuming $\pi_y = \pi_o = \frac{1}{2}$ as in the text, the steady state version of the asset market clearing condition can be written as:

$$2K = \sum_I \pi_i \Big(a_i^y((w(K), R(K))_{j\in(0,1)}, a_i^o, T_i) + a_i^o((w(K), R(K))_{j\in(-1,0)}, T_i) \Big)$$

Again, taking the total derivative,

$$\frac{d\bar{K}}{dT_L} = \frac{1}{2}\sum_I \pi_i \left(\frac{\partial a_i^y}{\partial T_i} + (\frac{\partial a_i^y}{\partial a_i^o}+1)\frac{\partial a_i^o}{\partial T_i}\right)\left(1 - \frac{1}{2}\sum \pi_i \left(\left(\frac{\partial a_i^y}{\partial w} + (\frac{\partial a_i^y}{\partial a_i^o}+1)\frac{\partial a_i^o}{\partial w}\right)\frac{\partial w}{\partial K} - \left(\frac{\partial a_i^y}{\partial R} + (\frac{\partial a_i^y}{\partial a_i^o}+1)\frac{\partial a_i^o}{\partial R}\right)\frac{\partial R}{\partial K}\right)^{-1}$$

Which can be written as:

$$\frac{d\bar{K}}{dT_L} = \frac{1}{2}\sum_I \pi_i \left(\frac{\partial a_i^y}{\partial T_i} + (\frac{\partial a_i^y}{\partial a_i^o} + 1)\frac{\partial a_i^o}{\partial T_i} \right)\left(1 - A_w w_K - A_r r_K\right)^{-1}$$

Here I've defined A_r and A_w, the elasticity of aggregate households asstes to the interest rate and wage respectively, in the following way:

$$A_w = \frac{1}{2}\sum \pi_i \left(\frac{\partial a_i^y}{\partial w} + (\frac{\partial a_i^y}{\partial a_i^o} + 1)\frac{\partial a_i^o}{\partial w} \right) \quad A_r = \frac{1}{2}\sum \pi_i \left(\frac{\partial a_i^y}{\partial R} + (\frac{\partial a_i^y}{\partial a_i^o} + 1)\frac{\partial a_i^o}{\partial R} \right)$$

Using the same logic as above, consumption smoothing and Cobb-Douglas production implies that $A_w w_K < 1$. Meanwhile, $r_K < 0$ and $A_r > 0$. Taken together this ensures that the denominator is positive.

Again, the numerator, is 0 whenever savings behavior is homothetic and negative whenever high productivity households have higher marginal propensities to save. Therefore, the derivative, $\frac{d\bar{K}}{dT_L} \leq 0$, and is negative whenever $\frac{\partial a_H^h}{\partial T_H} > \frac{\partial a_L^h}{\partial T_L}$ for $h \in \{y, o\}$.

A.2 Proof that $\omega_H^* = \omega_L^*$

The unconstrained planner's problem is to maximize social welfare, given by:

$$SW_s = \sum_I \pi_{ih}\lambda_i \left(\frac{(c_i^y)^{1-\sigma_y}}{1-\sigma_y} + \beta_i\gamma\frac{(c_i^o)^{1-\sigma_o}}{1-\sigma_o} + \beta_i\gamma\psi_a\frac{(a_i^o)^{1-\eta}}{1-\eta} \right)$$

subject only to the resource constraint (4) in steady state.

Let μ be the lagrange multiplier on the The first order conditions with respect to c_i^y are:

$$\pi_{iy}\lambda_i(c_i^y)^{-\sigma_y} = \pi_{iy}\mu$$

This implies that $\lambda_L(c_L^y)^{-\sigma_y} = \omega_L = \omega_H = \lambda_H(c_H^y)^{-\sigma_y}$.

A.3 Proof of Lemma 2

Again, social welfare in the steady state is defined as:

$$SW_s = \sum_I \pi_{ih}\lambda_i \left(\frac{(c_i^y)^{1-\sigma_y}}{1-\sigma_y} + \beta_i\gamma\frac{(c_i^o)^{1-\sigma_o}}{1-\sigma_o} + \beta_i\gamma\psi_a\frac{(a_i^o)^{1-\eta}}{1-\eta} \right)$$

The lifetime budget constraint is given by:

$$c_i^y + \frac{c_i^o + a_i^o}{R} = Ra_i^o + w\theta_i + T_i$$

I can express a change in SW_s following a change in T as:

$$dSW_s = \sum_i \pi_{ih}\lambda_i \bigg((c_i^y)^{-\sigma_y} dc_i^y + \beta_i \gamma (c_i^o)^{-\sigma_o} dc_i^o + \beta_i \gamma \psi_a (a_i^o)^{-\eta} da_i^o \bigg)$$

Using the household's budget constraint:

$$\begin{aligned} dc_i^y &= Rda_i^o + a_i^o dR + \theta_i dw + dT_i - da_i^y \\ dc_i^o &= Rda_i^y + a_i^y dR - da_i^o \end{aligned}$$

The household's optimality condition for bequests is $(c_i^o)^{-\sigma_o} = \psi(a_i^o)^{-\eta}$ and Euler equation is: $(c_i^y)^{-\sigma_y} = \beta_i R (c_i^o)^{-\sigma_o}$. Defining ω_i as in the text, setting $\gamma = 1$, and subbing in the bequest condition and Euler equation, this change can be written as:

$$\begin{aligned} dSW_s = \sum_I \pi_{ih}\omega_i \bigg(Rda_i^o + a_i^o dR + \theta_i dw + dT_i^y \bigg) + \\ \pi_{ih}\omega_{io} \bigg(a_i^y dR + dT_i^o \bigg) \end{aligned}$$

Note that $R\omega_{io} = \lambda_i R\beta (c_i^o)^{-\sigma_o} = \omega_i$. Therefore, $\omega_{io} = \omega_i / R$. Because $RT_i^y = T_i^o$, this is equivalent to:

$$dSW_s = \sum_I \pi_{ih}\omega_i dT_i^y + \sum_I \pi_{ih}\omega_i \bigg(Rda_i^o + a_i^o dR + \theta_i dw + (a_i^y)\frac{dR}{R} \bigg)$$

Let Γ_i^b be type-i households' bequests as a share of total capital. Then the above can be written:

$$\begin{aligned} dSW_s = \sum_I \pi_{ih}\omega_i dT_i^y + \sum_I \pi_{ih}\omega_{iy} RK d\Gamma_i^b + R \sum_I \pi_{ih}(\omega_i \Gamma_i^b) dK + \\ \sum_I \omega_i \pi_{ih} \bigg((a_i^o + \frac{a_i^y}{R}) dR + \theta_i dw \bigg) \end{aligned}$$

Note that for CES production functions, $dR/dw = -L/K$. Therefore, the above can be written as:

$$dSW_s = \sum_I \pi_{ih}\omega_i dT_i^y + \sum_I \pi_{ih}\omega_{iy} RKd\Gamma_i^b + R\sum_I \pi_{ih}(\omega_i\Gamma_i^b)dK + \frac{1}{2}\sum_I \omega_i\pi_i\left((-\frac{a_i^o}{K} + \frac{a_i^y}{KR}) + \frac{\theta_i}{L}\right)dwL$$

Defining Θ and K_{PI} as in the text,

$$\Theta = \frac{1}{2}\sum_I \omega_i\pi_i\left(\frac{\theta_i}{L} - \left(\frac{a_i^o}{K} + \frac{a_i^y}{KR}\right)\right) \tag{A.4}$$

Finally, multiply $R\frac{1}{2}\sum_I \pi_i(\omega_i\Gamma_i^b)dK$ by K/K, and note that $\Gamma_i^b K = a_i^o$ to get:

$$dSW_s = \sum_I \omega_{iy} dT_i^y + \frac{1}{2}\sum_I \omega_i RKd\Gamma_i^b + \left(R\frac{1}{2}\sum_I(\omega_i a_i^o) + wL\Theta w_K\right)K_{PI} \tag{A.5}$$

A.4 Proof of Proposition 1:

This proposition states that when $\frac{\partial a_H^h}{\partial PI_H} > \frac{\partial a_L^h}{\partial PI_L}$ for $h \in \{y, o\}$ and $R > 1$, then $\omega_L^* > \omega_H^*$. The Proof of Proposition 1 proceeds in the following way:

1. Prove that when $\frac{\partial a_H^h}{\partial PI_H} > \frac{\partial a_L^h}{\partial PI_L}$ for $h \in \{y, o\}$, and $R > 1$ and $\omega_L \geq \omega_H$, then $\Theta > 0$.
2. It is then immediate from equation A.5 and the result from Lemma 1 that $\frac{\partial \bar{K}}{\partial T_L} < 0$ when $\frac{\partial a_H^h}{\partial PI_H} > \frac{\partial a_L^h}{\partial PI_L}$ for $h \in \{y, o\}$ that when $\omega_H^* = \omega_L^*$, $dSW_s < 0$ and therefore such an allocation is not optimal.
3. Rule out the possibility that $\omega_H^* > \omega_L^*$, leaving $\omega_L^* > \omega_H^*$ as the only possibility for an optimum.

1. **When $\omega_L \geq \omega_H$, $\frac{\partial a_H^h}{\partial PI_H} > \frac{\partial a_L^h}{\partial PI_L}$ for $h \in \{y, o\}$, and $R > 1$, then $\Theta > 0$.**

Suppose there is a motive for redistribution ($\omega_L > \omega_H$). Whether additional capital contributes positively to welfare hinges on whether Θ is positive. Start with the case without bequests. In this case,

$$\Theta = \sum_I \omega_i\pi_{iy}\left(\frac{\theta_i}{L} - \frac{a_i^y}{KR}\right)\Bigg)$$

We have that $\sum_I \pi_{iy}a_i^y = K$ and $\sum_I \pi_{iy}\theta_i = L$.

Suppose $\pi_{Ly}\theta_L/L \geq \pi_{Ly}a_L^y/K$, meaning type-L households' share of labor total income is greater or equal to their share of total savings. By construction, type-H's share of labor income is less than or equal to their share of savings. If the economy is dynamically efficient and therefore $R > 1$, each types' labor income is weighted more heavily than their savings in Θ. To see this, simply note that $1 > \frac{1}{R}$.

Because $\pi_{Ly}a_L^y/K = 1 - \pi_{Hy}a_H^y/K$ and that $\pi_{Ly}\theta_L/L = 1 - \pi_{Hy}\theta_H/L$, we have that:

$$\Theta = \left(\omega_L \frac{\pi_{Ly}\theta_L}{L} + \omega_H(1 - \frac{\pi_{Ly}\theta_L}{L})\right) - \frac{1}{R}\left(\omega_L \frac{\pi_{Ly}a_L^y}{K} + \omega_H(1 - \frac{\pi_{Ly}a_L^y}{K})\right)$$

Which simplifies to:

$$\Theta = (\omega_L - \omega_H)\left(\left(\frac{\pi_{Ly}\theta_L}{L}\right) - \frac{1}{R}\left(\frac{\pi_{Ly}a_L^y}{K}\right)\right) + \omega_H(1 - R^{-1})$$

From this expression, it's clear that low-types having a higher labor share and $R > 1$ are sufficient conditions for Θ to be positive when $\omega_L > \omega_H$.

The logic for the $\psi_a > 0$ case is analogous. In this case, Θ is:

$$\begin{aligned}\Theta &= \sum_I \omega_i \left(\frac{\pi_{Ly}\theta_i}{L} - \frac{\pi_{Ly}a_i^y}{KR} - \frac{\pi_{Lo}a_i^o}{K}\right) + \omega_i \frac{\pi_{io}a_i^o}{KR} - \omega_i \frac{\pi_{io}a_i^o}{KR} \\ &= \left(\omega_L \frac{\pi_{Ly}\theta_L}{L} + \omega_H(1 - \frac{\pi_{Ly}\theta_L}{L})\right) - \frac{1}{R}\left(\omega_L \frac{a_L}{K} + \omega_H(1 - \frac{a_L}{K})\right) \\ &\quad + \omega_H \frac{\pi_{Ho}a_H^o}{KR} - \omega_H \frac{\pi_{Ho}a_H^o}{K} + \omega_L \frac{\pi_{Lo}a_L^o}{KR} - \omega_L \frac{\pi_{Lo}a_L^o}{K}\end{aligned}$$

Which simplifies to:

$$\begin{aligned}\Theta = (\omega_L - \omega_H)\left(\left(\frac{\pi_{Ly}\theta_L}{L}\right) - \frac{1}{R}\left(\frac{\pi_{Ly}a_L^y}{K}\right)\right) + \omega_H(1 - R^{-1}) + \\ \omega_H \frac{\pi_{Ho}a_H^o}{K}(R^{-1} - 1) + \omega_L \frac{\pi_{Lo}a_L^o}{K}(R^{-1} - 1)\end{aligned}$$

To see that low-productivity types have higher labor shares compared to savings shares when preferences are non-homothetic. To show this, I will prove that the derivative of a type's lifetime capital share, $\frac{a_i^y + a_i^o}{K}$ to their labor share is non-negative. That is, keeping L constant, by increasing θ_i you increase $\frac{a_i^y + a_i^o}{K}$. Define $a_i \equiv a_i^y + a_i^o$. The derivative of a_i/K

with respect to θ_i is:

$$\partial(a_i/K)/\partial\theta_i = \frac{\partial a_i/\partial\theta_i}{K} - \frac{a_i}{K}\frac{\partial K/\partial\theta_i}{K}$$

From Appendix A.1, we have that $\frac{\partial a_i^y}{\partial PI_i}$ and $\frac{\partial a_i^o}{\partial PI_i}$ increase as permanent incomes increase. Suppose $\theta_i < 1/2$ and therefore by construction $\theta_j > 1/2$.

In this case, because type-j households are losing permanent income but currently have higher permanent income, their loss in savings will outweigh the growth in savings of the type-i households. Therefore, we have that $\partial K/\partial\theta_i < 0$, and therefore the above derivative is positive.

Using identical logic, suppose $\theta_i > 1/2$. Then as θ_i increased, θ_j would decrease, decreasing type-j's share and increasing type-i's share by construction.

2. Using equation A.5, we see that at the hypothetical steady state with $K = \bar{K}$ and $\omega_L = \omega_H$, $\Theta > 0$ when $F_K > \delta$ (and $R > 1$) and preferences are non-homothetic, and therefore this cannot be a maximum.

3. Finally, to rule out the possibility that $\tilde{\omega}_H > \tilde{\omega}_L$ at the optimum, suppose at this allocation that $\tilde{c}_H^y < \tilde{c}_L^y$ and therefore $(\tilde{c}_H^y)^{-\sigma_y} > (\tilde{c}_L^y)^{-\sigma_y}$. At such an allocation, enough redistribution has been done that the permanent income of the high types is lower than that of the low types. Because labor is exogenous and taxes are lump-sum, this economy is isomorphic to one in which $\theta_L > \theta_H$.

By reducing T_L enough to interchange the types' permanent income, therefore, the economy would achieve the same level of aggregate capital, but would result in a new distribution of consumption, $c_L^{y'}\tilde{c}_H^y$ and $c_H^{y'} = \tilde{c}_L^y$. Because $\lambda_H \geq \lambda_L$, this allocation dominates the former in terms of social welfare, meaning the first cannot be optimal.

Suppose instead that $\tilde{c}_H^y > \tilde{c}_L^y$ and therefore $(\tilde{c}_H^y)^{-\sigma_y} < (\tilde{c}_L^y)^{-\sigma_y}$. Then redistributing more towards high would both increase utility directly and indirectly by raising capital, and therefore cannot be a solution.

A.5 Proof of Proposition 2

Proposition 2 states that when savings behavior is non-homothetic, and the steady state associated with $\omega_L = \omega_H$ is dynamically efficient ($F_K(\bar{K}) > \delta$), and γ is sufficiently high, then at the constrained optimum, $\omega_{Ht}^I > \omega_{Lt}^I$ for all t. Because assuming $\lambda_H \geq \lambda_L$ ruled out any allocation in which $\omega_{Lt} > \omega_{Ht}$, it is sufficient to show that $\omega_{Lt} = \omega_{Ht}$ is not optimal at any horizon. To do this, I proceed according to the following steps:

1. Show that any equilibrium allocation can be written as functions of fiscal policy and last period's assets, and characterize those functions. Show that these functions are necessary and sufficient conditions for an equilibrium.

2. Use these functions as constraints to solve for the constrained planner's first order conditions.

3. Show that if $\omega_{Lt} = \omega_{Ht}$, for sufficiently high γ the first order conditions are not satisfied.

Step 1: find necessary and sufficient conditions for an allocation to be an equilibrium. I again use the results from Section A.1 to show that any equilibrium path of capital, $\{K_{t+1}\}_{t\geq 0}$ can be expressed as an implicit function of fiscal policy and the previous period's capital. Repeating the results here:

Case 1: $\psi_a = 0$**:** In this case, K_{t+1} can be written as an implicit function of K_t and T_{it}, $K_{t+1}(K_t, T_{it})$, where $\frac{dK_{t+1}}{dK_t} > 0$ and $\frac{dK_{t+1}}{dT_{Lt}} < 0$ when savings behavior is non-homothetic and $\frac{dK_{t+1}}{dT_{Lt}} = 0$ when savings behavior is homothetic.

Next, define the following functions: (1) Use the resource constraint, $F(K_t, 1) = C_t + K_{t+1} - (1-\delta)K_t$ to write C_t as a function of K_t, K_{t+1}, $C_t(K_t, K_{t+1})$. (2) define Γ^h_{it} as type-i age-h generation-t households share of total consumption at time t. Using the household's Euler equation and lifetime budget constraint, and substituting in the firm's first order conditions for prices, c^y_{it} can be expressed as an implicit function of K_t, K_{t+1}, T_{it} and $c^o_{i,t+1}$ can be expressed as an implicit function of K_t, K_{t+1}, T_{it}. Finally, because $c^y_{it} = \Gamma^y_{it} C_t$ and $c^o_{it} = \Gamma^o_{it} C_t$, we can write $\Gamma^y_{it}(K_t, K_{t+1}, T_{it}, C_t) = c^y_{it}(K_t, K_{t+1}, T_{it)}/C_t$ and $\Gamma^o_t(K_{t-1}, K_t, T_{it-1}, C_t) = c^o_t(K_{t-1}, K_t, T_{it-1})/C_t$. Note that holding K_t and K_{t+1} (and therefore C_t) fixed, Γ^y_{it} is increasing in T_{it} and Γ^o_{it} is increasing in T_{it-1}.

With these functions so defined, the following conditions are necessary and sufficient for an equilibrium:

$$\Gamma^y_{it} = \Gamma^y_{it}(K_t, K_{t+1}, T_{it}, C_t) \text{ for } i \in \{L, H\} \tag{A.6}$$

$$\Gamma^o_{it+1} = \Gamma^y_{it}(K_t, K_{t+1}, T_{it}, C_{t+1}) \text{ for } i \in \{L, H\} \tag{A.7}$$

$$K_{t+1} = K_{t+1}(K_t, T_{Lt}) \tag{A.8}$$

$$C_t + K_{t+1} = F(K_t, 1) + (1-\delta)K_t \tag{A.9}$$

Necessary. Follows directly from the construction of the functions.

Sufficient. To show sufficiency, I find a set of prices and financial positions such that any allocation that satisfies equations (A.6)-(A.9) for policies, $\{T_{Lt}, T_{Ht}\}$ can be implemented as a competitive equilibrium, meaning that at these prices and policies, the allocation satisfies the firm's optimality conditions, the households' optimality conditions and budget constraints, and the government budget balances, the resource constraint, and asset market conditions clear. Set $w_t = F_L(K_t)$ and $R_t = 1-\delta+F_K(K_t)$. In doing so, the firm's first order conditions are satisfied. Set $a^y_{it} = \theta_i w_t + T_{it} - c_{it}$. By construction, at these prices the household's Euler equation and budget constraint are satisfied as long as equations A.6 and A.7 are satisfied. By construction, equation A.8 guarantees the asset market clears. Finally, equation A.9 is the resource constraint, implying that by Walras Law, the government budget is satisfied.

Case 2: $\psi_a > 0$**:** (1) In this case, using the results from the previous section, K_{t+1} can be written as an implicit function of K_t, K_{t-1},T_{Lt}, T_{Lt-1}, as well as $\{a^o_{i,t-1}, a^o_{i,t-2}\}_{i\in I}$. (2) Define Γ^h_{it} again as age-h type-i households' share of consumption in the same way as above. (3) Define Γ^b_{it} as type-i generation t-1 bequests as a share of total capital. So we have a function, $K_{t+1}((K_t - j, T_{it-j}, \{a^o_{i,t-1-j}\}_{i\in I})_{j=0,1})$, where $\frac{dK_{t+1}}{dK_t}, \frac{dK_{t+1}}{dK_{t-1}} > 0$ and $\frac{dK_{t+1}}{dT_{Lt}}, \frac{dK_{t+1}}{dT_{Lt-1}} < 0$ when savings behavior is non-homothetic and $\frac{dK_{t+1}}{dT_{Lt}}, \frac{dK_{t+1}}{dT_{Lt-1}} = 0$ when savings behavior is homothetic. Finally, $\frac{dK_{t+1}}{da^o_{i,t+j}} > 0$ for $i \in I$, and $j \in -1, -2$. In this case, necessary and sufficient conditions for an allocation,policy combination to be implementable are:

$$\Gamma^y_{it} = \Gamma^y_{it}(K_t, K_{t+1}, T_{it}, C_t, a^o_{t-1}) \text{ for } i \in \{L, H\} \tag{A.10}$$

$$\Gamma^o_{it+1} = \Gamma^y_{it}(K_t, K_{t+1}, T_{it}, C_{t+1}, a^o_{t-1}) \text{ for } i \in \{L, H\} \tag{A.11}$$

$$K_{t+1} = K_{t+1}(K_t, T_{Lt}, \{a^o_{it-1}, a^o_{it-2\}_{i\in I}}) \tag{A.12}$$

$$C_t + K_{t+1} = F(K_t, 1) + (1-\delta)K_t \tag{A.13}$$

$$a^o_{it+1} = a^o_{it+1}(K_{t+1}, K_t, a^o_{it-1}, T_{it}) \tag{A.14}$$

Necessary. Again, follows from the construction of the functions.

Sufficient. Again, I need a set of financial positions and prices, such that for policies $\{T_{Lt}, T_{Ht}\}$, the allocation that satisfies (A.10)-(A.14) can be implemented as a competitive equilibrium. Set $w_t = F_L(K_t)$ and $R_t = 1-\delta+F_K(K_t)$ as in the previous case. This ensures the firm's optimality conditions are satisfied. Set $a^y_{it} = w_t\theta_i + T_{it} - c_{it}$. By construction, at these prices, the household's Euler equation and budget constraint are satisfied as long as equations (A.10), (A.11), and (**??**) are. Equation (A.12) ensures the asset market condition clears, and equation (A.13) is the resources constraint, meaning the government's budget is satisfied by Walras Law.

Step 2: Solve for the optimal implementable allocation, x_I^*.

Case 1: ($\psi_a = 0$): In this case, social welfare is given by the following expression:

$$SW = \sum_I \lambda_i \pi_i \sum_{t=0}^{\infty} \gamma^t \left(\frac{\Gamma_{it}^y C_t^{1-\sigma_y}}{1-\sigma_y} + \beta_i \frac{\Gamma_{it+1}^o C_{t+1}^{1-\sigma_o}}{1-\sigma_o} \right) + \frac{1}{\gamma} \beta_i \frac{\Gamma_{i0}^o C_0^{1-\sigma_o}}{1-\sigma_o}$$

The planner's problem is to choose a sequence of T_{Lt} (which implies $T_{Ht} = -T_{Lt} \frac{\pi_L}{\pi_H}$ by the government's budget constraint) and a sequence $\{C_t\}_{t \geq 0}$ in order to maximize SW, subject to:

$$\begin{aligned} \Gamma_{it}^y &= \Gamma_{it}^y(K_t, K_{t+1}, T_{it}, C_t) \text{ for } i \in \{L, H\} \\ \Gamma_{it+1}^o &= \Gamma_{it}^y(K_t, K_{t+1}, T_{it}, C_{t+1}) \text{ for } i \in \{L, H\} \\ K_{t+1} &= K_{t+1}(K_t, T_{Lt}) \\ C_t + K_{t+1} &= F(K_t, 1) + (1-\delta) K_t \end{aligned}$$

Define λ_t as the lagrange multiplier with respect to the resource constraint. Define $\omega_{it} = \lambda_i \gamma^t (c_{it}^y)^{-\sigma_y}$ and $\omega_{it}^o = \beta_i \lambda_i \gamma^{t-1} (c_{it}^o)^{-\sigma_o} = \omega_{it-1}/R_t$ as in the text. Let The planner's first order condition with respect to T_{Lt} is:

$$\sum_{j \geq 0}^{\infty} C_{t+j} \left(\pi_L \omega_{Lt+j}^y \frac{\partial \Gamma_{Lt+j}^y}{\partial T_{Lt}} + \pi_H \omega_{Ht+j}^y \frac{\partial \Gamma_{Ht+j}^y}{\partial T_{Lt}} + \pi_L \omega_{Lt+j}^o \frac{\partial \Gamma_{Lt+j}^o}{\partial T_{Lt}} + \pi_H \omega_{Ht+j}^o \frac{\partial \Gamma_{Ht+j}^o}{\partial T_{Lt}} \right) + \sum_{t \geq 0}^{\infty} \frac{\partial K_{t+1}}{\partial T_{Lt}} \left(\lambda_{t+1} (F_K(K_{t+1}) + 1 - \delta) - \lambda_t \right) = 0$$

The first order condition with respect to C_t is:

$$\sum_I \lambda_i \pi_i \left(\gamma^{t-1} (\Gamma_{it}^o)^{1-\sigma_o} C_t^{-\sigma_o} + \gamma^t (\Gamma_{it}^y)^{1-\sigma_y} C_t^{-\sigma_y} \right) = \lambda_t$$

Homothetic savings: If savings behavior is homothetic ($\frac{\partial a_{Lt}^y}{\partial T_L} = \frac{\partial a_{Ht}^y}{\partial T_H}$) then by Lemma X, the derivative of capital with respect to the lump-sum tax, $\frac{\partial K_{t+1}}{\partial T} = 0$. In this case, $\frac{\partial \Gamma_{Lt}^h}{\partial T_L} = \frac{\partial \Gamma_{Ht}^h}{\partial T_L}$, and therefore the only solution to the problem is to set T_L such that $\omega_{Lt}^h = \omega_{Ht}^h$ for $h \in \{y, o\}$ and for all t.

Non-Homothetic savings: When savings behavior is non-homothetic, ($\frac{\partial a^y_{Lt}}{\partial T_L} < \frac{\partial a^y_{Ht}}{\partial T_H}$) then using the results in A.1, the derivative of capital with respect to the lump-sum tax, $\frac{\partial K_{t+1}}{\partial T} < 0$ for periods $t \geq 0$.

Step 3: Show that $\omega_{Lt} = \omega_{Ht}$ does not satisfy the FOC. Assume as in the Proposition a hypothetical steady state with capital level $\bar{K}$ such that $\pi_L \omega_{Lt} = \pi_H \omega_{Ht} = \pi_L \frac{\omega^o_{Lt}}{R(\bar{K})} = \pi_H \frac{\omega^o_{Ht}}{R(\bar{K})}$. That is, the hypothetical steady state corresponding to fiscal policy setting the population-weighted welfare weights are equal across types. Let $\bar{T}_L$ be the steady state fiscal policy that implements this allocation.

(a) Suppose $K_0 = \bar{K}$. Consider $\{\bar{T}_L\}_{t\geq 0}$ (and therefore $\{\bar{K}\}_{t\geq 0}$ as a potential solution to our planner's problem. In this case, the allocation would remain unchanged, and therefore, using the FOC with respect to C_t, $\lambda_t \gamma = \lambda_{t+1}$. Recall that we have assumed that $F_K(\bar{K} + 1 - \delta) > \frac{1}{\gamma}$. Plugging this into the FOC with respect to T_{Lt} :

$$\sum_{j\geq 0}^{\infty} C\left(\pi_L \omega_{Lt+j} \frac{\partial \Gamma^y_{Lt+j}}{\partial T_{Lt}} + \pi_H \omega_{Ht+j} \frac{\partial \Gamma^y_{Ht+j}}{\partial T_{Lt}} + \pi_L \omega^o_{Lt+j} \frac{\partial \Gamma^o_{Lt+j}}{\partial T_{Lt}} + \pi_H \omega^o_{Ht+j} \frac{\partial \Gamma^o_{Ht+j}}{\partial T_{Lt}}\right) = -\sum_{t\geq 0}^{\infty} \frac{\partial K_{t+1}}{\partial T_{Lt}} \left((F_K(\bar{K}) + 1 - \delta) - \frac{1}{\gamma}\right)\lambda_{t+1}$$

Using the fact that $\pi_L \omega_{Lt+j} = \pi_H \omega_{Ht+j}$ and $\omega^o_{it+j} = \omega_{it+j-1}/R(\bar{K}) = \omega_{it+j}\frac{1}{\gamma}/R(\bar{K})$, we can re-write the above as:

$$\pi_L \omega_{Lt}\left[\sum_{j\geq 1}^{\infty} C\gamma^j \left(\frac{\partial \Gamma^y_{Lt+j}}{\partial T_{Lt}} + \frac{\partial \Gamma^y_{Ht+j}}{\partial T_{Lt}} + \frac{1/\gamma}{R(\bar{K})}\frac{\partial \Gamma^o_{Lt+j}}{\partial T_{Lt}} + \frac{1/\gamma}{R(\bar{K})}\frac{\partial \Gamma^o_{Ht+j}}{\partial T_{Lt}}\right) + \left(\frac{\partial \Gamma^y_{Lt}}{\partial T_{Lt}} + \frac{\partial \Gamma^y_{Ht}}{\partial T_{Lt}}\right) = -\lambda_{t+1} \sum_{j\geq 0}^{\infty} \frac{\partial K_{t+1+j}}{\partial T_{Lt}} \left((F_K(\bar{K}) + 1 - \delta) - \frac{1}{\gamma}\right)\gamma^j\right.$$

Note that because, by assumption, $\frac{1/\gamma}{R(\bar{K})} < 1$ and therefore the left hand side of the equation is non-zero. T_{Lt} only affects the consumption share of generation t directly, but changes the consumption share of every subsequent generation by decreasing the capital stock and increasing the interest rate. Therefore, the first term above is negative, as the increasing interest rate decreases the consumption share of the young (which is weighted more heavily, as $\frac{1/\gamma}{R(\bar{K})} < 1$) than the increase in the consumption share when old. Note that the right hand side is positive, as $\lambda_t > 0$, $\frac{-\partial K_{t+1+j}}{\partial T_{Lt}} > 0$, and $F_K(\bar{K}) + 1 - \delta - \frac{1}{\gamma}$ is positive by assumption.

As $\gamma \to 1$, the first term gets larger in terms of absolute value and approaches $-\infty$. Therefore, for sufficiently large $\hat{\gamma}$, the first order condition is not satisfied, as the right hand side becomes an increasingly large positive number, while the left hand side becomes an increasingly large negative number.

(b) Now suppose $K_0 \neq \bar{K}$. Consider a path for fiscal policy, $\{\tilde{T}_{Lt}\}$, now with a time subscript, that implements the first best level of inequality at each period.

Prove that $\tilde{T}_{Lt}$ converges to $\bar{T}_L$ and therefore K_t converges to $\bar{K}$. To show that the path of fiscal policy converges to $\bar{T}_L$, the policy associated with the first-best steady state I show that the following are true:

1. The ratio $\frac{K_{t+1}}{K_t}$ is monotonically decreasing in K_t for a given fixed T_L. Therefore, if $K_t > \bar{K}$, $\frac{I(K_t)}{K_t} - \delta = \frac{K_{t+1}}{K_t} - 1 < 0$ and capital decreases, while the opposite is true if $K_t < \bar{K}$. Therefore, steady states are unique and stable.
2. If $K_t = \tilde{K}_t > \bar{K}$, then then T^*_{Lt} associated with first best equality in this case is higher that the $\tilde{T}_L$ associated with the steady state level of capital of $\tilde{K}$. Similarly, if $K_t = \tilde{K} < \bar{K}$, then $T^*_{Lt} < \tilde{T}_L$.
3. Using (1) and (2) Setting T^*_{Lt} at time t will bring K_{t+1} closer to $\bar{K}$. This implies that the T^*_{Lt+1} is closer to $\bar{T}_L$.

(1) Proof that $\frac{K_{t+1}}{K_t}$ is monotonically decreasing in K_t :

The derivative, $\frac{\partial K_{t+1}/K_t}{\partial K_t}$ is equal to:

$$\frac{\partial K_{t+1}/\partial K_t - K_{t+1}/K_t}{K_t} < 0$$

If we multiply by, $\frac{K_t}{K_{t+1}}$, we can see that this derivative is negative whenever:

$$\frac{\partial K_{t+1} K_t}{\partial K_t K_{t+1}} < 1$$

Using the implicit function theorem and the asset market clearing condition, this can be written in the following way, where A_r, A_w, and w_K are defined as in Section A.1:

$$\frac{\partial K_{t+1} K_t}{\partial K_t K_{t+1}} = \frac{\sum_I \pi_i \frac{a_{it}}{A_t} A_w w_K}{\frac{K_{t+1}}{A_t} - \sum_I \pi_i \frac{a_{it}}{A_t} A_r R_K} = \frac{A_w w_K}{1 - A_r R_K}$$

Where the final expression comes from the fact that $K_{t+1} = A_t$. Becuase $A_r > 0$ and $R_K < 0$, the denominator is above 1. Using the result from section A.1 that $A_w w_K < 1$, this fraction is less than one, meaning the derivative is negative, and the steady state is unique and stable.

(2) Take the first case where $K_t = \tilde{K} > \bar{K}$. Suppose that $T^*_{Lt} = \tilde{T}_L$. Then $K_{t+1} = K_t = \tilde{K}$. Using the results from Section A.1, in this case $\tilde{T}_L < \bar{T}_L$, which implies that $\tilde{c}^y_{Lt} < \bar{c}^y_L$ and $\tilde{c}^y_{Ht} > \bar{c}^y_H$ and therefore $\tilde{\omega}_L > \tilde{\omega}_H$. A contradiction. Suppose $T^*_{Lt} < \tilde{T}_L$. Using the convergence result from Part 1, this implies that $K_{t+1} > K_t > \bar{K}$. Again, using the results from Section A.1, this implies that $c^y_{Lt} < \bar{c}^y_L$ and $c^y_{Ht} > \bar{c}^y_H$ and therefore $\tilde{\omega}_L > \tilde{\omega}_H$. A contradiction. Therefore, it must be that $T^*_{Lt} > \tilde{T}_L$. By a symmetric argument, if $K_t = \tilde{K} < \bar{K}$, then $T^*_{Lt} < \tilde{T}_L$.

(3) Combining parts (1) and (2) implies that if $K_t > \bar{K}$, then by (2) $T^*_{Lt} > \tilde{T}_L$. Therefore, using (1), $K_{t+1} < K_t$ and $|K_{t+1} - \bar{K}| < |K_t - \bar{K}|$. This implies that K_t converges to $\bar{K}$ and therefore that T^*_{Lt} converges to $\bar{T}_L$.

Again consider the first order condition with respect to consumption:

$$\sum_I \lambda_i \pi_i \left(\gamma^{t-1} (\Gamma^o_{it})^{1-\sigma_o} C_t^{-\sigma_o} + \gamma^t (\Gamma^y_{it})^{1-\sigma_y} C_t^{-\sigma_y} \right) = \lambda_t$$

Note that as $K_{t+1} \to \bar{K}$, $C_{t+1} \to C$, and therefore for any arbitrary small ϵ, there exists a $\tau \geq 0$, such that for $t \geq 0$, $|\lambda_{t+1} - \lambda_t| < \epsilon$ for all $t \geq \tau$.

Again using the fact that $\pi_L \omega_{Lt+j} = \pi_H \omega_{Ht+j}$ by assumption, we can re-write the first order condition with respect to T_{Lt} as:

$$\pi_L \sum_{j \geq 1}^{\infty} C_{t+j} \gamma^j \left(\omega_{Lt} \frac{\partial \Gamma^y_{Lt+j}}{\partial T_{Lt}} + \omega_{Lt} \frac{\partial \Gamma^y_{Ht+j}}{\partial T_{Lt}} + \omega^o_{Lt} \frac{\partial \Gamma^o_{Lt+j}}{\partial T_{Lt}} + \omega^o_{Lt} \frac{\partial \Gamma^o_{Ht+j}}{\partial T_{Lt}} \right) +$$
$$\left(\frac{\partial \Gamma^y_{Lt}}{\partial T_{Lt}} + \frac{\partial \Gamma^y_{Ht}}{\partial T_{Lt}} \right) = -\lambda_{t+1} \sum_{j \geq 0}^{\infty} \frac{\partial K_{t+1+j}}{\partial T_{Lt}} \left(\lambda_{t+1} (F_K(\bar{K}) + 1 - \delta) - \lambda_t \right)$$

This term can be broken further into the (finite) portion before $t = \tau$ and the (infinite) portion for $t \geq \tau$.

$$\pi_L \sum_{j \geq \tau}^{\infty} C_{t+j}\gamma^j \left(\omega_{Lt}\frac{\partial \Gamma^y_{Lt+j}}{\partial T_{Lt}} + \omega_{Lt}\frac{\partial \Gamma^y_{Ht+j}}{\partial T_{Lt}} + \omega^o_{Lt}\frac{\partial \Gamma^o_{Lt+j}}{\partial T_{Lt}} + \omega^o_{Lt}\frac{\partial \Gamma^o_{Ht+j}}{\partial T_{Lt}}\right) +$$
$$\pi_L \sum_{j \geq 1}^{\tau-1} C_{t+j}\gamma^j \left(\omega_{Lt}\frac{\partial \Gamma^y_{Lt+j}}{\partial T_{Lt}} + \omega_{Lt}\frac{\partial \Gamma^y_{Ht+j}}{\partial T_{Lt}} + \omega^o_{Lt}\frac{\partial \Gamma^o_{Lt+j}}{\partial T_{Lt}} + \omega^o_{Lt}\frac{\partial \Gamma^o_{Ht+j}}{\partial T_{Lt}}\right) +$$
$$\left(\frac{\partial \Gamma^y_{Lt}}{\partial T_{Lt}} + \frac{\partial \Gamma^y_{Ht}}{\partial T_{Lt}}\right) = -\sum_{j \geq 0}^{\tau-1} \frac{\partial K_{t+1+j}}{\partial T_{Lt}}\left(\lambda_{t+1}(F_K(\bar{K}) + 1 - \delta) - \lambda_t\right)$$
$$-\sum_{j \geq \tau}^{\infty} \frac{\partial K_{t+1+j}}{\partial T_{Lt}}\left(\lambda_{t+1}(F_K(\bar{K}) + 1 - \delta) - \lambda_t\right)$$

Because after $t \geq \tau$, C_t, K_{t+1}, and λ_{t+t}/λ_t are all arbitrarily close to their steady state counterparts, the above can be written as:

$$\pi_L\omega_{Lt}\left[\sum_{j \geq 1}^{\infty} C\gamma^j \left(\frac{\partial \Gamma^y_{Lt+j}}{\partial T_{Lt}} + \frac{\partial \Gamma^y_{Ht+j}}{\partial T_{Lt}} + \frac{1/\gamma}{R(\bar{K})}\frac{\partial \Gamma^o_{Lt+j}}{\partial T_{Lt}} + \frac{1/\gamma}{R(\bar{K})}\frac{\partial \Gamma^o_{Ht+j}}{\partial T_{Lt}}\right) +\right.$$
$$\pi_L \sum_{j \geq 1}^{\tau-1} C_{t+j}\gamma^j \left(\omega_{Lt}\frac{\partial \Gamma^y_{Lt+j}}{\partial T_{Lt}} + \omega_{Lt}\frac{\partial \Gamma^y_{Ht+j}}{\partial T_{Lt}} + \omega^o_{Lt}\frac{\partial \Gamma^o_{Lt+j}}{\partial T_{Lt}} + \omega^o_{Lt}\frac{\partial \Gamma^o_{Ht+j}}{\partial T_{Lt}}\right) +$$
$$\left(\frac{\partial \Gamma^y_{Lt}}{\partial T_{Lt}} + \frac{\partial \Gamma^y_{Ht}}{\partial T_{Lt}}\right) = -\sum_{j \geq 0}^{\tau-1} \frac{\partial K_{t+1+j}}{\partial T_{Lt}}\left(\lambda_{t+1}(F_K(\bar{K}) + 1 - \delta) - \lambda_t\right)$$
$$-\lambda_{t+1}\sum_{j \geq 0}^{\infty} \frac{\partial K_{t+1+j}}{\partial T_{Lt}}\left((F_K(\bar{K}) + 1 - \delta) - \frac{1}{\gamma}\right)\gamma^j$$

Again, as $\gamma \to 1$, the right hand side becomes an (infinite large) positive number, while the left hand side is a negative number. Therefore, there exists a $\hat{\gamma} \in (0, 1)$ such that if $\gamma > \hat{\gamma}$ implementing the first best level of inequality violates the planner's first order conditions.

Case 2 ($\psi_a > 0$): The optimal implementable allocation x^*_I maximizes the following expression:

$$\max_{\{C_t\}_{t \geq 0}, T_L} \sum_I \lambda_i \sum_{t=0}^{\infty} \gamma^t \left(\frac{(\Gamma^y_{it}C_t)^{1-\sigma_y}}{1-\sigma_y} + \gamma^{-1}\beta_i\frac{(\Gamma^o_{it}C_t)^{1-\sigma_o}}{1-\sigma_o} + \beta_i\gamma^{-1}\psi_a\frac{(a^o_i)^{1-\eta}}{1-\eta}\right)$$

subject to:

$$\begin{aligned}
\Gamma^y_{it} &= \Gamma^y_{it}(K_t, K_{t+1}, T_{it}, C_t, a^o_{t-1}) \text{ for } i \in \{L, H\} \\
\Gamma^o_{it+1} &= \Gamma^y_{it}(K_t, K_{t+1}, T_{it}, C_{t+1}, a^o_{t-1}) \text{ for } i \in \{L, H\} \\
K_{t+1} &= K_{t+1}(K_t, T_{Lt}, \{a^o_{it-1}, a^o_{it-2\}_{i\in I}}) \\
C_t + K_{t+1} &= F(K_t, 1) + (1-\delta)K_t \\
a^o_{it+1} &= a^o_{it+1}(K_{t+1}, K_t, a^o_{it-1}, T_{it})
\end{aligned}$$

Let μ_t be the lagrange multiplier for the resource constraint. The first order condition with respect to C_t is:

$$\sum_I \left(\omega_{it}(\Gamma^y_{it}) + \frac{\omega_{it-1}}{R_t}(\Gamma^o_{it}) \right) = \mu_t \tag{A.15}$$

The first order condition with respect to T_{Lt} is:

$$\sum_{j\geq 0}^{\infty} C_t \Bigg(\omega_{Lt+j} \frac{\partial \Gamma^y_{Lt+j}}{\partial T_{Lt}} + \omega_{Ht+j} \frac{\partial \Gamma^y_{Ht+j}}{\partial T_{Lt}} + \omega^o_{Lt+j} \frac{\partial \Gamma^o_{Lt+j}}{\partial T_{Lt}} + \omega^o_{Ht+j} \frac{\partial \Gamma^o_{Ht+j}}{\partial T_{Lt}} = $$
$$+\omega^o_{Lt+j} \frac{\partial a^o_{Lt+j}}{\partial T_{Lt}} + \omega^o_{Ht+j} \frac{\partial a^o_{Ht+j}}{\partial T_{Lt}} \Bigg) - \sum_{t\geq 0}^{\infty} \frac{\partial K_{t+1}}{\partial T_{Lt}} \left((F_K(K_{t+j}) + 1 - \delta) - \frac{1}{\gamma} \right) \mu_{t+1} \tag{A.16}$$

Combining equations A.15 and A.16:

$$\sum_{j\geq 0}^{\infty} C_{t+j} \left(\omega_{Lt+j} \frac{\partial \Gamma^y_{Lt+j}}{\partial T_{Lt}} + \omega_{Ht+j} \frac{\partial \Gamma^y_{Ht+j}}{\partial T_{Lt}} + \omega^o_{Lt+j} \frac{\partial \Gamma^o_{Lt+j}}{\partial T_{Lt}} + \omega^o_{Ht+j} \frac{\partial \Gamma^o_{Ht+j}}{\partial T_{Lt}} + \omega^o_{Lt+j} \frac{\partial a^o_{Lt+j}}{\partial T_{Lt}} + \omega^o_{Ht+j} \frac{\partial a^o_{Ht+j}}{\partial T_{Lt}} \right)$$
$$= -\sum_{t\geq 0}^{\infty} \frac{\partial K_{t+j+1}}{\partial T_{Lt}} \left((F_K(K_{t+j}) + 1 - \delta) - \frac{1}{\gamma} \right) \left(\sum_I \omega_{it+j}(\Gamma^y_{it+j}) + \frac{\omega_{it+j-1}}{R_{t+j}}(\Gamma^o_{it+j}) \right)$$

Again, suppose $K_0 = \bar{K}$ and consider a solution where $\omega_{Lt+j} = \omega_{Ht+j} = \omega_L$ for all $j \geq 0$. Recall that $\omega^o_{it} = \frac{\omega^y_{it-1}}{R_t}$. Plugging this into the above, pulling out and canceling the ω_L gives you:

$$\sum_{j\geq 0}^{\infty} \bar{C}\gamma^j \left(\frac{\partial \Gamma^y_{Lt+j}}{\partial T_{Lt}} + \frac{\partial \Gamma^y_{Ht+j}}{\partial T_{Lt}} + \frac{1/\gamma}{R(\bar{K})} \left(\frac{\partial \Gamma^o_{Lt+j}}{\partial T_{Lt}} + \frac{\partial \Gamma^o_{Ht+j}}{\partial T_{Lt}} + \frac{\partial a^o_{Lt+j}}{\partial T_{Lt}} + \frac{\partial a^o_{Ht+j}}{\partial T_{Lt}} \right) \right)$$
$$= -\sum_{t\geq 0}^{\infty} \frac{\partial K_{t+j+1}}{\partial T_{Lt}} \left((F_K(K_{t+j}) + 1 - \delta) - \frac{1}{\gamma} \right) \gamma^j \left(\Gamma^y_{Lt+j} + \Gamma^y_{Ht+j} + \frac{1/\gamma}{R(\bar{K})}(\Gamma^o_{Lt+j} + \Gamma^o_{Ht+j}) \right)$$

This can be rewritten as:

$$\sum_{j\geq 1}^{\infty}\bar{C}\gamma^j\left(\frac{\partial\Gamma^y_{Lt+j}}{\partial T_{Lt}}+\frac{\partial\Gamma^y_{Ht+j}}{\partial T_{Lt}}+\frac{1/\gamma}{R(\bar{K})}\left(\frac{\partial\Gamma^o_{Lt+j}}{\partial T_{Lt}}+\frac{\partial\Gamma^o_{Ht+j}}{\partial T_{Lt}}+\frac{\partial a^o_{Lt+j}}{\partial T_{Lt}}+\frac{\partial a^o_{Ht+j}}{\partial T_{Lt}}\right)\right)$$

$$=-\sum_{t\geq 0}^{\infty}\frac{\partial K_{t+j+1}}{\partial T_{Lt}}\left((F_K(K_{t+j})+1-\delta)-\frac{1}{\gamma}\right)\gamma^j\left(\Gamma^y_{Lt+j}+\Gamma^y_{Ht+j}+\frac{1/\gamma}{R(\bar{K})}(\Gamma^o_{Lt+j}+\Gamma^o_{Ht+j})\right)$$

$$+\bar{C}\left(\frac{\partial\Gamma^y_{Lt}}{\partial T_{Lt}}+\frac{\partial\Gamma^y_{Ht}}{\partial T_{Lt}}+\frac{1/\gamma}{R(\bar{K})}\left(\frac{\partial\Gamma^o_{Lt}}{\partial T_{Lt}}+\frac{\partial\Gamma^o_{Ht}}{\partial T_{Lt}}+\frac{\partial a^o_{Lt}}{\partial T_{Lt}}+\frac{\partial a^o_{Ht}}{\partial T_{Lt}}\right)\right)$$

Note that by assumption $\frac{1/\gamma}{R(\bar{K})}<1$. Note also that T_{Lt} only affects the consumption share and bequests of generation t directly, but changes the consumption share and bequests of every subsequent generation by decreasing the capital stock and increasing the interest rate. Therefore, the first term above is negative. To see why, note that a lower capital stock at each horizon decreases income and therefore decreases bequests left for both low and high type (that is $\frac{\partial a^o_{it+j}}{\partial T_{Lt}}<0$). To see that the remaining part of this term is also negative, note that a decline in K_{t+1} increases R_{t+1}, increasing the consumption share when old, but because $\frac{1}{\gamma}/R(\bar{K})<1$ by assumption, the decline in the consumption share of the young is weighted more heavily. As before, the right hand side is positive, as $\mu_t>0$, $\frac{-\partial K_{t+1+j}}{\partial T_{Lt}}>0$, and $F_K(\bar{K})+1-\delta-\frac{1}{\gamma}$ is positive by assumption.

Therefore, as $\gamma\to 1$, the first term gets larger in terms of absolute value and approaches $-\infty$. Therefore, for sufficiently large $\hat{\gamma}$, the first order condition is not satisfied, as the right hand side becomes an increasingly large positive number, while the left hand side becomes an increasingly large negative number.

Suppose $K_0\neq\bar{K}$. By an identical argument as in the $\psi_a=0$ case, because the policy that would implement ideal equality, $\{T^*_{Lt}\}_{t\geq 0}$ eventually converges to $\bar{T}^*_L$, as $\gamma\to 1$, the planner's first order condition would have an ever large negative number on the left hand side and an ever larger positive number on the right hand side. Therefore, $\omega^*_{Lt}=\omega^*_{Ht}$ for all t is ruled out as a solution in this case as well.

A.6 Derivation of Unconstrained First Best

Define the unconstrained first-best allocation, $x^*\equiv\arg\max_{x\in\chi}\text{SW(x)}$

Let $\{c^{y*}_{it},c^{o*}_{it}\}_{i\in I,t\geq 0}$ be the sequence of consumption levels for the young and old associated with x^*. Define the first best social welfare weights, $\omega^{y*}_{it}=\pi_i\lambda_i(c^{y*}_{it})^{-\sigma_y}$ and $\omega^{o*}_{it}=\pi_i\lambda_i\beta_i(c^{o*}_{it})^{-\sigma_o}$. At the unconstrained first-best allocation, $\omega^{y*}_{Lt}=\omega^{y*}_{Ht}$ and $\omega^{o*}_{Lt}=\omega^{o*}_{Ht}$ for all $t\geq 0$.

At the unconstrained first-best allocation, social welfare weights are equal across both types at all points in time. Intuitively, if we supposed that the Pareto weights for each type were equal, this would imply a perfectly equal allocation of consumption when young and consumption when old across the two household types.

Proof: To find the first-best allocation, I solve the problem of benevolent social planner who discounts future *generations* at rate γ, and puts weight λ_i on the utility of type-i households. The planner aims to choose the allocation that maximizes the discounted infinite sum of the utility of all generations:

$$S(x) = \sum_I \frac{\pi_i}{2}\lambda_i\Bigg(\sum_{t=0}^{\infty}\gamma^t\Bigg(\frac{(c_t^{iy})^{1-\sigma_y}}{1-\sigma_y} + \beta\frac{(c_{t+1}^{i,o})^{1-\sigma_o}}{1-\sigma_o}\Bigg) + \beta\frac{(c_0^{i,o})^{1-\sigma_o}}{1-\sigma_o}\Bigg) \tag{A.17}$$

The unconstrained planner picks an allocation, $\{c_t^{it}\}, K_t$ that maximizes S(x) subject only to the resource constraint (A.18).

$$\sum_i \frac{\pi_i}{2}(c_t^{iy} + c_t^{io}) + K_{t+1} = (1-\delta)K_t + F(K_t, 1) \tag{A.18}$$

Let Γ_t^{iy} be the share of the total consumption of the young at time t, c_t^y consumed by type i households so that $c_t^{iy} = \Gamma_t^{iy}c_t^y$. Let Γ_o^{it} be define analogously. Define $\Phi_t^y = \sum_I \lambda_i(\Gamma_y^{it})^{1-\sigma_y}$ and $\Phi_t^o = \sum_I \lambda_i(\Gamma_o^{it})^{1-\sigma_o}$ Then (X) can be rewritten as

$$\frac{1}{4}\sum_t \gamma^t\Bigg(\Phi_t^y\frac{(c_t^y)^{1-\sigma_y}}{1-\sigma_y} + \beta\Phi_{t+1}^o\frac{(c_{t+1}^o)^{1-\sigma_o}}{1-\sigma_o}\Bigg) + \frac{1}{2}\beta\Phi_t^o\frac{(c_0^o)^{1-\sigma_o}}{1-\sigma_o}$$

The planner's FOC with respect to c_t^y, c_t^o, and K_{t+1} are

$$\gamma\frac{\phi_t^y}{\phi_t^o}(c_t^y)^{-\sigma_y} = \beta(c_t^o)^{-\sigma_o} \tag{A.19}$$

$$\frac{\phi_t^y}{\phi_{t+1}^o}(c_t^y)^{-\sigma_y} = \beta(c_{t+1}^o)^{-\sigma_o}(F_K(K_{t+1}, 1) + (1-\delta)) \tag{A.20}$$

$$\frac{1}{2}(c_t^y + c_t^o) + K_{t+1} = F(K_t, 1) + (1-\delta)K_t \tag{A.21}$$

The planner's first order condition with respect to c_0^o is given by:

$$\beta\Phi_0^o(c_0^o)^{-\sigma_o} = \gamma\Phi_0^y(c_0^y)^{-\sigma_y} \tag{A.22}$$

The planner's first order conditions with respect to Γ_t^{iy} and Γ_t^{io} imply that the optimal allocation satisfies (A.23) and (A.24).

$$c_t^{Ly} = c_t^{Hy}\left(\frac{\lambda_L}{\lambda_H}\right)^{\frac{1}{\sigma_y}} \tag{A.23}$$

$$c_t^{Lo} = c_t^{Ho}\left(\frac{\lambda_L}{\lambda_H}\right)^{\frac{1}{\sigma_o}} \tag{A.24}$$

Therefore, Γ^{iy}, Γ^{io}, Φ^{y}, and Φ^{o} are also constant, and an optimal steady state solution exists. Taking the steady state version of the above equations and combining them together shows that the optimal long-run capital stock corresponds to the well known 'modified golden rule' level.

$$\frac{1}{\gamma} = F_K(K^*, 1) + 1 - \delta \tag{A.25}$$

Equations (A.19)-(A.24) characterize the unconstrained planner's first best allocation, x^*. Because $\bar{K} \geq K^*$, the economy with $\tau_K = \tau_K^* = 0$ and $\lambda^m = \lambda^{m*} = \frac{\lambda_H}{\lambda_L}$ is currently over-investing in the long-run relative to the first best steady state level. Note that the optimal path of capital is **monotonically converging** to the steady state level.

To see this, suppose that $K_t, K_{t+1} < K^*$, but that $K_t > K_{t+1}$. Using equations (7) and (8) and the fact that $F_{K_t} - \delta > 0$ when $K_t < K^*$ we see that $c_t^0 < c_{t+1}^y$ and $c_t^o < c_{t+1}^o$ along the optimal path. This means that $c_t = \frac{1}{2}(c_t^y + c_t^o) < c_{t+1}$. Using the resource constraint, this implies that $F(K_{t+1}) + (1-\delta)K_{t+1} - K_{t+2} > F(K_t) + (1-\delta)K_t - K_{t+1}$. Because I've assumed $K_t > K_{t+1}$ this implies $K_{t+2} < K_{t+1}$. Applying this logic forward, this implies $K_{t+j+1} > K_{t+j}$ for all j, meaning K_t never increases, a contradiction.

A.6.1 Proof of Lemma 3 (i)

The steady state equilibrium conditions when $\tau_K = 0$ are the following.

$$\begin{aligned} \frac{1}{2}\sum T_i + \frac{1}{2}\sum T_a &= rB \\ (\theta_L w(K_{ss}) - a_L - T_L - T_y)^{-\sigma_y} &= \beta R(K_{ss})(\theta_L w(K_{ss}) - T_L - T_o + a_L R(K_{ss}))^{-\sigma_o} \\ (\theta_H w(K_{ss}) - a_H - T_H - T_y)^{-\sigma_y} &= \beta R(K_{ss})(\theta_H w(K_t) - T_H - T_o + a_H R(K_{ss}))^{-\sigma_o} \\ K_{ss} + B &= \frac{1}{2}a_L + a_H \end{aligned}$$

To keep T_H, T_H, and B constant, increasing T_o implies reducing T_y in order to satisfy the government budget constraint. In this case, the only way for the Euler equations to hold is for a_L and a_H to decline.

A.6.2 Proof of Lemma 3 (ii)

To be added

A.7 Proof of Proposition 3

Proposition 3 states that when the steady state associated with $\tau_{Kt} = 0$, $B_t = \bar{B}$ and $T_{yt} = T_{ot} = 0$ is dynamically efficient then the optimal policy implements a higher than first-best level of inequality. I first show that when the steady state is dynamically *inefficient*, the planner has all the tools necessary to implement first best. Then I show that when $F_K(\bar{K}) > \delta$, for sufficiently high γ the planner would never choose to increase debt or inter-generational transfers, and therefore the solution is identical to that in the previous section.

A.7.1 Proof that $x_I^* = x^*$ when $F_K(\bar{K}) + 1 - \delta < \frac{1}{\gamma}$

The first best allocation, x^* is characterized in section X.

Show that x^* can be implemented. To show that this allocation can be implemented as a non-binding equilibrium, one can simply (1) solve for a set of prices, assets, and tax instruments that implement it, and (2) show that these tax instruments do not violate the political constraint whenever $\frac{1}{\gamma} < F_K(K^s) + 1 - \delta$.

Take the first best allocation, x^*. Set $\tau_K = 0$ and set $w_t = F_L(K_t^*)$ and $R_t = (F_K(K_{t+1}^* + 1 - \delta)$ for all t. Note that we've assumed that $B_{-1}, a_{L,-1}, a_{H,-1}, R_0$ and K_0 are all exogenously given. Starting in period 0, use $R_0 a_{i0}$ along with $w_0 = F_L(K_0)$ and c_{i0}^{o*} to solve for T_{L0} and T_{H0} as functions of T_{o0} using the budget constraint of the initial old for both types. Plugging these functions into the initial young's budget constraint, along with $w_0(K_0)$ and c_{i0}^{y*} gives you a_{L0} and a_{H0} as functions of both T_{o0} and T_{y0}. Arbitrarily set $B_t = 0$ for all periods. Use the asset market clearing condition (5) to write T_{y0} as a function of T_{o0}. Finally use the government's budget - now in terms only of T_{o0} to solve for T_{o0}. This in turn pins down a_{L0} and a_{H0}. Given these asset levels, the process above can be repeated in all periods.

To see that the implied T_{yt} and T_{ot} never violate the political constraint () - i.e. that implementing x^* requires a policy in which $T_y > T_o$ for all t - suppose that there existed at least one period, t' in which $T_{yt'} < T_{ot'}$. Above I established that the planner's optimal solution, K_t converges monotonically to the steady state level, K^*. From Lemma 2, we know that given $\tau_K = 0$ and $\{T^h, T^\ell\}$ corresponding to λ^{m*}, $K_{t+1}(T_{ot'}, K_t) > K_t$ because $K_t < K_{ss}(T_{ot'})$, the steady state capital level associated with $\{T_{ot'}, T_{yt'}\}$. This violates monotonicity, which presents a contradiction.

Therefore, the optimal set of lump-sum taxes derived above will always satisfy the political constraint that $T_{yt} \geq T_{ot}$.

A.8 Proof of Proposition 4

The model is identical to the one in Section 2 with $\psi_a = 0$ and $\theta_i^o = 0$ for both types, except for the households' first order conditions are now:

$$f_\ell(\ell_i^y) = u_c^y(c_i^y)(1-\tau_\ell)\theta_{iy}w$$
$$u_c^y(c_i^y) = \beta R u_c^o(c_i^o)$$

Assuming generations are equally sized, the total change in social welfare:

$$\sum_I \lambda^i \pi_i \Bigg(u_c^y(c_i^h)\Bigg(d((1-\tau_\ell)\ell_i^y\theta_i^y w) + dT - da_i^y \Bigg) - f_\ell(\ell_i^h) \Bigg) + \beta u_c^o(c_i^o)\Bigg(R da_i^y + dR a_i^y \Bigg)$$

Expanding the first term and referring to $\theta_i^y = \theta_i$ and $\ell_i^y = \ell_i$:

$$d((1-\tau_\ell)\ell_i\theta_i w) = -\ell_i\theta_i w d\tau_\ell + \theta_i\ell_i(1-\tau_\ell)dw + \theta_i w(1-\tau_\ell)d\ell_i$$

Expanding the dT term:

$$dT = \tau_\ell L dw + \tau_\ell w dL + d\tau_\ell w L$$

Define $\omega_{ih} = \lambda_i\beta^{h-1}u_c^h(c_i^h)$, where $\sum_I\sum_H \omega_{ih} = 1$ and subbing in the labor supply condition and Euler equation, the total change becomes:

$$dSW = \sum_I \pi_i\omega_{iy}\Bigg(L - \ell_i\theta_i\Bigg)w d\tau_\ell + \sum_I \pi_i\Bigg(\omega_{iy}(\theta_i\ell_i(1-\tau_\ell)dw) + \omega_{io}a_i^y dR\Bigg) + \tau_\ell d(wL)$$

Defining Θ as:

$$\Theta = L^{-1}\sum_I \pi_i\omega_{ih}\Bigg(\theta_i\ell_i^h(1-\tau_\ell) + a_i^{h-1}\frac{\partial R}{\partial w}\Bigg)$$

Let $\omega_i = \omega_{iy}$. Recall that $\omega_{ih+1}/\omega_{ih} = \beta u_c^{h+1}(c_i^{h+1})/u_c^h(c_i^h) = R$, so Θ becomes:

$$\Theta = \sum_I \omega_i \pi_i \left(\theta_i \frac{\ell_i}{L}(1-\tau_\ell) + \frac{R a_i^y}{L}\frac{\partial R}{\partial w} \right)$$

Because we've assumed that households only earn and save in the first period of life, the change in social welfare is:

$$dSW = \underbrace{\sum_I \omega_i \left(L - \ell_i^h \theta_i \right) w d\tau_\ell}_{\text{Direct Effects}} + \underbrace{L\Theta dw + \tau_\ell d(wL)}_{\text{General Equilibrium Costs}}$$

Using a similar procedure as in Appendix A.3, I can solve for the general equilibrium change in steady state capital.

I use the government's budget constraint and the household's intra-temporal condition at each age to write household labor as a function of labor, assets, capital, and tax policy, $\ell_i^h(K, L, \tau_L, a_i^y)$.

$$f_\ell(\ell_i) = w(K,L)(1-\tau_L)\theta_i u_c(\theta_i w(K,L)\ell_i(1-\tau_\ell) + T(\tau_\ell) - a_i^y)$$

The firms' first order conditions now are:

$$w = F_L(K,L)$$
$$r + \delta = F_K(K,L)$$

The labor market clearing condition is:

$$L = \sum_I \pi_i \ell_i(w(K,L), \tau_L, T(\tau_\ell), a_i^y) \tag{A.26}$$

I use the household Euler equation and the firm's first order conditions to write consumption at each age, c_i^h as a function of R and of permanent income, PI_i, which is given by:

$$PI_i = (1-\tau_\ell)\theta_i w(K,L)\ell_i + T(\tau_\ell)$$

The households' budget constraints and the firm's first order conditions with respect to

R then allow me to write saving, a_i^y as a function of permanent income.

$$a_i^y = (1-\tau_\ell)\theta_i w(K,L)\ell_i + T(\tau_\ell wL) - c_i^y(R(K), PI_i)$$

From this, we can use A.26 to define aggregate L as a function of K, and τ_ℓ. Combining the

Using the the same procedure as in the fixed labor case, we can write the asset market clearing condition as a function of τ_ℓ and K.

$$K = \sum_I \pi_{iy} a_i^y(R(K), (1-\tau_\ell)\theta_i w(K, L(K,\tau_\ell))\ell_i + T(\tau_\ell w(K, L(\tau_\ell, K))L(K,\tau_\ell))$$

Assume π_{ih} is constant across types and ages and taking the total derivative of the above gives:

$$dK = \sum_I \pi_{iy}\Bigg[\frac{\partial a_i^y}{\partial R}\frac{\partial R}{\partial K}dK + \frac{\partial a_i^y}{\partial PI_i}\left(\frac{\partial PI_i}{\partial \tau_\ell}d\tau_\ell + \frac{\partial PI_i}{\partial w}\frac{\partial w}{\partial K}dK + \frac{\partial PI_i}{\partial w}\frac{\partial w}{\partial L}\left(\frac{\partial L}{\partial \tau_\ell}d\tau_\ell + \frac{\partial L}{\partial K}dK\right)$$
$$+\frac{\partial PI_i}{\partial \ell_i}(\frac{\partial \ell_i}{\partial \tau_\ell}d\tau_\ell + \frac{\partial \ell_i}{\partial K}dK) + \frac{\partial PI_i}{\partial T}\left(\frac{\partial T}{\partial \tau_\ell}d\tau_\ell + (\frac{\partial T}{\partial w}\frac{\partial w}{\partial L} + \frac{\partial T}{\partial L})\left(\frac{\partial L}{\partial \tau_\ell}d\tau_\ell + \frac{\partial L}{\partial K}dK\right) + \frac{\partial T}{\partial w}\frac{\partial w}{\partial K}dK\right)\Bigg]$$

Define the average derivative of savings to the interest rate (including substitution, income, and permanent income effects) as:

$$\frac{\partial A}{\partial R} = \sum_I \pi_{iy}\left(\frac{\partial a_i^y}{\partial R}\right)$$

Define the total average derivative of savings to the aggregate wage level:

$$\sum_I \pi_{iy}\left(\frac{\partial a_i^y}{\partial PI_i}\frac{\partial PI_i}{\partial w}\right) = M\bar{P}S(1-\tau_\ell)L + \mathcal{C}_1$$

Where $M\bar{P}S = \left(\sum_I \pi_{iy}\frac{\partial a_i^y}{\partial PI_i}\right)$ and $(1-\tau_\ell)L = \left(\sum_I \pi_{iy}(1-\tau_\ell)\theta_i\ell_i\right)$ and $\mathcal{C}_1 = \mathrm{Cov}(MPS, (1-\tau_\ell\theta_i\ell_i))$

We have that:

$$\sum_I \pi_{iy}\left(\frac{\partial a_i^y}{\partial PI_i}\frac{\partial PI_i}{\partial T}\frac{\partial T}{\partial w}\right) = M\bar{P}S\tau_\ell L$$

Similarly, we also have that:

$$\sum_I \pi_{iy}\left(\frac{\partial a_i^y}{\partial PI_i}\frac{\partial PI_i}{\partial \ell_i}\frac{\partial \ell_i}{\partial K}\right) = (1-\tau_\ell)w(M\bar{P}S\frac{\partial L}{\partial K}) + \mathcal{C}_2$$

$$\frac{\partial L}{\partial K} = \sum_I \pi_{iy}\frac{\theta_i \partial \ell_i}{\partial K} \text{ and } \mathcal{C}_2 = (1-\tau_\ell w)\text{Cov}(MPS, \theta_i\frac{\partial \ell_i}{\partial K})$$

We also have that:

$$\sum_I \pi_{iy}\left(\frac{\partial a_i^y}{\partial PI_i}\frac{\partial PI_i}{\partial T}\frac{\partial T}{\partial L}\right) = M\bar{P}S\tau_\ell w$$

Define the total derivative of the wage with respect to capital, $\frac{dw}{dK}$:

$$\frac{dw}{dK} = \frac{\partial w}{\partial K} + \frac{\partial w}{\partial L}\frac{\partial L}{\partial K}$$

Using all of these definitions and combining the 'dK' and $d\tau_\ell$ terms together:

$$dK\left(1 - \frac{\partial A}{\partial R}\frac{\partial R}{\partial K} - M\bar{P}S\left(L\frac{dw}{dK} + w\frac{\partial L}{\partial K}\right) - \mathcal{C}_1 - \mathcal{C}_2\right) = \sum_I \pi_{iy}\frac{\partial a_i^y}{\partial PI_i}(wL - w\theta_i\ell_i)d\tau_\ell + \sum_I \pi_{iy}\frac{\partial a_i^y}{\partial PI_i}\left(\left(\frac{\partial PI_i}{\partial \ell_i}\frac{\partial \ell_i}{\partial \tau_\ell}\right) + \left(\frac{\partial T}{\partial w}\frac{\partial w}{\partial L} + \frac{\partial T}{\partial L}\right)\frac{\partial L}{\partial \tau_\ell}\right)d\tau_\ell$$

Define the following term:

$$\mathcal{K}_1 = \sum_I \pi_{iy}\frac{\partial a_i^y}{\partial PI_i}\left(\left(\frac{\partial PI_i}{\partial \ell_i}\frac{\partial \ell_i}{\partial \tau_\ell}\right) + \left(\frac{\partial T}{\partial w}\frac{\partial w}{\partial L} + \frac{\partial T}{\partial L}\right)\frac{\partial L}{\partial \tau_\ell}\right)\frac{d\tau_\ell}{K} * \left(\frac{K_R}{K_R(1 - A_{wL}wL_K - \mathcal{C}_1 - \mathcal{C}_2) - A_R}\right)$$

Note that $\left(L\frac{dw}{dK} + w\frac{\partial L}{\partial K}\right) = \frac{\partial Lw}{\partial K}$ and $M\bar{P}S = \frac{\partial A}{\partial_{wL}}$. Define $X_Y \equiv \partial log(X)/\partial log(Y)$ as in the text. Isolating $\frac{dK}{K}$:

$$\frac{dK}{K} = \left(\sum_I \pi_{iy}\frac{\partial a_i^y}{\partial PI_i}\frac{(wL - w\theta_i\ell_i)}{K}d\tau_\ell\right)\left(\frac{K_R}{K_R(1 - A_{wL}wL_K - \mathcal{C}_1 - \mathcal{C}_2) - A_R}\right) + \mathcal{K}_1$$

As in the text, define K_{PI} as:

$$K_{PI} = \Bigg(\sum_I \pi_{iy}\frac{\partial a_i^y}{\partial PI_i}\frac{(wL - w\theta_i\ell_i)}{K}d\tau_\ell\Bigg)\Bigg(\frac{K_R}{K_R(1 - A_{wL}wL_K - \mathcal{C}_1 - \mathcal{C}_2) - A_R}\Bigg)$$

The total welfare effect is the direct effect plus:

$$(\frac{\partial L}{\partial \tau_\ell L}d\tau_\ell + \frac{\partial L}{\partial KL}dK)\Bigg(\tau_\ell wL + w(\Theta L + \tau_\ell L)w_L\Bigg) + w(\tau_\ell L + L\Theta)w_K\frac{dK}{Kd\tau_\ell}d\tau_\ell$$

Which can be simplified to,

$$wL(\tau_\ell + \Theta)\Bigg(w_K K_{PI} + (w_L + \frac{\tau_\ell}{\Theta + \tau_\ell})L_{\tau_\ell} + \mathcal{K} + \mathcal{L}\Bigg)$$

Here $\mathcal{K} = w_K\mathcal{K}_1$ capture the effects of labor distortions of aggregate capital. This term is defined as in equation (A.27).

$$\mathcal{K} \equiv w_L\sum_I \pi_{iy}\frac{\partial a_i^y}{\partial PI_i}\Bigg(\Bigg(\frac{\partial PI_i}{\partial \ell_i}\frac{\partial \ell_i}{\partial \tau_\ell}\Bigg) + \Bigg(\frac{\partial T}{\partial w}\frac{\partial w}{\partial L} + \frac{\partial T}{\partial L}\Bigg)\frac{\partial L}{\partial \tau_\ell}\Bigg)\frac{d\tau_\ell}{K} * \Bigg(\frac{K_R}{K_R(1 - A_{wL}wL_K - \mathcal{C}_1 - \mathcal{C}_2) - A_R}\Bigg) \quad \text{(A.27)}$$

The term $\mathcal{L}$ summarizes the effect of distorted labor supply on capital through its affect on household savings. Again, $\frac{\partial L}{\partial K}$ is implicitly defined by equation (A.26).

$$\mathcal{L} \equiv (w_L + \frac{\tau_\ell}{\tau_\ell + \Theta})\frac{\partial LK}{\partial KL}\frac{dK}{Kd\tau_\ell}d\tau_\ell \quad \text{(A.28)}$$

A.9 Derivation of Sufficient Statistic Formula

From the previous section, we saw that the semi-elasticity of K with respect to the direct change in permanent income was:

$$K_{PI} = \Bigg(\sum_I \pi_{iy}\frac{\partial a_i^y}{\partial PI_i}\frac{(wL - w\theta_i\ell_i)}{K}d\tau_\ell\Bigg)\Bigg(\frac{K_R}{K_R(1 - A_{wL}wL_K - \mathcal{C}_1 - \mathcal{C}_2) - A_R}\Bigg)$$

Here that $A_{wL}wL_K$ was defined as:

$$\begin{aligned} A_{wL}wL_K &\equiv \frac{\partial A}{\partial wL}\frac{wL}{wL}\frac{\partial wL}{\partial K} = \frac{\partial A}{\partial wL}\Bigg(L(\frac{\partial w}{\partial K} + \frac{\partial w}{\partial L}\frac{\partial L}{\partial K}) + w\frac{\partial L}{\partial K}\Bigg) \\ &= \frac{\partial A}{\partial wL}\frac{wL}{w}\frac{\partial w}{\partial K}\frac{K}{K} + \frac{\partial A}{\partial wL}\Bigg(\frac{wL}{w}\frac{\partial w}{\partial L}\frac{\partial L}{\partial K}\frac{K}{K} + w\frac{L}{L}\frac{\partial L}{\partial K}\frac{K}{K}\Bigg) \end{aligned}$$

Using the fact that $A = K$,

$$= A_{wL}(w_K + (w_L + 1)L_K)$$

Using the results from Appendix X, we know that $w_L < 0$ but $|w_L| < 1$, meaning $(1 + w_L)L_K > 0$. The bigger this value, the smaller the denominator of K_{PI}, and the larger K_{PI}. Therefore, setting this term equal to 0 generates a lower bound estimate of K_{PI}.

Recall from the previous section that $\mathcal{C}_1$ is the covariance between marginal propensities to save and after-tax labor income, while $\mathcal{C}_2$ is the covariance between MPS and the response of labor supply to an increase in capital. These two terms are almost certainly positive. I provide empirical evidence for the first as does Dynan et al. (2004). That capital improves labor income dis-proportionately for high income earners has been shown by Krusell et al. (2000). Therefore, setting these two covariances to 0 generates a lower-bound estimate of K_{PI}.

Finally, it is simpler to normalize the change in labor income by average labor income. Therefore, I multiply the expression for K_{PI} by $\frac{wL}{wL}$ to get the expression in the text.

$$K_{PI} = \left(\sum_I \pi_{iy} \frac{\partial a_i^y}{\partial PI_i} \frac{(wL - w\theta_i \ell_i)}{wL} d\tau_\ell \right) \left(\frac{K_R}{K_R(1 - A_{wL} w_K) - A_R} \right) \frac{wL}{K}$$

A.10 Elasticities in CES production company

Assume $\delta = 0$. The firm's first order condition is:

$$\alpha \left(\frac{K}{Y} \right)^{\frac{-1}{\rho}} = R - 1 \rightarrow \log(\alpha) - \frac{1}{\rho}\log(K) + \frac{1}{\rho}\log(Y) = \log(r)$$

Then $\frac{\partial \log(K)}{\partial \log(r)} = \left(-\frac{1}{\rho}(1 - \partial \log(Y)/\partial \log(K)) \right)^{-1} = \left(-\frac{1}{\rho}(1 - \alpha_K) \right)^{-1} = -\frac{\rho}{\alpha_L}$

Similarly,

$$w_K = \frac{\partial \log(w)}{\partial \log(K)} = \frac{(1 - \alpha_L)}{\rho} \quad w_L = \frac{\partial \log(w)}{\partial \log(L)} = \frac{-(1 - \alpha_L)}{\rho}$$

A.11 Extension with large open economy

In this case, the asset market clearing condition is:

$$K = A + NFA$$

Taking the total derivative:

$$dK = dNFA + dA$$

Proceeding as in the baseline case, this is approximately equal to:

$$dR = \Delta_{NH}\frac{K}{R}\left(K_R(1 - A_w w_R) - \frac{A}{K}A_R - \frac{NFA}{K}NFA_R\right)$$

Finally, multiplying by $\frac{\partial K}{\partial R}$ gives the final expression.

$$dK = K_R\Delta_{NH}\left(K_R(1 - A_w w_R) - \frac{A}{K}A_R - \frac{NFA}{K}NFA_R\right)^{-1}$$

A.12 Extension with H generations

The proof is identical to the $H = 2$ case except that total steady state assets, A are now given by:

$$A = \sum_I \sum_H \pi_{ih} a_i^h$$

The semi-elasticity of K with respect to a change in ther perment income distribution is now:

$$K_{PI} = \frac{\Omega}{K}\sum_{I,H}\frac{\partial a_i^h}{\partial PI_i}\left(\sum_H (wL - \theta_{ih}\ell_i^h w)\frac{1}{R^{h-1}}d\tau_\ell\right)$$

To take this to the data, I need an estimate of discounted permanent income for each quintile. A household's permanent income in this case is defined as the discounted weighted sum of their permanent income each period.

$$PI_{ij} = \sum_H \theta_i^h \ell_i^h w \frac{1}{R^h}$$

Therefore, the most straightforward approach is to choose an appropriate discount rate R, and to estimate PI_{ij} as:

$$\hat{PI}_{ij} = \sum_H \hat{PI}_{hijt} \frac{1}{R^h}$$

Due to data limitations, I split the data into 4 equally sized age groups, each spanning 10 years. I use 1.06 as my annual discount rate. Therefore, the appropriate discount rate in the formula is given by the following expression.

$$R = \frac{1}{10} \sum_{k=1}^{10} (1.06)^k$$

A.13 Extension with balanced growth path.

First, divide the young type-i's household budget constraint by the scaling factor, Z_t. Note I define $T_{it} = w_t L_t - w_t \theta_i \ell_{it}$

$$\frac{c_i^y}{Z_t} = \frac{\theta_i F_L(K_t)}{Z_t} + \frac{T_i}{Z_t} - \frac{a_i}{Z_t}$$

Doing the same for the type-i old households and plugging this expression into the household Euler equation gives you:

$$\left(\frac{\theta_i F_L(K_t)}{Z_t} + \frac{T_{it}}{Z_t} - \frac{a_i}{Z_t} \right)^{-\sigma_y} = \beta_i (F_K(K_{t+1}) + 1 - \delta) \left(\frac{a_{it} F_K(K_{t+1}) + 1 - \delta}{Z_{t+1}} \right)^{-\sigma_o}$$

Using the fact that $Z_{t+1} = (1+g) Z_t$:

$$\left(\frac{\theta_i F_L(K_t)}{Z_t} + \frac{T_{it}}{Z_t} - \frac{a_i}{Z_t} \right)^{-\sigma_y} = \frac{\beta_i}{1+g} (F_K(K_{t+1}) + 1 - \delta) \left(\frac{a_{it}(F_K(K_{t+1}) + 1 - \delta)}{Z_t} \right)^{-\sigma_o}$$

This implicitly defines $\frac{a_{it}}{Z_t}$ as a function of $\frac{T_{it}}{Z_t}$, K_t, and K_{t+1}.

$$\left(\theta_i (1-\alpha) \frac{K_t}{Z_t}^{\alpha} + \frac{T_{it}}{Z_t} - \frac{a_i}{Z_t} \right)^{-\sigma_y} = \frac{\beta_i}{1+g} (\alpha \frac{K_{t+1}}{Z_{t+1}}^{\alpha-1} + 1 - \delta)^{1-\sigma_o} \left(\frac{a_{it}}{Z_t} \right)^{-\sigma_o}$$

Take the asset market clearing condition and divide it by Z_{t+1} :

$$\frac{K_{t+1}}{Z_{t+1}} = \frac{a_{it}}{Z_t} \left(\frac{K_t}{Z_t}, \frac{T_{it}}{Z_t}, \frac{K_{t+1}}{Z_{t+1}} \right)$$

it's therefore possible to define a balanced growth path and solve for $\frac{K}{Z}$. Take the BGP version of the asset market clearing condition:

$$\frac{K}{Z} = \frac{a_i}{Z}\left(\frac{K}{Z}, \frac{T_i}{Z}\right)$$

Following the same steps as in the baseline model and take the total derivative of $\frac{K}{Z}$ with respect to $\frac{T}{Z}$. Defined all variables, $\tilde{x} \equiv \frac{x}{Z}$:

$$\hat{\tilde{K}}_{PI} = \frac{\tilde{K}_R}{\tilde{K}_R(1 - \tilde{A}_R\tilde{w}_K) - \tilde{A}_R} \sum_I \pi_i \frac{\partial \tilde{a}_i}{\partial \tilde{PI}_i}\left(\tilde{wL} - \tilde{w\theta}_i\right)$$

Note that the derivative $\frac{\partial \tilde{a}_i}{\partial \tilde{PI}_i} = \frac{\partial a_i}{\partial PI_i}\frac{Z}{Z} = \frac{\partial a_i}{\partial PI_i}$. Note also that the elasticity of x/Z with respect to y is $(\frac{\partial x}{Z}/\partial y)(yZ/x) = x_y$. Finally, divide by wL/Z to get an expression for $\hat{K}_{PI}$. Therefore, our expression for $\hat{K}_{PI}$ is identical to equation (14).

The firm's first order condition is:

$$F_K(K) = \alpha K^{-1/\rho} Y^{1/\rho} = r - \delta$$
$$\log(\alpha) + \frac{-1}{\rho}\log(K) + \frac{1}{\rho}(\log(Y)) = \log(r - \delta)$$

Again assuming $\delta = 0$, the above implies $K_r = \frac{\partial \log(K)}{\partial \log r} =$

$$\frac{\rho}{(-1 + \partial \log(Y)/\partial \log(K))} = \frac{\rho}{-(\alpha_L)}$$

The households' first order conditions in terms of the new normalized variables are:

$$f_\ell(\ell_i^y) = u_c^y(\tilde{c}_i^y)(1 - \tau_\ell)\theta_{iy} w$$
$$u_c^y(\tilde{c}_i^y) = \beta R u_c^o(\tilde{c}_i^o)$$

Again, assuming generations are equally sized, the total change in social welfare:

$$\sum_I \lambda^i \pi_i \left(u_c^y(\tilde{c}_i^h)\left(d((1 - \tau_\ell)\ell_i^y \theta_i^y \tilde{w}) + d\tilde{T} - d\tilde{a}_i^y\right) - f_\ell(\ell_i^h) d\ell^i + \beta u_c^o(\tilde{c}_i^o)\left(R d\tilde{a}_i^y + dR\tilde{a}_i^y\right)\right)$$

Expanding the first term and referring to $\theta_i^y = \theta_i$ and $\ell_i^y = \ell_i$:

$$d((1 - \tau_\ell)\ell_i \theta_i \tilde{w}) = -\ell_i \theta_i \tilde{w} d\tau_\ell + \theta_i \ell_i (1 - \tau_\ell) d\tilde{w} + \theta_i \tilde{w}(1 - \tau_\ell) d\ell_i$$

Expanding the dT term:

$$dT = \tau_\ell L d\tilde{w} + \tau_\ell \tilde{w} dL + d\tau_\ell \tilde{w} L$$

Define $\omega_{ih} = \lambda_i \beta^{h-1} u_c^h(\tilde{c}_i^h)$, where $\sum_I \sum_H \omega_{ih} = 1$ and subbing in the labor supply condition and Euler equation, the total change becomes:

$$dSW = \sum_I \pi_i \omega_{iy} \Big(L - \ell_i \theta_i \Big) \tilde{w} d\tau_\ell + \sum_I \pi_i \Big(\omega_{iy}(\theta_i \ell_i (1-\tau_\ell) d\tilde{w}) + \omega_{io} \tilde{a}_i^y dR \Big) + \tau_\ell d(\tilde{w} L)$$

Defining Θ as:

$$\Theta = L^{-1} \sum_I \pi_i \omega_{ih} \left(\theta_i \ell_i^h (1-\tau_\ell) + \tilde{a}_i^{h-1} \frac{\partial R}{\partial \tilde{w}} \right)$$

Let $\omega_i = \omega_{iy}$. Recall that $\omega_{ih+1}/\omega_{ih} = \beta u_c^{h+1}(\tilde{c}_i^{h+1})/u_c^h(\tilde{c}_i^h) = R$, so Θ becomes:

$$\Theta = \sum_I \omega_i \pi_i \left(\theta_i \frac{\ell_i}{L}(1-\tau_\ell) + \frac{R \tilde{a}_i^y}{L} \frac{\partial R}{\partial \tilde{w}} \right)$$

Because we've assumed that households only earn and save in the first period of life, the change in social welfare is:

$$dSW = \underbrace{\sum_I \omega_i \Big(L - \ell_i^h \theta_i \Big) \tilde{w} d\tau_\ell}_{\text{Direct Effects}} + \underbrace{L\tilde{\Theta} dw + \tilde{\tau}_\ell d(wL)}_{\text{General Equilibrium Costs}}$$

Expanding the GE costs as before:

$$dSW = \underbrace{\sum_I \omega_i \pi_i (\tilde{w} L - \tilde{w} \theta_i \ell_i) d\tau_\ell}_{\text{Direct Effect of Redistribution}} + \underbrace{wL(\tilde{\Theta} + \tilde{\tau}_\ell) w_K K_{PI}}_{\text{Direct Effect of NH Savings}}$$
$$\underbrace{wL(\tilde{\Theta} + \tilde{\tau}_\ell) \left(w_L + \frac{\tilde{\tau}_\ell}{\tilde{\Theta} + \tilde{\tau}_\ell} \right) \frac{dL}{L}}_{\text{Direct Effect of Labor Distortion}} + \underbrace{wL(\tilde{\Theta} + \tilde{\tau}_\ell)(w_K \mathcal{L} + \mathcal{C})}_{\text{Feedback Effects}}$$

Therefore, the ratio of the two channels remains the same as in baseline case.

A.14 Isolating Non-Homothetic Savings Channel

To isolate the effects of the NH savings channel, I solve for the direct effect of the 1 percentage point increase in the average tax rate + the budget balancing transfer on each household type's expected permanent income. I denote this change, ΔPI_i.

Households who are age h, type-i, and have drawn state s, earn $y_{hij} = \ell_{ij}^h \theta_i^h e_j w$. They pay labor income taxes equal to $\bar{\tau}_\ell \tau_{\ell hij}(y_{hij})$ in labor income taxes, where $\tau_{\ell hij}(y_{hij})$ is defined following Bénabou (2002). Then if we hold y_{hij} constant at the original steady state level, a one percentage point change in $\bar{\tau}_\ell$, $d\bar{\tau}_\ell$ will increase type-i households expected lifetime tax burden by:

$$d\tau_{\ell i} = d\bar{\tau}_\ell \sum_H \sum_J \mu_j e_j \theta_i^h \ell_{ij}^h w * \tau_{\ell hij}(e_j \theta_i^h \ell_{ij}^h w)$$

Here μ_j is the probability of realizing state e_j. Summing these changes together over all types and ages gives you the expected increase in the uniform transfer attributable to this change: $dT = \sum_I \pi_{ih} d\tau_{\ell i}$, which each household receives each period.

Using the original steady state interest rate r, the present value of this change is:

$$\Delta PI_i = \sum_H \sum_J \mu_j \frac{y_{hij} * \tau_{\ell hij}(y_{hij})}{(1+r)^h}$$

Then to construct the hypothetical lump-sum taxes for each type, as stated in the text, the following two conditions must hold:

$$\sum_H \frac{T_{ih}}{(1+r)^h} = \Delta PI_i$$

$$T_{ih} = \frac{T_{ih+1}}{1+r} \text{ for all h}$$