NBER WORKING PAPER SERIES

OPTIMAL POSITIVE CAPITAL TAXES AT INTERIOR STEADY STATES

Jess Benhabib
Bálint Szőke

Working Paper 25895
http://www.nber.org/papers/w25895

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2019

We are grateful to Marco Bassetto, Raquel Fernandez, Sarolta Laczo, Albert Marcet, Benjamin Moll, Alain Trannoy, and two anonymous referees for helpful suggestions and criticism. We owe special thanks to Leslie Reinhorn for his detailed comments and careful reading of an earlier draft. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Optimal Positive Capital Taxes at Interior Steady States
Jess Benhabib and Bálint Szőke
NBER Working Paper No. 25895
May 2019
JEL No. E62,H21,H23

ABSTRACT

We generalize recent results of Bassetto and Benhabib (2006) and Straub and Werning (2018) in a model with endogenous labor-leisure choice where all agents are allowed to save and accumulate capital. In particular, using a neoclassical infinite horizon model with standard balanced growth preferences and agents heterogeneous in their initial wealth holdings, we provide a sufficient condition under which optimal redistributive capital taxes can remain at their allowed upper bound forever, even if the resulting equilibrium trajectory converges to a unique steady state with positive and finite consumption, capital, and labor. We first generate some simple parametric examples which satisfy our sufficient condition and for which closed form solutions exist. We then provide an interpretation of our sufficient condition for equilibria induced by general constant returns neo-classical production functions. Using recent evidence on wealth distribution in the United States, we argue that our sufficient condition is empirically plausible.

Jess Benhabib
Department of Economics
New York University
19 West 4th Street, 6th Floor
New York, NY 10012
and NBER
jess.benhabib@nyu.edu

Bálint Szőke
Department of Economics
New York University
19 West 4th Street, 6th Floor
New York, NY 10012
balint.szoke@nyu.edu

1 Introduction

The seminal work of Chamley (1986) shows that when the social planner raises revenues for government expenditures, optimal capital tax rates may remain positive in transition, but at steady-states they must be set to zero. In other words, at the steady-state, the social planner commits to zero capital taxes and raises revenues by taxing labor earnings instead. Judd (1985) demonstrates that the same optimal tax policy applies in an economy where taxes are chosen for redistributive purposes, solely according to preferences of hand-to-mouth workers whose income is composed of labor earnings and government transfers. The reason is that positive capital taxes distort savings, which in the limit shrink the capital tax base too much, while also depressing the marginal product of labor. The general conclusion from these studies is that committing to positive capital taxes forever is a bad idea.

An initial counterexample to Judd (1985) was given by Lansing (1999) showing that with log preferences, optimal capital taxes could be positive forever in some equilibria.[1] As later established by Reinhorn (2019), however, this example turned out to be a knife-edge case. Later, Bassetto and Benhabib (2006) studied a model with a continuum of agents that differed in their initial capital stocks, but are otherwise all allowed to choose their savings optimally. Assuming inelastic labor supply, they established a condition under which a sufficiently wealth-poor household would choose to tax capital at the maximally allowed rate forever, and would redistribute taxes lump-sum and equally to all.[2] More recently, Straub and Werning (2018) obtained very similar results in the frameworks of Chamley (1986) and Judd (1985). In particular, they showed that under certain conditions, optimal capital taxes can remain positive forever for capital trajectories that converge to extremes, but not those that converge to an interior steady-state.

In this paper, we show that when the government raises revenue via capital and labor taxes optimally in a model with infinitely lived agents, who are heterogenous only with respect to their initial capital holdings, taxing capital forever for redistributive purposes can in fact be consistent with equilibrium trajectories of consumption, leisure and capital converging to *interior* steady-states, while the steady-state optimal tax rate $\tau^* \in [0, \bar{\tau}]$ is set at its maximum $\tau^* = \bar{\tau}$. First, we generalize the Bassetto and Benhabib (2006) condition for maximal capital taxes to a neoclassical growth model with endogenous labor-leisure choice and Gorman aggregable balanced growth preferences. Under this condition, if the sequence of tax rates is optimally chosen according to the preferences of the median household that is sufficiently wealth-poor relative to the average household, or alternatively, if the planner assigns relatively more weight to wealth-poor households, the implemented policy will feature capital tax rates that are kept at their upper bounds forever.

The Bassetto-Benhabib condition, however, involves the equilibrium value function of the household with mean wealth, so it is not immediately obvious how to generate examples satisfying this condition at interior steady-states. The example that Bassetto and Benhabib (2006) gave is for an AK

[1] While capital, consumption, and the net return to capital converge to an interior steady state in the example of Lansing (1999), multipliers do *not* converge, thereby violating the assumptions of Judd (1985). The multipliers diverge, but they satisfy transversality (see Reinhorn (2019), pages 9-10).

[2] Bassetto and Benhabib (2006) also showed that if preferences are Gorman aggregable, a Condorcet winner exists, so the median household's preferred policy gets implemented if the sequence of tax rates is chosen by majority voting.

model where the equilibrium capital stock, depending on parameters, perpetually contracts or grows.[3] To demonstrate that this outcome is just a special case, we define a class of models with constant relative risk-averse (CRRA) preferences and constant elasticity of substitution (CES) production functions for which closed form solutions exist. This allows us to provide examples with positive long-run capital taxation *and* interior steady-states. To show that this finding does not depend on stringent parametric assumptions, we complement our examples with more standard model calibrations requiring numerical solutions. We then provide an interpretation of our sufficient condition implying that optimal capital taxes will remain at their upper bound forever for equilibria generated under arbitrary constant returns neoclassical production functions. Finally, we investigate the empirical validity of the Bassetto-Benhabib condition by using recent data from Wolff (2017) on the wealth distribution in the United States and find that the sufficient condition is empirically plausible.

We should note that our model, where optimal capital taxes finance lump-sum redistribution and are maximal forever at interior steady states differs from the model of Judd (1985) or Straub and Werning (2018) where workers are not allowed to save. Our wealth-poor agents are not required to immediately consume the wages and transfers they receive. Therefore, even though everyone dislikes having to pay capital taxes, wealth-poor households find the implied redistributive transfers more valuable in our setting, where they have the option to save them either fully or in part, than in an economy where they are constrained to be hand-to-mouth consumers.

There are also a number of studies that deviate from the original Chamley or Judd models in which optimal capital taxes can remain positive forever. For example in OLG models with realistic life cycle profiles having non-constant labor endowments, if labor taxes cannot depend on age, they may not be able provide an optimal intertemporal redistribution across households that maximizes the social welfare function of a newborn, especially if labor supply is endogenous. Then a positive capital income tax that mimics a labor income tax rate that can vary with age can be optimal despite its intertemporal distortion of accumulation (see Erosa and Gervais (2002)). Similarly, in incomplete market models with endogenous labor, borrowing constraints, and idiosyncratic earnings risk, if labor taxes are restricted to be proportional and cannot be progressive, relying on labor income taxes alone translates directly into low consumption for poor households at the constraint. Such labor taxes may not be optimal, and require instead a redistributive positive capital tax forever if progressive labor taxes are ruled out (see Hubbard and Judd (1987)). Comprehensive theoretical and quantitative analyses of such cases are studied and illustrated in detail by Conesa, Kitao, and Krueger (2009). On the other hand, Atkeson, Chari, and Kehoe (1999) and Chari, Teles, and Nicolini (2016) show that in such macroeconomic models with an enlarged tax system that also includes consumption taxes, capital should not be taxed in steady-state, either in representative agent models, or in models with heterogeneous agents differing in their initial wealth.

In the next section we describe the model environment and derive the value functions of households in competitive equilibria. Using these value functions, section 3 discusses how different households rank the available tax policies. Theorem 4 contains our main result providing a sufficient condition

[3] In their explicit example, Bassetto and Benhabib (2006) have linear production $y = rk$ and CRRA preferences, and they set the discount factor β so that $\beta r = 1$ and the capital stock contracts to zero. Nevertheless, it is easy to see that if we increased r slightly, the capital stock would grow forever.

under which certain households prefer to keep capital taxes at their upper bounds forever. Proposition 5 shows that positive optimal long-run capital tax rate can be consistent with an interior steady-state. Section 4 computes our sufficient condition under various functional assumptions that allow us to derive key equilibrium objects in terms of model primitives. Section 4.2 then presents a rough calibration illustrating that our sufficient condition is empirically plausible. Section 5 concludes.

2 The model

Consider a deterministic neoclassical growth model with a continuum of households of unit measure, indexed by i, who differ only in their initial wealth level. Household preferences are given by

$$\sum_{t=0}^{\infty} \beta^t u\left(c_t^i, 1-n_t^i\right) = \sum_{t=0}^{\infty} \beta^t \frac{\left[\left(c_t^i\right)^{\xi}\left(1-n_t^i\right)^{1-\xi}\right]^{1-\sigma}}{1-\sigma} \qquad \xi \in [0,1], \quad \sigma > 0 \tag{1}$$

where c_t^i and n_t^i are period t consumption and labor supply by agent i, respectively.[4] This functional form is a popular choice in the business cycle literature (see e.g. Kydland and Prescott (1982)) and it has been used by Chari, Christiano, and Kehoe (1994) to study optimal fiscal policy in an economy with homogeneous households. That u is a homogeneous function ensures that Gorman aggregation holds,[5] that is, there exists a representative agent endowed with the average initial wealth level and preferences of the form (1) over average consumption and leisure plans $\{(c_t, 1-n_t)\}_{t=0}^{\infty}$. An important special case of u, associated with $\xi = 1$, is CRRA preferences with inelastic labor supply.

Output y_t at time t is produced by competitive firms using capital k_t and labor n_t according to the linearly homogeneous production function

$$y_t = F(k_t, n_t)$$

with partial derivatives $F_k > 0$, $F_n \geq 0$, $F_{kn} \geq 0$, $F_{nn} \leq 0$, and $F_{kk} \leq 0$. Firms rent capital and labor from the competitive factor markets at rates r_t and w_t, respectively.

In each period t, the government levies proportional taxes on labor income $\nu_t \in [0,1)$[6] and capital income τ_t, subject to exogenous bounds $\tau_t \in [0, \bar{\tau}]$,[7,8] and provides lump-sum transfers (or taxes) tr_t

[4] Appealing to the Inada conditions we abstract from the nonnegativity constraints on consumption and leisure.

[5] An alternative (separable) form satisfying this condition would be $u^i = (1-\sigma)^{-1}\left[(c_t^i)^{1-\sigma} + (1-n_t^i)^{1-\sigma}\right]$, where Gorman aggregation is ensured by assuming *identical* elasticities of consumption and labor.

[6] With endogenous labor supply, if $\nu = 1$, households choose not to work, hence output is zero. We rule out this case.

[7] The upper bound can be justified by the fact that households can avoid renting out their capital stock, or by the presence of a "black market technology" that allows households to hide their capital income from the tax collector at a proportional cost $\bar{\tau}$. The zero lower bound will not be binding for the wealth distributions that we will consider.

[8] The upper bound on capital income tax rates, $\bar{\tau}$, can be unrestricted, and in principle capital taxes paid can exceed capital income. However, we later impose $\bar{\tau} < 1$ to permit the existence of an interior steady state with positive capital stocks and with capital income taxes set at their upper bound forever. For this result $\bar{\tau}$ has to be less than unity because with positive discounting the steady state after-tax return on capital has to be positive. See Proposition 5 and the discussion in section 3.4.

to all households. The period-by-period government budget constraint is

$$R_t b_t + g_t + tr_t = \tau_t r_t k_t + \nu_t w_t n_t + b_{t+1}, \qquad t \geq 0 \tag{2}$$

where R_t is the gross rate of return on one-period bonds held from $t-1$ to t. In general, the government uses $\{\tau_t, \nu_t, tr_t\}_{t=0}^{\infty}$ and one-period debt $\{b_{t+1}\}_{t=0}^{\infty}$ for the following purposes: (i) to pay for spending on a public good at an exogenous rate $\{g_t\}_{t=0}^{\infty}$, (ii) to redistribute wealth among households, and (iii) to pay back its initial debt b_0.

To simplify algebra, we impose the no arbitrage condition by stipulating that the gross return on bonds and capital are equal:

$$R_t = 1 + (1 - \tau_t) r_t - \delta,$$

where δ is the rate of depreciation. That said, we reformulate the constraints for τ_t in terms of bounds on the gross after-tax rate of return on capital:

$$\underline{R}_t \coloneqq 1 + (1 - \bar{\tau}) r_t - \delta \;\leq\; R_t \;\leq\; 1 + r_t - \delta =: \overline{R}_t \tag{3}$$

For convenience, we also define time 0 after-tax prices $q_t \coloneqq \prod_{s=1}^{t} R_s^{-1}$ for all $t \geq 1$ with q_0 being normalized to one. Moreover, let a_t^i be the wealth of household i at time t, consisting of capital k_t^i and maturing government bonds b_t^i. The average wealth level is then $a_t = \int a_t^i di$. The period-by-period budget constraint of household i is

$$c_t^i + a_{t+1}^i \leq R_t a_t^i + (1 - \nu_t) w_t n_t^i + tr_t, \qquad t \geq 0, \quad i \in [0, 1], \tag{4}$$

and we assume that households cannot run Ponzi schemes:

$$\lim_{T \to \infty} \; q_T a_{T+1}^i \geq 0 \qquad i \in [0, 1].$$

Optimality requires that the limit cannot be positive, so the household budget constraints in present discounted value form can be written as

$$\sum_{t=0}^{\infty} q_t c_t^i \leq R_0 a_0^i + \sum_{t=0}^{\infty} q_t \left(tr_t + (1 - \nu_t) \, w_t n_t^i \right) \qquad i \in [0, 1]. \tag{5}$$

Importantly, tr_t is lump-sum and it is independent of the household's type i. For simplicity, let S^{tr} denote the present discounted value of all transfers, $S^{tr} \coloneqq \sum_{t=0}^{\infty} q_t tr_t$. The standard representative-agent optimal tax problem—the so called *Ramsey problem*—rules out lump-sum components from tax policy, because they allow the government to always achieve the first best rendering the problem uninteresting. In contrast, in our heterogeneous agent setting, as stressed by Werning (2007), lump-sum transfers allow for redistributive effects of the tax system, so it is important to include them as a possible policy tool.

2.1 Competitive equilibria

When designing an optimal policy, we consider only those allocations and prices that constitute competitive equilibria for given budget-feasible government policies.

Definition 1. *A budget feasible policy is an expenditure plan* $\{g_t\}_{t=0}^{\infty}$*, a tax plan* $\{\tau_t, \nu_t, tr_t\}_{t=0}^{\infty}$*, and a debt issuance plan* $\{b_t\}_{t=0}^{\infty}$ *that satisfy* (2) *and* (3) *for all* $t \geq 0$*, with given* b_0 *and*

$$\lim_{T\to\infty} q_T b_{T+1} = 0.$$

Definition 2. *A competitive equilibrium consists of a budget-feasible policy* $\{\tau_s, \nu_s, tr_s, b_s, g_s\}_{s=0}^{\infty}$*, an allocation* $\left\{c_s, n_s, k_s, \left\{c_s^i, n_s^i, a_s^i\right\}_{i\in[0,1]}\right\}_{s=0}^{\infty}$*, and a price system* $\{r_s, w_s\}_{s=0}^{\infty}$ *that satisfy*

1. *For* $\forall i \in [0,1]$*, the sequences* $\left\{\left(c_s^i, n_s^i\right)\right\}_{s=0}^{\infty}$ *maximize household utilities* (1) *subject to* (5) *and given* a_0^i*. The sequence* $\{a_s^i\}_{s=0}^{\infty}$ *can be recovered from* (4) *satisfied with equality.*

2. *Factor prices equal their marginal products:*

$$r_t = F_k\left(k_t, n_t\right), \qquad w_t = F_n\left(k_t, n_t\right) \qquad t \geq 0$$

3. *Markets clear*

$$\int c_t^i di = c_t, \qquad \int n_t^i di = n_t \qquad t \geq 0$$
$$c_t + g_t + k_{t+1} = F(k_t, n_t) + (1-\delta)k_t \qquad t \geq 0. \tag{6}$$

Let $\mathcal{C}$ be the set of all competitive equilibria indexed by alternative budget-feasible government policies. Without extra restrictions on the sequence $\{g_s\}_{s=0}^{\infty}$, however, it is possible that the aggregate feasibility condition (6) cannot be satisfied for some period t, in which case $\mathcal{C}$ is empty. To avoid this, we require that the sequence $\{g_s\}_{s=0}^{\infty}$ is not "too high":

Assumption 1. *The upper bound* $\bar{\tau}$*, government expenditure plan* $\{g_t\}_{t=0}^{\infty}$*, and initial government debt* b_0 *are such that setting* $\nu_t = 0$ *and* $\tau_t = \bar{\tau}$*, for all* $t \geq 0$ *gives rise to an equilibrium with non-negative transfers* $S^{tr} \geq 0$ *through the government budget constraint. This implies*

$$R_0 b_0 + \sum_{t=0}^{\infty} q_t g_t \leq \sum_{t=0}^{\infty} \bar{\tau} q_t r_t k_t.$$

In other words, taxing only capital at the maximum rate forever generates enough revenue to fully cover the exogenous expenditure plan. Assumption 1 guarantees that the set $\mathcal{C}$ is nonempty.

2.2 Household i's utility in competitive equilibria

The form of the utility function allows us to express the present discounted value of each household for any given competitive equilibrium as a function of the representative agent's value function. Naturally,

features of this function inform us about how household i values the different tax policies and implied competitive equilibria. To derive this function, we start with household i's first-order conditions:

$$q_t = \beta^t \frac{u_c\left(c_t^i, 1-n_t^i\right)}{u_c\left(c_0^i, 1-n_0^i\right)} \qquad \text{and} \qquad (1-\nu_t)w_t = \frac{u_{1-n}\left(c_t^i, 1-n_t^i\right)}{u_c\left(c_t^i, 1-n_t^i\right)}. \tag{7}$$

Due to the aggregable utility function, the same necessary conditions hold for aggregate consumption, c_t, and aggregate labor, n_t, implying that in any competitive equilibrium, household i's marginal utilities are proportional to the representative household's marginal utilities. As a result,

$$c_t^i = \alpha^i c_t \qquad \text{and} \qquad 1-n_t^i = \alpha^i\left(1-n_t\right) \qquad t \geq 0, \tag{8}$$

where the nonnegativity restrictions on consumption and labor imply that the endogenous constant α^i must satisfy $0 < \alpha_i \leq 1/(1-n_t)$ for all t and for almost all i. Using the aggregate versions of (7), we derive household i's implementability condition (IC) from its budget constraint (5):

$$\sum_{t=0}^{\infty} \beta^t \left[u_c\left(c_t, 1-n_t\right) c_t^i - u_{1-n}\left(c_t, 1-n_t\right) n_t^i\right] = u_c\left(c_0, 1-n_0\right)\left(R_0 a_0^i + S^{tr}\right). \tag{9}$$

We define the value of the *average* household's after-tax initial wealth measured in units of utility:

$$\mathsf{A}(c_0, n_0, \tau_0) \coloneqq u_c(c_0, 1-n_0)\left[1 + (1-\tau_0)F_k(k_0, n_0) - \delta\right] a_0. \tag{10}$$

This variable turns out to summarize completely how each household's equilibrium utility is affected by the triple (c_0, n_0, τ_0). To see this, derive the equilibrium value of α^i by subtracting the average household's IC from (9) and using the equilibrium relationships in (8) to substitute out $(c_t^i - c_t)$ and $(n_t^i - n_t)$. In an equilibrium indexed by the pair (V, A), the value of α^i is

$$\alpha^i = \begin{cases} 1 + \frac{\mathsf{A}(c_0, n_0, \tau_0)}{V(1-\sigma)}\left(\frac{a_0^i}{a_0} - 1\right) & , \quad \text{if } \sigma \neq 1 \\ 1 + \frac{\mathsf{A}(c_0, n_0, \tau_0)}{(1-\beta)^{-1}}\left(\frac{a_0^i}{a_0} - 1\right) & , \quad \text{if } \sigma = 1 \end{cases} \tag{11}$$

where $V \coloneqq \sum_{t=0}^{\infty} \beta^t u\left(c_t, 1-n_t\right)$ is the present discounted utility of the agent with average initial wealth a_0. Upon substituting (8) into (1), the present discounted utility of household i is

$$V^i\left(V, \mathsf{A};\ \Delta a_0^i\right) \coloneqq \begin{cases} \left(\alpha^i\right)^{1-\sigma} V & , \quad \text{if } \sigma \neq 1 \\ \frac{\log \alpha^i}{1-\beta} + V & , \quad \text{if } \sigma = 1 \end{cases} \tag{12}$$

where we define the term entering α^i as

$$\Delta a_0^i \coloneqq \frac{a_0^i - a_0}{a_0} = \frac{a_0^i}{a_0} - 1 \tag{13}$$

measures the relative position (relative to the average) of household i in the initial wealth distribution. Function V^i represents household i's equilibrium utility in a remarkably compact way. In particular,

the V^i-relevant features of any equilibrium can be summarized by two variables: the average household's value V and the utility value of the average household's after-tax wealth A. These variables embody a rich set of possible tax policies, allocations, and prices.[9] Given that households are indifferent between equilibria that lead to the same (V, A), we will henceforth denote the elements of $\mathcal{C}$ by simply using the induced pairs (V, A).

2.3 Subsets of $\mathcal{C}$

We define two subsets of the set of competitive equilibria $\mathcal{C}$ that will prove to be useful. The first subset $\mathcal{C}^* \subset \mathcal{C}$ includes those equilibria that are induced by "eventually time-invariant" policies:

$$\mathcal{C}^* := \{\ (V, \mathsf{A}) \in \mathcal{C} \quad : \quad \exists t_F \geq 0, \quad \text{s.t.} \quad g_t = g^*,\ \nu_t = \nu^*,\ \tau_t = \tau^* \leq \bar{\tau}, \quad \forall t \geq t_F\ \}\,.$$

The second subset $\overline{\mathcal{T}} \subset \mathcal{C}$ includes those equilibria that feature maximal capital taxation forever:

$$\overline{\mathcal{T}} := \{\ (V, \mathsf{A}) \in \mathcal{C} \quad : \quad \tau_t = \bar{\tau}, \quad \forall t \geq 0\ \}\,.$$

In addition, we will be interested in capital tax policies with the "bang-bang" property:

Definition 3. *The capital tax sequence $\{\tau_t\}_{t=0}^{\infty}$ has the bang-bang property if $\tau_t < \bar{\tau}$ implies $\tau_s = 0$ for $s > t$. That is, there exists a time T, s.t. $\tau_t = \bar{\tau}$ for $t < \mathsf{T}$ and $\tau_t = 0$ for $t > \mathsf{T}$.*

Figure 1 illustrates these objects for a particular example economy.[10] The green and orange areas represent the set $\mathcal{C}^*$ and the intersection of $\mathcal{C}^*$ and $\overline{\mathcal{T}}$ in the (V, A)-space, respectively. That is, the orange set contains equilibria induced by policies with indefinite maximal capital taxes and "eventually time-invariant" labor taxes. Certain policies can be readily identified in Figure 1: (i) equilibria induced by eventually zero labor taxes (and no capital taxes) are denoted by the dashed blue line, (ii) those with bang-bang capital tax policies (without labor taxes) for different T values are denoted by the dotted blue line.[11] In addition, the big colored circles represent two equilbria of particular importance: the black dot, $\left(\overline{V}, \overline{\mathsf{A}}\right)$, shows the allocation induced by the policy using only lump-sum *taxes*, whereas the red dot, $(\underline{V}, \underline{\mathsf{A}})$, represents the equilibrium induced by the policy in Assumption 1, i.e., $(\tau_t, \nu_t) = (\bar{\tau}, 0)$, $t \geq 0$. Since in our example government expenditures are positive, the equilibrium $\left(\overline{V}, \overline{\mathsf{A}}\right)$ is supported by lump-sum *taxes*, that is, the present discounted value of transfers S^{tr} is negative. The grey dots represent other equilibria where this property holds even if distorting taxes are also used. Loosely speaking, while capital taxes tend to decrease both V and A, labor taxes have an opposite effect on the two equilibrium objects: they increase A and decrease V.

That the boundary of the set $\mathcal{C}$ consists of equilibria with well-defined tax policies holds true more generally. To show these properties formally, we first define iso-A sets in the space of competitive

[9] It might be surprising that S^{tr} (transfer) does not appear in V^i. This follows from the fact that S^{tr} is independent of i, so its effect on household i can be captured by the average household's value function and choices.

[10] We set the preference parameters $(\beta, \sigma, \xi) = (0.96, 5, 0.8)$. Suppose that $b_0 = 0$ and the government expenditure plan is time invariant, $g_t = g^*$, $t \geq 0$, with the values $(\bar{\tau}, g^*) = (0.25, 0.05)$ being chosen to make Assumption 1 hold. The production function is Cobb-Douglas with captial share $\rho = 1/3$ and depreciation rate $\delta = 0$.

[11] The definition of a "bang-bang" capital tax policy is silent about the value of τ_{T}. For simplicity, we set $\tau_{\mathsf{T}} = 0$ in Figure 1 and confirm numerically that "bang-bang" policies with $\tau_{\mathsf{T}} > 0$ lie in the interior of the plotted set.

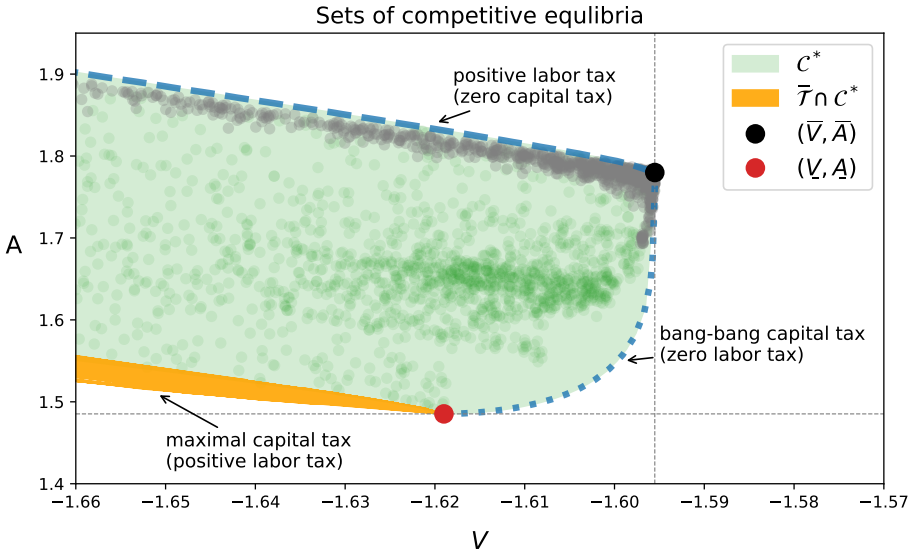

Figure 1: Subsets of competitive equilibria in the (V, A)-space. Light green region represents equilibria induced by eventually time-invariant policies, while the orange region shows those equilibria among these that are induced by a policy with maximal capital taxes forever (for an arbitrary labor tax sequence). Transparent green dots show equilibria induced by eventually time-invariant, *random* tax paths $\{(\tau_t, \nu_t)\}_{t\geq 0}$ for various (τ^*, ν^*) and t_F-values. Grey dots denote equilibria with lump-sum *taxes*, i.e. $S^{tr} \leq 0$. The horizontal and vertical dashed gray lines represent the lowest attainable A and the highest attainable V, respectively.

equilibria. An iso-A set of value $\widetilde{\mathsf{A}}$, denoted by $\mathcal{C}(\widetilde{\mathsf{A}})$, consists of all $(V, \mathsf{A}) \in \mathcal{C}$ with $\mathsf{A} = \widetilde{\mathsf{A}}$.[12] That said, Lemma 1 shows that policies with maximal capital taxes forever tend to induce (V, A) pairs in the "bottom left" corner of $\mathcal{C}$ (orange region). More precisely, the "western boundary" of $\mathcal{C}$ consists of equilibria featured by $\tau_t = \bar{\tau}$ forever irrespective of labor tax policy.

Lemma 1. *If $\sigma > 1$ and $\xi < 1$,*[13] *the equilibrium $(\underline{V}, \underline{\mathsf{A}})$, induced by the policy $(\tau_t, \nu_t) = (\bar{\tau}, 0)$ for all $t \geq 0$, is associated with the lowest attainable A value. Moreover, for all feasible $\mathsf{A} \geq \underline{\mathsf{A}}$ such that $\mathcal{C}(\mathsf{A}) \neq \emptyset$, the V-minimizing equilibrium over $\mathcal{C}(\mathsf{A})$ belongs to $\overline{\mathcal{T}}$.*

Proof. See Appendix A.1. □

In addition, by mimicking the argument of Bassetto and Benhabib (2006), Lemma 2 shows that the "eastern boundary" of $\mathcal{C}$ must consist of equilibria that are induced by bang-bang capital tax policies.

Lemma 2. *For all feasible $\mathsf{A} \geq \underline{\mathsf{A}}$ such that $\mathcal{C}(\mathsf{A}) \neq \emptyset$, the V-maximizing equilibrium over $\mathcal{C}(\mathsf{A})$ is induced by a tax policy with bang-bang capital taxes and eventually zero labor taxes, i.e., if $\tau_t < \bar{\tau}$ then $\tau_s = \nu_s = 0$ for $s > t$.*

Proof. See Appendix A.2. □

[12] Clearly, these sets can be represented by horizontal slices in Figure 1.

[13] The case $\xi = 1$ was treated in Bassetto and Benhabib (2006). See their Theorem 3.

3 Preferences over tax policies

When faced with the choice, household i prefers to implement the equilibrium that maximizes V^i.[14] Remarkably, the only i-dependent term in the V^i function is Δa_0^i, so for any given relative wealth position, household i's attitude toward the alternative tax policies can be represented by indifference curves in the (V, A)-space. The way in which these curves are positioned relative to the $\mathcal{C}$-set in Figure 1 determines the relationship between the interests of household i and those of the average household. In this section we show that in our economy, households' interests are not necessarily aligned, in fact, agents with different initial wealth levels prefer very different tax policies. For ease of reference, in what follows, we call household i "wealth-poor" if $\Delta a_0^i < 0$ and "wealth-rich" if $\Delta a_0^i \geq 0$.

Preferences over tax policies are shaped by a tension between two effects of taxation. First, capital taxes distort prices, thereby altering all households' optimal decisions. In a world without utility-enhancing government purchases, this distortion has a negative effect on everyone's welfare. In fact, because our economy features Gorman aggregation and a common discount factor, this effect is proportional, independent of the initial wealth levels. This does not mean, however, that every household prefers the same tax policy. With the availability of transfers, wealth inequality among households brings about a second role of taxation: *redistribution*. One can see this by combining the government's budget constraint with those of household i's:

$$\sum_{t=0}^{\infty} \bar{q}_t c_t^i = \overline{R}_0(a_0^i - b_0) + \sum_{t=0}^{\infty} \bar{q}_t \left[w_t n_t^i - g_t\right] - \underbrace{\sum_{t=0}^{\infty} \bar{q}_t \nu_t w_t (n_t^i - n_t)}_{\text{redistribution through } \nu_t} + \underbrace{\sum_{t=0}^{\infty} \bar{q}_t \tau_t r_t (a_t - a_t^i)}_{\text{redistribution through } \tau_t}$$

where $\bar{q}_t := \prod_{s=1}^{t} \overline{R}_s^{-1}$ denotes *before-tax* time 0 prices. The redistributive effect of capital taxation is captured by the last term on the right hand side: when $\tau_t > 0$, wealth is being redistributed from wealth-rich households with $a_t^i > a_t$ to the wealth-poor households with $a_t^i < a_t$. Evidently, labor taxation also has a redistributive effect, captured by the third term on the right hand side, because households who work more pay more labor tax as well. Notice, however, that if leisure is a normal good, it is the wealth-poor households who work relatively more, thus labor taxes induce redistribution from the wealth-poor to the wealth-rich households.

Households determine their preferred tax policy by trading-off these two effects of capital taxation: (1) the inefficiency caused by the distorted inter-temporal margins and (2) the induced wealth redistribution. There is consensus on the harmfulness of the former, but households with different initial wealth levels naturally disagree on the latter. In Theorem 4 we provide a condition under which the benefits from redistribution for household i are so large that the household's preferred tax policy features maximum capital taxation forever.

[14] To reiterate, we restrict the choice set of household i to competitive equilibria as in Definitions 1 and 2 above, consistent with optimizing agents, competitive market clearing, and the intertemporal budget constraint of the government.

3.1 Indifference curves

We want to maximize V^i with respect to (V, A) over the "budget set" $\mathcal{C}$. Similar to standard optimal choice problems, this constitutes finding a point of tangency between the indifference curves of V^i and the boundary of the set $\mathcal{C}$. Appendix B shows that V^i is a (weakly) concave function of its two arguments. Moreover, by using the implicit function theorem, the slope of the indifference curve of V^i in our (V, A)-space can be written as

$$-\frac{\partial V^i}{\partial V} \Big/ \frac{\partial V^i}{\partial \mathsf{A}}.$$

Appendix B shows that the sign of the denominator hinges only on the relative wealth level:

$$\operatorname{sign}\left(\frac{\partial V^i}{\partial \mathsf{A}}\right) = \operatorname{sign}\left(\Delta a_0^i\right) \qquad \forall (V, \mathsf{A}) \in \mathcal{C}. \tag{14}$$

Intuitively, for a given level of V, a wealth-poor household wants to shrink the spread of the cross-sectional distribution of utilities, thereby bringing its own equilibrium utility closer to V. Since the initial wealth level is the only source of heterogeneity, the reduction of the spread Δa_0^i can be achieved by making the utility value of the after-tax return on initial wealth lower. A similar argument applies for the wealth-rich households, but with opposite signs.

As for the derivative with respect to V, we obtain the following formula for $\sigma \neq 1$:

$$\frac{\partial V^i}{\partial V} = \left(\alpha^i\right)^{-\sigma} \left[1 + \sigma \underbrace{\frac{\mathsf{A}(c_0, n_0, \tau_0)}{V\,(1-\sigma)}}_{=:\mathsf{D}} \Delta a_0^i\right] = \frac{1 + \sigma \mathsf{D} \Delta a_0^i}{\left(1 + \mathsf{D}\Delta a_0^i\right)^{\sigma}}. \tag{15}$$

Evidently, the sign of this function depends on parameters and the *competitive equilibrium* in which the partial derivative is being evaluated. Nonetheless, Lemma 3 discusses two cases in which the sign is unambiguously positive: (i) household i is wealth-rich, or (ii) preferences are such that the substitution effect dominates the income effect.

Lemma 3. *The partial derivative* $\frac{\partial V^i}{\partial V}$ *is positive, if* $\sigma \leq 1$ *or* $\Delta a_0^i \geq 0$.

Proof. Since $(1-\sigma)V > 0$, the term D is positive, so for $a_0^i \geq a_0$ the partial derivative is always positive regardless of the equilibrium allocation. Likewise, from $\alpha^i = 1 + \mathsf{D}\Delta a_0^i > 0$ it follows that as long as $\sigma < 1$, we have $\frac{\partial V^i}{\partial V} > 0$. The log case, $\sigma = 1$, trivially follows from (12). □

Tax policy preferred by the average household

Before turning to our main case, we briefly discuss the average household's problem, which, due to Gorman aggregation, can be viewed as a standard representative-agent optimal tax problem. The existence of lump-sum taxes, however, renders this problem trivial, because the first best (from the representative agent's point of view) is always achievable. Nonetheless, even if lump-sum *taxes* were not allowed, the well-known result by Chamley (1986) for a representative agent would hold in our

setting for the average household, that is, the capital tax sequence $\{\tau_t\}_{t=0}^{\infty}$ that maximizes the average household's welfare would have the bang-bang property. While the finiteness of T could be non-trivial,[15] in our setting the average household clearly prefers to set capital taxes to zero as soon as possible, hence $\mathsf{T} = 0$.

Tax policy preferred by wealth-rich households

Lemma 3 implies that for wealth-rich households, the slope of the indifference curves is unambiguously negative and they prefer high V and high A values. As a result, their preferred tax policies lie on the "northern boundary" of $\mathcal{C}$, so from Lemma 2 it follows that they want capital tax policies with the bang-bang property. In fact, since their preferred equilibrium (V^*, A^*) must feature $V^* \leq \overline{V}$, the corresponding after-tax return on capital must be $\mathsf{A}^* \geq \overline{\mathsf{A}}$. As Figure 1 illustrates, such equilibria are featured by $\mathsf{T}^i = 0$.

Figure 2 shows the set $\mathcal{C}^*$ along with indifference curves (thin grey curves) for households with various initial wealth levels. For each panel, the grey arrows in the bottom right corner shows the direction in which V^i increases. The top right panel displays the case of a wealth-rich household with the blue filled dot denoting its preferred equilibrium pair (V^*, A^*). While we provide no formal proof, Figure 1 and 2 clearly support the intuition that because labor taxes redistribute resources from poor to rich households, wealth-rich households tend to prefer policies that use distorting labor taxes, at least for some periods.

Tax policy preferred by wealth-poor households

Our main case of interest is when household i's preferred tax policy features maximal capital taxation forever, that is, when some household i prefers a competitive equilibrium (V, A) that belongs to $\overline{\mathcal{T}}$. As we saw above, a necessary condition for this result is that household i is wealth-poor. In an environment similar to ours, but with inelastic labor supply, Bassetto and Benhabib (2006) provided a sufficient condition for indefinite maximal capital taxation, namely, that V^i is decreasing in V at the allocation preferred by household i. Theorem 4 shows that a slightly altered version of the Bassetto-Benhabib-condition applies in our setting as well.

Theorem 4. *If at the equilibrium induced by $(\tau_t, \nu_t) = (\bar{\tau}, 0)$ for all $t \geq 0$ the partial derivative is non-positive, $\frac{\partial V^i}{\partial V} \leq 0$, then the capital tax sequence preferred by household i features $\mathsf{T}^i = \infty$.*

Proof. From the property $\frac{\partial V^i}{\partial V} \leq 0$ and Lemma 3, we can conclude that $\sigma > 1$ and $\Delta a_0^i < 0$. (14) in section 3.1 then implies that the slope of the indifference curves of household i in the (V, A)-space is non-positive (recall that V is negative when $\sigma > 1$). Therefore, the preferred equilibrium must be on the "western boundary" of the set $\mathcal{C}$. Then, if $\xi < 1$, Lemma 1 implies $\mathsf{T}^i = \infty$. The case $\xi = 1$ is covered by Theorem 3 in Bassetto and Benhabib (2006). □

[15] Indeed, abstracting from lump-sum transfers or taxes, Straub and Werning (2018) show that if the initial government debt $b_0 > 0$ is sufficiently large and $\sigma > 1$, the representative household could find it optimal to set $\mathsf{T} = \infty$. They argue that in this case no interior steady state exists; both capital and consumption must converge to zero asymptotically.

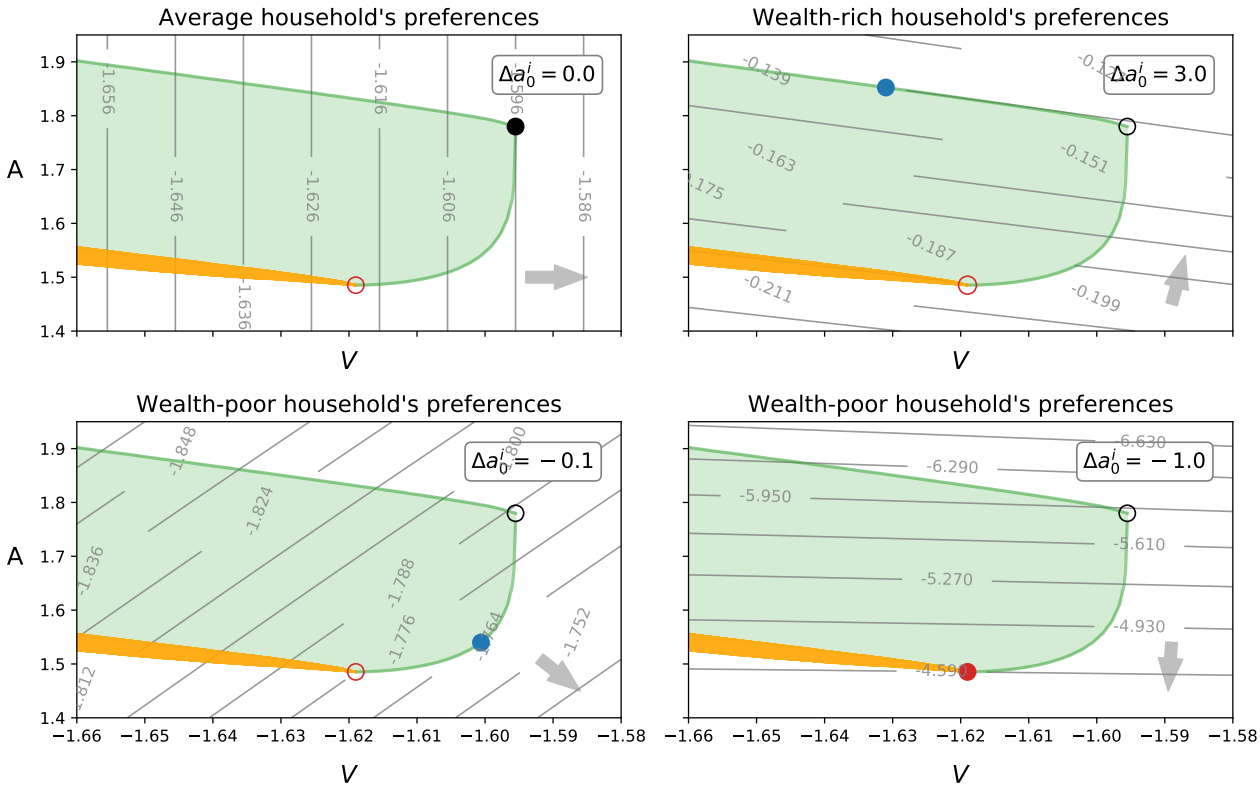

Figure 2: Effect of initial wealth on the preferred tax policy. Thin gray curves represent indifference (iso-V^i) curves of household i with initial wealth position indicated in the top right corner of each panels. The arrows in the bottom right corner of the four panels show the direction in which V^i increases. In each panel, the filled dot denotes the equilibrium preferred by the respective household. Parameters are as described in footnote 10.

The two bottom panels of Figure 2 illustrate how the preferences of wealth-poor households are shaped by their relative wealth position. The bottom left panel displays the case of a household that is not significantly worse-off than the average household. The partial derivative is $\partial V^i/\partial V > 0$ and as the filled blue dot suggests, the preferred tax policy features a finite stopping time for capital taxes. In contrast, the bottom right panel shows the case of a household whose wealth is so much lower than the average that the partial derivative $\partial V^i/\partial V$ becomes negative over large part of the set $\mathcal{C}$. As a result, the household would want to implement a policy with maximal capital taxation forever.

3.2 Interior steady-state

Because of Gorman aggregation, we can view every $(V, \mathsf{A}) \in \mathcal{C}$ as an equilibrium induced by a representative-agent neoclassical growth model for a particular—not necessarily optimal—feasible government policy. As such, the analysis of the long-run properties of aggregate consumption, capital and labor is standard, provided that the given equilibrium has a steady-state. Recent findings of Straub and Werning (2018) render the question of "interiority" of such steady-states non-trivial. By revisiting the setting of Judd (1985), they show that if positive long-run capital taxation is optimal and the cor-

responding allocation converges, then consumption must converge to zero, i.e. the steady-state cannot be interior. In our context, however, where everyone can save, we obtain the following proposition:

Proposition 5. *Consider an equilibrium in the set $\overline{\mathcal{T}} \cap \mathcal{C}^*$ induced by an eventually time-invariant government policy with maximal capital taxation forever. The steady-state consumption, c^*, capital, k^* and labor, n^*, are all positive as long as $\bar{\tau} < 1$ and $\nu^* < 1$.*

Proof. Let c^*, k^*, and n^* denote the steady-state value of consumption, capital, and labor. Using the linear homogeneity of F and the average household's Euler equation in the steady-state, we obtain the following condition for the net return on capital:

$$F_k\left(\frac{k^*}{n^*}, 1\right) = \frac{\beta^{-1} - 1 + \delta}{1 - \bar{\tau}} \tag{16}$$

Equation (16) determines the steady-state capital-labor ratio, which is positive as long as $\bar{\tau} < 1$. Using the capital-labor ratio we can then use the steady-state intratemporal first-order condition

$$\frac{c^*}{1 - n^*} \frac{1 - \xi}{\xi} = (1 - \nu^*) F_n\left(\frac{k^*}{n^*}, 1\right)$$

and the resource constraint

$$c^* + g^* + \delta k^* = F(k^*, n^*)$$

to solve for c^*, k^*, and n^* separately, which are nonzero as long as the long-run labor tax is $\nu^* < 1$. □

In our economy, positive long-run capital taxation can be both optimal and lead to a steady-state where consumption, capital, and labor are all nonzero while the capital tax remains at its upper bound, $\bar{\tau}$.[16] The apparent difference of this result from those in Judd (1985) and Straub and Werning (2018) is due to the fact that the latter papers analyze a setting in which some agents ("workers") cannot save, so redistribution is valuable to them only to the extent that it adds to their consumption streams. The possibility of savings in our setting makes redistribution relatively more valuable, because poor households can decide how to allocate the extra resources over time. In fact, they may value redistribution so much more so as to prefer maximal capital taxes at a steady-state with positive capital stock.

3.3 Inelastic labor supply

Interestingly, if labor supply is inelastic, Theorem 4 can be strengthened. In this case, the condition is not only sufficient, but also necessary.

Theorem 6. *Suppose that $\xi = 1$, that is, the period utility function is $u(c_t^i, 1 - n_t^i) = \frac{(c_t^i)^{1-\sigma}}{1-\sigma}$ for all i. The capital tax sequence preferred by household i features $\mathsf{T}^i = \infty$ if and only if at the equilibrium induced by $(\tau_t, \nu_t) = (\bar{\tau}, 0)$ for all $t \geq 0$ the partial derivative is non-positive, $\frac{\partial V^i}{\partial V} \leq 0$.*

[16] Importantly, we are not looking at steady states induced by policies preferred by the *average* household. Indeed, because lump-sum taxes are available, equilibria in $\overline{\mathcal{T}}$ cannot be optimal for the average household.

Proof. For the if part see case 2 and 3 of Theorem 3 in Bassetto and Benhabib (2006).[17] For the other direction, note that if labor supply is inelastic, taxing labor is non-distortionary. Moreover, it is straightforward to see that labor taxes and government transfers cancel each other in the average household's budget constraint. Because the average household's decisions and disposable income are both independent of ν_t, the values V and A do not depend on the particular $\{\nu_t\}_{t=0}^{\infty}$ sequence either. As a result, all elements of $\overline{\mathcal{T}}$ induce $(\underline{V}, \underline{\mathsf{A}})$ with the lowest attainable V and A. Appendix B shows that the function V^i is concave, implying that in order for V^i to be maximized at $(\underline{V}, \underline{\mathsf{A}})$, we need downward sloping indifference curves that increase in the direction of lower V and A. As we saw in section 3.1, this requires $\frac{\partial V^i}{\partial V} \leq 0$ at $(\underline{V}, \underline{\mathsf{A}})$. □

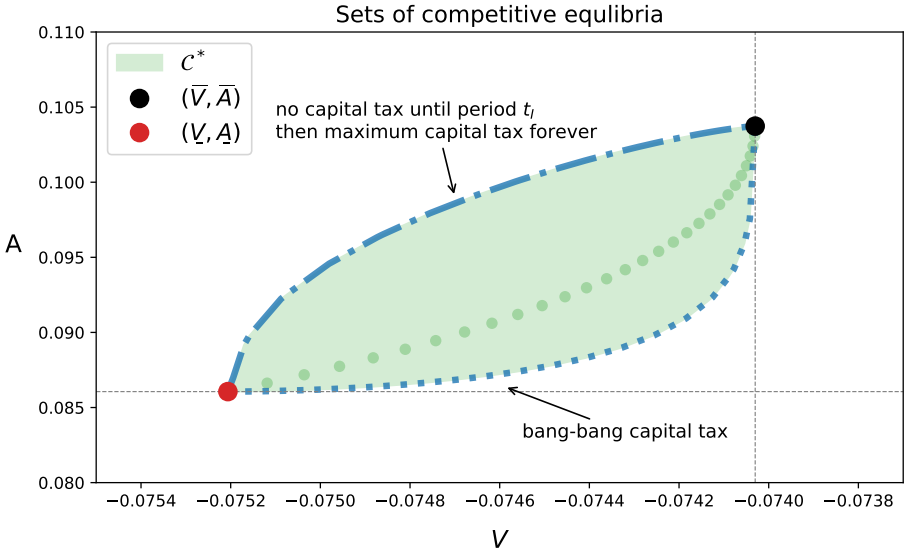

Figure 3: Subsets of competitive equilibria in the (V, A)-space with inelastic labor supply. Light green region represents equilibria induced by eventually time-invariant policies. Transparent green dots show equilibria induced by $(\tau_t, \nu_t) = (\tau^*, 0)$, $t \geq 0$ for various $\tau^* \leq \bar{\tau}$ values. The horizontal and vertical dashed gray lines represent the lowest attainable A and the highest attainable V, respectively. Parameters are as described in footnote 10, except for $\xi = 1$.

Figure 3 illustrates how the case with inelastic labor supply differs from our general setup. The key difference is that labor taxes cease to have any effect on the equilibrium objects V and A, so the set $\overline{\mathcal{T}}$ becomes a singleton consisting only of $(\underline{V}, \underline{\mathsf{A}})$. In addition, the set $\mathcal{C}^*$ is now bounded with its upper boundary being made up of equilibria such that the capital tax rate is zero up to a certain period t_I, then it is set at its maximum value forever. The dash-dotted blue line represents such equilibria for various t_I values with t_I increasing as we move from $(\underline{V}, \underline{\mathsf{A}})$ to $(\overline{V}, \overline{\mathsf{A}})$. Just like in Figure 1 the lower boundary of the set is determined by bang-bang capital tax policies. The dotted blue line represents such equilibria for various T values with T increasing as we move from $(\overline{V}, \overline{\mathsf{A}})$ to $(\underline{V}, \underline{\mathsf{A}})$. The particular shape of $\mathcal{C}^*$ implies that the wealth-rich households' preferred equilibrium coincides with that of the

[17] There are two offsetting inconsequential typos in Bassetto and Benhabib (2006) that can be corrected as follows: i) the second inequality on the top of page 220 should be reversed, and ii) in the 7th line of page 220 the whole expression for the change in the utility index preceeding the inequality sign on the right should be multiplied by a negative sign.

average household: because labor taxes have no redistributive effect, wealth-rich households have no reason to tolerate distortionary taxes.

3.4 Discussion

At first glance, our finding that wealth-poor households can prefer equilibria with positive steady-state capital tax rates seems to be at odds with some results in Judd (1985). Notice, however, that our economy differs from the model in Judd (1985) in which: (i) some households are excluded from the capital market, (ii) government is *not* allowed to issue debt.

Point (i) turns out to be the critical difference. Without the option to save, households have no choice but to consume their wages and transfers every period. Because both sources of their disposable income depend positively on the aggregate capital stock, the non-saving households' interests are naturally aligned with those of the wealth-rich "capitalists". The situation is different when everyone can save. In this case, wealth-poor households' can save their transfers with the aim of bootstrapping themselves out of poverty. A necessary condition for this channel to be operative is that savings respond *negatively* to permanent increases in future interest rates, that is, the income effect dominates the substitution effect because the intertemporal elasticity of substitution is lower than one, i.e., $\sigma > 1$.[18] In this case, households react to lower after-tax interest rates by choosing a steeper consumption path that implies faster capital accumulation and relatively more redistribution as the faster capital accumulation results in a larger tax base.[19]

As for point (ii), the lack of government debt in Judd (1985) implies that tax revenues cannot be saved, so the government is not tempted to impose a huge capital levy at the beginning of time. As a result, in Judd's model the upper bound on capital taxes does not play any critical role. When heterogeneity among households arise from differential capital holdings, however, the upper bound on capital taxes becomes essential to keep the problem nontrivial. For instance, if τ_0 were unrestricted, wealth-poor households would all prefer to confiscate the aggregate initial capital stock and redistribute it equally, thereby eliminating wealth inequality completely at the initial period.

Unlike us, Straub and Werning (2018) revisit the model in Judd (1985), in which some households are prohibited from saving, but similar to us, they find that if $\sigma > 1$, the optimal long-run capital tax can be positive. In this case, however, the non-saving household's consumption must converge to zero, so the steady-state cannot be interior. In contrast, Proposition 5 shows that if no one is excluded from the financial markets, taxing capital at the maximum rate forever does not rule out the possibility of an *interior* steady-state.

As we saw above, however, this interior steady-state result requires that $\bar{\tau} < 1$. This follows from the fact that with positive discounting, a steady-state cannot be compatible with a nonpositive after-tax interest rate, $(1 - \bar{\tau})r$ (see (16)). While in principle our main finding does not depend on the exact value of $\bar{\tau} < 1$, low $\bar{\tau}$ values tend to make our condition in Theorem 4 easier to be satisfied. Intuitively, the higher the upper bound the easier it is for the planner to concentrate all redistribution

[18] More precisely, the intertemporal elasticity of substitution associated with (1) is $IES = [1 - \xi(1 - \sigma)]^{-1}$. For any $\xi \in (0, 1]$, IES is lower (larger) than one if and only if $\sigma > 1$ ($\sigma < 1$).

[19] Faster capital accumulation also benefits the wealth-poor households by increasing the marginal product of labor.

at the early periods. The desirability of capital taxation is independent of $\bar{\tau}$, but with a more relaxed upper bound, the necessary amount of redistribution might be achieved with a lower T^i.

Because preferences over capital tax policies are directly linked to the household's initial wealth level, our model predicts that the implemented capital tax policy hinges on the wealth inequality in the economy. A simple measure of wealth inequality, or more precisely, the skewness of the wealth distribution, is the difference between the median and the average households' wealth levels. This difference is captured by (13) being applied to the median household. Denote this with Δa_0^m.[20] The partial derivative in our sufficient condition for $\mathsf{T}^m = \infty$ is a function of Δa_0^m: the higher the level of wealth inequality, the more likely it is that positive long-run capital taxation is optimal. Using recent Census Bureau data in section 4.2.2, we will show that U.S. wealth inequality seems to be high enough so that the sufficient condition for $\mathsf{T}^m = \infty$ is satisfied for a wide range of parameter values.

4 Quantitative examples

In general, the condition in Theorem 4 is hard to check, because it depends on the value function and optimal policies of the average household. To help us understand its content better, in Section 4.1 we provide a simple parametric example for which closed form solution exists, and therefore the condition can be expressed in terms of primitives. Moreover, to investigate the condition's empirical plausibility, Section 4.2 presents an alternative approach exploiting the fact that in equilibrium the value function of the representative agent in a neoclassical economy can be represented by a linear intertemporal budget constraint using time varying return and wage sequences consistent with the existence of interior steady-state equilibria. In a simplified version of our Section 2 economy, it is possible to express the value function V in terms of initial values and primitive parameters. That said, the condition in Theorem 4 can be evaluated using recent estimates of US wealth data. Finally, in Section 4.3 we solve the equilibrium objects numerically with standard functional forms and provide parameter ranges for which the condition is satisfied.

4.1 Special cases

CRRA utility with CES production

To obtain closed form solutions for the equilibrium objects (V, A) we follow Benhabib and Rustichini (1994) and assume that (i) the production function is of the CES form with parameter η, (ii) labor supply is inelastic ($\xi = 1$), and (iii) the households' intertemporal elasticity of substitution (IES) is reciprocal to the CES parameter of production. These assumptions give rise to saving policies that are linear in current income, hence, the average household's value function V can be solved in closed form. In more detail, let the production function be

$$F(k_t, n_t) = z\left(\rho k_t^{1-\eta} + (1-\rho) n_t^{1-\eta}\right)^{\frac{1}{1-\eta}}$$

[20] While the *tail* of the U.S. wealth distribution can be approximated by a Pareto distribution with tail index 1.5, the full distribution is certainly not Pareto. Nevertheless, if we were to use a full Pareto distribution with tail index $\varphi = 1.5$, we would get $\Delta a_0^m = 2^{1/\varphi}(\varphi - 1)/\varphi - 1 \approx -0.47$. For a very rough calibration to the U.S. economy, see section 4.2.2.

with $z > 0$, $\eta > 0$, and $\rho \in [0,1]$, and suppose that $\eta = \sigma > 1$. To simplify algebra, we assume that $(g_t, b_t) = (0,0)$ for all $t \geq 0$, implying $a_0 = k_0$ and a balanced government budget every period:

$$tr_t = \tau_t r_t k_t + \nu_t w_t$$

In Appendix C.1 we show that with full depreciation, $\delta = 1$,[21] the tax policy with constant rates $(\tau_t, \nu_t) = (\bar{\tau}, 0)\ \forall t \geq 0$, induces an equilibrium in which the optimal consumption is $c_t = \lambda F(k_t, n_t)$ with the marginal propensity to consume being:

$$\lambda = 1 - \left[\beta(1-\bar{\tau})z^{1-\sigma}\rho\right]^{\frac{1}{\sigma}}.$$

Plugging in the implied functions V, c_0, and n_0 into (12), taking the partial derivate with respect to V, and rearranging terms yield the following sufficient condition for $\mathsf{T}^m = \infty$:

$$\frac{1}{\sigma} \leq \mathsf{D}\left(-\Delta a_0^m\right) = \left(1 - \frac{\bar{\tau}}{\lambda}\right) \frac{\rho k_0^{1-\sigma}}{\rho k_0^{1-\sigma} + \frac{1-\rho}{1-\beta}} \left(\frac{a_0 - a_0^m}{a_0}\right)$$

where m denotes the household with median income, $a_0^m \leq a_0$. Intuitively, in order for this condition to be satisfied, the following objects should be relatively large: (1) wealth inequality measured by $\Delta a_0^m < 0$, (2) capital share in production technology, $\rho \in [0,1]$, and (3) marginal propensity to consumption, λ. With these objects at hand, computing the evolution of the growth rate of average capital is straightforward:

$$\frac{k_{t+1}}{k_t} = (1-\lambda)z\left(\rho + (1-\rho)\left(\frac{1}{k_t}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$$

One can obtain the corresponding steady-state value k^* by setting the growth rate equal to one. Figure 4 represents a particular example such that the (i) the sufficient condition for $\mathsf{T}^m = \infty$ is satisfied and (ii) the steady-state is interior, that is $0 < k^* < \infty$, $0 < c^* < \infty$. In addition, dashed lines (computed numerically) in Figure 4 illustrate that the steady state values change continuously with small perturbations of the parameters. In particular, even if we deviate from the case $\sigma = \eta$ for which closed form solution exists, the steady-state is still interior.

Linear production – sustained growth with time consistency

Another example of interest is when the production function is linear implying an endogenously growing economy. This involves setting $\xi = 1$ and $\rho = 1$, i.e., the production function is $y_t := F(k_t, n_t) = zk_t$. Households do not work, $n_t = 0$, and the optimal consumption becomes

$$c_t = \left(1 - \left[\beta(1-\bar{\tau})z^{1-\sigma}\right]^{\frac{1}{\sigma}}\right) y_t.$$

[21] A depreciation scheme can easily be introduced into this formulation as in Benhabib and Rustichini (1994) if current and past investments, depreciated over their lifetime according to a general depreciation profile, are aggregated within a CES production function.

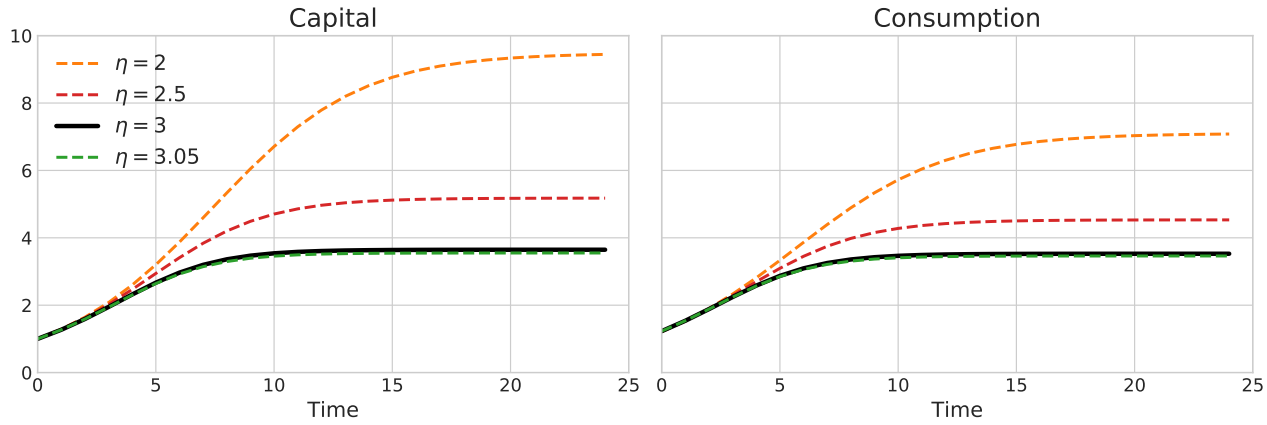

Figure 4: Equilibrium paths of capital and consumption induced by the policy $(\tau_t, \nu_t) = (\bar{\tau}, 0)$ for all $t \geq 0$. Black solid lines show the knife-edge case $\sigma = \eta$ for which closed form solution exists. Dashed lines illustrate how the equilibrium paths change when $\sigma \neq \eta$. As we move η from 2 to 3.05, the corresponding values of the partial derivative, $\partial V^i / \partial V$, are: -6.87, -1.57, -0.11, and -0.04. Parameter values are: $\beta = .96$, $\sigma = 3$, $\xi = 1$, $z = 2.5$, $\rho = 0.95$, $\delta = 1$, $\bar{\tau} = 0.1$, $\Delta a_0^m = -1$, $k_0 = 1$.

The sufficient condition in Theorem 4 simplifies to

$$\frac{1}{\sigma} \leq \left(1 - \frac{\bar{\tau}}{\lambda}\right) \left(\frac{a_0 - a_0^m}{a_0}\right)$$

whereas the constant growth rate of the economy is

$$\frac{k_{t+1}}{k_t} = \frac{y_{t+1}}{y_t} = \frac{c_{t+1}}{c_t} = (1 - \lambda) z = (\beta z)^{\frac{1}{\sigma}} (1 - \bar{\tau})^{\frac{1}{\sigma}}.$$

Clearly, there is nothing in this specification that would prevent $\frac{k_{t+1}}{k_t} > 1$. For given values of $(\beta, \bar{\tau})$, different z values can lead to either perpetual growth or perpetual contraction. This example illustrates that even if there is no interior steady-state, this does not imply that aggregate capital must converge to zero. Indeed, taxing capital at its maximum rate is consistent with sustained growth.

In this paper we follow the standard approach to studying optimal taxation by assuming full commitment. With linear production function, however, this is not necessary. With capital taxes being fixed at their upper bound, both the interest rate and the wealth shares, and thus our key sufficient condition $dV^i/dV < 0$, are invariant through time. As a result, the solution is time consistent. No commitment is required, the wealth-poor household chooses to stick with its initial plan.

4.2 Alternative characterization

In the examples of Section 4.1 we imposed strict parametric restrictions and focused on simple special cases in order to derive a condition—expressed in terms of primitives—under which $\mathsf{T}^i = \infty$. In this subsection we follow an alternative strategy. Using a simplified version of our Section 2 economy, we assume a general constant return to scale neoclassical production function with the existence of an

equilibrium characterized by returns $\{r_s\}_{s=0}^{\infty}$ and wages $\{w_s\}_{s=0}^{\infty}$, such that the function V is well-behaved. By manipulation the average household's budget constraint, we derive a formula for V and express our condition in terms of $\{r_s\}_{s=0}^{\infty}$ and $\{w_s\}_{s=0}^{\infty}$. To obtain a slight generalization of our previous findings, this subsection uses a hyperbolic absolute risk aversion (HARA) utility specification:[22]

$$u\left(c_t^i\right) = \left(\frac{\sigma}{1-\sigma}\right)\left[\frac{\psi}{\sigma}c_t^i + \bar{u}\right]^{1-\sigma} \tag{17}$$

4.2.1 Simplified setting

There are N households, each owning a share ω_t^i of the period-t aggregate capital stock, a_t, such that $\sum_{i=1}^{N}\omega_i = 1$ and $a_t^i = \omega_t^i a_t$, for $t \geq 0$. Suppose that labor is inelastically supplied ($\xi = 1$) and all households receive 'labor income' $\{e_t\}_{t\geq 0}$ irrespective of their wealth level. To guarantee finite budget as $t \to \infty$, we restrict the growth rate of labor income as

$$\chi_t \coloneqq e_t + \sum_{j=t+1}^{\infty}\frac{e_j}{\prod_{s=t+1}^{j} R_s} < \infty.$$

For simplicity, we assume that $\nu_t = 0$ and $g_t = 0$, but given that labor supply is inelastic and lump-sum taxes are allowed, this is without much loss of generality. More importantly, we assume balanced government budget every period, that is, $b_t = 0$, $t \geq 0$. As a result the only motive for using distorting capital taxes is wealth redistribution yielding the transfer: $tr_t = \tau_t r_t a_t$.

With these simplifications, household i's period budget constraint becomes

$$a_{t+1}^i = R_t a_t^i + e_t + tr_t - c_t^i =: R_t a_t^i - (c_t^i - d_t) \tag{18}$$

where d_t denotes the non-capital income agent i receives at time t. Iterating this constraint forward and combining it with the Euler equation, $\left(\frac{\psi}{\sigma}c_{t+1}^i + \bar{u}\right) = \left(\frac{\psi}{\sigma}c_t^i + \bar{u}\right)(\beta R_{t+1})^{\frac{1}{\sigma}}$, and the transversality condition, we can solve for

$$c_t^i = \lambda_t\left(R_t a_t^i + d_t + \sum_{j=t+1}^{\infty}\frac{d_j}{\prod_{s=t+1}^{j} R_s} - \frac{\sigma}{\psi}\bar{u}\zeta_t\right) \tag{19}$$

where

$$\zeta_t \coloneqq \sum_{j=t+1}^{\infty}\prod_{s=t+1}^{j}\frac{(\beta R_s)^{\frac{1}{\sigma}} - 1}{R_s} \qquad \text{and} \qquad \lambda_t \coloneqq \left(1 + \sum_{j=t+1}^{\infty}\prod_{s=t+1}^{j}\left(\beta R_s^{1-\sigma}\right)^{\frac{1}{\sigma}}\right)^{-1}.$$

The following assumption assures that there are $\lambda^l, \lambda^r \in \mathbb{R}$, such that $0 < \lambda^l \leq \lambda_t \leq \lambda^h < 1$ for all $t \geq 0$. Note that the assumption places no further restrictions on the tax rate in the initial period.

Assumption 2. $\beta < R_t^{\sigma-1}$ *for all* $t \geq 1$. *(Sufficient but not necessary for* $\lambda_t > 0$*)*

[22] The specification used throughout the paper is a special subclass associated with $\bar{u} = 0$ and $\psi = \sigma$.

Using the Euler equation and (19), the value function of household i can be written as

$$V^i = \frac{\sigma}{1-\sigma}\left[\sum_{t=0}^{\infty}\beta^t\left(\frac{\psi}{\sigma}c_t^i+\bar{u}\right)^{1-\sigma}\right] = \frac{\sigma}{1-\sigma}\left(\frac{\psi}{\sigma}c_0^i+\bar{u}\right)^{1-\sigma}\lambda_0^{-1} \tag{20}$$

We then postulate the law of motion of aggregate capital as

$$a_{t+1} = \varepsilon_{t+1}a_t + \gamma_{t+1} \tag{21}$$

and using this transition rule, we derive an expression for d_t and plug it into (19) to obtain

$$c_t^i = \lambda_t\left(\frac{R_t\omega_t^i}{r_t}+\tau_t\frac{1}{N}+\frac{1}{N}\sum_{j=t+1}^{\infty}\tau_j\prod_{s=t+1}^{j}\frac{\varepsilon_s}{R_s}\right)r_ta_t+\lambda_t\frac{f_t}{N}+\lambda_t\zeta_t.$$

Appendix C.2 contains formulas for the equilibrium processes $\{\varepsilon_t\}_{t\geq 0}$, $\{\gamma_t\}_{t\geq 0}$, and $\{f_t\}_{t\geq 0}$, but they will not be needed for our purposes.

4.2.2 Condition for $\mathsf{T}^m = \infty$

Let $\bar{a}_0$ and a_0^m denote the average and median initial wealth levels, respectively. Using the above formulas, we can write the derivative of the median household's value function with respect to V as

$$\frac{\partial V^m}{\partial V} = 1+\sigma\frac{u_c(0)R_0}{(1-\sigma)V}(a_0^m-\bar{a}_0) = 1+\psi\frac{R_0}{\left(\frac{\psi}{\sigma}c_0+\bar{u}\right)\lambda_0^{-1}}(a_0^m-\bar{a}_0). \tag{22}$$

In the isoelastic case, $\bar{u}=0$ and $\psi=\sigma$, this becomes

$$\frac{\partial V^m}{\partial V} = 1-\sigma\frac{R_0(\bar{a}_0-a_0^m)}{R_0\bar{a}_0+N^{-1}\left(\tau_0+\sum_{j=1}^{\infty}\tau_j\prod_{s=1}^{j}\frac{\varepsilon_s}{R_s}\right)r_0a_0+N^{-1}f_0},$$

where the denominator is the lifetime wealth of the household with average initial wealth level. It is the sum of the value of the after-tax return on average capital, the discounted transfers due to growth factor of capital, the discounted present value of labor income (via the term $N^{-1}f_0$), and the discounted value transfers accruing through the additive growth in capital.

To get a rough idea about what this partial derivative would be in the data, we use information from Table 1 in Wolff (2017) for the year 2016 to obtain:

$$\frac{\partial V^i}{\partial V} = 1-\sigma\left(\frac{[1+r(1-\bar{\tau})]\,(\$667,600-\$78,100)}{[1+r(1-\bar{\tau})]\,\$667,600+\$1,662,000}\right)$$

where $\$1,662,000$ in the denominator is lifetime individual mean earnings plus transfers (mean income of $\$83,100$ capitalized at 5%). The mean and median physical wealth levels are given by $\$667,600$ and $\$78,100$, respectively. These values allow us to compute $\sigma_{\min}$: the minimum σ that makes the

partial derivative equal to 0. Using the above expression with $r = 0.06$ and $\bar{\tau} = 0.3$[23] we obtain $\sigma_{\min} \approx 3.9$. Although the exact number depends on the interest rate and the upper bound on capital taxes, $\sigma_{\min}$ changes only slightly due to the high observed inequality in the data.[24] As a result, as long as $\sigma \geq \sigma_{\min}$, the sufficient condition in Theorem 4—applied to the median household—is satisfied.[25]

While this calibration assumes isoelastic utility and we derive a bound on σ, (22) shows that the sufficient condition for maximal capital taxation forever does not require constant IES.

4.3 Numerical example

To get a sense of how our sufficient condition for $\mathsf{T}^i = \infty$ depends on key model parameters, we turn to numerical methods and compute the equilibrium pair $(\underline{V}, \underline{\mathsf{A}})$ for a range of parameter values that is deemed plausible in the literature. In particular, we specify $F(k_t, n_t)$ to be Cobb-Douglas with capital share parameter ρ and use our isoelastic utility specification (1) parametrized by (σ, ξ). We then combine the computed equilibrium pairs with wealth inequality measured by Wolff (2017):

$$\Delta a_0^m = \frac{\$78,100 - \$667,600}{\$667,600} \approx -0.88$$

to obtain estimates for the partial derivative $\partial V^i / \partial V$.

As default parameterization, we use $\xi = 0.357$ to make the representative household work one-third of its time and $\rho = 1/3$ to get the standard capital income share. In addition, we set $\sigma = 4$ implying IES= 0.5. Regarding the other parameters, we use $\beta = 0.96$, $\bar{\tau} = 0.25$, and $\delta = 0.02$. Table 1 shows how the value of $\partial V^i / \partial V$ changes as we deviate from this default parametrization by varying σ (first column), ξ (second column), or ρ (third column). Recall that the condition in Theorem 4 is satisfied when the the partial derivative is nonpositive.

$(\sigma,\ \xi,\ \rho)$	$\partial V^i/\partial V$	$(\sigma,\ \xi,\ \rho)$	$\partial V^i/\partial V$	$(\sigma,\ \xi,\ \rho)$	$\partial V^i/\partial V$
(3, 0.4, 0.3)	0.471	(4, 0.2, 0.3)	0.714	(4, 0.4, 0.2)	-0.211
(4, 0.4, 0.3)	-0.045	(4, 0.4, 0.3)	-0.045	(4, 0.4, 0.3)	-0.045
(5, 0.4, 0.3)	-0.613	(4, 0.7, 0.3)	-1.904	(4, 0.4, 0.7)	0.736

Table 1: Sensitivity of the sufficient condition with respect to some key parameters. First column varies σ, second column varies ξ, third column varies ρ. Second row shows the default parametrization.

5 Concluding Remarks

We showed that in our heterogeneous agent economy preferred capital tax policies can be ranked according to the households' initial wealth level. Why does this matter? At the very least, depending

[23] Recall that $\bar{\tau}$ is the tax on capital income. The corresponding value if the tax rate applies to both capital and its income would be around 2%.

[24] In fact $\sigma_{\min}$ is an overestimate, because mean earnings is lower than mean income.

[25] These computations assume that the median agent is decisive. If the decisive agent were poorer than the median, the required minimum σ would be lower.

on the social welfare function, we can obtain quite different optimal capital tax policies. In general, the more weight the planner assigns to wealth-poor households, the longer the capital tax should remain at its maximum level, and under the condition of Theorem 4, it remains at its maximum forever.

Our results on interior steady-states with maximal capital taxes stem from redistributive considerations arising from heterogeneous capital holdings, and therefore differ from the representative agent model considered in Chamley (1986), in which the key motive for taxation is government-spending. In addition, we use Gorman aggregagable preferences where all agents, irrespective of their initial wealth, are allowed to save, so we also differ from the model of Judd (1985), in which the workers are not allowed to save. Finally, our model also differs from Werning (2007), who studies a similar economy except that households differ in their labor productivity (but not in their initial wealth) and finds that the optimal capital tax is always zero, while labor taxation is used to reduce inequality by channeling wealth from the more productive to the less productive households. In contrast, in our setting with identical labor productivities, taxing labor increases wealth inequality because leisure is a normal good.

Appendices

A Proofs

A.1 Proof of Lemma 1

Proof. We prove the lemma in two steps:

Part I: The equilibrium associated with the lowest A value must be induced by $(\tau_t, \nu_t) = (\bar{\tau}, 0)$ for all $t \geq 0$. The proof is by contradiction. Consider an equilibrium (V^*, A^*) induced by tax policies such that for some $N \geq 0$ either $\tau_N < \bar{\tau}$ or $\nu_N > 0$ (or both). We show that this equilibrium cannot have the lowest A value by constructing a *feasible* perturbation $(V^{**}, \mathsf{A}^{**})$ with $\mathsf{A}^{**} < \mathsf{A}^*$.

Let N be the first period in which either $\nu_N > 0$ or $\tau_N < \bar{\tau}$ (or both) and let M be the first period after N with $\tau_{M+1} > 0$. Using the intratemporal FOC in (7), the Euler equation between period $N-1$ and N can be written as

$$u_c\left(c_{N-1}, 1-n_{N-1}\right) = \underbrace{\frac{\beta[1+(1-\tau_N)F_k(k_N, n_N)-\delta]}{(1-\nu_N)F_n(k_N, n_N)}}_{\searrow \text{ in } k \qquad \nearrow \text{ in } n} u_{1-n}(c_N, 1-n_N) \tag{23}$$

where the arrows below the underbrace indicate that for a given pair (τ_N, ν_N), the first term on the right hand side is decreasing in k_N and increasing in n_N. Moreover, using the intratemporal FOC

$$c_t = \left(\frac{\xi}{1-\xi}\right)(1-\nu_t)F_n(k_t, n_t)(1-n_t)$$

accompanied with the fact that if $\sigma > 1$ and $\xi < 1$ then $u_{c(1-n)} < 0$ implies that for given k_t value, both marginal utilities u_c and u_{1-n} are increasing in n_t.

Note that $N > 0$, because otherwise we can easily decrease A by setting $(\tau_0, \nu_0) = (\bar{\tau}, 0)$. Now consider the following perturbation to the candidate equilibrium: in period $N = 0$, decrease $u_c(0)$ by decreasing n_0 (and increasing c_0) so that it leads to lesser capital accumulation, i.e., $dk_t < 0$, for $t \leq N$. This is possible, because $(\tau_t, \nu_t) = (\bar{\tau}, 0)$ for all $t < N$ and the reduced k_t necessarily reduces the first term on the right hand side of (23), so in order to keep the Euler equation satisfied $u_{1-n}(t)$ must decrease. The source of this perturbation is the increase in τ_N or the decreasing in ν_N (or both), which is feasible by assumption. To undo the effect on capital accumulation, we increase u_c in periods $N, \ldots, M$, by decreasing $\tau_M > 0$. The household reacts to this change in tax policy by decreasing capital accumulation before N leading to a first-order decrease in A^* which is a contradiction.

Part II: Within the set of equilibria with a particular $\mathsf{A} \geq \underline{\mathsf{A}}$, the one that minimizes the average household's value features maximum capital taxation forever. The proof is by contradiction. Consider an equilibrium (V^*, A^*) induced by a tax policy with $\tau_N < \bar{\tau}$ for some $N > 0$. We show that this equilibrium cannot minimize V over $\mathcal{C}(\mathsf{A}^*)$ by constructing a *feasible* perturbation $(V^{**}, \mathsf{A}^{**})$ such that $\mathsf{A}^{**} = \mathsf{A}^*$ and $V^{**} < V^*$.

Let N be the first period in which $\tau_N < \bar{\tau}$. Then reduce $u_c(N-1)$ proportionately by a factor

$d\Psi$ and increase $u_c(N), \ldots, u_c(M)$ by a corresponding (constant) factor $d\Theta$ so that feasibility remains satisfied, where M is the first period after N such that $\tau_M > 0$. This perturbation entails raising τ_N and reducing τ_M. Using the functional form assumptions, the required adjustments are

$$d\Psi = \frac{u_{cc}dc_{N-1} + u_{c(1-n)}d(1-n_{N-1})}{u_c} = -\left[1-\xi(1-\sigma)\right]\frac{dc_{N-1}}{c_{N-1}} + (1-\xi)(1-\sigma)\frac{d(1-n_{N-1})}{1-n_{N-1}}$$
$$d\Theta = -\left[1-\xi(1-\sigma)\right]\frac{dc_t}{c_t} + (1-\xi)(1-\sigma)\frac{d(1-n_t)}{1-n_t} \qquad N \leq t \leq M.$$

Because the perturbed allocation must be an equilibrium, the intratemporal FOC requires

$$\frac{dc_t}{c_t} = \frac{F_{nk}}{F_n}dk_t + \left[1 - \frac{F_{nn}(1-n_t)}{F_n}\right]\left(\frac{d(1-n_t)}{1-n_t}\right) - \left(\frac{d\nu_t}{1-\nu_t}\right) \qquad \forall t \geq 0 \tag{24}$$

with $dk_{N-1} = 0$. As a result, $d\Psi < 0$ implies $dc_{N-1} > 0$ and $d(1-n_{N-1}) > 0$, hence $dk_N < 0$. Moreover, we choose $d\Theta$ so that the perturbation leads to $dk_{M+1} = 0$. To this end, we pick dc_t and $d(1-n_t)$ such that for all $N \leq t \leq M$

(i) $F_n(t)d(1-n_t) + dc_t < 0$ [positive capital accumulation]

(ii) $dc_t < 0$, $d(1-n_t) > 0$ [lower labor supply due to reduced wages]

The latter is feasible due to $dk_t < 0$, or alternatively, we can *increase* labor taxes (see (24)) to ensure that both properties hold. In more detail, the implied change in capital is

$$dk_N = dF(N-1) + (1-\delta)dk_{N-1} - dc_{N-1} = -\left[F_n(N-1)d(1-n_{N-1}) + dc_{N-1}\right]$$
$$dk_{t+1} = -\sum_{j=N-1}^{t}\left(\prod_{s=j+1}^{t}\left[1+F_k(s)-\delta\right]\right)\left[F_n(j)d(1-n_j) + dc_j\right] \qquad N \leq t \leq M$$

Therefore, $dk_{M+1} = 0$ requires

$$0 = \left[F_n(N-1)d(1-n_{N-1}) + dc_{N-1}\right] + \sum_{j=N}^{M}\left(\prod_{s=N}^{j}\left[1+F_k(s)-\delta\right]^{-1}\right)\left[F_n(j)d(1-n_j) + dc_j\right].$$

Using the Euler equation between period $N-1$ and N:

$$\frac{1}{1+F_k(N)-\delta} \leq \frac{1}{1+(1-\tau_N)F_k(N)-\delta} = \frac{\beta u_c(N)}{u_c(N-1)}$$

along with (i), we obtain

$$0 \geq \left[F_n(N-1)d(1-n_{N-1}) + dc_{N-1}\right] + \sum_{j=N}^{M}\left(\frac{\beta^j u_c(j)}{\beta^{N-1}u_c(N-1)}\right)\left[F_n(j)d(1-n_j) + dc_j\right] \tag{25}$$

with the right hand side being strictly negative unless $\tau_N = 0$.

The effect of the perturbation on the average household's value is:

$$dV = \sum_{j=N-1}^{M} \beta^j \left[u_c(j)dc_j + u_{1-n}(j)d(1-n_j)\right] = \sum_{j=N-1}^{M} \beta^j u_c(j) \left[dc_j + (1-\nu_j)F_n(j)d(1-n_j)\right]$$
$$= \underbrace{\sum_{j=N-1}^{M} \beta^j u_c(j) \left[dc_j + F_n(j)d(1-n_j)\right]}_{\leq 0} - \underbrace{\sum_{j=N-1}^{M} \beta^j \left(\frac{\nu_j}{1-\nu_j}\right) u_{1-n}(j)d(1-n_j)}_{\geq 0}$$

where for the second equality we use the intratemporal FOC. The first term in the last line is non-positive due to (25), while the second term is non-negative because $d(1-n_t) > 0$ for all $N-1 \leq t \leq M$ by construction and it is strictly positive unless $\nu_t = 0$ for $N-1 \leq t \leq M$. As a result, the only case when dV is not strictly negative is when $\tau_N = 0$ and $\nu_t = 0$ for $N-1 \leq t \leq M$, i.e. when the unperturbed allocation maximizes $\sum_{t=N-1}^{M} \beta^t u(c_t)$ subject to the resource constraint and the initial and terminal values of capital, k_{N-1} and k_{M+1}. In this case, by strict concavity of utility, the perturbed allocation has a negative second-order effect on V^*, hence $V^{**} < V^*$.

By construction, the proposed perturbation keeps the intratemporal FOC and the resource constraint satisfied in all periods. The Euler equation is also satisfied, since the perturbation changes marginal utilities, in a way that leaves the marginal rate of substitution (MRS) between period j and $j+1$ unaffected when $N \leq j \leq M$. The change in the MRS between $N-1$ and N is achieved by raising τ_N, while the change between M and $M+1$ is achieved by reducting τ_M. Then, from Bellman's optimality principle it follows that with k_0, k_{N-1}, and k_{M+1} unchanged, the segment between 0 and N and the segment after M (with fixed tax policies) remain unperturbed. The only case when this does not imply $\mathsf{A}^{**} = \mathsf{A}^*$ is when $N = 1$. Nonetheless, in this case we can increase ν_0 to ensure $d\mathsf{A} = 0$:

$$\frac{d\nu_0}{1-\nu_0} = \underbrace{\left[1 - \frac{F_{nn}(1-n_0)}{F_n} + \frac{(1-\xi)(\sigma-1)}{[1-\xi(1-\sigma)]} + \frac{(1-\tau_0)(1-n_0)F_{kn}(0)}{[1+(1-\tau_0)F_k(0)-\delta][1-\xi(1-\sigma)]}\right]}_{>0} \left(\frac{d(1-n_0)}{1-n_0}\right)$$

so that the conclusion $dV < 0$ does not change. This is because the labor taxes that the average agent pays are exactly and fully returned as transfers, so the distortion induced by higher ν_0 further reduces V^*. As a result, we have $\mathsf{A}^{**} = \mathsf{A}^*$ and $V^{**} < V^*$ which is a contradiction.

□

A.2 Proof of Lemma 2

Proof. Consider first the equilibrium sequences of consumption, labor, and capital that maximize V and determine $(\overline{V}, \overline{\mathsf{A}})$. The solution must solve the first order conditions and resource constraint

$$u_c(c_t, 1-n_t) = \beta u_c(c_{t+1}, 1-n_{t+1})[1 + F_k(k_{t+1}, n_{t+1}) - \delta] \tag{26}$$
$$u_{1-n}(c_t, 1-n_t) = u_c(c_t, 1-n_t)F_n(k_t, n_t) \tag{27}$$
$$F(k_t, n_t) = c_t + g_t + k_{t+1} - (1-\delta)k_t \tag{28}$$

and hence involves setting $(\tau_t, \nu_t) = (0, 0)$ for all $t \geq 0$. Absent concern for redistribution, the average household has no incentive to distort the economy since lump-sum taxes are available.

For given (feasible) A^*, let $\{c_t^*, n_t^*, k_t^*\}_{t\geq 0}$ be a sequence such that (i) $\mathsf{A}(c_0^*, n_0^*, \tau_0) = \mathsf{A}^*$ and (ii) it maximizes V over $\mathcal{C}(\mathsf{A}^*)$. An equivalent statement of the lemma is

$$\begin{aligned} \text{if} \quad u_c(c_t^*, 1-n_t^*) > \beta \underline{R}_{t+1} u_c(c_{t+1}^*, 1-n_{t+1}^*) \quad \text{then} \\ u_c(c_s^*, 1-n_s^*) = \beta \overline{R}_{s+1} u_c(c_{s+1}^*, 1-n_{s+1}^*), \qquad \text{and} \\ u_{1-n}(c_s^*, 1-n_s^*) = F_n(k_s^*, n_s^*) u_c(c_s^*, 1-n_s^*) \qquad \forall s > t. \end{aligned}$$

Suppose this were not true. Then the sequence $\{c_s^*, n_s^*\}_{s=t+1}^{\infty}$ does not satisfy the necessary first-order conditions of maximizing $\sum_{s=t+1}^{\infty} \beta^s u(c_s, 1-n_s)$ subject to k_{t+1}^*. Because the proposed sequence makes the upper bound constraint for τ_{t+1} slack, this implies the existence of an alternative $\{c_s^{**}, n_s^{**}\}_{s=t+1}^{\infty}$, such that $\left(\{c_s^*, n_s^*\}_{s=0}^t, \{c_s^{**}, n_s^{**}\}_{s=t+1}^{\infty}\right)$ is a competitive equilibrium $(V^{**}, \mathsf{A}^{**})$, but such that, for a sufficiently small $\epsilon > 0$,

$$\sum_{s=t+1}^{\infty} \beta^s u(c_s^{**}, 1-n_s^{**}) = \sum_{s=t+1}^{\infty} \beta^s u(c_s^*, 1-n_s^*) + \epsilon$$

This implies that the new equilibrium has $\mathsf{A}^{**} = \mathsf{A}^*$, but $V^{**} > V^*$ which is a contradiction.

□

B Properties of V^i

Clearly, if $\sigma \neq 1$ and $a_0^i \neq a_0$, the function V^i is smooth. The first derivatives are:

$$\begin{aligned} \frac{\partial V^i}{\partial V} &= (1-\sigma)\left(\alpha^i\right)^{-\sigma}\left(-\Delta a_0^i \mathsf{D}\right) + \left(\alpha^i\right)^{1-\sigma} = \\ &= \left(\alpha^i\right)^{-\sigma}\left[-(1-\sigma)\Delta a_0^i \mathsf{D} + \alpha^i\right] = \left(\alpha^i\right)^{-\sigma}\left[1 + \sigma \mathsf{D} \Delta a_0^i\right] \end{aligned}$$

and

$$\frac{\partial V^i}{\partial \mathsf{A}} = (1-\sigma)\left(\alpha^i\right)^{-\sigma} \frac{\Delta a_0^i}{(1-\sigma)V} V = \left(\alpha^i\right)^{-\sigma} \Delta a_0^i \quad \Rightarrow \quad \operatorname{sign}\left(\frac{\partial V^i}{\partial \mathsf{A}}\right) = \operatorname{sign}\left(\Delta a_0^i\right).$$

The second derivatives are

$$\begin{aligned} \frac{\partial^2 V^i}{\partial V^2} &= (-\sigma)\left(\alpha^i\right)^{-\sigma-1}\left(\frac{-\mathsf{D}\Delta a_0^i}{V}\right)\left(1 + \sigma \mathsf{D} \Delta a_0^i\right) + \sigma\left(\alpha^i\right)^{-\sigma}\left(\frac{-\mathsf{D}\Delta a_0^i}{V}\right) = \\ &= \left(\alpha^i\right)^{-\sigma-1} \sigma\left[-(1+\sigma \mathsf{D}\Delta a_0^i) + \alpha^i\right]\left(\frac{-\mathsf{D}\Delta a_0^i}{V}\right) = \\ &= -\left(\alpha^i\right)^{-\sigma-1} \sigma(1-\sigma)^2 \frac{\mathsf{D}^2\left(\Delta a_0^i\right)^2}{(1-\sigma)V} < 0, \end{aligned}$$

and

$$\frac{\partial^2 V^i}{\partial \mathsf{A}^2} = -\left(\alpha^i\right)^{-\sigma-1}\left(\frac{\sigma\left(\Delta a_0^i\right)^2}{(1-\sigma)V}\right) < 0,$$

and

$$\begin{aligned}\frac{\partial^2 V^i}{\partial V \partial \mathsf{A}} &= (-\sigma)\left(\alpha^i\right)^{-\sigma-1}\left(\frac{\Delta a_0^i}{(1-\sigma)V}\right)(1+\sigma \mathsf{D}\Delta a_0^i) + \sigma\left(\alpha^i\right)^{-\sigma}\left(\frac{\Delta a_0^i}{(1-\sigma)V}\right) = \\ &= \left(\alpha^i\right)^{-\sigma-1}\sigma\left[-(1+\sigma \mathsf{D}\Delta a_0^i)+\alpha^i\right]\left(\frac{\Delta a_0^i}{(1-\sigma)V}\right) = \\ &= \left(\alpha^i\right)^{-\sigma-1}\left(\frac{\sigma\left(\Delta a_0^i\right)^2}{(1-\sigma)V}\right)(1-\sigma)\mathsf{D} \qquad \Rightarrow \qquad \operatorname{sign}\left(\frac{\partial^2 V^i}{\partial V \partial \mathsf{A}}\right) = \operatorname{sign}(1-\sigma).\end{aligned}$$

To show that the Hessian of V^i is negative semi-definite, we compute

$$\left(\frac{\partial^2 V^i}{\partial V^2}\right)\left(\frac{\partial^2 V^i}{\partial \mathsf{A}^2}\right) - \left(\frac{\partial^2 V^i}{\partial V \partial \mathsf{A}}\right)^2 = \left[\left(\alpha^i\right)^{-\sigma-1}\right]^2\left(\frac{\sigma\left(\Delta a_0^i\right)^2}{(1-\sigma)V}\right)^2\left[(1-\sigma)^2\mathsf{D}^2 - (1-\sigma)^2\mathsf{D}^2\right] = 0$$

C Derivations for Section 4

C.1 CES production function with CRRA utility

The production function is

$$F(k_t, n_t) = z\left(\rho k_t^{1-\eta} + (1-\rho)n_t^{1-\eta}\right)^{\frac{1}{1-\eta}}$$

with $z, \eta > 0$ and $\rho > 0$ and suppose that $\eta = \sigma > 1$. This production function implies the following competitive factor prices

$$\begin{aligned} r_t &= F_k(k_t, n_t) = z\rho k_t^{-\eta}\left(\rho k_t^{1-\eta} + (1-\rho)n_t^{1-\eta}\right)^{\frac{\eta}{1-\eta}} = \rho z^{1-\eta}\left(\frac{y_t}{k_t}\right)^{\eta} \\ w_t &= F_l(k_t, n_t) = z(1-\rho)n_t^{-\eta}\left(\rho k_t^{1-\eta} + (1-\rho)n_t^{1-\eta}\right)^{\frac{\eta}{1-\eta}} = (1-\rho)z^{1-\eta}\left(\frac{y_t}{n_t}\right)^{\eta}\end{aligned}$$

It is well-known that the CES production function satisfies the Inada conditions if $\sigma = \eta = 1$ which corresponds to the case of Cobb-Douglas production function with logarithmic utility. We require $\sigma > 1$, which implies

$$\lim_{k\to 0} F_k(k, n) = z\rho^{\frac{1}{1-\eta}} > 0 \qquad\qquad \lim_{k\to\infty} F_k(k, n) = 0.$$

For simplicity, let $\delta = 1$ and suppose that $g_t = 0$ and $b_t = 0$ for $t \geq 0$, implying that $a_0 = k_0$, so that the only motive for taxing is wealth redistribution. As a result, the government must keep balanced budget every period:

$$tr_t = \tau_t r_t k_t + \nu_t w_t$$

Consider the tax policy with constant rates $\tau_t = \bar{\tau}$ and $\nu_t = 0$, $\forall t \geq 0$. Guess that the optimal consumption is linear $c_t = \lambda y_t$. In this case, $k_{t+1} = (1-\delta)k_t + (1-\lambda)y_t$. Substituting this into the Euler equation implies

$$\begin{aligned}
c_t^{-\sigma} &= \beta c_{t+1}^{-\sigma}\left(1 + (1-\bar{\tau})F_k(k_{t+1}, n_{t+1}) - \delta\right) \\
\lambda^{-\sigma} y_t^{-\sigma} &= \beta(\lambda y_{t+1})^{-\sigma}(1-\bar{\tau})\rho z^{1-\eta}\left(\frac{y_{t+1}}{k_{t+1}}\right)^{\eta} \\
y_t^{-\sigma} &= \beta(1-\bar{\tau})\rho z^{1-\eta} y_{t+1}^{\eta-\sigma}\left((1-\lambda)y_t\right)^{-\eta} \\
\lambda &= 1 - \left[\beta(1-\bar{\tau})\rho z^{1-\sigma}\right]^{\frac{1}{\sigma}}
\end{aligned}$$

This leads to the following form of the value function

$$V(k_0) = \frac{z^{1-\sigma}\left(\upsilon_1 k_0^{1-\sigma} + \upsilon_2 n_0^{1-\sigma}\right)}{1-\sigma}.$$

We find υ_1 and υ_2 by plugging the guesses for c_t, n_t, and V into the Bellman equation

$$\begin{aligned}
z^{1-\sigma}(\upsilon_1 k_t^{1-\sigma} + \upsilon_2) &= (\lambda y_t)^{1-\sigma} + \beta z^{1-\sigma}\left(\upsilon_1[(1-\lambda)y_t]^{1-\sigma} + \upsilon_2\right) \\
\upsilon_1 k_t^{1-\sigma} + \upsilon_2 &= \lambda^{1-\sigma}(\rho k_t^{1-\sigma} + (1-\rho)) + \beta\left[\upsilon_1(1-\lambda)^{1-\sigma} z^{1-\sigma}(\rho k_t^{1-\sigma} + (1-\rho)) + \upsilon_2\right]
\end{aligned}$$

and matching coefficients to obtain:

$$\begin{aligned}
\upsilon_1 &= \frac{\lambda^{1-\sigma}\rho}{1 - (\beta\rho)\left[(1-\lambda)z\right]^{1-\sigma}} \\
\upsilon_2 &= \frac{1}{1-\beta}\left(\lambda^{1-\sigma}(1-\rho) + \beta\upsilon_1(1-\rho)[(1-\lambda)z]^{1-\sigma}\right)
\end{aligned}$$

Plugging in V, c_0, and n_0 into the partial derivative formula

$$\frac{\partial V^i}{\partial V} \leq 0 \qquad \Leftrightarrow \qquad 1 + \sigma\frac{\mathsf{A}}{(1-\sigma)V}\Delta a_0^i = 1 + \sigma\frac{c_0^{-\sigma}(1-\bar{\tau})F_k a_0}{z^{1-\sigma}\left(\upsilon_1 k_0^{1-\sigma} + \upsilon_2\right)}\Delta a_0^i \leq 0$$

so the sufficient condition for $\mathsf{T}^m = \infty$ becomes

$$\begin{aligned}
\frac{1}{\sigma} \leq \mathsf{D}\left(-\Delta a_0^m\right) &= \frac{(1-\bar{\tau})}{\lambda}(1 - \beta\rho[(1-\lambda)z]^{1-\sigma})\frac{\rho k_0^{1-\sigma}}{\rho k_0^{1-\sigma} + \frac{1-\rho}{1-\beta}}\left(-\Delta a_0^m\right) \\
&= \left(1 - \frac{\bar{\tau}}{\lambda}\right)\frac{\rho k_0^{1-\sigma}}{\rho k_0^{1-\sigma} + \frac{1-\rho}{1-\beta}}\left(\frac{a_0 - a_0^m}{a_0}\right)
\end{aligned}$$

where m denotes the household with median income, $a_0^m \leq a_0$.

C.2 Alternative characterization

We postulate the law of motion of aggregate capital as

$$a_{t+1} = \varepsilon_{t+1} a_t + \gamma_{t+1}. \tag{29}$$

Using this transition rule, we derive a formula for d_t and plug it into (19) to get

$$c_t^i = \lambda_t \left(\frac{R_t \omega_t^i}{r_t} + \tau_t \frac{1}{N} + \frac{1}{N} \sum_{j=t+1}^{\infty} \tau_j \prod_{s=t+1}^{j} \frac{\varepsilon_s}{R_s} \right) r_t a_t + \lambda_t \frac{f_t}{N}$$

where

$$f_t = e_t + \sum_{j=t+1}^{\infty} \left[e_j + \tau_j r_j \sum_{s=t+1}^{j} \gamma_s \left(\prod_{k=s+1}^{j} \varepsilon_k \right) \right] \prod_{s=t+1}^{j} R_s^{-1} \tag{30}$$

is discounted present value of labor income plus transfers accruing through the additive accumulation in capital. Plugging c_t^i into (18) and summing over all agents imply

$$a_{t+1} = \left[(1-\lambda_t) \frac{\overline{R}_t}{r_t} - \lambda_t \left(\sum_{j=t+1}^{\infty} \tau_j \prod_{s=t+1}^{j} \frac{\varepsilon_s}{R_s} \right) \right] r_t a_t + e_t - \lambda_t f_t$$

Therefore, the equilibrium relation describing growth rates for our economy is:

$$\varepsilon_{t+1} = (1-\lambda_t)(1 + F_k(k_t, 1) - \delta) - \lambda_t r_t \left(\sum_{j=t+1}^{\infty} \tau_j \prod_{s=t+1}^{j} \frac{\varepsilon_s}{R_s} \right) \tag{31}$$

and given $\{\varepsilon_t\}_{t\geq 0}$ in principle we could solve for $\{\gamma_t\}_{t\geq 0}$ and $\{f_t\}_{t\geq 0}$,[26] but their explicit solutions are not needed for out purposes.

[26] The analytical solution requires to use continued fractions. See Benhabib (2007)

References

Atkeson, Andrew, V. V. Chari, and Patrick J. Kehoe. 1999. "Taxing capital income: a bad idea." *Quarterly Review, Federal Reserve Bank of Minneapolis* 23 (3):3–17.

Bassetto, Marco and Jess Benhabib. 2006. "Redistribution, taxes, and the median voter." *Review of Economic Dynamics* 9 (2):211 – 223.

Benhabib, Jess. 2007. "Dynamic economic equilibrium under redistribution." *Journal of Difference Equations and Applications* 13 (2-3):107–119.

Benhabib, Jess and Aldo Rustichini. 1994. "A note on a new class of solutions to dynamic programming problems arising in economic growth." *Journal of Economic Dynamics and Control* 18 (3):807 – 813.

Chamley, Christophe. 1986. "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives." *Econometrica* 54 (3):607–622.

Chari, V. V., Lawrence J. Christiano, and Patrick J. Kehoe. 1994. "Optimal Fiscal Policy in a Business Cycle Model." *Journal of Political Economy* 102 (4):617–652.

Chari, V. V., Pedro Teles, and Juan Nicolini. 2016. "More on the Optimal Taxation of Capital." Tech. rep.

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

Erosa, Andrés and Martin Gervais. 2002. "Optimal Taxation in Life-Cycle Economies." *Journal of Economic Theory* 105 (2):338 – 369.

Hubbard, R. Glenn and Kenneth L Judd. 1987. "Liquidity Constraints, Fiscal Policy, and Consumption." Working Paper 943, National Bureau of Economic Research.

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

Kydland, Finn E. and Edward C. Prescott. 1982. "Time to Build and Aggregate Fluctuations." *Econometrica* 50 (6):1345–1370.

Lansing, Kevin J. 1999. "Optimal redistributive capital taxation in a neoclassical growth model." *Journal of Public Economics* 73 (3):423 – 453.

Reinhorn, Leslie J. 2019. "On optimal redistributive capital taxation." *Journal of Public Economic Theory* 0 (0):1–28.

Straub, Ludwig and Iván Werning. 2018. "Positive Long-Run Capital Taxation: Chamley-Judd Revisited." Harvard University.

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

Wolff, Edward N. 2017. "Household Wealth Trends in the United States, 1962 to 2016: Has Middle Class Wealth Recovered?" Working Paper 24085, National Bureau of Economic Research.