Positive Long Run Capital Taxation: Chamley-Judd Revisited*

Ludwig Straub
MIT

Iván Werning
MIT

June 2018

According to the Chamley-Judd result, capital should not be taxed in the long run. In this paper, we overturn this conclusion, showing that it does not follow from the very models used to derive it. For the main model in Judd (1985), we prove that the long run tax on capital is positive and significant, whenever the intertemporal elasticity of substitution is below one. For higher elasticities, the tax converges to zero but may do so at a slow rate, after centuries of high tax rates. The main model in Chamley (1986) imposes an upper bound on capital taxes. We provide conditions under which these constraints bind forever, implying positive long run taxes. When this is not the case, the long-run tax may be zero. However, if preferences are recursive and discounting is locally non-constant (e.g., not additively separable over time), a zero long-run capital tax limit must be accompanied by zero private wealth (zero tax base) or by zero labor taxes (first best). Finally, we explain why the equivalence of a positive capital tax with ever increasing consumption taxes does not provide a firm rationale against capital taxation.

Keywords: Optimal taxation; Capital taxation

1 Introduction

One of the most startling results in optimal tax theory is the famous finding by Chamley (1986) and Judd (1985). Although working in somewhat different settings, their conclu-

*This paper benefited from comments by Fernando Alvarez, VV Chari, Peter Diamond, Mikhail Golosov, Tom Phelan, Stefanie Stantcheva, the editor and four anonymous referees, and was very fortunate to count on research assistance from Greg Howard, Lucas Manuelli and Andrés Sarto. We also thank seminar participants at Harvard and the 8th Banco de Portugal Conference on Monetary Economics. All remaining errors are our own. This paper supersedes "A Reappraisal of Chamley-Judd Zero Capital Taxation Results" presented by Werning at the Society of Economic Dynamics' 2014 meetings in Toronto.

sions were strikingly similar: capital should go untaxed in any steady state. This implication, dubbed the Chamley-Judd result, is commonly interpreted as applying in the long run, taking convergence to a steady state for granted.[1] The takeaway is that taxes on capital should be zero, at least eventually.

Economic reasoning sometimes holds its surprises. The Chamley-Judd result was not anticipated by economists' intuitions, despite a large body of work at the time on the incidence of capital taxation and on optimal tax theory more generally. It represented a major watershed from a theoretical standpoint. One may even say that the result is puzzling, as witnessed by the fact that economists have continued to take turns putting forth various intuitions to interpret it, none definitive nor universally accepted.

Interpretation aside, a crucial issue is the result's applicability. Many have questioned the model's assumptions, especially that of infinitely-lived agents (e.g. Golosov and Werning, 2006; Banks and Diamond, 2010). Still others have set up alternative models, searching for different conclusions. These efforts notwithstanding, opponents and proponents alike acknowledge Chamley-Judd as one of the most important benchmarks in the optimal tax literature.

In this paper, we do not propose a new model or seek to take a stand on the appropriate model. Instead, we question the Chamley-Judd results by arguing that a zero long-run tax result does not follow even within the logic of these models. For both the models in Chamley (1986) and Judd (1985), we provide results showing a positive long-run tax when the intertemporal elasticity of substitution is less than or equal to one. We conclude that these models do not actually provide an unambiguous argument against long-run capital taxation. We discuss what went wrong with the original results, their interpretations and proofs.

Before summarizing our results in greater detail, it is useful to briefly recall the setups in Chamley (1986) and Judd (1985), where in the latter case we will specifically focus on the model in Judd (1985, Section 3).[2]

[1] To quote from a few examples, Judd (2002): "[...] setting τ_k equal to zero in the long run [...] various results arguing for zero long-run taxation of capital; see Judd (1985, 1999) for formal statements and analyses." Atkeson et al. (1999): "By formally describing and extending Chamley's (1986) result [...] This approach has produced a substantive lesson for policymakers: In the long run, in a broad class of environments, the optimal tax on capital income is zero." Phelan and Stacchetti (2001): "A celebrated result of Chamley (1986) and Judd (1985) states that with full commitment, the optimal capital tax rate converges to zero in the steady state." Saez (2013): "The influential studies by Chamley (1986) and Judd (1985) show that, in the long-run, optimal linear capital income tax should be zero."

[2] Judd (1985) also provides extensions to the model in Judd (1985, Section 3) that generally bring the setup somewhat closer to that in Chamley (1986). In particular, Judd (1985, Section 4-5) allow workers to save, capitalists to work and considers non-constant discounting a la Uzawa (1968). However, throughout the formal analysis in Judd (1985) the government is assumed to run a balanced budget, i.e. no government bonds are allowed. Interpolating our results for Judd (1985, Section 3) and Chamley (1986), we believe

Start with the similarities. Both papers assume infinitely-lived agents and take as given an initial stock of capital. Taxes are basically restricted to proportional taxes on capital and labor—lump-sum taxes are either ruled out or severely limited. To prevent expropriatory capital levies, the tax rate on capital is constrained by an upper bound.[3] Turning to differences, Chamley (1986) focused on a representative agent and assumed perfect financial markets, with unconstrained government debt. Judd (1985) emphasizes heterogeneity and redistribution in a two-class economy, with workers and capitalists. In addition, the model in Judd (1985) features financial market imperfections: workers do not save and the government balances its budget, i.e. debt is restricted to zero. As emphasized by Judd (1985), it is most remarkable that a zero long-run tax result obtains despite the restriction to budget balance.[4] Although extreme, imperfections of this kind may capture relevant aspects of reality, such as the limited participation in financial markets, the skewed distributions of wealth and a host of difficulties governments may face managing their debts or assets.[5]

We begin with the model in Judd (1985) and focus on situations where desired redistribution runs from capitalists to workers. Working with an isoelastic utility over consumption for capitalists, $U(C) = \frac{C^{1-\sigma}}{1-\sigma}$, we establish that when the intertemporal elasticity of substitution (IES) is below one, $\sigma > 1$, taxes rise and converge towards a positive limit tax, instead of declining towards zero. This limit tax is significant, driving capital to its lowest feasible level. Indeed, with zero government spending the lowest feasible capital stock is zero and the limit tax rate on wealth goes to 100%. The long-run tax is not only not zero, it is far from that.

The economic intuition we provide for this result is based on the anticipatory savings effects of future tax rates. When the IES is less than one, any anticipated increase in taxes leads to higher savings today, since the substitution effect is relatively small and dominated by the income effect. When the day comes, higher tax rates do eventually lower capital, but if the tax increase is sufficiently far off in the future, then the increased savings generate a higher capital stock over a lengthy transition. This is desirable, since it increases wages and tax revenue. To exploit such anticipatory effects, the optimum

similar conclusions apply for these variant models in Judd (1985, Section 4–5).

[3]Consumption taxes (Chamley, 1980; Coleman II, 2000) and dividend taxes with capital expenditure (investment) deductions (Abel, 2007) can mimic initial wealth expropriation. Both are disallowed.

[4]Because of the presence of financial restrictions and imperfections, the model in Judd (1985) does not fit the standard Arrow-Debreu framework, nor the optimal tax theory developed around it such as Diamond and Mirrlees (1971).

[5]Another issue may arise on the other end. Without constraints on debt, capitalists may become highly indebted or not own the capital they manage. The idea that investment requires "skin in the game" is popular in the finance literature and macroeconomic models with financial frictions (see Brunnermeier et al., 2012; Gertler and Kiyotaki, 2010, for surveys).

involves an increasing path for capital tax rates. This explains why we find positive tax rates that rise over time and converge to a positive value, rather than falling towards zero.

When the IES is above one, $\sigma < 1$, we verify numerically that the solution converges to the zero-tax steady state.[6] This also relies on anticipatory savings effects, working in reverse. However, we show that this convergence may be very slow, potentially taking centuries for wealth taxes to drop below 1%. Indeed, the speed of convergence is not bounded away from zero in the neighborhood of a unitary IES, $\sigma = 1$. Thus, even for those cases where the long-run tax on capital is zero, this property provides a misleading summary of the model's tax prescriptions.

We confirm our intuition based on anticipatory effects by generalizing our results for the Judd (1985) economy to a setting with arbitrary savings behavior of capitalists. Within this more general environment we also derive an inverse elasticity formula for the steady state tax rate, closely related to one in Piketty and Saez (2013). However, our derivation stresses that the validity of this formula requires sufficiently fast convergence to an interior steady state, a condition that we show fails in important cases.

We then turn to the representative agent Ramsey model studied by Chamley (1986). As is well appreciated, in this setting upper bounds on the capital tax rate are imposed to prevent expropriatory levels of taxation. We provide two sets of results.

Our first set of results show that in cases where the tax rate does converge to zero, there are other implications of the model, hitherto unnoticed. These implications undermine the usual interpretation against capital taxation. Specifically, if the optimum converges to a steady state where the bounds on tax rates are slack, we show that the tax is indeed zero. However, for recursive non-additive utility, we also show that this zero-tax steady state is necessarily accompanied by either zero private wealth—in which case the tax base is zero—or a zero tax on labor income—in which case the first best is achieved. This suggests that zero taxes on capital are attained only after taxes have obliterated private wealth or allowed the government to proceed without any distortionary taxation. Needless to say, these are not the scenarios typically envisioned when interpreting zero long-run tax results. Away from additive utility, the model simply does not justify a steady state with a positive tax on labor, a zero tax on capital and positive private wealth.

Our second set of results show that the tax rate may not converge to zero. In particular, we show that the upper bounds imposed on the tax rate may bind forever, implying a positive long-run tax on capital. We prove that this is guaranteed if the IES is below one and debt is high enough. Importantly, the debt level required is below the peak of the

[6] We complement these numerical results by proving a local convergence result around the zero-tax steady state when $\sigma < 1$.

Laffer curve, so this result is not driven by budgetary necessity: the planner chooses to tax capital indefinitely, but is not compelled to do so. Intuitively, higher debt leads to higher labor taxes, making capital taxation attractive to ease the labor tax burden. However, because the tax rate on capital is capped, the only way to expand capital taxation is to prolong the time spent at the bound. At some point, for high enough debt, indefinite taxation becomes optimal.

All of these results run counter to what is certainly by now established wisdom, cemented by a significant follow-up literature, extending and interpreting long-run zero tax results. In particular, Judd (1999) presents an argument against positive capital taxation without requiring convergence to a steady state, using a representative agent model without financial market imperfections, similar in this regard to Chamley (1986). However, as we explain, these arguments invoke assumptions on endogenous multipliers that may be violated at the optimum. We also explain why the intuition offered in that paper, based on the observation that a positive capital tax is equivalent to a rising tax on consumption, does not provide a rationale against indefinite capital taxation.

To conclude, we present a hybrid model that combines heterogeneity and redistribution as in Judd (1985), but allows for government debt as in Chamley (1986). Capital taxation turns out to be especially potent in this setting: whenever the IES is less than one, the optimal policy sets the tax rate at the upper bound forever. This suggests that positive long-run capital taxation should be expected for a wide range of models that are descendants of Chamley (1986) and Judd (1985).

Related Literature. Aside from a long literature finding different kinds of zero capital tax results,[7] our paper is part of a strand of papers that find positive or negative long-run capital taxes can be optimal.[8] Almost all of these papers find this result by modifying the environment, moving away from the setups in Chamley (1986) and Judd (1985), with the following exceptions: Lansing (1999) considered a special case of the setup in Judd (1985, Section 3) and found positive long-run capital taxes are possible (see our discussion in Section 2); and Reinhorn (2002 and 2013) clarified the discrepancy with Judd (1985, Section 3). Bassetto and Benhabib (2006) study capital taxation in a median-voter political

[7]For papers with exogenous growth see e.g. Chamley (1980, 1986); Judd (1985, 1999); Atkeson et al. (1999); Chari and Kehoe (1999) . For papers with endogenous growth see e.g. Lucas (1990); Jones et al. (1993, 1997). For results with uncertainty see e.g. Zhu (1992); Judd (1993); Chari et al. (1994). For results with heterogeneous agents see e.g. Werning (2007); Greulich et al. (2016).

[8]For results on capital taxation in OLG models, see e.g. Erosa and Gervais (2002). For models with social weights on future periods/generations, see e.g. Farhi and Werning (2010, 2013). For results with limited commitment, see e.g. Chari and Kehoe (1990); Stokey (1991); Farhi et al. (2012). For models with incomplete markets and idiosyncratic risk, see e.g. Aiyagari (1995); Conesa et al. (2009).

economy model where agents are heterogeneous with respect to initial wealth. Their main result provides a median-voter theorem and a "bang-bang property" for capital taxes. For a case with AK technology, linear in capital and no labor, they give a condition for the median voter to prefer indefinite capital taxation. Finally, several authors study a variant of the Chamley (1986) economy where capital tax bounds are only imposed in the initial period, to limit expropriation, but not imposed in later periods, see e.g. Chari et al. (1994), Chari and Kehoe (1999), Sargent and Ljungqvist (2004) and Werning (2007). Our analysis does not apply in these cases; indeed, as these studies correctly show, with additively separable and isoelastic preferences over consumption, the capital tax is zero after the second period.

2 Capitalists and Workers

We start with the two-class economy without government debt laid out in Judd (1985). Time is indefinite and discrete, with periods labeled by $t = 0, 1, 2, \ldots$.[9] There are two types of agents, workers and capitalists. Capitalists save and derive all their income from the returns to capital. Workers supply one unit of labor inelastically and live hand to mouth, consuming their entire wage income plus transfers. The government taxes the returns to capital to pay for transfers targeted to workers.

Preferences. Both capitalists and workers discount the future with a common discount factor $\beta < 1$. Workers have a constant labor endowment $n = 1$; capitalists do not work. Consumption by workers will be denoted by lowercase c, consumption by capitalists by uppercase C. Capitalists have utility

$$\sum_{t=0}^{\infty} \beta^t U(C_t) \qquad \text{with} \qquad U(C) = \frac{C^{1-\sigma}}{1-\sigma}$$

for $\sigma > 0$ and $\sigma \neq 1$, and $U(C) = \log C$ for $\sigma = 1$. Here $1/\sigma$ denotes the (constant) intertemporal elasticity of substitution (IES). Workers have utility

$$\sum_{t=0}^{\infty} \beta^t u(c_t)$$

where u is increasing, concave, continuously differentiable and $\lim_{c \to 0} u'(c) = \infty$.

[9] Judd (1985) formulates the model in continuous time, but this difference is immaterial. As usual, the continuous-time model can be thought of as a limit of the discrete time one as the length of each period shrinks to zero.

Technology. Output is obtained from capital and labor using a neoclassical constant returns production function $F(k_t, n_t)$ satisfying standard conditions.[10] Capital depreciates at rate $\delta > 0$. In equilibrium $n_t = 1$, so define $f(k) = F(k, 1)$. The government consumes a constant flow of goods $g \geq 0$. We normalize both populations to unity and abstract from technological progress and population growth. The resource constraint in period t is then

$$c_t + C_t + g + k_{t+1} \leq f(k_t) + (1 - \delta)k_t.$$

There is some given positive level of initial capital, $k_0 > 0$.

Markets and Taxes. Markets are perfectly competitive, with labor being paid wage $w_t^* = F_n(k_t, n_t)$ and the before-tax return on capital being given by

$$R_t^* = f'(k_t) + 1 - \delta.$$

The after-tax return equals R_t and can be parameterized as either

$$R_t = (1 - \tau_t)(R_t^* - 1) + 1 \qquad \text{or} \qquad R_t = (1 - \mathcal{T}_t)R_t^*,$$

where τ_t is the tax rate on the net return to wealth and $\mathcal{T}_t$ the tax rate on the gross return to wealth, or wealth tax for short. Whether we consider a tax on net returns or on gross returns is irrelevant and a matter of convention. We say that capital is taxed whenever $R_t < R_t^*$ and subsidized whenever $R_t > R_t^*$.

Capitalist and Worker Behavior. Capitalists solve

$$\max_{\{C_t, a_{t+1}\}} \sum_{t=0}^{\infty} \beta^t U(C_t) \quad \text{s.t.} \quad C_t + a_{t+1} = R_t a_t \quad \text{and} \quad a_{t+1} \geq 0,$$

for some given initial wealth a_0. The associated Euler equation and transversality conditions,

$$U'(C_t) = \beta R_{t+1} U'(C_{t+1}) \qquad \text{and} \qquad \beta^t U'(C_t) a_{t+1} \to 0,$$

are necessary and sufficient for optimality.

[10] We assume that F is increasing and strictly concave in each argument, continuously differentiable, and satisfying the standard Inada conditions $F_k(k, 1) \to \infty$ as $k \to 0$ and $F_k(k, 1) \to 0$ as $k \to \infty$. Moreover assume that capital is essential for production, that is, $F(0, n) = 0$ for all n.

Workers live hand to mouth, their consumption equals their disposable income

$$c_t = w_t^* + T_t = f(k_t) - f'(k_t)k_t + T_t,$$

which uses the fact that $F_n = F - F_k k$. Here $T_t \in \mathbb{R}$ represent government lump-sum transfers (when positive) or taxes (when negative) to workers.[11]

Government Budget Constraint. As in Judd (1985), the government cannot issue bonds and runs a balanced budget. This implies that total wealth equals the capital stock $a_t = k_t$ and that the government budget constraint is

$$g + T_t = (R_t^* - R_t)\, k_t.$$

Planning Problem. Using the Euler equation to substitute out R_t, the planning problem can be written as[12]

$$\max_{C_{-1},\{c_t,C_t,k_{t+1}\}} \sum_{t=0}^{\infty} \beta^t (u(c_t) + \gamma U(C_t)), \tag{1a}$$

subject to

$$c_t + C_t + g + k_{t+1} = f(k_t) + (1-\delta)k_t, \tag{1b}$$

$$\beta U'(C_t)(C_t + k_{t+1}) = U'(C_{t-1})k_t, \tag{1c}$$

$$\beta^t U'(C_t)k_{t+1} \to 0. \tag{1d}$$

The government maximizes a weighted sum of utilities with weight γ on capitalists. By varying γ one can trace out points on the constrained Pareto frontier and characterize their associated policies. We often focus on the case with no weight on capitalists, $\gamma = 0$, to ensure that desired redistribution runs from capitalists towards workers. Equation (1b) is the resource constraint. Equation (1c) combines the capitalists' first-order condition and budget constraint and (1d) imposes the transversality condition; together conditions (1c) and (1d) ensure the optimality of the capitalists' saving decision.

[11] Equivalently, one can set up the model without lump-sum transfers/taxes to workers, but allowing for a proportional tax or subsidy on labor income. Such a tax perfectly targets workers without creating any distortions, since labor supply is perfectly inelastic in the model.

[12] Judd (1985) includes upper bounds on the taxation of capital, which we have omitted because they do not play any important role. As we shall see, positive long run taxation is possible even without these constraints; adding them would only reinforce this conclusion. Upper bounds on taxation play a more crucial role in Chamley (1986).

The necessary first-order conditions are

$$\mu_0 = 0, \tag{2a}$$

$$\lambda_t = u'(c_t), \tag{2b}$$

$$\mu_{t+1} = \mu_t \left(\frac{\sigma - 1}{\sigma \kappa_{t+1}} + 1 \right) + \frac{1}{\beta \sigma \kappa_{t+1} \upsilon_t} \left(1 - \gamma \upsilon_t\right), \tag{2c}$$

$$\frac{u'(c_{t+1})}{u'(c_t)} (f'(k_{t+1}) + 1 - \delta) = \frac{1}{\beta} + \upsilon_t(\mu_{t+1} - \mu_t), \tag{2d}$$

where $\kappa_t \equiv k_t/C_{t-1}$, $\upsilon_t \equiv U'(C_t)/u'(c_t)$ and the multipliers on constraints (1b) and (1c) are $\beta^t \lambda_t$ and $\beta^t \mu_t$, respectively.[13] Here, (2a) follows from the first order condition with respect to C_{-1}.

2.1 Previous Steady State Results

Judd (1985, pg. 72, Theorem 2) provided a zero-tax result, which we adjust in the following statement to stress the need for the steady state to be interior and for multipliers to converge.

Theorem 1 (*Judd, 1985*)**.** *Suppose quantities and multipliers converge to an interior steady state, i.e. c_t, C_t, k_{t+1} converge to positive values, and μ_t converges. Then the tax on capital is zero in the limit: $\mathcal{T}_t = 1 - R_t/R^*_t \to 0$.*

The proof is immediate: from equation (2d) we obtain $R^*_t \to 1/\beta$, while the capitalists' Euler equation requires that $R_t \to 1/\beta$. The simplicity of the argument follows from strong assumptions placed on endogenous outcomes. This raises obvious concerns. By adopting assumptions that are close relatives of the conclusions, one may wonder if anything of use has been shown, rather than assumed. We elaborate on a similar point in Section 3.3.

In our rendering of Theorem 1, the requirement that the steady state be interior is important: otherwise, if $c_t \to 0$ one cannot guarantee that $u'(c_{t+1})/u'(c_t) \to 1$ in equation (2d). Likewise, even if the allocation converges to an interior steady state but μ_t does not converge, then $\upsilon_t(\mu_{t+1} - \mu_t)$ may not vanish in equation (2d). Thus, the two situations that prevent the theorem's application are: (i) non-convergence to an interior steady state; or (ii) non-convergence of $\mu_{t+1} - \mu_t$ to zero. In general, one expects that (i) implies (ii). The literature has provided an example of (ii) where the allocation does converge to an interior steady state.

[13] We chose the sign of λ_t in the conventional way and the sign of μ_t such that the term in the current value Lagrangian is given by $\mu_t \left(\beta U'(C_t)(C_t + k_{t+1}) - U'(C_{t-1})k_t\right)$.

Theorem 2. (Lansing, 1999; Reinhorn, 2002 and 2013) *Assume $\sigma = 1$. Suppose the allocation converges to an interior steady state, so that c_t, C_t and k_{t+1} converge to strictly positive values. Then,*

$$\mathcal{T}_t \to \frac{1-\beta}{1+\gamma v\beta/(1-\gamma v)},$$

where $v = \lim v_t$ and the multiplier μ_t in the system of first-order conditions (2c) does not converge. This implies a positive long-run tax on capital if redistribution towards workers is desirable, $1 - \gamma v > 0$.

The result follows easily by combining (2c) and (2d) for the case with $\sigma = 1$ and comparing it to the capitalist's Euler equation, which requires $R_t = \frac{1}{\beta}$ at a steady state. Lansing (1999) first presented the logarithmic case as a counterexample to Judd (1985). Reinhorn (2002 and 2013) correctly clarified that in the logarithmic case the Lagrange multipliers explode, explaining the difference in results.[14]

Lansing (1999) depicts the result for $\sigma = 1$ as a knife-edge case: "the standard approach to solving the dynamic optimal tax problem yields the wrong answer in this (knife-edge) case [...]" (from the Abstract, page 423) and "The counterexample turns out to be a knife-edge result. Any small change in the capitalists' intertemporal elasticity of substitution away from one (the log case) will create anticipation effects [...] As capitalists' intertemporal elasticity of substitution in consumption crosses one, the trajectory of the optimal capital tax in this model undergoes an abrupt change." (page 427) Lansing (1999) suggests that whenever $\sigma \neq 1$ the long-run tax on capital is zero. We shall show that this is not the case.

2.2 Main Result: Positive Long-Run Taxation

Logarithmic Utility. Before studying $\sigma > 1$, our main case of interest, it is useful to review the special case with logarithmic utility, $\sigma = 1$. We assume $\gamma = 0$ to guarantee that desired redistribution runs from capitalists to workers.

When $U(C) = \log C$ capitalists save at a constant rate $s > 0$,

$$C_t = (1-s)R_t k_t \qquad \text{and} \qquad k_{t+1} = sR_t k_t.$$

Although $s = \beta$ with logarithmic preferences, nothing we will derive depends on this

[14] Lansing (1999) suggests a technical difficulty with the argument in Judd (1985) that is specific to $\sigma = 1$. Indeed, at $\sigma = 1$ one degree of freedom is lost in the planning problem, since C_{t-1} must be proportional to k_t. However, since equations (2) can still satisfied by the optimal allocation for some sequence of multipliers, we believe the issue can be framed exactly as Reinhorn (2002 and 2013) did, emphasizing the non convergence of multipliers.

fact, so we can interpret s as a free parameter that is potentially divorced from β.[15]

The planning problem becomes

$$\max_{\{c_t, k_{t+1}\}} \sum_{t=0}^{\infty} \beta^t u(c_t) \quad \text{s.t.} \quad c_t + \frac{1}{s} k_{t+1} + g = f(k_t) + (1-\delta) k_t,$$

with k_0 given. This amounts to an optimal neoclassical growth problem, where the price of capital equals $\frac{1}{s} > 1$ instead of the actual unit cost. The difference arises from the fact that capitalists consume a fraction $1-s$. The government and workers must save indirectly through capitalists, entrusting them with resources today by holding back on current taxation, so as to extract more tomorrow. From their perspective, technology appears less productive because capitalists feed off a fraction of the investment. Lower saving rates s increase this inefficiency.[16]

Since the planning problem is equivalent to a standard optimal growth problem, we know that there exists a unique interior steady state and that it is globally stable. The modified golden rule at this steady state is $\beta s R^* = 1$. A steady state also requires $sRk = k$, or simply $sR = 1$. Putting these conditions together gives $R/R^* = \beta < 1$.

Proposition 1. *Suppose $\gamma = 0$ and that capitalists have logarithmic utility, $U(C) = \log C$. Then the solution to the planning problem converges monotonically to a unique steady state with a positive tax on capital given by $\mathcal{T} = 1 - \beta$.*

This proposition echoes the result in Lansing (1999), as summarized by Theorem 2, but also establishes the convergence to the steady state. Interestingly, the long-run tax rate depends only on β, not on the savings rate s or other parameters.

Although Lansing (1999) and the subsequent literature interpreted this result as a knife-edge counterexample, we will argue that this is not the case, that positive long run taxes are not special to logarithmic utility. One way to proceed would be to exploit continuity of the planning problem with respect to σ to establish that for any fixed time t, the tax rate $\mathcal{T}_t(\sigma)$ converges as $\sigma \to 1$ to the tax rate obtained in the logarithmic case (which we know is positive for large t). While this is enough to dispel the notion that the logarithmic utility case is irrelevant for $\sigma \neq 1$, it has its limitations. As we shall see, the convergence is not uniform and one cannot invert the order of limits: $\lim_{t\to\infty} \lim_{\sigma\to 1} \mathcal{T}_t(\sigma)$

[15] This could capture different discount factors between capitalists and workers or an ad hoc behavioral assumption of constant savings, as in the standard Solow growth model. We pursue this line of thought in Section 2.3 below.

[16] This kind of wedge in rates of return is similar to that found in countless models where there are financial frictions between "experts" able to produce capital investments and "savers". Often, these models are set up with a moral hazard problem, whereby some fraction of the investment *returns* must be kept by experts, as "skin in the game" to ensure good behavior.

does not equal $\lim_{\sigma\to 1}\lim_{t\to\infty}\mathcal{T}_t(\sigma)$. Therefore, arguing by continuity does not help characterize the long run tax rate $\lim_{t\to\infty}\mathcal{T}_t(\sigma)$ as a function of σ. We proceed by tackling the problem with $\sigma\neq 1$ directly.

Positive Long-Run Taxation: IES < 1. We now consider the case with $\sigma > 1$ so that the intertemporal elasticity of substitution $\frac{1}{\sigma}$ is below unity. We continue to focus on the situation where no weight is placed on capitalists, $\gamma = 0$. Section 2.4 shows that the same results apply for other value of γ, as long as redistribution from capitalists to workers is desired.

Towards a contradiction, suppose the allocation were to converge to an interior steady state $k_t \to k$, $C_t \to C$, $c_t \to c$ with $k, C, c > 0$. This implies that κ_t and v_t also converge to positive values, κ and v. Combining equations (2c) and (2d) and taking the limit for the allocation, we obtain

$$f'(k) + 1 - \delta = \frac{1}{\beta} + v(\mu_t - \mu_{t-1}) = \frac{1}{\beta} + \mu_t \frac{\sigma - 1}{\sigma\kappa} v + \frac{1}{\beta\sigma\kappa}.$$

Since $\sigma > 1$, this means that μ_t must converge to

$$\mu = -\frac{1}{(\sigma - 1)\beta v} < 0. \tag{3}$$

Now consider whether $\mu_t \to \mu < 0$ is possible. From the first-order condition (2a) we have $\mu_0 = 0$. Also, from equation (2c), whenever $\mu_t \geq 0$ then $\mu_{t+1} \geq 0$. It follows that $\mu_t \geq 0$ for all $t = 0, 1, \ldots$, a contradiction to $\mu_t \to \mu < 0$.[17] This proves that the solution cannot converge to any interior steady state, including the zero-tax steady state.

Proposition 2. *If $\sigma > 1$ and $\gamma = 0$, no solution to the planning problem converges to the zero-tax steady state, or any other interior steady state.*

It follows that *if the optimal allocation converges,* then either $k_t \to 0$, $C_t \to 0$ or $c_t \to 0$. With positive spending $g > 0$, $k_t \to 0$ is not feasible; this also rules out $C_t \to 0$, since capitalists cannot be starved while owning positive wealth.

Thus, provided the solution converges, $c_t \to 0$. This in turn implies that either $k_t \to k_g$ or $k_t \to k^g$ where $k_g < k^g$ are the two solutions to $\frac{1}{\beta}k + g = f(k) + (1-\delta)k$, using the fact that (1c) implies $C = \frac{1-\beta}{\beta}k$ at any steady state.[18] We next show that the solution does

[17] Notice that this reasoning only relies on the necessity of the first-order conditions. Therefore, it also applies when the planning problem (1a) is non-convex.

[18] Here we assume that government spending g is feasible, that is, $g < \max_k\{f(k) + (1-\delta)k - \frac{1}{\beta}k\}$.

indeed converge, and that it does so towards the lowest sustainable value of capital, k_g, so that the long-run tax on capital is strictly positive. The proof uses the fact that $\mu_t \to \infty$ and $c_t \to 0$, as argued above, but requires many other steps detailed in the appendix.

Proposition 3. *If $\sigma > 1$ and $\gamma = 0$ then any solution to the planning problem converges to $c_t \to 0$, $k_t \to k_g$, $C_t \to \frac{1-\beta}{\beta} k_g$, with a positive limit tax on wealth: $\mathcal{T}_t = 1 - \frac{R_t}{R_t^*} \to \mathcal{T}^g > 0$. The limit tax $\mathcal{T}^g$ is decreasing in spending g, with $\mathcal{T}^g \to 1$ as $g \to 0$.*

The zero-tax interpretation of Judd (1985) is invalidated here because the allocation does not converge to an interior steady state and multipliers do not converge. According to our result, the tax rate not only does not converge to zero, it reaches a sizable level. Perhaps counterintuitively, the long-run tax on capital, $\mathcal{T}^g$, is inversely related to the level of government spending, since k_g is increasing with spending g. This underscores that long-run capital taxation is not driven by budgetary necessity.

As the proposition shows, optimal taxes may reach very high levels. Up to this point, we have placed no limits on tax rates. It may be of interest to consider a situation where the planner is further constrained by an upper bound on the tax rate for net returns (τ) or gross wealth ($\mathcal{T}$), perhaps due to evasion or political economy considerations. If these bounds are sufficiently tight to be binding, it is natural to conjecture that the optimum converges to these bounds, and to an interior steady state allocation with a positive limit for worker consumption, $\lim_{t\to\infty} c_t > 0$.

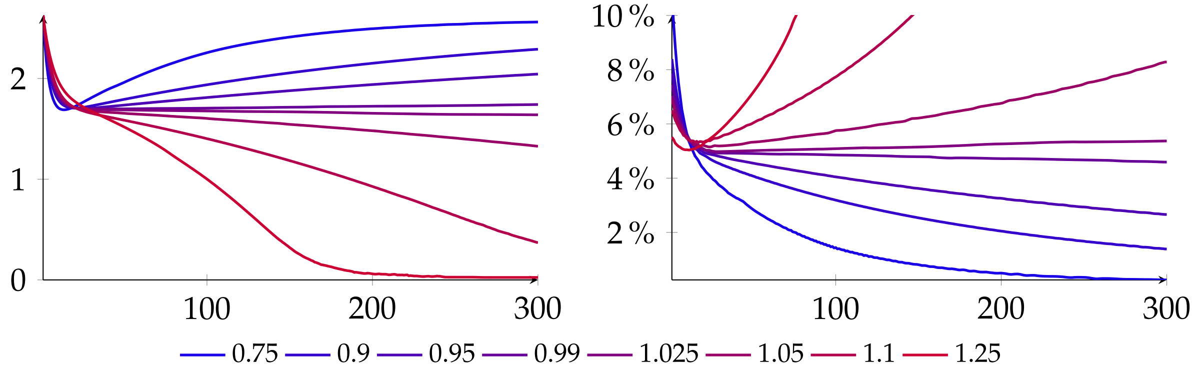
Solution for IES near 1. Figure 1 displays the time path for the capital stock and the tax rate on wealth, $\mathcal{T}_t = 1 - R_t/R_t^*$, for a range of σ that straddles the logarithmic $\sigma = 1$ case. We set $\beta = 0.95$, $\delta = 0.1$, $f(k) = k^\alpha$ with $\alpha = 0.3$ and $u(c) = U(c)$. Spending g is chosen so that $\frac{g}{f(k)} = 20\%$ at the zero-tax steady state. The initial value of capital, k_0, is set at the zero-tax steady state. Our numerical method is based on a recursive formulation of the problem described in the appendix.

To clarify the magnitudes of the tax on wealth, $\mathcal{T}_t$, consider an example: If $R^* = 1.04$ so that the before-tax net return is 4%, then a tax on wealth of 1% represents a 25% tax on the net return; a wealth tax of 4% represents a tax rate of 100% on net returns, and so on.

A few things stand out in Figure 1. First, the results confirm what we showed theoretically in Proposition 3, that for $\sigma > 1$ capital converges to $k_g = 0.0126$. In the figure this convergence is monotone[19], taking around 200 years for $\sigma = 1.25$. The asymptotic tax rate is very high, approximately $\mathcal{T}^g = 1 - R/R^* = 85\%$, lying outside the figure's

[19] This depends on the level of initial capital. For lower levels of capital the path first rises then falls.

Figure 1: Optimal time paths for capital (left) and wealth taxes (right).



Note. This figure shows the optimal time paths of capital k_t (left panel) and wealth taxes $\mathcal{T}_t$ (right panel) for various values of the inverse IES σ.

range, and, since the after-tax return equals $R = 1/\beta$ in the long run, this implies that the before-tax return $R^* = f'(k_g) + 1 - \delta$ is exorbitant.

Second, for $\sigma < 1$, the path for capital is non-monotonic[20] and eventually converges to the zero-tax steady state. However, the convergence is relatively slow, especially for values of σ near 1. This makes sense, since, by continuity, for any period t, the solution should converge to that of the logarithmic utility case as $\sigma \to 1$, with positive taxation as described in Proposition 1. By implication, for $\sigma < 1$ the rate of convergence to the zero-tax steady state must be zero as $\sigma \uparrow 1$. To further punctuate this point, Figure 2 shows the number of years it takes for the tax on wealth to drop below 1% as a function of $\sigma \in (\frac{1}{2}, 1)$. As σ rises, it takes longer and longer and as $\sigma \uparrow 1$ it takes an eternity.

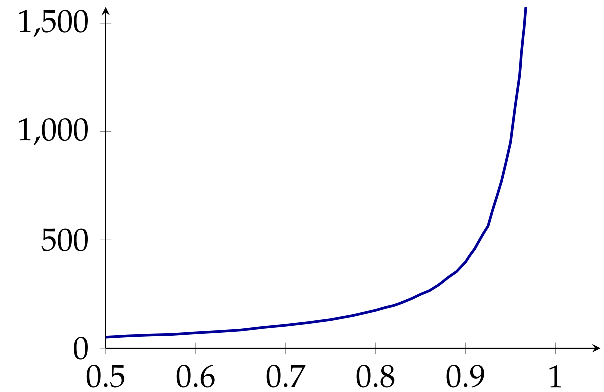
The logarithmic case leaves other imprints on the solutions for $\sigma \neq 1$. Returning to Figure 1, for both $\sigma < 1$ and $\sigma > 1$ we see that over the first 20-30 years, the path approaches the steady state of the logarithmic utility case, associated with a tax rate around $\mathcal{T} = 1 - \beta = 5\%$. The speed at which this takes place is relatively quick, which is explained by the fact that for $\sigma = 1$ it is driven by the standard rate of convergence in the neoclassical growth model. The solution path then transitions much more slowly either upwards or downwards, depending on whether $\sigma > 1$ or $\sigma < 1$.

Intuition: Anticipatory Effects of Future Taxes on Current Savings. Why does the optimal tax eventually rise for $\sigma > 1$ and fall for $\sigma < 1$? Why are the dynamics relatively slow for σ near 1?

To address these normative questions it helps to back up and review the following

[20] This is possible because the state variable has two dimensions, (k_t, C_{t-1}). At the optimum, for the same capital k, consumption C is initially higher on the way down than it is on the way up.

Figure 2: Slow speed of convergence to zero taxes for σ close to but below 1.



Note. This plot shows the time it takes until the wealth tax $\mathcal{T}_t$ falls below 1% for an inverse IES $\sigma \in (\frac{1}{2}, 1)$.

positive exercise. Start from a constant tax on wealth and imagine an unexpected announcement of higher future taxes on capital. How do capitalists react today? There are substitution and income effects pulling in opposite directions. When $\sigma > 1$, the substitution effect is muted compared to the income effect, and capitalists lower their consumption to match the drop in future consumption. As a result, capital rises in the short run and falls in the long run.[21] When instead $\sigma < 1$, the substitution effect is stronger and capitalists increase current consumption. In the logarithmic case, $\sigma = 1$, the two effects cancel out, so that current consumption and savings are unaffected.

Returning to the normative questions, lowering capitalists' consumption and increasing capital is desirable for workers. When $\sigma < 1$, this can be accomplished by promising lower tax rates in the future. This explains why a declining path for taxes is optimal. In contrast, when $\sigma > 1$, the same is accomplished by promising higher tax rates in the future; explaining the increasing path for taxes. These incentives are absent in the logarithmic case, when $\sigma = 1$, explaining why the tax rate converges to a constant.

When $\sigma < 1$ the rate of convergence to the zero-tax steady state is also driven by these anticipatory effects. With σ near 1, the potency of these effects is small, explaining why the rate of convergence is low and indeed becomes vanishingly small as $\sigma \uparrow 1$.

In contrast to previous intuitions offered for zero long-run tax results, the intuition we provide for our results—zero and nonzero long-run taxes alike, depending on σ—is not about the desired *level* for the tax. Instead, we provide a rationale for the desired *slope* in the path for the tax: an upward path when $\sigma > 1$ and a downward path when $\sigma < 1$.

[21] It is important to note that $\sigma > 1$ does *not* imply that the supply for savings "bends backward". Indeed, as a positive exercise, if taxes are raised permanently within the model, then capital falls over time to a lower steady state for any value of σ, including $\sigma > 1$. Higher values of σ imply a less elastic response over any *finite* time horizon, and thus a slower convergence to the lower capital stock. The case with $\sigma > 1$ is widely considered more plausible empirically.

The conclusions for the optimal long-run tax then follow from these desired slopes, rather than the other way around.

Our intuition based on slopes has an interesting implication for the effects of limited commitment in this economy. Since the planner promises higher future taxation when $\sigma > 1$, renegotiation by the planner might lead to *lower* rather than higher capital taxes. This is the polar opposite of the conventional wisdom, according to which limited commitment leads to higher capital taxation.

2.3 General Savings Functions and Inverse Elasticity Formula

The intuition suggests that the essential ingredient for positive long run capital taxation in the model of Judd (1985, Section 3) is that capitalists' savings decrease in future interest rates. To make this point even more transparently, we now modify the model and assume capitalists behave according to a general "ad-hoc" savings rule,

$$k_{t+1} = S(R_t k_t; R_{t+1}, R_{t+2}, \dots),$$

where $S(I_t; R_{t+1}, R_{t+2}, \dots) \in [0, I_t]$ is a continuously differentiable function taking as arguments current wealth $I_t = R_t k_t \geq 0$ and future interest rates $\{R_{t+1}, R_{t+2}, \dots\} \in \mathbb{R}_+^{\mathbb{N}}$. We assume that savings increase with income, $S_I > 0$. This savings function encompasses the case where capitalists maximize an additively separable utility function, as in Judd (1985), but is more general. For example, the savings function can be derived from the maximization of a recursive utility function, or even represent behavior that cannot be captured by optimization, such as hyperbolic discounting or self-control and temptation.

Again, we focus on the case $\gamma = 0$. The planning problem is then

$$\max_{\{c_t, R_t, k_{t+1}\}} \sum_{t=0}^{\infty} \beta^t u(c_t),$$

subject to

$$\begin{aligned} c_t + R_t k_t + g &= f(k_t) + (1-\delta)k_t, \\ k_{t+1} &= S(R_t k_t; R_{t+1}, R_{t+2}, \dots), \end{aligned}$$

with k_0 given.

We can show that, consistent with the intuition spelled out above, long-run capital taxes are positive whenever savings decrease in future interest rates.

Proposition 4. *Suppose $\gamma = 0$ and assume the savings function is decreasing in future rates, so that $S_{R_t}(I; R_1, R_2, \dots) \leq 0$ for all $t = 1, 2, \dots$ and all arguments $\{I, R_1, R_2, \dots\}$. If the optimum converges to an interior steady state in c, k, and R, and at the steady state $\beta R S_I \neq 1$, then the limit tax rate is positive and $\beta R S_I < 1$.*

This generalizes Proposition 2, since the case with iso-elastic utility and IES less than one is a special case satisfying the hypothesis of the proposition. Once again, the intuition here is that the planner exploits anticipatory effects by raising tax rates over time to increase present savings.

The result requires $\beta R S_I < 1$ at the steady state, which is satisfied when savings are linear in income, since then $S_I R = 1$ at a steady state. Note that savings are linear in income in the isoelastic utility case. More generally, $R S_I < 1$ is natural, as it ensures local stability for capital given a fixed steady-state return, i.e. the dynamics implied by the recursion $k_{t+1} = S(Rk_t, R, R, \cdots)$ for fixed R.

Inverse Elasticity Formula. There is a long tradition relating optimal tax rates to elasticities. In the context of our general savings model, spelled out above, we derive the following "inverse elasticity rule"

$$\mathcal{T} = 1 - \frac{R}{R^*} = \frac{1 - \beta R S_I}{1 + \sum_{t=1}^{\infty} \beta^{-t+1} \epsilon_{S,t}}, \tag{4}$$

where $\epsilon_{S,t} \equiv \frac{R_t}{S} \frac{\partial S}{\partial R_t}(R_0 k_0; R_1, R_2, \dots)$ denotes the elasticity of savings with respect to future interest rates evaluated at the steady state in c, k, and R. Although the right hand side is endogenous, equation (4) is often interpreted as a formula for the tax rate. Our inverse elasticity formula is closely related to a condition derived by Piketty and Saez (2013, see their Section 3.3, equation 16).[22]

We wish to make two points about our formula. First, note that the relevant elasticity in this formula is *not* related to the response of savings to *current*, transitory or permanent, changes in interest rates. Instead, the formula involves a sum of elasticities of savings with respect to *future* changes in interest rates. Thus, it involves the anticipatory effects discussed above. Indeed, the variation behind our formula changes the after-tax interest rate at a single future date T, and then takes the limit as $T \to \infty$. For any finite T, the

[22] Their formula is derived under the special assumptions of additively separable utility, an exogenously fixed international interest rate and an exogenous wage. None of this is important, however. The two formulas remain different because of slightly different elasticity definitions; ours is based on partial derivatives of the primitive savings function S with respect to a single interest rate change, while theirs is based on the implicit total derivative of the capital stock sequence with respect to a permanent change in the interest rate.

term $\sum_{t=1}^{T} \beta^{-t+1}\epsilon_{S,t}$ represents the sum of the anticipatory effects on capitalists' savings behavior in periods 0 up to $T-1$; while $\sum_{t=1}^{\infty} \beta^{-t+1}\epsilon_{S,t}$ captures the limit as $T \to \infty$. It is important to keep in mind that, precisely because it is anticipatory effects that matter, the relevant elasticities are negative in standard cases, e.g. with additive utility and IES below one.

Second, the derivation we provide in the appendix requires convergence to an interior steady state as well as additional conditions (somewhat cumbersome to state) to allow a change in the order of limits and obtain the simple expression $\sum_{t=1}^{\infty} \beta^{-t+1}\epsilon_{S,t}$. These latter conditions seem especially hard to guarantee ex ante, with assumptions on primitives, since they may involve the endogenous speed of convergence to the presumed interior steady state.[23] As we have shown, in this model one cannot take these properties for granted, neither the convergence to an interior steady state (Proposition 3) nor the additional conditions. Indeed, Proposition 4 already supplies counterexamples to the applicability of the inverse elasticity formula.

Corollary. *Under the conditions of Proposition 4, the inverse elasticity formula (4) cannot hold if* $1 + \sum_{t=1}^{\infty} \beta^{-t+1}\epsilon_{S,t} < 0$.

This result provides conditions under which the formula (4) cannot characterize the long run tax rate. Whenever the discounted sum of elasticities with respect to future rates, $\sum_{t=1}^{\infty} \beta^{-t+1}\epsilon_{S,t}$, is negative and less than -1, the formula implies a negative limit tax rate. Yet, under the same conditions as in Proposition 4, this is not possible since this result shows that if convergence takes place, the tax rate is positive.

The case with additive and iso-elastic utility is an extreme example where the sum of elasticities $\sum_{t=1}^{\infty} \beta^{-t+1}\epsilon_{S,t}$ diverges. As it turns out, in this case $\beta^{-t}\epsilon_{S,t} = -\frac{\sigma-1}{\sigma}\frac{1-\beta}{\beta}$ at a steady state and the sum of elasticities diverges. It equals $+\infty$ if the IES is greater than one, or $-\infty$ if the IES is less than one.[24] In both cases, formula (4) suggests a zero steady state tax rate. Piketty and Saez (2013) use this to argue that this explains the Chamley-Judd result of a zero long-run tax. However, as we have shown, when the IES is less than one the limit tax rate is not zero. This counterexample to the applicability of the inverse elasticity formula (4) assumes additive utility and, thus, an infinite sum of elasticities. However, the problem may also arise for non-additive preferences or with ad hoc saving functions. Indeed, the conditions for the corollary may be met in cases where the sum of elasticities is finite, as long as its value is sufficiently negative.

[23]Unfortunately, one cannot ignore transitions by choice of a suitable initial condition. For example, even in the additive utility case with $\sigma < 1$ and even if we start at the zero capital tax steady state, capital does not stay at this level forever. Instead, capital first falls and then rises back up at a potentially slow rate.

[24]Proposition 12 in the appendix shows that the infinite sum $\sum_{t=1}^{\infty} \beta^{-t+1}\epsilon_{S,t}$ also diverges for general recursive, non-additive preferences.

It should be noted that our corollary provides sufficient conditions for the formula to fail, but other counterexamples may exist outside its realm. Suggestive of this is the fact that when the denominator is positive but small the formula may yield tax rates above 100%, which seems nonsensical, requiring $R < 0$. More generally, very large tax rates may be inconsistent with the fact that steady state capital must remain above $k_g > 0$.

To summarize, the inverse elasticity formula (4) fails in important cases, providing misleading answers for the long run tax rate. This highlights the need for caution in the application of steady state inverse elasticity rules.

2.4 Redistribution Towards Capitalists

In the present model, a desire to redistribute towards workers, away from capitalists, is a prerequisite to create a motive for positive wealth taxation. Proposition 3 assumes no weight on capitalists, $\gamma = 0$, to ensure that desired redistribution runs in this direction. When $\gamma > 0$ the same results obtain as long as the desire for redistribution continues to run from capitalists towards workers. In contrast, when γ is high enough the desired redistribution flips from workers to capitalists. When this occurs, the optimum naturally involves negative tax rates, to benefit capitalists.

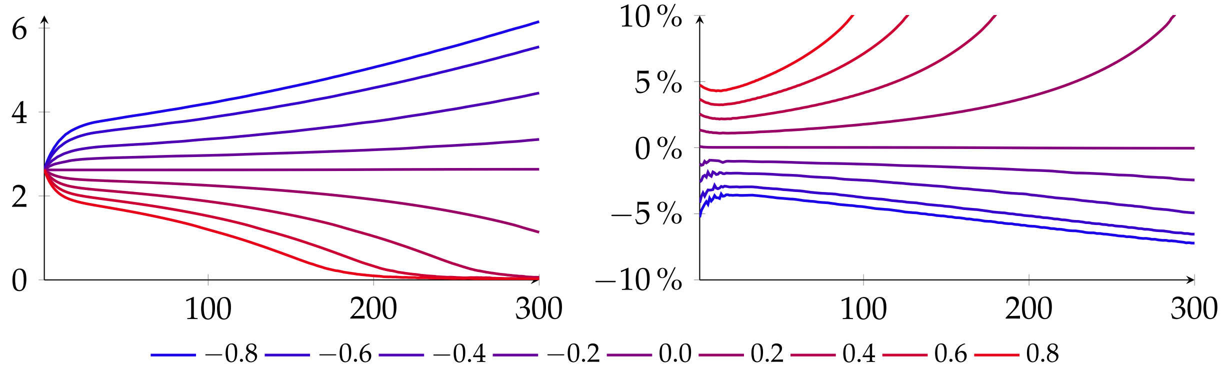
We verify these points numerically. Figure 3 illustrates the situation by fixing $\sigma = 1.25$ and varying the weight γ. Since initial capital is set at the zero-tax steady state, k^*, the direction of desired redistribution flips exactly at $\gamma^* = u'(c^*)/U'(C^*)$. At this value of γ, the planner is indifferent between redistributing towards workers or capitalists at the zero-tax steady state (k^*, c^*, C^*).[25] When $\sigma > 1$ and $\gamma > \gamma^*$ the solution converges to the highest sustainable capital k^g, the highest solution to $\frac{1}{\beta}k + g = f(k) + (1-\delta)k$, rather than k_g, the lowest solution to the same equation.

A deeper understanding of the dynamics can be grasped by noting that the planning problem is recursive in the state variable (k_t, C_{t-1}). It is then possible to study the dynamics for this state variable locally, around the zero tax steady state, by linearizing the first-order conditions (2). We do so for a continuous-time version of the model, to ensure that our results are comparable to Kemp et al. (1993). The details are contained in the appendix. We obtain the following characterization.

Proposition 5. *For a continuous-time version of the model,*

[25] Rather than displaying γ in the legend for Figure 3, we perform a transformation that makes it more easily interpretable: we report the proportional change in consumption for capitalists that would be desired at the steady state, e.g. -0.4 represents that the planner's ideal allocation of the zero-tax output would feature a 40% reduction in the consumption of capitalists, relative to the steady state value $C = \frac{1-\beta}{\beta}k$. The case $\gamma = \gamma^*$ corresponds to 0 in this transformation.

Figure 3: Wealth taxes diverge as long as the planner has a desire for redistribution.



Note. This figure shows the optimal time paths over 300 years for the capital stock (left panel) and wealth taxes (right panel), for various redistribution preferences (zero represents no desire for redistribution; see footnote 25).

(a) *if* $\sigma > 1$*, the zero-tax steady state is locally saddle-path stable;*

(b) *if* $\sigma < 1$ *and* $\gamma \leq \gamma^*$*, the zero-tax steady state is locally stable;*

(c) *if* $\sigma < 1$ *and* $\gamma > \gamma^*$*, the zero-tax steady state may be locally stable or unstable and the dynamics may feature cycles.*

The first two points confirm our theoretical and numerical observations. For $\sigma > 1$ the solution is saddle-path stable, explaining why it does not converge to the zero-tax steady state—except for the knife-edged cases where there is no desire for redistribution, in which case the tax rate is zero throughout. For $\sigma < 1$ the solution converges to the zero tax steady state whenever redistribution towards workers is desirable. This lends theoretical support to our numerical findings for $\sigma < 1$, discussed earlier and illustrated in Figure 1.

The third point raises a distinct possibility which is not our focus: the system may become unstable or feature cyclical dynamics. This is consistent with Kemp et al. (1993), who also studied the linearized system around the zero-tax steady state. They reported the potential for local instability and cycles, applying the Hopf Bifurcation Theorem. Proposition 5 clarifies that a necessary condition for this dynamic behavior is $\sigma < 1$ and $\gamma > \gamma^*$. The latter condition is equivalent to a desire to redistribute away from workers towards capitalists. We have instead focused on low values of γ that ensure that desired redistribution runs from capitalists to workers. For this reason, our results are completely distinct to those in Kemp et al. (1993).

3 Representative Agent Ramsey

In the previous section we worked with the two-class model without government debt in Judd (1985, Section 3). Chamley (1986), in contrast, studied a representative agent Ramsey model with unconstrained government debt; Judd (1999) adopted the same assumptions. This section presents results for such representative agent frameworks.

We first consider situations where the upper bounds on capital taxation do not bind in the long run (Section 3.1). We then prove, for additively separable preferences, that these bounds may, in fact, bind indefinitely (Section 3.2). Readers mainly interested in the latter result may skip Section 3.1.

3.1 First Best or Zero Taxation of Zero Wealth?

In this subsection, we first review the discrete-time model and zero capital tax steady state result in Chamley (1986, Section 1) and then present a new result. We show that if the economy settles down to a steady state where the bounds on the capital tax are not binding, then the tax on capital must be zero. This result holds for general recursive preferences that, unlike time-additive utility, allow the rate of impatience to vary. Non-additive utility constituted an important element in Chamley (1986, Section 1), to ensure that zero-tax results were not driven by an "infinite long-run elasticity of savings".[26] However, we also show that other implications emerge away from additive utility. In particular, if the economy converges to a zero-tax steady state there are two possibilities. Either private wealth has been wiped out, in which case nothing remains to be taxed, or the tax on labor also falls to zero, in which case capital income and labor income are treated symmetrically. These implications paint a very different picture, one that is not favorable to the usual interpretation of zero capital tax results.

Preferences. We write the representative agent's utility as $\mathcal{V}(U_0, U_1, \ldots)$ with per period utility $U_t = U(c_t, n_t)$ depending on consumption c_t and labor supply n_t. Assume that utility $\mathcal{V}$ is increasing in every argument and satisfies a Koopmans (1960) recursion

$$V_t = W(U_t, V_{t+1}) \tag{5a}$$

$$V_t = \mathcal{V}(U_t, U_{t+1}, \ldots) \tag{5b}$$

$$U_t = U(c_t, n_t). \tag{5c}$$

[26] At any steady state with additive utility one must have $R = 1/\beta$ for a fixed parameter $\beta \in (0, 1)$. This is true regardless of the wealth or consumption level. In this sense, the supply of savings is infinitely elastic at this rate of interest.

Here $W(U, V')$ is an aggregator function. We assume that both $U(c, n)$ and $W(U, V')$ are twice continuously differentiable, with $W_U, W_V, U_c > 0$ and $U_n < 0$. Consumption and leisure are taken to be normal goods,

$$\frac{U_{cc}}{U_c} - \frac{U_{nc}}{U_n} \leq 0 \qquad \text{and} \qquad \frac{U_{cn}}{U_c} - \frac{U_{nn}}{U_n} \leq 0,$$

with at least one strict inequality.

Regarding the aggregator function, the additively separable utility case amounts to the particular linear choice $W(U, V') = U + \beta V'$ with $\beta \in (0, 1)$. Nonlinear aggregators allow local discounting to vary with U and V', as in Koopmans (1960), Uzawa (1968) and Lucas and Stokey (1984). Of particular interest is how the discount factor varies across potential steady states. Define $\bar{U}(V)$ as the solution to $V = W(\bar{U}(V), V)$ and let $\bar{\beta}(V) \equiv W_V(\bar{U}(V), V)$ denote the steady state discount factor. It will prove useful below to note that the strict monotonicity of $\mathcal{V}$ immediately implies that $\bar{\beta}(V) \in (0, 1)$ at any steady state with utility V.[27]

Technology. The economy is subject to the sequence of resource constraints

$$c_t + k_{t+1} + g_t \leq F(k_t, n_t) + (1 - \delta)k_t \qquad t = 0, 1, \ldots \tag{6}$$

where F is a concave, differentiable and constant returns to scale production function taking as inputs labor n_t and capital k_t, and the parameter $\delta \in [0, 1]$ is the depreciation rate of capital. The sequence for government consumption, $\{g_t\}$, is given exogenously.

Markets and Taxes. Labor and capital markets are perfectly competitive, yielding before tax wages and rates of return given by $w_t^* = F_n(k_t, n_t)$ and $R_t^* = F_k(k_t, n_t) + 1 - \delta$.

The agent maximizes utility subject to the sequence of budget constraints

$$\begin{aligned} c_0 + a_1 &\leq w_0 n_0 + R_0 k_0 + R_0^b b_0, \\ c_t + a_{t+1} &\leq w_t n_t + R_t a_t \qquad t = 1, 2, \ldots, \end{aligned}$$

and the No Ponzi condition $\frac{a_{t+1}}{R_1 R_2 \cdots R_t} \to 0$. The agent takes as given the after-tax wage w_t and the after-tax gross rates of return, R_t. Total assets $a_t = k_t + b_t$ are composed of capital k_t and government debt b_t; with perfect foresight, both must yield the same return

[27] A positive marginal change dU in the constant per period utility stream increases steady state utility by some constant $d\mathcal{V}$. By virtue of (5a) this implies $d\mathcal{V} = W_U dU + W_V d\mathcal{V}$, which yields a contradiction unless $W_V < 1$.

in equilibrium for all $t = 1, 2, \ldots$, so only total wealth matters for the agent; this is not true for the initial period, where we allow possibly different returns on capital and debt. The after-tax wage and return relate to their before-tax counterparts by $w_t = (1 - \tau_t^n) w_t^*$ and $R_t = (1 - \tau_t)(R_t^* - 1) + 1$ (here it is more convenient to work with a tax rate on net returns than on gross returns).

Importantly, we follow Chamley (1986) and allow for an indirect constraint on the capital tax rate given by $R_t \geq 1$. For positive before-tax interest rates $R^* - 1$ this is precisely equivalent to assuming $\tau_t \leq 1$.[28] As is well understood, without constraints on capital taxation the solution involves extraordinarily high initial capital taxation, typically complete expropriation, unless the first best is achieved first. Taxing initial capital mimics the missing lump-sum tax, which has no distortionary effects. We note that our main result in this section, Proposition 6, does not depend on the specific form of the capital tax constraint.

Planning problem. The *implementability condition* for this economy is

$$\sum_{t=0}^{\infty} (\mathcal{V}_{ct} c_t + \mathcal{V}_{nt} n_t) = \mathcal{V}_{c0} \left(R_0 k_0 + R_0^b b_0 \right), \tag{7}$$

whose derivation is standard. In the additive separable utility case $\mathcal{V}_{ct} = \beta^t U_{ct}$ and $\mathcal{V}_{nt} = \beta^t U_{nt}$ and expression (7) reduces to the standard implementability condition popularized by Lucas and Stokey (1983) and Chari et al. (1994). Given R_0 and R_0^b, any allocation satisfying the implementability condition and the resource constraint (6) can be sustained as a competitive equilibrium for some sequence of prices and taxes.[29]

To enforce the constraints on the taxation of capital in periods $t = 1, 2, \ldots$ we impose

$$\mathcal{V}_{ct} = R_{t+1} \mathcal{V}_{ct+1}, \tag{8a}$$

$$R_t \geq 1. \tag{8b}$$

The planning problem maximizes $\mathcal{V}(U_0, U_1, \ldots)$ subject to (6), (7) and (8). In addition, we take R_0^b as given. The constraint $R_t \geq 1$ may or may not bind forever. In this subsection we are interested in situations where the constraint does not bind asymptotically, i.e. it is slack after some date $T < \infty$. In the next subsection we discuss the possibility of the constraint binding forever.

[28] When $R^* - 1$ is negative, however, an upper bound directly imposed on taxes τ_t allows arbitrarily low after-tax interest rates R_t.

[29] The argument is identical to that in Lucas and Stokey (1983) and Chari et al. (1994).

Chamley (1986) provided the following result—slightly adjusted here to make explicit the need for the steady state to be interior, for multipliers to converge and for the bounds on taxation to be asymptotically slack.

Theorem 3 (Chamley, 1986, Theorem 1)**.** *Suppose the optimum converges to an interior steady state where the constraints on capital taxation are asymptotically slack. Let $\tilde{\Lambda}_t = \mathcal{V}_{ct}\Lambda_t$ denote the multiplier on the resource constraint (6) in period t. Suppose further that the multiplier Λ_t converges to an interior point $\Lambda_t \to \Lambda > 0$. Then the tax on capital converges to zero $\frac{R_t}{R_t^*} \to 1$.*

The proof is straightforward. Consider a sufficiently late period t, so that the bounds on the capital tax rate are no longer binding. Then the first-order condition for k_{t+1} includes only terms from the resource constraint (6) and is simply $\tilde{\Lambda}_t = \tilde{\Lambda}_{t+1}R^*_{t+1}$. Equivalently, using that $\tilde{\Lambda}_t = \mathcal{V}_{ct}\Lambda_t$ we have

$$\mathcal{V}_{ct}\Lambda_t = \mathcal{V}_{ct+1}\Lambda_{t+1}R^*_{t+1}.$$

On the other hand the representative agent's Euler equation (8a) is

$$\mathcal{V}_{ct} = \mathcal{V}_{ct+1}R_{t+1}.$$

The result follows from combining these last two equations.

With the specific constraint $R_t \geq 1$ on capital taxation assumed here and in Chamley (1986), there would be no need to require the constraints on capital taxation not to bind. The reason is that in this case the constraints imposed by (8) do not involve k_{t+1}, so the argument above goes through unchanged. In fact, this is essentially the form that Theorem 1 in Chamley (1986) takes, although the assumption of converging multipliers is not stated explicitly, but imposed within the proof. We chose to explicitly assume the capital tax constraints to be no longer binding to allow a broader applicability of the theorem to situations without the specific constraints in (8).[30]

The main result of this subsection is stated in the next proposition. Relative to Theorem 3, we make no assumptions on multipliers and prove that the steady-state tax rate is zero. More importantly, we derive new implications of reaching an interior steady state.

Proposition 6. *Suppose the optimal allocation converges to an interior steady state and assume the bounds on capital tax rates are asymptotically slack. Then the tax on capital is asymptotically*

[30]Note that as long as the multiplier Λ_t converges, one does not even need to assume the allocation converges to arrive at the zero-tax conclusion. This is essentially the argument used by Judd (1999). However, the problem is that one cannot guarantee that the multiplier converges. We shall discuss this in subsection 3.3.

zero. In addition, if the discount factor is locally non-constant at the steady state, so that $\bar{\beta}'(V) \neq 0$, then either

(a) private wealth converges to zero, $a_t \to 0$; or

(b) the allocation converges to the first-best, with a zero tax rate on labor.

This result shows that at any interior steady state where the bounds on capital taxes do not bind, the tax on capital is zero; this much basically echoes Chamley (1986), or our rendering in Theorem 3. However, as long as the rate of impatience is not locally constant, so that $\bar{\beta}'(V) \neq 0$, the proposition also shows that this zero tax result comes with other implications. There are two possibilities. In the first possibility, the capital income tax base has been driven to zero—perhaps as a result of heavy taxation along the transition. In the second possibility, the government has accumulated enough wealth—perhaps aided by heavy taxation of wealth along the transition—to finance itself without taxes, so the economy attains the first best. Thus, capital taxes are zero, but the same is true for labor taxes.

To sum up, if the economy converges to an interior steady state, then either both labor and capital are treated symmetrically or there remains no wealth to be taxed. Both of these implications do not sit well with the usual interpretation of the zero capital tax result. To be sure, in the special (but commonly adopted) case of additive separable utility one can justify the usual interpretation where private wealth is spared from taxation and labor bears the entire burden. However, this is no longer possible when the rate of impatience is not constant. In this sense, the usual interpretation describes a knife edged situation.

3.2 Long Run Capital Taxes Binding at Upper Bound

We now show that the bounds on capital tax rates may bind forever, contradicting a claim by Chamley (1986). This claim has been echoed throughout the literature, e.g. by Judd (1999), Atkeson et al. (1999) and others.

For our present purposes, and following Chamley (1986) and Judd (1999), it is convenient to work with a continuous-time version of the model and restrict attention to additively separable preferences,[31]

$$\int_0^\infty e^{-\rho t} U(c_t, n_t) dt. \tag{9a}$$

[31] Continuous time allowed Chamley (1986) to exploit the bang-bang nature of the optimal solution. Since we focus on cases where this is not the case it is less crucial for our results. However, we prefer to keep the analyses comparable.

$$U(c,n) = u(c) - v(n) \qquad \text{with} \qquad u(c) = \frac{c^{1-\sigma}}{1-\sigma}, \quad v(n) = \frac{n^{1+\zeta}}{1+\zeta}, \tag{9b}$$

where $\sigma, \zeta > 0$. Following Chamley (1986), we adopt an iso-elastic utility function over consumption; this is important to ensure the bang-bang nature of the solution. We also assume iso-elastic disutility from labor, but we believe similar results to ours can be shown for arbitrary convex disutility functions $v(n)$. The resource constraint is

$$c_t + \dot{k}_t + g = f(k_t, n_t) - \delta k_t, \tag{10}$$

where f has constant returns to scale with $f(0,n) = f(k,0) = 0$, is differentiable and strictly concave in each argument, and satisfies the usual Inada conditions. For simplicity, government consumption is taken to be constant at $g > 0$. We denote the before-tax net interest rate by $r_t^* = f_k(k_t, n_t) - \delta$. The implementability condition is now

$$\int_0^\infty e^{-\rho t} \left(u'(c_t)c_t - v'(n_t)n_t\right) = u'(c_0)a_0, \tag{11}$$

where $a_0 = k_0 + b_0$ denotes initial private wealth, consisting of capital k_0 and government bonds b_0. Since unrestricted subsidies to capital act as lump-sum tax when initial private wealth is negative, we focus on the case where initial private wealth is positive, $a_0 > 0$.[32] To enforce bounds on capital taxation we follow Chamley (1986) and impose

$$\dot{\theta}_t = \theta_t(\rho - r_t), \tag{12a}$$

$$r_t \geq 0, \tag{12b}$$

where $\theta_t = u'(c_t)$ denotes the marginal utility of consumption, and r_t denotes the after-tax interest rate. Whenever the before-tax return on capital $r_t^* \equiv f_k(k_t, n_t) - \delta$ is positive, constraint (12b) corresponds to a capital tax constraint $\tau_t = 1 - r_t/r_t^* \leq \bar{\tau}$ with $\bar{\tau} \equiv 1$. The planning problem maximizes (9a) subject to (10), (11) and (12).

Chamley (1986, Theorem 2, pg. 615) formulated the following claim regarding the path for capital tax rates.[33]

Claim. *There exists a time T with the following three properties:*

(a) for $t < T$, the constraint (12b) is binding, that is, $r_t = 0$ and $\tau_t = 1$;

(b) for $t > T$ capital income is untaxed, that is, $r_t = r_t^*$ and $\tau_t = 0$;

[32]Observe that, in Proposition 7, $a_0 > 0$ is always satisfied if $b_0 \in [\underline{b}, \bar{b}]$ for $\sigma > 1$, so the focus on $a_0 > 0$ does not affect our main result.

[33]Similar claims are made in Atkeson et al. (1999), Judd (1999) and many other papers.

(c) $T < \infty$.

At a crucial juncture in the proof of this claim, Chamley (1986) states in support of part (c) that "The constraint $r_t \geq 0$ cannot be binding forever (the marginal utility of private consumption [...] would grow to infinity [...] which is absurd)."[34] Our next result shows that there is nothing absurd about this within the logic of the model and that, quite to the contrary, part (c) of the above claim is incorrect: indefinite taxation, $T = \infty$, may be optimal.

Before presenting our result, some definitions are in order. Given a path for government spending, the tax burden the government must impose varies with initial government debt b_0. As with a regular, "static Laffer curve" there exists a maximum burden of taxes agents can finance, here given by a threshold level for initial government debt, $\bar{b}$. When $b_0 > \bar{b}$, no feasible allocation exists, while there are always feasible allocations if $b_0 < \bar{b}$. Naturally, at the peak of this "Laffer curve" when $b_0 = \bar{b}$ the tax on capital must be set to its upper bound indefinitely. Crucially, however, it may be optimal to set the tax on capital at its upper bound indefinitely when $b_0 < \bar{b}$, even not doing so is feasible.

Proposition 7. *Suppose preferences are given by (9). Fix any initial capital stock $k_0 > 0$ and assume initial private wealth $k_0 + b_0$ is positive.*

Then, the bang-bang property *holds, so that at any optimum of the planning problem there exists a time $T \in [0, \infty]$, such that capital taxes τ_t are set at their upper bound $\bar{\tau} = 1$ before T and set to zero thereafter. Whenever the economy is not at its first best, T is strictly positive. Moreover:*

A. For $\sigma > 1$, there exists a lower bound on debt $\underline{b} < \bar{b}$, such that:

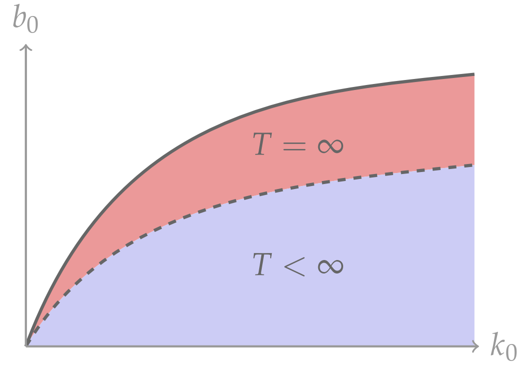
- *(a) If $b_0 = \bar{b}$, the unique optimum has $T = \infty$ and there is no feasible allocation with $T < \infty$.*
- *(b) If $b_0 \in [\underline{b}, \bar{b})$, the unique optimum has $T = \infty$ but there exist feasible allocations with $T < \infty$.*
- *(c) If $b_0 < \underline{b}$, any optimum has $T < \infty$.*

B. For $\sigma = 1$:

- *(a) If $b_0 = \bar{b}$, the unique optimum has $T = \infty$ and there is no feasible allocation with $T < \infty$.*
- *(b) If $b_0 < \bar{b}$, any optimum has $T < \infty$.*

[34] It is worth pointing out, however, that although Chamley (1986) claims $T < \infty$ it never states that T is small. Indeed, it cautions to the possibility that it is quite large saying "the length of the period with capital income taxation at the 100 per cent rate can be significant."

Figure 4: Graphical representation of the case $\sigma > 1$ in Proposition 7.



C. *For $\sigma < 1$: Any optimum has $T < \infty$.*

Proposition 7 offers a full characterization of the optimal capital tax policy in this economy. First, we prove a bang-bang property of capital taxes, according to which capital taxes are binding at their upper bound, $\tau_t = 1$, until some time T and drop to zero thereafter. It turns out that previous proofs of the bang-bang property (see, e.g., Chamley (1986) or Atkeson et al. (1999)) heavily relied on the false premise that capital taxes cannot be positive forever. We provide a new proof that avoids this issue.

Using the bang-bang property of capital taxes, we then characterize optimal capital taxes, distinguishing by the position of σ relative to 1. For $\sigma > 1$, we prove that it is optimal to tax capital indefinitely for a positive-measure interval of b_0. Crucially, for $b_0 < \bar{b}$ indefinite taxation is not driven by budgetary need—there are feasible plans with $T < \infty$; however, the plan with $T = \infty$ is simply better. This is illustrated in Figure 4 with a qualitative plot of the set of states (k_0, b_0) for which indefinite capital taxation is optimal if $\sigma > 1$. By contrast, for $\sigma < 1$ we show that at any optimum, $T < \infty$, so $T = \infty$ is never optimal. The case $\sigma = 1$ lies in between, in that $T = \infty$ is optimal only if $b_0 = \bar{b}$.

The basic idea behind our proof of part A of Proposition 7 is simple. To illustrate it, let λ_t denote the multiplier on the resource constraint (10) at time t and μ be the multiplier on the IC constraint (11). Both can be proven to be non-negative. Using this notation, if the period T of positive capital taxation is finite, the first order condition for consumption c_t after time T reads

$$\lambda_t = (1 - \mu(\sigma - 1))u'(c_t),$$

which requires $\mu \leq 1/(\sigma - 1)$. Yet, as initial government debt b_0 becomes large, $b_0 > \underline{b}$, so does μ, to the point where it crosses $1/(\sigma - 1)$, making it impossible for finite capital taxation to be optimal. Therefore, a sufficiently large burden of taxation due to high b_0, coupled with an intertemporal elasticity σ^{-1} less than 1 points to indefinite capital taxation. To make this approach watertight, we specifically construct allocations with $T = \infty$ and show that they satisfy the first order conditions whenever $b_0 \geq \underline{b}$. Since, as we show, the planning problem can be recast into a concave maximization problem, the first order conditions (together with transversality conditions) are sufficient for an optimum.

Our next result assumes $g = 0$ and constructs the solution for a set of initial conditions that allow us to guess and verify its form.

Proposition 8. *Suppose that preferences are given by* (9) *with* $\sigma > 1$, *and that* $g = 0$. *There exist* $\underline{k} < \bar{k}$ *and* $b_0(k_0)$ *such that: for any* $k_0 \in (\underline{k}, \bar{k}]$ *and initial debt* $b_0(k_0)$ *the optimum satisfies* $\tau_t = 1$ *for all* $t \geq 0$ *and* $c_t, k_t, n_t \to 0$ *exponentially with constant* n_t/k_t *and* c_t/k_t.

Under the conditions stated in the proposition the solution converges to zero in a homogeneous, constant growth rate fashion. This explicit example illustrates that convergence takes place, but not to an interior steady state. It turns out that this latter property is more general: at least with additively separable utility, whenever indefinite taxation of capital is optimal, $T = \infty$, no interior steady state exists, even if capital taxes are constrained by tax bounds $\bar{\tau} < 1$, that is, if we impose $r_t \geq r_t^*(1 - \overline{\tau})$.

To see why this is the case consider first the case with $\bar{\tau} = 1$. Then the after tax interest rate is zero whenever the bound is binding. Since the agent discounts the future positively this prevents a steady state. In contrast, when $\bar{\tau} < 1$ the before-tax interest rate may be positive and the after tax interest rate equal to the discount rate, $(1 - \bar{\tau})r^* = \rho$, the condition for constant consumption. This suggests the possibility of a steady state. However, we must also verify whether labor, in addition to consumption, remains constant. This, in turn, requires a constant labor tax. Yet, one can show that under the assumptions of Proposition 7, but allowing $\bar{\tau} < 1$, we must have

$$\partial_t \tau_t^n = (1 - \tau_t^n)\tau_t r_t^*,$$

implying that the labor tax strictly rises over time whenever the capital tax is positive, $\tau_t > 0$. This rules out an interior steady state. Intuitively, the capital tax inevitably distorts the path for consumption, but the optimum attempts to undo the intertemporal distortion in labor by varying the tax on labor. We conjecture that the imposition of an upper bound on labor taxes solves the problem of an ever-increasing path for labor taxes, leading to the existence of *interior* steady states with positive capital taxation.

3.3 Revisiting Judd (1999)

Up to this point we have focused on the Chamley-Judd zero-tax results. A follow-up literature has offered both extensions and interpretations. One notable case doing both is Judd (1999). This paper is related to Chamley (1986) in that it studies a representative agent economy with perfect financial markets and unrestricted government bonds. It also allows for other state variables, such as human capital, and in that sense builds on Judd (1985, Section 5) and Jones et al. (1993). At its core, Judd (1999) provides a zero capital tax result without requiring the allocation to converge to a steady state. The paper also offers a connection between capital taxation and rising consumption taxes to provide an intuition for zero-tax results. Let us consider each of these two points in turn.

Bounded Multipliers and Zero Average Capital Taxes. Abstracting away from some of the additional ingredients in Judd (1999), the essence of the main result in Judd (1999) can be restated using our continuous-time setup from Section 3.2. With $\bar{\tau} = 1$, the planning problem maximizes (9a) subject to (10), (11), (12a), and (12b). Let $\hat{\Lambda}_t = \theta_t \Lambda_t$ denote the co-state for capital, that is, the current value multiplier on equation (10), satisfying $\dot{\hat{\Lambda}}_t = \rho \hat{\Lambda}_t - r_t^* \hat{\Lambda}_t$. Using that $\dot{\hat{\Lambda}}_t / \hat{\Lambda}_t = \dot{\theta}_t/\theta_t + \dot{\Lambda}_t/\Lambda_t$ and $\dot{\theta}_t/\theta_t = \rho - r_t$ we obtain

$$\frac{\dot{\Lambda}_t}{\Lambda_t} = r_t - r_t^*.$$

If Λ_t converges then $r_t - r_t^* \to 0$. Thus, the Chamley (1986) steady state result actually follows by postulating the convergence of Λ_t, without assuming convergence of the allocation. Judd (1999, pg. 13, Theorem 6) goes down this route, but assumes that the endogenous multiplier Λ_t remains in a bounded interval, instead of assuming that it converges.

Theorem 4 (Judd, 1999)**.** *Let $\theta_t \Lambda_t$ denote the (current value) co-state for capital in equation* (10) *and assume*

$$\Lambda_t \in [\underline{\Lambda}, \bar{\Lambda}],$$

for $0 < \underline{\Lambda} \leq \bar{\Lambda} < \infty$. Then the cumulative distortion up to t is bounded,

$$\log\left(\frac{\Lambda_0}{\bar{\Lambda}}\right) \leq \int_0^t (r_s - r_s^*) ds \leq \log\left(\frac{\Lambda_0}{\underline{\Lambda}}\right),$$

and the average distortion converges to zero,

$$\frac{1}{t}\int_0^t (r_s - r_s^*) ds \to 0.$$

In particular, under the conditions of this theorem, the optimum cannot converge to a steady state with a positive tax on capital.[35] More generally, the condition requires departures of r_t from r_t^* to average zero.

Note that our proof proceeded without any optimality condition except the one for capital k_t.[36] In particular, we did not invoke first-order conditions for the interest rate r_t nor for the tax rate on capital τ_t. Naturally, this poses two questions. Do the bounds on Λ_t essentially assume the result? And are the bounds on Λ_t consistent with an optimum?

Regarding the first question, we can say the following. The multiplier $e^{-\rho t}\hat{\Lambda}_t$ represents the planner's (time 0) social marginal value of resources at time t. Thus,

$$\mathrm{MRS}^{\mathrm{Social}}_{\mathrm{t,t+s}} = e^{-\rho s}\frac{\hat{\Lambda}_{t+s}}{\hat{\Lambda}_t} = e^{-\int_0^s r^*_{t+\tilde{s}}\,d\tilde{s}}$$

represents the marginal rate of substitution between t and $t+s$, which, given the assumption $\bar{\tau} = 1$, is equated to the marginal rate of transformation. The private agent's marginal rate of substitution is

$$\mathrm{MRS}^{\mathrm{Private}}_{\mathrm{t,t+s}} = e^{-\rho s}\frac{\theta_{t+s}}{\theta_t} = e^{-\int_0^s r_{t+\tilde{s}}d\tilde{s}},$$

where θ_t represents marginal utility. It follows, by definition, that

$$\mathrm{MRS}^{\mathrm{Social}}_{\mathrm{t,t+s}} = \frac{\Lambda_{t+s}}{\Lambda_t}\cdot \mathrm{MRS}^{\mathrm{Private}}_{\mathrm{t,t+s}}.$$

This expressions shows that the rate of growth in Λ_t is, by definition, equal to the wedge between social and private marginal rates of substitution. Thus, the wedge $\frac{\Lambda_{t+s}}{\Lambda_t} = e^{\int_0^s (r_{t+\tilde{s}} - r^*_{t+\tilde{s}})d\tilde{s}}$ is the only source of nonzero taxes. Whenever Λ_t is constant, social and private MRSs coincide and the intertemporal wedge is zero, $r_t = r_t^*$; if Λ_t is enclosed in a bounded interval, the same conclusion holds on average.

These calculations afford an answer to the first question posed above: assuming the (average) rate of growth of Λ_t is zero is tantamount to assuming the (average) zero long-run tax conclusion. We already have an answer to the second question, whether the bounds are consistent with an optimum, since Proposition 7 showed that indefinite taxation may be optimal.

Corollary. *At the optimum described in Proposition 7 we have that $\Lambda_t \to 0$ as $t \to \infty$. Thus, in*

[35]The result is somewhat sensitive to the assumption that $\bar{\tau} = 1$; when $\bar{\tau} \neq 1$ and technology is nonlinear, the co-state equation acquires other terms, associated with the bounds on capital taxation.

[36]In this continuous time optimal control formulation, the costate equation for capital is the counterpart to the first-order condition with respect to capital in a discrete time formulation. Indeed, the same result can be easily formulated in a discrete time setting.

this case the assumption on the endogenous multiplier Λ_t adopted in Judd (1999) is violated.

There is no guarantee that the endogenous object Λ_t remains bounded away from zero, as assumed by Judd (1999), making Theorem 4 inapplicable.

Exploding Consumption Taxes. Judd (1999) also offers an intuitive interpretation for the Chamley-Judd result based on the observation that an indefinite tax on capital is equivalent to an ever-increasing tax on consumption. This casts indefinite taxation of capital as a villain, since rising and unbounded taxes on consumption appear to contradict standard commodity tax principles, as enunciated by Diamond and Mirrlees (1971), Atkinson and Stiglitz (1972) and others.

The equivalence between capital taxation and a rising path for consumption taxes is useful. It explains why prolonging capital taxation comes at an efficiency cost, since it distorts the consumption path. If the marginal cost of this distortion were increasing in T and approached infinity as $T \to \infty$ this would give a strong economic rationale against indefinite taxation of capital. We now show that this is not the case: the marginal cost remains bounded, even as $T \to \infty$. This explains why a corner solution with $T = \infty$ may be optimal.

We proceed with a constructive argument and assume, for simplicity, that technology is linear, so that $f(k, n) - \delta k = r^* k + w^* n$ for fixed parameters $r^*, w^* > 0$.

Proposition 9. *Suppose utility is given by (9), with $\sigma > 1$. Suppose technology is linear. Then the solution to the planning problem can be obtained by solving to the following static problem:*

$$\begin{aligned} \max_{T,c,n} \quad & u(c) - v(n), \\ s.t. \quad & (1 + \psi(T))\, c + G = k_0 + \omega n, \\ & u'(c)c - v'(n)n = (1 - \tau(T))u'(c)a_0, \end{aligned} \tag{13}$$

where $\omega > 0$ is proportional to w^*; G is the present value of government consumption; and, c and n are measures of lifetime consumption and labor supply, respectively. The functions ψ and τ are increasing with $\psi(0) = \tau(0) = 0$; ψ is bounded away from infinity and τ is bounded away from 1. Moreover, the marginal trade-off between costs (ψ) and benefits (τ) from extending capital taxation

$$\frac{d\psi}{d\tau} = \frac{\psi'(T)}{\tau'(T)}$$

is bounded away from infinity.

Given c, n and T we can compute the paths for consumption c_t and labor n_t. Behind the scenes, the static problem solves the dynamic problem. In particular, it optimizes over the path for labor taxes. In this static representation, $1+\psi(T)$ is akin to a production cost of consumption and $\tau(T)$ to a non-distortionary capital levy. On the one hand, higher T increases the efficiency cost from the consumption path. On the other hand, it increases revenue in proportion to the level of initial capital. Prolonging capital taxation requires trading off these costs and benefits.

Importantly, despite the connection between capital taxation and an ever increasing, unbounded tax on consumption, the proposition shows that the tradeoff between costs and benefits is bounded, $\frac{d\psi}{d\tau} < \infty$, even as $T \to \infty$. In other words, indefinite taxation does not come at an infinite marginal cost and helps explain why this may be optimal.

Should we be surprised that these results contradict commodity tax principles, as enunciated by Diamond and Mirrlees (1971), Atkinson and Stiglitz (1972) and others? No, not at all. As general as these frameworks may be, they do not consider upper bounds on taxation, the crucial ingredient in Chamley (1986) and Judd (1999). Their guiding principles are, therefore, ill adapted to these settings. In particular, formulas based on local elasticities do not apply, without further modification.

Effectively, a bound on capital taxation restricts the path for the consumption tax to lie below a straight line going through the origin. In the short run, the consumption tax is constrained to be near zero; to compensate, it is optimal to set higher consumption taxes in the future. As a result, it may be optimal to set consumption taxes as high as possible at all times. This is equivalent to indefinite capital taxation.

4 A Hybrid: Redistribution and Debt

Throughout this paper we have strived to stay on target and remain faithful to the original models supporting the Chamley-Judd result. This is important so that our own results are easily comparable to those in Judd (1985) and Chamley (1986). However, many contributions since then offer modifications and extensions of the original Chamley-Judd models and results. In this section we depart briefly from our main focus to show that our results transcend their original boundaries and are relevant to this broader literature.

To make this point with a relevant example, we consider a hybrid model, with redistribution between capitalists and workers as in Judd (1985), but sharing the essential feature in Chamley (1986) of unrestricted government debt. It is very simple to modify the model in Section 2 in this way. We add bonds to the wealth of capitalists $a_t = k_t + b_t$, modifying equation (1c) to

$$\beta U'(C_t)(C_t + k_{t+1} + b_{t+1}) = U'(C_{t-1})(k_t + b_t)$$

and the transversality condition to $\beta^t U'(C_t)(k_{t+1} + b_{t+1}) \to 0$. Together, these two conditions imply a present value implementability condition, which with $U(C) = C^{1-\sigma}/(1-\sigma)$ and initial returns on capital and bonds of R_0 and R_0^b is given by

$$(1-\sigma)\sum_{t=0}^{\infty}\beta^t U(C_t) = U'(C_0)(R_0 k_0 + R_0^b b_0). \tag{14}$$

Anticipated Confiscatory Taxation. For $\sigma > 1$ the left hand side in equation (14) is decreasing in C_t and the right hand side is decreasing in C_0. In particular, the values of C_t, for all $t = 0, 1, \ldots$, can be set infinitesimally small without violating (14). Since (14) is strictly speaking not defined for $C_t = 0$, the problem without weight on capitalists ($\gamma = 0$) has a supremum that can only be approximated as $C_t \to 0$. Given $\sigma > 1$, this limit can be implemented by making R_t infinitesimally small in some period $t \geq 1$, or, equivalently, setting the wealth tax (i.e. tax on gross returns) $\mathcal{T}_t$ in that period arbitrarily close to 100%. This same logic applies if the tax is temporarily restricted for periods $t \leq T - 1$ for some given T, but is unrestricted in period T.

Proposition 10. *Consider the two-class model from Section 2 but with unrestricted government bonds. Suppose $\sigma > 1$ and $\gamma = 0$. If capital taxation is unrestricted in at least one period, then the optimum (a supremum) features a wealth tax $\mathcal{T}_t \to 100\%$ in some period t and $C_t \to 0$ for all $t = 0, 1, \ldots$*

This result exemplifies how extreme the tax on capital may be without bounds. In addition to this result, even when $\sigma < 1$, if no constraints are imposed on taxation except at $t = 0$, then in the continuous time limit as the length of time periods shrinks to zero, taxation tends to infinity. This point was also raised in Chamley (1986) for the representative agent Ramsey model, and served as a motivation for imposing a stationary constraint, $R_t \geq 1$.

Long Run Taxation with Constraints. We now impose upper bounds on capital taxation and show that these constraints may bind forever, just as in Section 3.2.

Proposition 11. *Consider the two-class model from Section 2 but with unrestricted government bonds. Suppose $\sigma > 1$ and $\gamma = 0$. If capital taxation is restricted by the constraint $R_t \geq 1$, then at the optimum $R_t = 1$ in all periods t, i.e. capital should be taxed indefinitely.*

Intuitively, $\sigma > 1$ is enough to ensure indefinite taxation of capital in this model because $\gamma = 0$ makes it optimal to tax capitalists as much as possible. Similar results hold for positive but low enough levels of γ, so that redistribution from capitalists to workers is desired. The results also hold for less restrictive constraints than $R_t \geq 1$.

Proposition 11 assumes that transfers are perfectly targeted to workers and capitalists do not work. However, indefinite taxation, $T = \infty$, is also possible when these assumptions are relaxed, so that capitalists work and receive equal transfers.

Overall, these results suggest that indefinite taxation can be optimal in a range of models that are descendants of Chamley-Judd, with a wide range of assumptions regarding the environment, heterogeneity, social objectives and policy instruments.

5 Conclusions

This study revisited two closely related models and results, Chamley (1986) and Judd (1985). Our findings contradict well-known results or their standard interpretations. We showed that, provided the intertemporal elasticity of substitution (IES) is less than one, the long run tax on capital can actually be positive. Empirically, an IES below one is considered most plausible.

Why were the proper conclusions missed by Judd (1985), Chamley (1986) and many others? Among other things, these papers assume that the endogenous multipliers associated with the planning problem converge. Although this seems natural, we have shown that this is not necessarily true at the optimum. In fact, on closer examination it is evident that presuming the convergence of multipliers is equivalent to the assumption that the intertemporal rates of substitution of the planner and the agent are equal. This then implies that no intertemporal distortion or tax is required. Consequently, analyses based on these assumptions amount to little more than assuming zero long-run taxes.

In this paper, we have stayed away from evaluating the realism of the existing Chamley-Judd models or proposing an alternative model. Instead, we explored the implications of their assumptions. Different models offer different prescriptions and we should settle the mapping from models to prescriptions, on the one hand, and discuss the applicability of one model versus another, on the other hand. The scope of this paper has been concerned with the former, not the latter.

Even within the two models, it may well be the case that one finds a zero long-run tax on capital, e.g. for the model in Judd (1985) one may set $\sigma < 1$, and in Chamley (1986) the bounds may not bind forever if debt is low enough.[37] In this paper, we refrain from

[37] Any quantitative exercise could also evaluate the welfare gains from different policies. For example, even when $T < \infty$ is optimal, the optimal value of T may be very high and indefinite taxation, $T = \infty$, may closely approximate the optimum. One can also compare various non-optimal simple policies, such as

making any such claim, one way or another. We confined our attention to the original theoretical zero-tax results, widely perceived as delivering ironclad conclusions that are independent of parameter values or initial conditions. Based on our results, we have found little basis for such an interpretation.

References

Abel, Andrew B., "Optimal Capital Income Taxation," Working Paper 13354, National Bureau of Economic Research 2007.

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

Atkeson, Andrew, Varadarajan V Chari, and Patrick J Kehoe, "Taxing capital income: a bad idea," *Federal Reserve Bank of Minneapolis Quarterly Review*, 1999, *23*, 3–18.

Atkinson, Anthony B and Joseph E Stiglitz, "The structure of indirect taxation and economic efficiency," *Journal of Public Economics*, 1972, *1* (1), 97–119.

Banks, James and Peter A Diamond, "The base for direct taxation," in "Dimensions of Tax Design: The Mirrlees Review," Oxford University Press for The Institute for Fiscal Studies, 2010, chapter 6, pp. 548–648.

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

Benveniste, L.M. and J.A. Scheinkman, "On the Differentiability of the Value Function in Dynamic Models of Economics," *Econometrica*, 1979, *47* (3), 727–732.

Brunnermeier, Markus K, Thomas M Eisenbach, and Yuliy Sannikov, "Macroeconomics with financial frictions: A survey," Technical Report, National Bureau of Economic Research 2012.

Chamley, Christophe, "Optimal Intertemporal Taxation and the Public Debt," Cowles Foundation Discussion Papers 554 (available at http://ideas.repec.org/p/cwl/cwldpp/554.html), Cowles Foundation for Research in Economics, Yale University April 1980.

never taxing capital versus always taxing capital at a fixed rate.

—, “Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives,” *Econometrica*, 1986, *54* (3), pp. 607–622.

Chari, V.V. and Patrick J Kehoe, “Sustainable Plans,” *Journal of Political Economy*, 1990, *98* (4), 783–802.

— **and** —, “Optimal Fiscal and Monetary Policy,” *Handbook of Macroeconomics*, 1999, *1.*

—**, Lawrence J. Christiano, and Patrick J. Kehoe**, “Optimal Fiscal Policy in a Business Cycle Model,” *Journal of Political Economy*, 1994, *102* (4), 617–52.

Coleman II, Wilbur John, “Welfare and optimum dynamic taxation of consumption and income,” *Journal of Public Economics*, 2000, *76* (1), 1–39.

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

Diamond, Peter A and James A Mirrlees, “Optimal Taxation and Public Production: I–Production Efficiency,” *American Economic Review*, March 1971, *61* (1), 8–27.

E, Jr Lucas Robert, “Supply-Side Economics: An Analytical Review,” *Oxford Economic Papers*, April 1990, *42* (2), 293–316.

Erosa, Andres and Martin Gervais, “Optimal Taxation in Life-Cycle Economies,” *Journal of Economic Theory*, 2002, (105), 338–369.

Farhi, Emmanuel and Ivan Werning, “Progressive Estate Taxation,” *The Quarterly Journal of Economics*, May 2010, *125* (2), 635–673.

— **and Iván Werning**, “Estate Taxation with Altruism Heterogeneity,” *American Economic Review*, 2013, *103* (3), 489–495.

—**, Christopher Sleet, Iván Werning, and Sevin Yeltekin**, “Non-linear Capital Taxation Without Commitment,” *Review of Economic Studies*, October 2012, *79* (4), 1469–1493.

Gelfand, I. M. and S. V. Fomin, *Calculus of Variations*, Prentice-Hall Inc., 2000.

Gertler, Mark and Nobuhiro Kiyotaki, “Chapter 11 - Financial Intermediation and Credit Policy in Business Cycle Analysis,” in Benjamin M. Friedman and Michael Woodford, eds., *Handbook of Monetary Economics*, Vol. 3, Elsevier, 2010, pp. 547 – 599.

Greulich, Katharina, Sarolta Laczo, and Albert Marcet, “Pareto-Improving Optimal Capital and Labor Taxes,” *Working Paper*, 2016.

Jones, Larry E, Rodolfo E Manuelli, and Peter E Rossi, "Optimal Taxation in Models of Endogenous Growth," *Journal of Political Economy*, 1993, *101* (3), 485–517.

_ , _ , and _ , "On the Optimal Taxation of Capital Income," *Journal of Economic Theory*, 1997.

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

_ , "Optimal Taxation in Dynamic Stochastic Economies: Theory and Evidence," *Working Paper*, 1993.

_ , "Optimal taxation and spending in general competitive growth models," *Journal of Public Economics*, 1999, *71* (1), 1–26.

_ , "Capital-Income Taxation with Imperfect Competition," *American Economic Review Papers and Proceedings*, 2002, *92* (2), pp. 417–421.

Kemp, Murray C, Ngo Van Long, and Koji Shimomura, "Cyclical and Noncyclical Redistributive Taxation," *International Economic Review*, May 1993, *34* (2), 415–29.

Koopmans, Tjalling C, "Stationary ordinal utility and impatience," *Econometrica: Journal of the Econometric Society*, 1960, pp. 287–309.

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

Lucas, Robert E., Jr. and Nancy L. Stokey, "Optimal Fiscal and Monetary Policy in an Economy without Capital," *Journal of Monetary Economics*, 1983, *12*, 55–93.

_ and _ , "Optimal growth with many consumers," *Journal of Economic Theory*, 1984, *32* (1), 139 – 171.

Michel, Phillippe, "On the Transversality Condition in Infinite Horizon Optimal Problems," *Econometrica*, July 1982, *50* (4), 975–985.

Mikhail, Aleh Tsyvinski Golosov and Iván Werning, "New Dynamic Public Finance: A User?s Guide," *NBER Macroeconomics Annual*, 2006.

Phelan, Christopher and Ennio Stacchetti, "Sequential Equilibria in a Ramsey Tax Model," *Econometrica*, 2001, *69* (6), 1491–1518.

Piketty, Thomas and Emmanuel Saez, "A Theory of Optimal Inheritance Taxation," *Econometrica*, 2013, *81* (5), 1851–1886.

Reinhorn, Leslie J, "On optimal redistributive capital taxation," *mimeo*, 2002 and 2013.

Saez, Emmanuel, "Optimal progressive capital income taxes in the infinite horizon model," *Journal of Public Economics*, 2013, *97*, 61–74.

Sargent, Thomas J. and Lars Ljungqvist, *Recursive Macroeconomic Theory*, Vol. 2, MIT Press, 2004.

Seierstad, Atle and Knut Sydsaeter, *Optimal Control Theory with Economic Applications*, Elsevier, 1987.

Stokey, Nancy L., "Credible Public Policy," *Journal of Economic Dynamics and Control*, 1991, *15*, 627–656.

Uzawa, Hirofumi, "Time preference, the consumption function, and optimum asset holdings," *Value, Capital, and Growth: Papers in Honour of Sir John Hicks (University of Edinburgh Press, Edinburgh)*, 1968, pp. 485–504.

Werning, Ivan, "Optimal fiscal policy with redistribution," *The Quarterly Journal of Economics*, 2007, *122* (3), 925–967.

Zhu, Xiaodong, "Optimal Fiscal Policy in a Stochastic Growth Model," *Journal of Economic Theory*, 1992, *58*, 250–89.

Appendix (For Online Publication)

A Recursive Formulation of (1a)

In our numerical simulations, we use a recursive representation of the Judd (1985) economy. The two constraints in the planning problem feature the variables $C_{t-1}, k_t, C_t, k_{t+1}$ and c_t. This suggests a recursive formulation with (k_t, C_{t-1}) as the state and c_t as a control. The associated Bellman equation is then

$$\begin{aligned} V(k, C_-) = \max_{c \geq 0, (k', C) \in A} & \{u(c) + \gamma U(C) + \beta V(k', C)\} \\ & c + C + k' + g = f(k) + (1 - \delta)k \\ & \beta U'(C)(C + k') = U'(C_-)k \\ & c, C, k' \geq 0. \end{aligned} \tag{15}$$

Here, A is the *feasible* set, that is, states (k_0, C_{-1}) such that there exists a sequence $\{k_{t+1}, C_t\}$ satisfying all the constraints in (1) including the transversality condition. At $t = 0$, capital k_0 is given, so there is no need to impose $\beta U'(C_0)(C_0 + k_1) = U'(C_{-1})k_0$. Thus, the planner maximizes $V(k_0, C_{-1})$ with respect to C_{-1}. If V is differentiable, the first order condition is

$$V_C(k_0, C_{-1}) = 0.$$

Since one can show that $\mu_t = V_C(k_t, C_{t-1})U''(C_{t-1})k_t$, this is akin to the condition $\mu_0 = 0$ in equation (2a).[38]

B Proof of Proposition 3

The proof of Proposition 3 consists of three parts. In the first part, we provide a few definitions that are necessary for the proof. In particular, we define the *feasible* set of states. In the second part, we characterize the feasible set of states geometrically. The proofs for the results in that part are somewhat cumbersome and lengthy, so they are relegated to the end of this section to ensure greater readability. Finally, in the third part,

[38] Alternatively, we may impose that R_0 is taken as given, with $R_0 = R_0^*$ for example, to exclude an initial capital tax. In that case the planner solves

$$\max_{k_1, c_0, C_0} \{u(c_0) + \gamma U(C_0) + \beta V(k_1, C_0)\}$$

subject to

$$\begin{aligned} C_0 + k_1 &= R_0 k_0 \\ c_0 + C_0 + k_1 &= f(k_0) + (1 - \delta)k_0 \\ c_0, C_0, k_1 &\geq 0. \end{aligned}$$

This alternative gives rise to similar results.

we use our geometric results to prove Proposition 3. Readers interested only in the main steps of the proof are advised to jump straight to the third part.

B.1 Definitions

For the proof of Proposition 3 we make a number of definitions, designed to simplify the exposition. A state (k, C_-) as in the recursive statement (15) of problem (1a) will sometimes be abbreviated by z, and a set of states by Z. The total state space is denoted by $Z_{\text{all}} \subset \mathbb{R}_+^2$ and is defined below. It will prove useful at times to express the set of constraints in (15) as

$$k' = x - C_- \left(\frac{\beta x}{k}\right)^{1/\sigma} \tag{16a}$$

$$C = C_- \left(\frac{\beta x}{k}\right)^{1/\sigma} \tag{16b}$$

$$C_-^{\sigma/(\sigma-1)} \left(\frac{\beta}{k}\right)^{1/(\sigma-1)} \leq x \leq f(k) + (1-\delta)k - g, \tag{16c}$$

where $x = k' + C$ replaces $c = f(k) + (1-\delta)k - g - x$ as control. In the last equation, the first inequality ensures non-negativity of k' while the second inequality is merely the resource constraint. Substituting out x, we can also write the law of motion for capital as $k' = \frac{1}{\beta}\frac{k}{C_-^\sigma}C^\sigma - C$, which we will be using below.

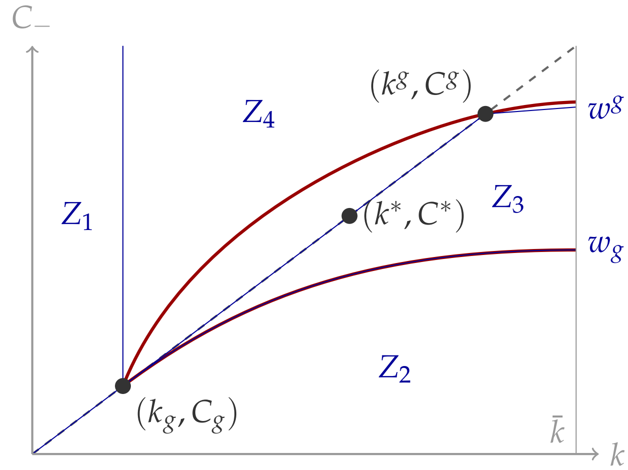
The whole set of future states z' which can follow a given state $z = (k, C_-)$ is denoted by $\Gamma(z)$, which can be the empty set. We will call a path $\{z_t\}$ *feasible* if (a) $z_{t+1} \in \Gamma(z_t)$ for all $t \geq 0$, which precludes $\Gamma(z_t)$ from being empty; and (b) if the transversality condition holds along the path, $\beta^t C_t^{-\sigma} k_{t+1} \to 0$. Similarly, a state z will be called *feasible*, if there exists a feasible (infinite) path $\{z_t\}$ starting at $z_0 = z$. In this case, z is *generated by* $\{z_t\}$. Because $z_1 \in \Gamma(z)$, we also say z is *generated by* z_1. A *steady state* $z = (k, C_-) \in \mathbb{R}_+^2$ is defined to be a state with $C_- = (1-\beta)/\beta k$. For very low and high capital levels k, steady states turn out to be infeasible, but all others are *self-generating*, $z \in \Gamma(z)$, as we argue below. Similarly, a set Z is called *self-generating* if every $z \in Z$ is generated by a sequence in Z. Denote by Z^* ($= A$ in the notation above) the set of all feasible states. An integral part of the proof will be to characterize Z^*.

It will be important to specify between which capital stocks the economy is moving. For this purpose, define k_g and $k^g > k_g$ to be the two roots to the equation

$$k = \underbrace{f(k) + (1-\delta)k - g}_{\equiv F(k)} - \frac{1-\beta}{\beta}k. \tag{17}$$

Demanding that $k^g > k_g$ is tantamount to specifying $F'(k^g) < 1/\beta < F'(k_g)$. Equation (17) was derived from the resource constraint, demanding that capitalists' consumption is at the steady state level of $C = \frac{1-\beta}{\beta}k$ and workers' consumption is equal to zero. Equation (17) need not have two solutions, not even a single one, in which case government con-

Figure 5: The state space of the Judd (1985) planning problem.



Note. This figure shows the two-dimensional state space of the Judd (1985) model. The entire state space is denoted by Z_{all}, which includes the feasible set Z^* (between the two red curves), and all sets Z_i (separated by the blue curves). The point (k^*, C^*) is the zero-tax steady state. Showing that this is the qualitative shape of the feasible set Z^* is an integral part of the proof of Proposition 3

sumption is unsustainably high for *any* capital stock. Such values for g are uninteresting and therefore ruled out. Corresponding to k_g and k^g, we define $C_g \equiv (1-\beta)/\beta\, k_g$ and $C^g \equiv (1-\beta)/\beta\, k^g$ as the respective steady state consumption of capitalists. The steady states (k_g, C_g) and (k^g, C^g) represent the lowest and highest feasible steady states, respectively. The reason for this is that the steady state resource constraint (17) is violated for any $k \notin [k_g, k^g]$.

As in the Neoclassical Growth Model, the set of feasible states of this model is easily seen to allow for arbitrarily large capital stocks. This is why we cap the state space for high values of capital, and we take the total state space to be $Z_{\text{all}} = [0, \bar{k}] \times \mathbb{R}_+$ for states (k, C_-), where $\bar{k} \equiv \max\{k_{\max}, k_0\}$ and $k = k_{\max}$ solves $k = f(k) + (1-\delta)k - g$. This way, the set of capital stocks that are resource feasible given an initial capital stock of k_0 must necessarily lie in the interval $[0, \bar{k}]$, so the restriction for $\bar{k}$ is without loss of generality for any given initial capital stock k_0. Note that with this state space, the set of feasible states Z^* is also capped at $\bar{k}$ in its k-component.

We now characterize the geometry of the set of feasible states Z^*. The results derived there are essential for the actual proof of Proposition 3 in Section B.3.

B.2 Geometry of Z^*

For better guidance through this section, we refer the reader to figure 5, which shows the typical shape of Z^*. The main results in this section are characterizations of the bottom and top boundaries of Z^*. We proceed by splitting up the state space, $Z_{\text{all}} = [0, \bar{k}] \times \mathbb{R}_+$, into four pieces and characterizing the feasible states in each of the four pieces.

Define

$$w_g(k) \equiv \begin{cases} \frac{1-\beta}{\beta}k & \text{for } 0 \leq k \leq k_g \\ C_g \left(\frac{k}{k_g}\right)^{1/\sigma} & \text{for } k_g \leq k \leq \bar{k} \end{cases}$$

$$w^g(k) \equiv \begin{cases} \frac{1-\beta}{\beta}k & \text{for } 0 \leq k \leq k^g \\ C^g \left(\frac{k}{k^g}\right)^{1/\sigma} & \text{for } k^g \leq k \leq \bar{k}, \end{cases}$$

and split up the state space as follows (see figure 5)

$$Z_{\text{all}} = \underbrace{\left\{k < k_g, C_- \geq \frac{1-\beta}{\beta}k\right\}}_{Z_1} \cup \underbrace{\{C_- < w_g(k)\}}_{Z_2}$$
$$\cup \underbrace{\{k \geq k_g, w_g(k) \leq C_- \leq w^g(k)\}}_{Z_3} \cup \underbrace{\{k \geq k_g, C_- \geq w^g(k)\}}_{Z_4}.$$

Lemma 1 characterizes the feasible states in sets Z_1 and Z_2.

Lemma 1. *$Z^* \cap Z_1 = Z^* \cap Z_2 = \emptyset$. All states with $k < k_g$ or $C_- < w_g(k)$ are infeasible.*

Proof. See Subsection B.4.1. □

In particular, Lemma 1 shows that all states with $C_- < w_g(k)$ are infeasible. Lemma 2 below complements this result stating that all states with $w_g(k) \leq C_- \leq w^g(k)$ (and $k \geq k_g$) in fact are feasible, that is, lie in Z^*. This means, $\{C_- = w_g(k), k \geq k_g\}$ constitutes the lower boundary of the feasible set Z^*.

Lemma 2. *$Z_3 \subseteq Z^*$, or equivalently, all states with $w_g(k) \leq C_- \leq w^g(k)$ and $k \geq k_g$ are feasible and generated by a feasible steady state. Moreover, states on the boundary $\{C_- = w_g(k), k > k_g\}$ can only be generated by a single feasible state, (k_g, C_g). Thus, there is only a single "feasible" control for those states, $c > 0$.*

Proof. See Subsection B.4.2. □

Lemma 2 finishes the characterization of all feasible states with $C_- \leq w^g(k)$. What remains is a characterization of feasible states with $C_- > w^g(k)$, or in terms of the $k - C_-$ diagram of Figure 5, the characterization of the red top boundary. This boundary is inherently more difficult than the bottom boundary because it involves states that are not merely one step away from a steady state. Rather, paths might not reach a steady state at all in finite time. The goal of the next set of lemmas is an iterative construction to show that the boundary takes the form of an increasing function $\bar{w}(k)$ such that states with $C_- > w^g(k)$ are feasible if and only if $C_- \leq \bar{w}(k)$.

For this purpose, we need to make a number of new definitions: Let $\psi(k, C_-) \equiv (k + C_-)/C_-^\sigma$. Applying the ψ function to the successor (k', C) of a state (k, C_-) and using the IC constraint (1c) gives $\psi(k', C) = \beta^{-1}k/C_-^\sigma$, a number that is *independent of the control x*. Hence, for every state (k, C_-) there exists an iso-ψ curve containing all its potential successor states.

In some situations it will be convenient to abbreviate the laws of motion for capitalists' consumption and capital, equations (16a) and (16b), as $k'(x, k, C_-)$ and $C(x, k, C_-)$.

Finally, define an operator T on the space of continuous, increasing functions $v : [k_g, \bar{k}] \to \mathbb{R}_+$, as,

$$Tv(k) = \sup\{C_- \mid \exists x \in (0, F(k)] : v(k'(x, k, C_-)) \geq C(x, k, C_-)\}, \tag{18}$$

where recall that $F(k) = f(k) + (1-\delta)k - g$, as in (17). The operator is designed to extend a candidate top boundary of the set of feasible states by one iteration. To make this formal, let $Z^{(i)}$ be the set of states with $C_- \geq w^g(k)$ which are i steps away from reaching $C_- = w^g(k)$. For example, $Z^{(0)} = \{C_- = w^g(k)\}$. Lemma 3 proves some basic properties of the operator T.

Lemma 3. *T maps the space of continuous, strictly increasing functions $v : [k_g, \bar{k}] \to \mathbb{R}_+$ with $\psi(k, v(k))$ strictly decreasing in k and $v(k_g) = C_g$, $v(k^g) = C^g$, into itself.*

Proof. See Subsection B.4.3. □

Lemma 4 uses the operator T to describe the sets $Z^{(i)}$.

Lemma 4. *$Z^{(i)} = \{w^g(k) \leq C_- \leq T^i w^g(k)\}$. In particular $T^i w^g(k) \geq T^j w^g(k) \geq w^g(k)$ for $i \geq j$.*

Proof. See Subsection B.4.4. □

The next two lemmas characterize the limit function $\bar{w}(k)$, whose graph will describe the top boundary of the set of feasible states.

Lemma 5. *There exists a continuous limit function $\bar{w}(k) \equiv \lim_{i\to\infty} T^i w^g(k) = T\bar{w}(k)$, with $\bar{w}(k_g) = C_g$ and $\bar{w}(k^g) = C^g$. All states with $C_- = \bar{w}(k)$ are feasible, but only with policy $c = 0$.*

Proof. See Subsection B.4.5. □

Lemma 6. *No state with $C_- > \bar{w}(k)$ (and $k_g \leq k \leq \bar{k}$) is feasible.*

Proof. See Subsection B.4.6. □

Finally, Lemma 7 shows an auxiliary result which is both used in the proof of Lemma 6 and in Lemma 9 below.

Lemma 7. *Let $\{k_{t+1}, C_t\}$ be a path starting at (k_0, C_{-1}) with controls $c_t = 0$. Let $k_g < k_0 \leq \bar{k}$. Then:*

(a) If $C_{-1} = \bar{w}(k_0)$, $(k_{t+1}, C_t) \to (k^g, C^g)$.

(b) If $C_{-1} > \bar{w}(k_0)$, $(k_{t+1}, C_t) \not\to (k^g, C^g)$.

Proof. See Subsection B.4.7. □

B.3 Proof of Proposition 3

Armed with the results from Section B.2 we now prove Proposition 3 in a series of intermediate results. For all statements in this section, we consider an economy with an initial capital stock of $k_0 \in [k_g, \bar{k}]$. We call a path $\{k_{t+1}, C_t\}$ *optimal path*, if the initial C_{-1} was optimized over given the initial capital stock k_0. Analogously, we call a path $\{k_{t+1}, C_t\}$ *locally optimal path*, if the initial C_{-1} was not optimized over but rather taken as given at a certain level, respecting the constraint that (k_0, C_{-1}) be feasible. If $\{k_{t+1}, C_t\}$ is a locally optimal path, with control c_{t+1} at some point $\{k_{t+1}, C_t\}$ we say this control is *optimal at* $\{k_{t+1}, C_t\}$. Notice that along both optimal and locally optimal paths, first order conditions are *necessary*, as long as paths are interior; they need not be sufficient, in the sense that there could be multiple optima that satisfy our characterization below.

The first lemma proves that the multiplier on the capitalists' IC constraint explodes along an optimal path, and at the same time, workers' consumption drops to zero.

Lemma 8. *Along any optimal path, $c_t \to 0$.*

Proof. Let $\{k_{t+1}, C_t\}$ be the optimal path. Suppose first the optimal path hits the boundary of the feasible set Z^* at some finite time. Given that no path can hit the $k = \bar{k}$ boundary after $t = 0$, and given Lemma 2 this means the path hits the top boundary—the graph of $\bar{w}$—after finite time. Lemma 5 showed that along that boundary, the control is necessarily zero, $c = 0$.

Now suppose the optimal path is interior at all times. In that case, the first order conditions are necessary. Using the notation from problem (1a) the necessary first order conditions are equations (2a)–(2d). In particular, the one for μ_t states

$$\mu_{t+1} = \mu_t \left(\frac{\sigma - 1}{\sigma \kappa_{t+1}} + 1 \right) + \frac{1}{\beta \sigma \kappa_{t+1} v_t}.$$

From Lemma 1 we know that $\kappa_{t+1} = k_{t+1}/C_t$ is bounded away from ∞. Since $\mu_0 = 0$ by (2a) and $\sigma > 1$, it follows that $\mu_t \geq 0$ and $\mu_t \to \infty$. To show that $c_t \to 0$, suppose to the contrary that $c_t \not\to 0$. In this case, there exists $\underline{c} > 0$ and an infinite sequence of indices (t_s) such that $c_{t_s} \geq \underline{c}$ for all s. Along these indices, the FOC for capital (2d) implies

$$\underbrace{u'(c_{t_s})}_{\leq u'(\underline{c})}(f'(k_{t_s}) + (1-\delta)) = \frac{1}{\beta} \underbrace{u'(c_{t_s - 1})}_{\geq 0} + \underbrace{U'(c_{t_s - 1})}_{\text{bounded away from } 0} \cdot \underbrace{(\mu_{t_s} - \mu_{t_s - 1})}_{\geq \text{const} \cdot \mu_{t_s - 1} \to \infty},$$

and so $k_{t_s} \to 0$ for $s \to \infty$, which is impossible within the feasible set Z^* because it violates $k \geq k_g$ (see Lemma 1). This proves that also for interior optimal paths, $c_t \to 0$. ☐

Lemma 8 is important because it shows that workers' consumption drops to zero. Together with the following lemma, this gives us a crucial geometric restriction of where an optimal path goes in the long run.

Lemma 9. *The set of states where $c = 0$ is an optimal control is the top boundary, the graph of $\bar{w}$. It follows that an optimal path approaches either (k_g, C_g) or (k^g, C^g).*

Proof. First, we show that any state in the interior of Z^* can be generated by a path with positive controls $c > 0$. Any state in the interior of Z^* is element of some $Z^{(i)}$, $i < \infty$, and can thus reach the set $\{C_- \leq w^g(k)\} \setminus \{(k_g, C_g), (k^g, C^g)\}$ in finite time. From there, at most two steps are necessary to reach a interior steady state (k_{ss}, C_{ss}) with $k_g < k_{ss} < k^g$ and hence positive consumption $c_{ss} > 0$. Note that such an interior steady state can be reached without leaving the interior of the feasible set, since by Lemmas 2 and 7, hitting the upper or lower boundary once means convergence to a non-interior steady state.[39] This proves that any state in the interior is generated by such an interior path, with positive controls $c > 0$.

Now take an interior state (k_0, C_{-1}). We prove that any optimal control at that state is positive. Suppose to the contrary, $c_0 = 0$ is an optimal control at (k_0, C_{-1}). This means, (k_0, C_{-1}) is generated by a locally optimal path $\{k_{t+1}, C_t\}$, where (k_1, C_0) is precisely linked to (k_0, C_{-1}) using control $c_0 = 0$, or equivalently, $x_0 = F(k_0)$. Since (k_0, C_{-1}) is interior, any state $(k'(\tilde{x}_0, k_0, C_{-1}), C(\tilde{x}_0, k_0, C_{-1}))$ with slightly positive controls, that is, $\tilde{x}_0 < F(k_0)$, has to be feasible too. Therefore, we find the following first order necessary condition for local optimality of c_0,[40]

$$\frac{u'(c_1)}{u'(c_0)}(f'(k_1) + 1 - \delta) \geq \frac{1}{\beta} + v_0(\mu_1 - \mu_0),$$

where the inequality is there due to the (implicit) boundary condition $c_0 \geq 0$. This condition can only be satisfied if $c_1 = 0$ as well. We can iterate this logic: If (k_1, C_0) is interior, it must be that $c_2 = 0$ is optimal at (k_2, C_1). If (k_1, C_0) is not interior, then it must be on the top boundary of Z^*, that is, on the graph of $\bar{w}$,[41] where it has policy $c = 0$ forever after. This proves, by induction, that if any interior state (k, C_-) has $c = 0$ as an optimal policy, any locally optimal path starting at (k, C_-) with $c = 0$ as initial optimal policy must have $c = 0$ forever, yielding utility $u(0)/(1-\beta)$. This, however, contradicts local optimality of such a path: We showed above that any interior state (k_0, C_{-1}) is generated by a path with strictly positive controls. Therefore, any optimal control at an interior state (k_0, C_{-1}) is positive.

Finally, notice that states (k, C_-), $k > k_g$, along the bottom boundary of Z^* only admit a single feasible control, which is positive (see Lemma 2). Thus, by Lemma 5, the set where $c = 0$ is an optimal control is precisely the top boundary $\{(k, C_-) \mid k \in [k_g, \bar{k}], C_- = \bar{w}(k)\}$. It follows that an optimal path either hits the boundary of Z^* at some point, in which case it converges either to (k_g, C_g) or (k^g, C^g) (by Lemma 7), or it remains interior forever and thus (by Lemma 8) approaches the set $\{c = 0\}$ of all states where $c = 0$ is an optimal control, that is, the graph of $\bar{w}$.[42] Then it must share the same limiting behavior

[39] Note that hitting the right boundary at $k = \bar{k}$ (other than with k_0) is of course not feasible due to depreciation.

[40] A *locally* optimal path still satisfies the first order conditions (2b)–(2d), just not (2a) which comes from the optimal choice of C_{-1}.

[41] On the lower boundary of Z^* (excluding (k_g, C_g)), a policy of $c = 0$ would not be feasible, see Lemma 2.

[42] By the Maximum Theorem, the control c is upper hemicontinuous in the state, so its graph is closed. Hence, if along a path $\{k_{t+1}, C_t\}$ it holds that $c_t \to 0$, then $\{k_{t+1}, C_t\}$ necessarily approximates the set $\{c = 0\}$, in the sense that the distance between $\{k_{t+1}, C_t\}$ and the set shrinks to zero (or else you could take

as states in the set $\{c=0\}$.[43] By virtue of Lemma 7, it can then either converge to (k_g, C_g) or (k^g, C^g). $\square$

Lemma 10. *If an optimal path $\{k_{t+1}, C_t\}$ converges to (k^g, C^g), then the value function V is locally decreasing in C at each point (k_{t+1}, C_t), for all $t > T$, with T large enough.*

Proof. Let $x_t \equiv F(k_t) - c_t$ and consider the following variation: Suppose that at a point T, (k_{T+1}, C_T) is not at the lower boundary (in which case it cannot converge to (k^g, C^g) anyway) and that $c_t < F(k_t) - F'(k_t)k_t$ for all $t \geq T$.[44] For simplicity, call this $T=-1$. Do the perturbation $\hat{C}_{-1} \equiv C_{-1} - \epsilon$, $\hat{k}_0 = k_0$, but keep the controls c_t at their optimal level for (k_0, C_{-1}), that is $\hat{c}_t = c_t$. Denote the perturbed capital stock and capitalists' consumption by $\hat{k}_{t+1} = k_{t+1} + dk_{t+1}$ and $\hat{C}_t = C_t + dC_t$. Then the control x changes by $dx_t = F'_t dk_t$ to first order. We want to show that $dk_{t+1} > 0$ and $dC_t < 0$ for all $t \geq 0$, knowing that $dC_{-1} = -\epsilon$ and $dk_0 = 0$.

From the constraints we find,

$$dk_{t+1} = \underbrace{F'(k_t)dk_t}_{\geq 0} \underbrace{-\frac{C_t}{C_{t-1}}dC_{t-1}}_{>0} + \underbrace{\frac{1}{\sigma}\frac{C_t}{x_t}\frac{F(k_t) - F'(k_t)k_t - c_t}{k_t}dk_t}_{\geq 0} > 0$$

$$dC_t = \underbrace{\frac{C_t}{C_{t-1}}dC_{t-1}}_{\leq 0} \underbrace{-\frac{1}{\sigma}\frac{C_t}{x_t}\frac{F(k_t) - F'(k_t)k_t - c_t}{k_t}dk_t}_{\leq 0} < 0.$$

Using matrix notation, this local law of motion can be written as

$$\begin{pmatrix} dk_{t+1} \\ dC_t \end{pmatrix} = \begin{pmatrix} a_t + b_t & -d_t \\ -b_t & d_t \end{pmatrix} \begin{pmatrix} dk_t \\ dC_{t-1} \end{pmatrix},$$

with $a_t = F'(k_t)$, $d_t = C_t/C_{t-1}$, $b_t = \frac{1}{\sigma}\frac{C_t}{x_t}\frac{F(k_t)-F'(k_t)k_t-c_t}{k_t}$. Close to (k^g, C^g), this matrix has $d \approx 1$. Suppose for one moment that a was zero; the fact that $a > 0$ only works in favor of the following argument. With $a = 0$, the matrix has a single nontrivial eigenvalue of $b + d$, which exceeds 1 strictly in the limit, and the associated eigenspace is spanned by $(1, -1)$. The trivial eigenvalue's eigenspace is spanned by (d, b). Notice that the latter eigenvector is not collinear with the initial perturbation $(0, -1)$, implying that $dk_\infty > 0$ and $dC_\infty < 0$. Hence, $\hat{k}_\infty > k_\infty = k^g$ and $\hat{C}_\infty < C_\infty = C^g$.

But notice that to the bottom right of (k^g, C^g), the new point is interior, which implies a continuation value strictly larger than $u(0)/(1-\beta)$ (see proof of Lemma 9). More

a subsequence $\{k_{n_t+1}, C_{n_t}, c_{n_t}\}$ in the graph of c whose limit is not in the graph, contradicting the graph being closed).

[43]The formal reason for this is as follows: Suppose the optimal path $\{k_{t+1}, C_t\}$ did not share the limiting behavior of the set $\{c=0\}$, that is, suppose the path had a convergent subsequence $\{k_{n_t+1}, C_{n_t}\} \to \{k^*, C^*\} \in \{c=0\} \setminus \{(k_g, C_g), (k^g, C^g)\}$. Suppose $k^* \in (k_g, k^g)$, the case $k^* > k^g$ is analogous. Because $\bar{w}(k^*) > \frac{1-\beta}{\beta}k^*$, $h(k_{n_t+1}, C_{n_t})$ is eventually strictly decreasing in t (see logic around equation (30)) and converges to $h(k^*, C^*)$. But convergence of $h(k_{n_t+1}, C_{n_t})$ implies $C^* = \frac{1-\beta}{\beta}k^*$—a contradiction.

[44]Such a finite $T > 0$ exists for two reasons: (a) because $c_t \to 0$; and (b) because $F(k) - F'(k)k$ which is positive in a neighborhood around $k = k^g$ since k^g was defined by $F(k^g) = k^g/\beta$ and $F'(k^g) < 1/\beta$.

formally, this means there must exist a time $T' > 0$ for which the continuation value of $(k_{T'+1}, C_{T'})$ is strictly dominated by the one for $(\hat{k}_{T'+1}, \hat{C}_{T'})$, that is, $V(k_{T'+1}, C_{T'}) < V(\hat{k}_{T'+1}, \hat{C}_{T'})$. Because all controls were equal up until time T', this implies that $V(k_{T+1}, C_T) < V(k_{T+1}, C_T - \epsilon)$ for ϵ small (Recall that we had set $T = -1$ during the proof). Thus, the value function must increase if C_T is lowered, for a path starting at (k_{T+1}, C_T), for large enough T. This proves that the value function is locally decreasing in C at that point. □

And finally, Lemma 11 proves Proposition 3.

Lemma 11. *An optimal path converges to (k_g, C_g).*

Proof. By Lemma 9 it is sufficient to prove that an optimal path does not converge to (k^g, C^g). Suppose the contrary held and there was an optimal path converging to (k^g, C^g). By Lemma 10, this means that the value function is locally decreasing around the optimal path (k_{t+1}, C_t) for $t \geq T$, with $T > 0$ sufficiently large. Consider the following feasible variation for $t = -1, 0, \ldots, T$, $\hat{C}_t = C_t(1 - d\epsilon_t)$, $\hat{k}_{t+1} = k_{t+1}$, $\hat{x}_t = x_t - C_t d\epsilon_t$ where[45]

$$d\epsilon_t = \left(1 - \frac{1}{\sigma}\frac{C_t}{x_t}\right)^{-1} d\epsilon_{t-1}. \tag{19}$$

Observe that (19) is precisely the relation which ensures that the variation satisfies all the constraints of the system (in particular (16b) of which (19) is the linearized version). Workers' consumption increases with this variation by $dc_t = C_t d\epsilon_t > 0$. Therefore, the value of this path changes by

$$dV = \underbrace{\sum_{t=0}^{T} \beta^t u'(c_t) dc_t}_{>0} + \beta^{T+1} \underbrace{\left(V(k_{T+1}, C_T - C_T d\epsilon_T) - V(k_{T+1}, C_T)\right)}_{>0,\ \text{by Lemma 10}} > 0,$$

which is contradicting the optimality of $\{k_{t+1}, C_t\}$. An optimal path converges to (k_g, C_g). □

B.4 Proofs of Auxiliary Lemmas

B.4.1 Proof of Lemma 1

Proof. Focus on Z_1 first and consider a state $(k_1, C_0) \in Z_1$, that is, $k_1 < k_g$ and $C_0 \geq \frac{1-\beta}{\beta} k_1$. Suppose (k_1, C_0) was feasible, and as such generated by a path of states $\{(k_{t+1}, C_t)\}_{t \geq 0}$, each of which compatible with (16a)–(16c). We now show by induction the claim that $(k_{t+1}, C_t) \in Z_1$ and $k_{t+1} \leq \beta F(k_t)$ for any $t \geq 0$. This will lead to a contradiction since $\beta F(k)$ is a concave and increasing function with $\beta F(0) < 0$ and smallest fixed point $\beta F(k_g) = k_g$. Thus, any sequence of capital stocks $\{k_{t+1}\}$ satisfying $k_{t+1} \leq \beta F(k_t)$, starting at any $k_1 < k_g$, necessarily drops below zero in finite time, contradicting feasibility.

[45] Notice that $x_t = C_t + k_{t+1} \geq C_t$ by definition of x_t, and $\sigma > 1$. Hence this expression is well defined.

Pick a point (k_t, C_{t-1}) of the sequence and assume $(k_t, C_{t-1}) \in Z_1$. Then, $x_{t+1} \equiv k_{t+1} + C_t \leq F(k_t)$ by (16c), and so

$$k_{t+1} = x_{t+1} - C_t \underbrace{\left(\frac{\beta x_{t+1}}{k_t}\right)^{1/\sigma}}_{\geq \beta x_{t+1}/k_t} \leq \beta x_{t+1}\left(\frac{1}{\beta} - \frac{C_t}{k_t}\right) \leq \beta x_{t+1} \leq \beta F(k_t), \tag{20}$$

where in the first inequality we used the fact that $\beta x_{t+1}/k_t \leq \beta F(k_t)/k_t < 1$ which holds since $k_t < k_g$; and in the second inequality we used that $C_{t-1} \geq \frac{1-\beta}{\beta} k_t$. Building on (20), the fact that $k_{t+1} \leq \beta x_{t+1}$ proves that

$$C_t = x_{t+1} - k_{t+1} \geq \frac{1-\beta}{\beta} k_{t+1}. \tag{21}$$

To sum up, this implies that $k_{t+1} \leq \beta F(k_t) < k_g$ and that $C_t \geq \frac{1-\beta}{\beta} k_{t+1}$, so $(k_{t+1}, C_t) \in Z_1$. Moreover, $k_{t+1} \leq \beta F(k_t)$. This proves the aforementioned claim and hence the desired contradiction. No state in Z_1 is feasible.

Now consider a state $(k_1, C_0) \in Z_2$. Again, suppose it was generated by a path of feasible states $\{(k_{t+1}, C_t)\}$. Define $h(k, C_-) \equiv k/C_-^\sigma$ for any state (k, C_-). The proof idea is to show the claim that $(k_{t+1}, C_t) \in Z_2$ for all t and that $h(k_{t+1}, C_t)$ is strictly increasing and diverges to $+\infty$. Since k_{t+1} is bounded from above by $\bar{k}$, this will mean that $C_t \to 0$. Moreover, k_{t+1} is bounded away from zero since feasibility requires $\beta F(k) \geq 0$ and $\beta F(k)$ turns negative for k sufficiently close to zero. Lemma 12 below proves that this combination of convergence of C_t to zero and k_{t+1} bounded away from zero violates the transversality condition.

We now prove the aforementioned claim by induction. Take a state $(k_t, C_{t-1}) \in Z_2$ from the sequence. By construction of Z_2, it holds that $C_{t-1} < w_g(k_t)$, or in particular, $(C_{t-1}/C_g)^\sigma < k_t/k_g$.[46] Notice that if the next state in the sequence, (k_{t+1}, C_t), satisfied $C_t \geq \frac{1-\beta}{\beta} k_{t+1}$, we must have $(k_{t+1}, C_t) \in Z_1$ which is infeasible according to the above.[47] Therefore, $C_t < \frac{1-\beta}{\beta} k_{t+1}$. Then,

$$h(k_{t+1}, C_t) = \frac{k_{t+1}}{C_t^\sigma} = \frac{k_{t+1}}{C_{t-1}^\sigma \beta x_{t+1}/k_t} = \frac{k_t}{C_{t-1}^\sigma} \underbrace{\frac{k_{t+1}}{\beta(k_{t+1} + C_t)}}_{>1} > h(k_t, C_{t-1}), \tag{22}$$

[46]This inequality even holds if $k_t < k_g$ because there, $C_g(k_t/k_g)^{1/\sigma} > (1-\beta)/\beta k_t$. To see this recall that $C_g = (1-\beta)/\beta k_g$ and so $C_g(k_t/k_g)^{1/\sigma}/((1-\beta)/\beta k_t) = (k_t/k_g)^{1/\sigma - 1} > 1$, where we used $\sigma > 1$.

[47]Note that if $C_t \geq (1-\beta)/\beta\, k_{t+1}$, then $k_{t+1} < k_g$. The reason is as follows: The constraints (16a) and (16b) can be rewritten as $k_{t+1} = (C_t/C_{t-1})^\sigma k_t/\beta - C_t$. Because $(C_{t-1}/C_g)^\sigma < k_t/k_g$, this implies that $k_{t+1} > (C_t/C_g)^\sigma k_g/\beta - C_t$. Note that the right hand side of this inequality is increasing in C_t as long as it is positive (which is the only interesting case here). Substituting in $C_t \geq (1-\beta)/\beta\, k_{t+1}$, this gives $k_{t+1} > (k_{t+1}/k_g)^\sigma k_g/\beta - (1-\beta)/\beta k_{t+1}$. Rearranging, $k_{t+1}/k_g > (k_{t+1}/k_g)^\sigma$, a condition which can only be satisfied if $k_{t+1}/k_g < 1$ (recall that $\sigma > 1$).

which, together with $C_t < \frac{1-\beta}{\beta} k_{t+1}$ implies that both $(k_{t+1}, C_t) \in Z_2$ and $h(k_{t+1}, C_t)$ is strictly increasing in t. To show that $h(k_{t+1}, C_t)$ diverges to $+\infty$, suppose it were the case that $h(k_{t+1}, C_t)$ converged to some $H > 0$. Using (22), convergence of $h(k_{t+1}, C_t)$ would imply that $k_{t+1} / \left(\beta(k_{t+1} + C_t)\right) \to 1$, or equivalently that $k_{t+1}/C_t \to \beta/(1-\beta)$. Since k_{t+1} is bounded away from zero (see argument in previous paragraph), this can only be the case if (k_{t+1}, C_t) converges to a feasible steady state,[48] that is some $\left(k, \frac{1-\beta}{\beta} k\right)$ with $k_g \leq k \leq k^g$. However, as $(k_{t+1}, C_t) \in Z_2$ for any t, it is the case that $(C_t/C_g)^\sigma < k_{t+1}/k_g$, or,

$$h(k_{t+1}, C_t) > h(k_g, C_g) = \sup_{k_g \leq k \leq k^g} h(k, (1-\beta)/\beta k),$$

where the equality follows because $k/((1-\beta)/\beta k)^\sigma$ is decreasing in k. This shows that $h(k_{t+1}, C_t) \to \infty$ and hence completes the proof by contradiction. No state in Z_2 is feasible. □

Lemma 12. *Suppose that $C_t \to 0$ and k_{t+1} bounded away from zero for a given path of states (k_{t+1}, C_t). Then, this path is not feasible.*

Proof. Suppose the path (k_{t+1}, C_t) is feasible. In particular, this necessitates that the IC condition $\beta U'(C_t)(C_t + k_{t+1}) = U'(C_{t-1})k_t$ and the transversality condition $\beta^t U'(C_t) k_{t+1} \to 0$ hold. We back out (after tax) interest rates from the allocation as $R_t \equiv U'(C_{t-1})/(\beta U'(C_t))$. Thus we can recover the capitalists' per period budget constraint $C_t + k_{t+1} = R_t k_t$, and, using the transversality condition, also present value budget constraints starting at any given time $t_0 \geq 0$,

$$\sum_{t=t_0}^{\infty} \frac{1}{\overline{R}_{t_0,t}} C_t = R_{t_0} k_{t_0}, \tag{23}$$

where we denote $\overline{R}_{t_0,t} \equiv R_{t_0+1} \cdots R_t$. Also, by construction of R_t, consumption can be expressed as

$$C_t = \beta^{(t-t_0)/\sigma} \left(\overline{R}_{t_0,t}\right)^{1/\sigma} C_{t_0}. \tag{24}$$

Define $\underline{K} \equiv \inf_t k_{t+1} > 0$ and $\overline{K} \equiv \sup_t k_{t+1} > 0$. Using the per period budget constraints, we then have

$$R_t = \frac{C_t + k_{t+1}}{k_t} \geq \frac{k_{t+1}}{k_t}$$

and similarly,

$$\overline{R}_{t_0,t} \geq \frac{k_{t+1}}{k_{t_0+1}} \geq \frac{\underline{K}}{\overline{K}}. \tag{25}$$

Combining (23), (24) and (25), we find

$$k_{t_0} = \frac{1}{R_{t_0}} \sum_{t=t_0}^{\infty} \beta^{(t-t_0)/\sigma} \underbrace{\left(\overline{R}_{t_0,t}\right)^{-(1-1/\sigma)}}_{\leq \left(\underline{K}/\overline{K}\right)^{-(1-1/\sigma)}} C_{t_0} \leq C_{t_0} \frac{\left(\underline{K}/\overline{K}\right)^{-(2-1/\sigma)}}{1-\beta^{1/\sigma}}.$$

[48] Notice that, if $k_{t+1}/C_t^\sigma \to H > 0$ and $k_{t+1}/C_t \to \beta/(1-\beta)$ then convergence of k_{t+1} and C_{t+1} themselves immediately follows.

Since t_0 was arbitrary, this implies that $k_t \to 0$, leading to the desired contradiction. Thus, the path (k_{t+1}, C_t) cannot be feasible. $\square$

B.4.2 Proof of Lemma 2

Proof. Consider a state (k, C_-) with $w_g(k) \le C_- \le w^g(k)$ and $k \ge k_g$. In particular, $C_- \le (1-\beta)/\beta k$, $(C_-/C_g)^\sigma \ge k/k_g$ and $(C_-/C^g)^\sigma \le k/k^g$.[49] The idea of the proof is to show that in fact such a state can be generated by a steady state (k_{ss}, C_{ss}) (with $C_{ss} = (1-\beta)/\beta k_{ss}$ and $k_g \le k_{ss} \le k^g$). By definition of k_g and k^g, such a steady state is always self-generating.

Guess that the right steady state has $k_{ss} = (\beta C_-/(1-\beta))^{\sigma/(\sigma-1)} k^{-1/(\sigma-1)}$ and $C_{ss} = (1-\beta)/\beta k_{ss}$. It is straightforward to check that this steady state can be attained with control $x = (C_{ss}/C_-)^\sigma k/\beta$. This steady state is self-generating because $k_g \le k_{ss} \le k^g$, which follows from $(C_-/C_g)^\sigma \ge k/k_g$ and $(C_-/C^g)^\sigma \le k/k^g$. Finally, the control x is resource-feasible because $C_- \le (1-\beta)/\beta k$ and thus,

$$x = \frac{1}{\beta}\left[\frac{\left(\frac{\beta}{1-\beta}C_-\right)^\sigma}{k}\right]^{1/(\sigma-1)} \le \frac{k}{\beta} \le f(k) + (1-\delta)k - g,$$

where the latter inequality follows from the fact that $k_g \le k \le k^g$ and the definition of k_g and k^g. This concludes the proof that all states with $w_g(k) \le C_- \le w^g(k)$ and $k \ge k_g$ are feasible.

Now regard a state on the boundary $\{C_- = w_g(k), k > k_g\}$, so we also have that $C_- < (1-\beta)/\beta k$.[50] Such a state is generated by $(k_{ss}, C_{ss}) = (k_g, C_g)$. Moreover, the unique control which moves (k, C_-) to (k_g, C_g) is $x < k/\beta \le f(k) + (1-\delta)k - g$, or in terms of c, $c > 0$.

To show that (k_g, C_g) is in fact the only feasible state generating (k, C_-), let (k', C) be a state generating (k, C_-). If $k' < k_g$, then (k', C) is not feasible by Lemma 1, and if $k' = k_g$ only (k_g, C_g) generates (k, C_-). Suppose $k' > k_g$. Then, $C < (1-\beta)/\beta k'$,[51] and so we can recycle equation (22) to see $h(k', C) > h(k, C_-)$. Because $h(k, C_-) = h(k_g, C_g)$ however, this implies that $h(k', C) > h(k_g, C_g)$, or put differently, $C < w_g(k')$. Again by Lemma 1 such a (k', C) is not feasible. Therefore, the only state that can generate a state on the boundary $\{C_- = w_g(k), k > k_g\}$ is (k_g, C_g), and the associated unique control involves positive c. $\square$

[49] These inequalities hold for all $k \ge k_g$. The proofs are analogous to the proofs in footnotes 46 and 50.

[50] This holds because $C_- = w_g(k) = C_g(k/k_g)^{1/\sigma}$ and thus $C_-/((1-\beta)/\beta k) = (k/k_g)^{1/\sigma - 1} < 1$.

[51] This holds because by the IC constraint (1c), $\beta(k' + C)/C^\sigma = k_g/C_g^\sigma$ or equivalently $(k' + C)/C = 1/(1-\beta)\,(C/C_g)^\sigma$. Thus, letting $\kappa = k'/C$, $(\kappa + 1)\kappa^\sigma = (1-\beta)^{-1} \cdot (\beta/(1-\beta))^\sigma \cdot (k'/k_g)^\sigma$. Since the right hand side is increasing in κ, the fact that $k' > k_g$ tells us that $\kappa > \beta/(1-\beta)$, which is what we set out to show.

B.4.3 Proof of Lemma 3

Proof. Let $\mathcal{V}$ ($\tilde{\mathcal{V}}$) be the space of all continuous, weakly (strictly) increasing functions $v : [k_g, \bar{k}] \to \mathbb{R}_+$ with $\psi(k, v(k))$ weakly (strictly) decreasing in k, and $v(k_g) = C_g, v(k^g) = C^g$. For these functions, T is well-defined since for small values of C_-, $k'(F(k), k, C_-)$ tends to $F(k) \in (k_g \bar{k}]$. Moreover, the supremum in (18) is attained for all $k \in [k_g, \bar{k}]$ since the set of C_- in (18) is closed and bounded. We next show that (a) instead of considering all possible controls x, it is sufficient to consider $x = F(k)$; and (b) instead of looking for C_- that satisfy the inequality in (18), it suffices to look for solutions to the corresponding relation with equality. This will allow us to write

$$Tv(k) = \max\{C_- \,|\, v(k'(F(k), k, C_-)) = C(F(k), k, C_-)\}, \tag{26}$$

The formal arguments behind these two steps are:

(a) Fix $k \in [k_g, \bar{k}]$ and $v \in \mathcal{V}$. Suppose the supremum in (18) is attained by C_-, with control $x_0 < F(k)$. Define $\Phi_{v,k,C_-} : [0, F(k)] \to \mathbb{R}$ by

$$\Phi_{v,k,C_-}(x) = \underbrace{\psi\left(k'(x,k,C_-), C(x,k,C_-)\right)}_{\text{constant in } x} - \underbrace{\psi\left(k'(x,k,C_-), v(k'(x,k,C_-))\right)}_{\text{decreasing in } x} \tag{27}$$

and notice that $v(k'(x_0, k, C_-)) \geq C(x_0, k, C_-)$ is equivalent to $\Phi_{v,k,C_-}(x_0) \geq 0$. Since $\Phi_{v,k,C_-}(x)$ is weakly increasing in x due to $v \in \mathcal{V}$, $\Phi_{v,k,C_-}(F(k)) \geq \Phi_{v,k,C_-}(x_0)$ and so $v(k'(F(k), k, C_-)) \geq C(F(k), k, C_-)$. Therefore, focusing on controls $x = F(k)$ is without loss in (18).

(b) Now argue that equality (rather than inequality) is without loss in (18). Suppose the supremum were attained by C_- with control $x = F(k)$ and *strict* inequality, $v(k'(F(k), k, C_-)) > C(F(k), k, C_-)$. Since both sides of this inequality are continuous in C_-, it follows that slightly increasing C_- still satisfies the inequality and hence C_- could not have attained the supremum in the first place.
Notice also that the equation $v(k'(F(k), k, C_-)) = C(F(k), k, C_-)$ can never have more than one solution since raising C_- weakly decreases the left hand side and strictly increases the right hand side.

Now we argue that T maps $\mathcal{V}$ into $\tilde{\mathcal{V}}$. Take $v \in \mathcal{V}$. To show Tv is continuous and strictly increasing, define first the auxiliary function $\Psi_v : [k_g, \bar{k}] \times \mathbb{R}_{++} \to \mathbb{R}$ by

$$\Psi_v : (k, C_-) = \underbrace{\psi\left(k'(F(k), k, C_-), C(F(k), k, C_-)\right)}_{\nearrow \text{ in } k \text{ and } \searrow \text{ in } C_-} - \underbrace{\psi\left(k'(F(k), k, C_-), v(k'(F(k), k, C_-))\right)}_{\searrow \text{ in } k \text{ and } \nearrow \text{ in } C_-}.$$

The function Ψ_v is continuous and consists of two terms: The first term is equal to $\beta^{-1}k/C_-^\sigma$, using the definition of ψ, and hence strictly increasing in k and strictly decreasing in C_-.

For the second term, recall that

$$k'(F(k), k, C_-) = F(k)\left(1 - C_-\left(\frac{\beta}{kF(k)^{\sigma-1}}\right)^{1/\sigma}\right)$$

is strictly increasing in k and strictly decreasing in C_-, and v is such that $\psi(k, v(k))$ is weakly decreasing in k. Thus, the second term is weakly decreasing in k and weakly increasing in C_-. Putting both terms together gives us that $\Psi_v(k, C_-)$ is continuous, strictly increasing in k, and strictly decreasing in C_-. We can rewrite Tv as

$$Tv(k) = C_- \quad \text{where } C_- \text{ is the unique number with } \Psi_v(k, C_-) = 0.$$

Since Ψ_v is continuous, strictly increasing in k, strictly decreasing in C_- and admits a unique solution $C_- = Tv(k)$ to the equation $\Psi_v(k, C_-) = 0$, it follows that $Tv(k)$ is continuous and strictly increasing.[52]

To prove that $k \mapsto \psi(k, Tv(k))$ is strictly decreasing, pick $k_1 < k_2$ in $[k_g, \bar{k}]$. Suppose $\psi(k_1, Tv(k_1)) \le \psi(k_2, Tv(k_2))$. Since $Tv(k)$ is strictly increasing, it follows that

$$\frac{k_1}{Tv(k_1)^\sigma} - \frac{k_2}{Tv(k_2)^\sigma} < \underbrace{\frac{k_1}{Tv(k_1)^\sigma} + Tv(k_1)^{1-\sigma}}_{\psi(k_1, Tv(k_1))} \underbrace{- \frac{k_2}{Tv(k_2)^\sigma} - Tv(k_2)^{1-\sigma}}_{-\psi(k_2, Tv(k_2))} \le 0.$$

Defining $k_i' \equiv k'(F(k_i), k_i, Tv(k_i))$ and $C_i \equiv C(F(k_i), k_i, Tv(k_i))$, we find

$$\psi(k_1', C_1) = \beta^{-1}\frac{k_1}{Tv(k_1)^\sigma} < \beta^{-1}\frac{k_2}{Tv(k_2)^\sigma} = \psi(k_2', C_2). \tag{28}$$

This, however, implies that $Tv(k_2)$ cannot have been optimal: Pick an alternative consumption level $C_{2,-}$ as $C_{2,-} = Tv(k_1)(k_2/k_1)^{1/\sigma}$, which exceeds $Tv(k_2)$ by (28). Moreover, pick the policy $x_2 \equiv F(k_1)$, which is feasible, $x_2 \le F(k_2)$. Since $k_1/Tv(k_1)^\sigma = k_2/C_{2,-}^\sigma$ by construction, it follows that $(k'(x_2, k_2, C_{2,-}), C(x_2, k_2, C_{2,-})) = (k_1', C_1)$, which lies on the graph of v. Hence $Tv(k_2)$ cannot have been optimal and so $\psi(k, Tv(k))$ is decreasing in k.

Finally, we prove that $Tv(k_g) = C_g$. Note that $k'(F(k_g), k_g, C_g) = k_g$ and $C(F(k_g), k_g, C_g) = C_g$. Because $k'(F(k_g), k_g, C_-)$ is strictly decreasing in C_- and so $k'(F(k_g), k_g, C_-) < k_g$ for $C_- > C_g$ (for $k < k_g$, $v(k)$ is not even defined), this implies that $Tv(k_g) = C_g$, concluding the proof that $T(\mathcal{V}) \subset \tilde{\mathcal{V}}$. □

B.4.4 Proof of Lemma 4

Proof. Note that any state (k, C_-) reaches the space $\{C_- \le v(k)\}$ in one step if and only if $C_- \le Tv(k)$ (provided that v satisfies the regularity properties in Lemma 3). Thus, by iteration, $Z^{(i)} = \{w^g(k) \le C_- \le T^i w^g(k)\}$. Because $Z^{(i)} \supseteq Z^{(j)}$ for $i \ge j$, it holds that

[52] This is a fact that holds more generally: If $I_1, I_2 \subset \mathbb{R}$ are intervals and $f : I_1 \times I_2 \to \mathbb{R}$ is continuous, strictly increasing in x, and strictly decreasing in y with the property that for each x there exists a unique $y^*(x)$ s.t. $f(x, y^*(x)) = 0$, then $y^*(x)$ must be continuous and strictly increasing in x.

$T^i w^g(k) \geq T^j w^g(k)$.[53] □

B.4.5 Proof of Lemma 5

Proof. The existence of the limit $\lim_{i\to\infty} T^i w^g(k)$ is straightforward for every k (monotone sequence, bounded above because for large values of C_-, $k'(F(k), k, C_-) < k_g$ for any k). It can easily be verified that $w^g \in \mathcal{V}$. Thus, using Lemma 3, ϖ must be weakly increasing, $\varpi(k_g) = C_g$, $\varpi(k^g) = C^g$, and $\psi(k, \varpi(k))$ must be weakly decreasing. To show $\varpi \in \mathcal{V}$, suppose now that ϖ were not continuous. Then, there would have to be two arbitrarily close values of k, $k_1 < k_2$ with a significant gap between $T^N w^g(k_1)$ and $T^N w^g(k_2) > T^N w^g(k_1)$ for some large N. Since $k'(...)$ and $C(...)$ are both continuous, k_1 and k_2 can be chosen sufficiently close so that

$$k_1' \equiv k'(F(k_1), k_1, T^N w^g(k_1)) > k'(F(k_2), k_2, T^N w^g(k_2)) \equiv k_2',$$

yet the inequality is reversed for $C(...)$, $C_1 \equiv C(F(k_1), k_1, T^N w^g(k_1)) < C(F(k_2), k_2, T^N w^g(k_2)) \equiv C_2$. However, this contradicts the definition of $T^N w^g$ since both pairs (k_1', C_1) and (k_2', C_2) have to lie on the graph of the same increasing function $T^{N-1} w^g$ but the latter is to the top left of the former. Therefore, ϖ is continuous and $\varpi \in \mathcal{V}$.

Applying Dini's Theorem, the convergence of $T^n w^g$ to ϖ is also uniform, and by interchanging limits we find that

$$\begin{aligned}\varpi(k'(F(k), k, \varpi(k)) &= \lim_{n\to\infty} T^n w^g(k'(F(k), k, T^{n+1} w^g(k))) \\ &= \lim_{n\to\infty} C(F(k), k, T^{n+1} w^g(k)) = C(F(k), k, \varpi(k)), \quad (29)\end{aligned}$$

and thus, by the representation of T in (26), $\varpi = T\varpi$. This also means that $\varpi \in \tilde{\mathcal{V}}$, so ϖ is strictly increasing and $\psi(k, \varpi(k))$ strictly decreasing. Hence, for any given k, the only feasible policy at point $(k, \varpi(k))$ is $x = F(k)$ (or equivalently $c = 0$) since for any feasible policy x, $\Phi_{\varpi,k,\varpi(k)}(x)$ from (27) needs to be non-negative; but by $\varpi \in \tilde{\mathcal{V}}$ and (29), $\Phi_{\varpi,k,\varpi(k)}$ is strictly increasing with $\Phi_{\varpi,k,\varpi(k)}(F(k)) = 0$, so $x = F(k)$ is the only feasible policy. □

B.4.6 Proof of Lemma 6

Proof. Define h as before, $h(k', C) \equiv k'/C^\sigma$. Fix a state (k_1, C_0) with $C_0 > \varpi(k_1)$. First, consider the case $C_0 \geq (1-\beta)/\beta k_1$ and suppose it were generated by a feasible path $\{(k_{t+1}, C_t)\}$. As an intermediate result we now establish that $C_t > (1-\beta)/\beta k_{t+1}$ along such a path. We do this by distinguishing the following two cases:

(a) If $k_{t+1} \leq k^g$, this follows directly from $C_t > \varpi(k_{t+1}) \geq (1-\beta)/\beta k_{t+1}$. The former inequality holds by construction of ϖ,[54] the latter by Lemma 4.

[53] A subtlety here is that $Z^{(i)} \supseteq Z^{(j)}$ only holds because states in the set $\{C_- = w^g(k)\}$ is "self-generating", that is, if a path hits the set $\{C_- = w^g(k)\}$ after j steps, it can stay in that set forever. In particular, it can hit the set after $i \geq j$ steps as well. This explains why $Z^{(i)} \supseteq Z^{(j)}$.

[54] If it were violated, $C_0 \leq T^t \varpi(k_1) = \varpi(k_1)$ by construction of ϖ. This would contradict our assumption that $C_0 > \varpi(k_1)$.

(b) If instead $k_{t+1} > k^g$, it must be the case that $k_{s+1} > k^g$ for all $s < t$ as well.[55] But then, using that $x_t \leq F(k_t) < k_t/\beta$ for $k_t > k^g$,

$$\frac{k_{t+1}}{C_t} = \frac{x_t}{C_t} - 1 < \frac{k_t/\beta}{C_{t-1}} - 1 = \frac{\beta}{1-\beta}.$$

We use our intermediate result as follows (still for the case $C_0 \geq (1-\beta)/\beta k_1$). Consider

$$h(k_{t+1}, C_t) = \frac{k_{t+1}}{C_t^\sigma} = \frac{k_t}{C_{t-1}^\sigma} \underbrace{\frac{k_{t+1}}{\beta(k_{t+1} + C_t)}}_{<1} < h(k_t, C_{t-1}). \tag{30}$$

If $h(k_{t+1}, C_t)$ converges to zero, then either $k_{t+1} \to 0$ or $C_t \to \infty$ (in which case $k_{t+1} \to 0$ by the law of motion for capital and the fact that $k_t \leq \bar{k}$). Such a path is not feasible because then $F(k_{t+1})$ drops below zero in finite time (see the proof of Lemma 1 for a similar argument). Hence, suppose $h(k_{t+1}, C_t) \to \underline{h} > 0$. Then, $k_{t+1}/(\beta(k_{t+1} + C_t)) \to 1$, so the path must approximate the steady state line described by $\{(k, C_-) \,|\, C_- = (1-\beta)/\beta k\}$. Because $C_t > \bar{w}(k_{t+1})$ along the path, (k_{t+1}, C_t) must be converging to (k^g, C^g).

Next we show that this convergence is still true if we take c_t to be zero. Suppose there were times with $c_t > 0$. Then, define a new path $\{(\hat{k}_{t+1}, \hat{C}_t)\}$, starting at the same initial state (k_1, C_0) but with controls $c_t = 0$. Observe that

$$h(\hat{k}_{t+1}, \hat{C}_t) = \psi(\hat{k}_{t+1}, \hat{C}_t) - \hat{C}_t^{1-\sigma} = \beta^{-1} h(\hat{k}_t, \hat{C}_{t-1}) - h(\hat{k}_t, \hat{C}_{t-1})^{(\sigma-1)/\sigma} (\beta F(\hat{k}_t))^{-(\sigma-1)/\sigma}$$

$$\hat{k}_{t+1} = F(\hat{k}_t) - \left(\frac{\beta F(\hat{k}_t)}{h(\hat{k}_t, \hat{C}_{t-1})} \right)^{1/\sigma},$$

where the first equation is increasing in $h(\hat{k}_t, \hat{C}_{t-1})$ for the relevant parameters for which $h(\hat{k}_{t+1}, \hat{C}_t) \geq 0$, and similarly the second equation is increasing in $F(\hat{k}_t)$ if $\hat{k}_{t+1} \geq 0$. By induction over t, if $h(\hat{k}_t, \hat{C}_{t-1}) \geq h(k_t, C_{t-1})$ and $\hat{k}_t \geq k_t$ (induction hypothesis), then, because $F(\hat{k}_t) \geq x_t$,

$$h(\hat{k}_{t+1}, \hat{C}_t) \geq \beta^{-1} h(k_t, C_{t-1}) - h(k_t, C_{t-1})^{(\sigma-1)/\sigma} (\beta x_t)^{-(\sigma-1)/\sigma} = h(k_{t+1}, C_t)$$

$$\hat{k}_{t+1} \geq F(k_t) - \left(\frac{\beta F(k_t)}{h(k_t, C_{t-1})} \right)^{1/\sigma},$$

confirming that $\hat{k}_t \geq k_t$ and $h(\hat{k}_t, \hat{C}_{t-1}) \geq h(k_t, C_{t-1})$ for all t. Given that $h(k_{t+1}, C_t) \to$

[55] The reason for this is that for any state (k, C_-) with $k \leq k^g$ and $C_- > \bar{w}(k)$ we have that $k' \equiv k'(x, k, C_-) \leq k^g$ for any control $x \leq F(k)$. First, if $\psi(k', C) \geq \psi(k^g, C^g)$, then the curve $\{(k'(x, k, C_-), C(x, k, C_-)),\ x > 0\}$ and the graph of $\bar{w}$ necessarily intersect at a state $\tilde{k}$ with capital less than k^g. The intersection is unique since $\psi(k, \bar{w}(k))$ is strictly increasing. Since $C_- > \bar{w}(k)$ it cannot be that $\tilde{k} = k'(x, k, C_-)$ for a feasible $x \leq F(k)$ and therefore, any $k'(x, k, C_-)$ with a feasible $x \leq F(k)$ is necessarily less than $\tilde{k} \leq k^g$. Second, if $\psi(k', C) < \psi(k^g, C^g)$, that is, $k/C_-^\sigma < k^g/(C^g)^\sigma$, then $k' \leq F(k) - C_- \left(\frac{\beta F(k)}{k}\right)^{1/\sigma} < F(k^g) - C^g \left(\frac{\beta F(k^g)}{k^g}\right)^{1/\sigma} = k^g$.

$\underline{h} > 0$, either $\left(\hat{k}_{t+1}, \hat{C}_t\right) \to (k^g, C^g)$ as well, or $\{(\hat{k}_{t+1}, \hat{C}_t)\}$ converges to some steady state between k_g and k^g. The latter cannot be because of $\hat{C}_t > \varpi(\hat{k}_{t+1})$ along the path. But the former is precluded by Lemma 7 below. This provides a contradiction, proving that a state (k_1, C_0) with $C_0 > \varpi(k_1)$ and $C_0 > (1-\beta)/\beta k_1$ cannot be feasible.

Now, consider the case $C_0 < (1-\beta)/\beta k_1$. Due to $C_0 > \varpi(k_1)$, this can only be the case if $k_1 > k^g$. Again, suppose (k_1, C_0) were generated by a feasible path $\{(k_{t+1}, C_t)\}$. Given the first half of this proof, if at any point (k_{t+1}, C_t) lies above the steady state line, we have the desired contradiction. Therefore, suppose $C_t < (1-\beta)/\beta k_{t+1}$ for all t. In that case,

$$h(k_{t+1}, C_t) = \frac{k_{t+1}}{C_t^\sigma} = \frac{k_t}{C_{t-1}^\sigma} \underbrace{\frac{k_{t+1}}{\beta(k_{t+1} + C_t)}}_{>1} > h(k_t, C_{t-1}).$$

Note that $h(k_{t+1}, C_t)$ is bounded from above, for example by $h(k_g, C_g)$ (because all states below the steady state line with h equal to $h(k_g, C_g)$ are below the graph of w^g and thus below ϖ as well). So, $h(k_{t+1}, C_t)$ converges and $k_{t+1}/(\beta(k_{t+1} + C_t)) \to 1$. The state approximates the steady state line. Because the only feasible steady state with below the steady state line but above the graph of ϖ is (k^g, C^g) it follows that $(k_{t+1}, C_t) \to (k^g, C^g)$. Following the same steps as before, it can be shown that without loss of generality, controls c_t can be taken to be zero along the path. By Lemma 7 below this is a contradiction, concluding our proof that no state (k_1, C_0) with $C_0 > \varpi(k_1)$ is feasible. □

B.4.7 Proof of Lemma 7

Proof. We prove each of the results in turn.

(a) Notice that $c = 0$ takes any state on the graph of ϖ to another state on the graph of ϖ (because $T\varpi = \varpi$). Suppose $k_1 < k^g$ (the case $k_1 > k^g$ is analogous). Then, no future capital stock k_{t+1} can exceed k^g. Because if it did, there would have to be a capital stock $k \in (k_g, k^g)$ with $k'(F(k), k, \varpi(k)) = k^g$, by continuity of $k \mapsto k'(F(k), k, \varpi(k))$. But this is impossible by definition of k^g.[56] Thus, along the path, $C_t > (1-\beta)/\beta k_{t+1}$ and so $h(k_{t+1}, C_t)$ is decreasing. As $h(k_g, C_g) > h(k, \varpi(k))$ for all $k > k_g$,[57] this means $(k_{t+1}, C_t) \to (k^g, C^g)$.

(b) For simplicity, focus on the case $k_0 < k^g$. Again, the case $k_0 > k^g$ is completely analogous. Suppose (k_{t+1}, C_t) were converging to (k^g, C^g). Note that at k^g, $F(k)/k$ is decreasing.[58] Thus, there exists a time $T > 0$ for which the capital stock k_T is sufficiently close to k^g that $F(k)/k$ is decreasing for all k in a neighborhood of k^g which includes $\{k_t\}_{t \geq T}$. Let $\{\hat{k}_{t+1}, \hat{C}_t\}$ denote the path with $c_t = 0$, starting from $(k_T, \varpi(k_T))$. We already know that $\{\hat{k}_{t+1}, \hat{C}_t\}$ does indeed converge to (k^g, C^g) from the first part of this proof. Also, observe that both (k_{t+1}, C_t) and $(\hat{k}_{t+1}, \hat{C}_t)$ have

[56] By definition of k^g, $F(k^g) = k^g + C^g$, and so, $F(k) < k^g + C^g$ for $k < k^g$.
[57] Note that $\varpi(k) > w_g(k)$ and $h(k, w_g(k)) = \text{const}$, see Lemmas 1 and 2 above.
[58] This holds because $F'(k^g) < 1/\beta$ and $F(k^g) = 1/\beta k^g$, and so $\frac{d}{dk}F(k)/k < 0$.

controls $c_t = 0$ here, unlike in the proof of Lemma 6.
In the remainder of this proof, we denote the "zero control $c = 0$" laws of motion for capital and capitalists' consumption by $L_k(k, C_-) \equiv k'(F(k), k, C_-)$ and $L_C(k, C_-) \equiv C(F(k), k, C_-)$ (only for this proof). Since $F(k)/k$ is locally decreasing, it follows that $dL_k/dk > 0$, $dL_k/dC_- < 0$ and $dL_C/dk < 0$, $dL_C/dC_- > 0$. This implies that because $C_{T-1} > \bar{w}(k_T)$ (which must hold or else $C_0 \leq \bar{w}(k_1)$ by construction of $\bar{w}$), $C_t > \hat{C}_t$ and $k_{t+1} > \hat{k}_{t+1}$ for all $t \geq T$. Moreover, borrowing from equation (22), we know that

$$h(k_{t+1}, C_t) = h(k_t, C_{t-1}) \left(\frac{1}{\beta} - \left(\frac{1}{h(k_t, C_{t-1})} \right)^{1/\sigma} \frac{1}{(\beta F(k_t))^{1-1/\sigma}} \right),$$

which implies that by induction $h(k_{t+1}, C_t) \leq h(\hat{k}_{t+1}, \hat{C}_t)$, that is,

$$\begin{aligned}
&\log h(k_{t+T}, C_{t+T-1}) \\
&\quad = \log h(k_T, C_{T-1}) + \sum_{s=0}^{t-1} \log \left(\frac{1}{\beta} - \left(\frac{1}{h(k_{T+s}, C_{T+s-1})} \right)^{1/\sigma} \frac{1}{(\beta F(k_{T+s}))^{1-1/\sigma}} \right) \\
&\quad \leq \log h(k_T, C_{T-1}) + \sum_{s=0}^{t-1} \log \left(\frac{1}{\beta} - \left(\frac{1}{h(\hat{k}_{T+s}, \hat{C}_{T+s-1})} \right)^{1/\sigma} \frac{1}{(\beta F(\hat{k}_{T+s}))^{1-1/\sigma}} \right) \\
&\qquad\qquad = \log h(\hat{k}_{t+T}, \hat{C}_{t+T-1}) + \log h(k_T, C_{T-1}) - \log h(\hat{k}_T, \hat{C}_{T-1}).
\end{aligned}$$

As $t \to \infty$, this equation yields

$$\log h(k^g, C^g) \leq \log h(k^g, C^g) + \underbrace{\log h(k_T, C_{T-1}) - \log h(\hat{k}_T, \hat{C}_{T-1})}_{= -k_T(\hat{C}_{T-1}^{-\sigma} - C_{T-1}^{-\sigma}) < 0},$$

which is a contradiction. Therefore, $(k_{t+1}, C_t) \not\to (k^g, C^g)$. □

C Numerical Method

To solve the Bellman equation (15) we must first compute the feasible set Z^*. We restrict the range of capital to a closed interval $[\underline{k}, \bar{k}]$ with $\underline{k} \geq k_g$. This leads us to seek a subset $Z^{*\underline{k}} \subset Z^*$ of the feasible set Z^*. We compute this set numerically as follows.

Start with the set $Z^*_{(0)}$ defined by $C_- = \frac{1-\beta}{\beta} k$ and $k \in [\underline{k}, \bar{k}]$. This set is self generating and thus $Z^*_{(0)} \subset Z^{*\underline{k}}$. We define an operator that finds all points (k, C_-) for which one can find c, K', C satisfying the constraints of the Bellman equation and $(k', C) \in Z^*_{(0)}$. This gives a set $Z^*_{(1)}$ with $Z^*_{(0)} \subset Z^*_{(1)}$. Iterating on this procedure we obtain $Z^*_{(0)}, Z^*_{(1)}, Z^*_{(2)}, \ldots$ and we stop when the sets do not grow much. We then solve the Bellman equation by value function iteration. We start with a guess for V_0 that uses a feasible policy to evaluate utility. This ensures that our guess is below the true value function. Iterating on the Bellman equation then leads to a monotone sequence $V_0, V_1, \ldots$ and we stop when iteration n

yields a V_n that is sufficiently close to V_{n-1}. Our procedure uses a grid that is defined on a transformation of (k, C_-) that maps Z^* into a rectangle. We linearly interpolate between grid points.

The code was programmed in Matlab and executed with parallel 'parfor' commands, to improve speed and allow denser grids, on a cluster of 64-128 workers. Grid density was adjusted until no noticeable difference in the optimal paths were observed.

D Proof of Proposition 4

As in Appendix B we use the notation that $F(k) = f(k) + (1-\delta)k$. The derivatives of S evaluated at some time τ are denoted by $S_{I,\tau} \equiv \frac{\partial S_\tau}{\partial I_\tau}$ and $S_{\tau,R_t} \equiv \frac{\partial S_\tau}{\partial R_t}$, for $t > \tau$.

Define the following object,

$$\omega_\tau = \frac{dW_\tau}{dk_{\tau+1}} = \sum_{\tau' \geq \tau+1} \beta^{\tau'-\tau} u'(c_{\tau'})(F'(k_{\tau'}) - R_{\tau'}) \left(\prod_{s=\tau+1}^{\tau'-1} S_{I,s} R_s \right), \tag{31}$$

which corresponds to the response in welfare W_τ, measured in units of period τ utility, of a change in savings by an infinitesimal unit between periods τ and $\tau+1$. Now consider the effect of a one-time change in the capital tax, effectively changing R_t to $R_t + dR$ in period t. This has three types of effects on total welfare: It changes savings behavior in all periods $\tau < t$ through the effect of R_t on S_τ. It changes capitalists' income in period t through the effect of R_t on $R_t k_t$. And finally it changes workers' income in period t directly through the effect of R_t on $F(k_t) - R_t k_t$. Summing up these three effects, one obtains a total effect of

$$\begin{aligned} dW = \sum_{\tau=0}^{t-1} \beta^{\tau-t} \omega_\tau \quad & \underbrace{S_{\tau,R_t} dR}_{\text{change in savings in period } \tau<t} \\ & + \omega_t \quad \underbrace{S_{I,t} k_t dR}_{\text{change in savings in period } t} \quad -u'(c_t) \quad \underbrace{k_t dR}_{\text{change in workers' income in period } t} \quad . \end{aligned}$$

The total effect needs to net out to zero along the optimal path, that is,

$$\omega_t S_{I,t} - u'(c_t) = -\frac{1}{k_t} \sum_{\tau=0}^{t-1} \beta^{\tau-t} \omega_\tau S_{\tau,R_t}. \tag{32}$$

By optimization over the initial interest rate R_0, we find the condition

$$\omega_0 S_{I,0} k_0 - u'(c_0) k_0 = 0. \tag{33}$$

This shows that $S_{I,0} > 0$ and so $\omega_0 \in (0, \infty)$. By their definition (31), the ω_τ satisfy the recursion

$$\omega_\tau = u'(c_{\tau+1})(F'(k_{\tau+1}) - R_{\tau+1}) + \beta S_{I,\tau+1} R_{\tau+1} \omega_{\tau+1}.$$

Since it is easy to see that $R_{\tau+1} > 0$ for all τ,[59] it follows that ω_τ is finite for all τ. Then, due to the recursive nature of (32), if $\omega_\tau > 0$ for $\tau < t$,

$$\omega_t S_{I,t} - u'(c_t) = -\frac{1}{k_t}\sum_{\tau=0}^{t-1}\beta^{\tau-t}\underbrace{\omega_\tau}_{>0}\underbrace{S_{\tau,R_t}}_{\leq 0} \geq 0.$$

In particular, using the initial condition (33), this proves by induction that

$$\omega_t S_{I,t} - u'(c_t) \geq 0 \quad \text{for all } t > 0. \tag{34}$$

Now suppose the economy were converging to an interior steady state with non-positive limit tax (either zero or negative), that is, $\Delta_t \equiv F'(k_t) - R_t$ converges to a non-positive number, $c_t \to c > 0$ and $S_{I,t}R_t \to S_I R > 0$. It is immediate by (31) that if Δ_t converges to a negative number, then ω_t must eventually become negative—contradicting (34). Hence suppose $\Delta_t \to 0$. Distinguish two cases.

Case I: Suppose first that $\beta S_I R > 1$. Thus, $\prod_{s=1}^{\tau}(\beta S_{I,s}R_s)$ is unbounded and diverges to ∞. Then, because ω_0 is finite, we have that the partial sums in the expression for ω_0 coming from (31) have to converge to zero,

$$\bar{\omega}_\tau \equiv \sum_{\tau' \geq \tau+1} \beta u'(c_{\tau'})(F'(k_{\tau'}) - R_{\tau'}) \prod_{s=1}^{\tau'-1}(\beta S_{I,s}R_s) \to 0, \quad \text{as } \tau \to \infty.$$

Hence,

$$\omega_\tau = \left(\prod_{s=1}^{\tau}(\beta S_{I,s}R_s)\right)^{-1} \bar{\omega}_\tau \to 0,$$

contradicting the fact that ω_t is bounded away from zero by $u'(c)/S_I$. Therefore, $\beta S_I R > 1$ is not compatible with any interior steady state. (This argument does not use the fact that we focus on $\Delta_t \to 0$.)

Case II: Suppose $\beta S_I R < 1$. In this case, we show convergence of ω_τ to zero directly. Fix $\epsilon > 0$. Let τ be large enough such that $\beta S_{I,s}R_s < b$ for some $b < 1$ and that $|u'(c_{\tau'})\Delta_{\tau'}| < \epsilon(1-b)$. Then,

$$|\omega_\tau| \leq \sum_{\tau'=\tau+1} \epsilon(1-b)b^{\tau'-1-\tau} = \epsilon.$$

Again, this contradicts the fact that ω_t is bounded away from zero by $u'(c)/S_I$.

This concludes our proof, establishing that the capital tax $\mathcal{T}_t = \Delta_t/F'(k_t)$ must converge to a positive number at the interior steady state.

[59] Otherwise capital would be zero forever after due to $S(0,\dots) = 0$, a contradiction to the allocation converging to an interior steady state.

E Derivation of the Inverse Elasticity Rule (4) and Proof of the Corollary

Derivation of the Inverse Elasticity Rule. In this section, we continue using the notation and results of Section D. Consider equation (32). Because $\beta S_I R < 1$, ω_τ converges to

$$\omega = \frac{\beta}{1 - \beta S_I R}(F'(k) - R)u'(c).$$

We make the additional convergence assumption

$$\sum_{\tau=1}^{t} \beta^{-\tau} \frac{\omega_{t-\tau} k_{t-\tau}}{\omega_t k_t} \epsilon_{S_{t-\tau},R_t} \to \sum_{\tau=1}^{\infty} \beta^{-\tau} \epsilon_{S,\tau} \in [-\infty, \infty], \quad \text{as } t \to \infty, \tag{35}$$

which amounts to first taking the limit of the summands as $t \to \infty$, and then taking the limit of the series, instead of considering both limits simultaneously. Under this order of limits assumption, we can characterize the limit of equation (32) as $t \to \infty$,

$$\underbrace{S_{I,t} - \frac{u'(c_t)}{\omega_t}}_{\to S_I - \frac{u'(c)}{\omega}} = -\underbrace{\sum_{\tau=1}^{t} \beta^{-\tau} \frac{\omega_{t-\tau} k_{t-\tau}}{\omega_t k_t} \epsilon_{S_{t-\tau},R_t}}_{\to \sum_{\tau=1}^{\infty} \beta^{-\tau} \epsilon_{S,\tau}}. \tag{36}$$

Distinguish two cases according to whether $\omega = 0$ or $\omega \neq 0$. First, if $\omega = 0$, or equivalently the limit tax $\mathcal{T}$ is zero, then (36) reveals that $\sum_{\tau=1}^{\infty} \beta^{-\tau} \epsilon_{S,\tau}$ is either plus or minus infinity. Therefore, the inverse elasticity formula holds in this case as both sides of (4) converge to zero.

Second, if $\omega \neq 0$, then by taking the limit of (32) as $t \to \infty$ and using the condition (35), we find

$$S_I - \frac{u'(c)}{\omega} = -\sum_{\tau=1}^{\infty} \beta^{-\tau} \epsilon_{S,\tau},$$

which can be rewritten as

$$\frac{\beta S_I R}{1 - \beta S_I R}(F'(k) - R) - R = -\frac{1}{1 - \beta S_I R}(F'(k) - R) \sum_{\tau=1}^{\infty} \beta^{-\tau+1} \epsilon_{S,\tau}.$$

Note that $F'(k) - R = \frac{\mathcal{T}}{1-\mathcal{T}} R$. Therefore, we can rearrange the condition to

$$\begin{aligned}
\frac{\beta S_I R}{1 - \beta S_I R} - \frac{1 - \mathcal{T}}{\mathcal{T}} &= -\frac{1}{1 - \beta S_I R} \sum_{\tau=1}^{\infty} \beta^{-\tau+1} \epsilon_{S,\tau} \\
\Rightarrow \quad \mathcal{T} &= \frac{1 - \beta R S_I}{1 + \sum_{t=1}^{\infty} \beta^{-t+1} \epsilon_{S,t}}.
\end{aligned}$$

This is precisely the inverse elasticity formula (4).

Proof of the Corollary. Notice that by Proposition 4 the limit tax rate is positive, $\mathcal{T} > 0$, conditional on convergence to an interior steady state. If now the inverse elasticity formula implies a negative tax rate, then either the regularity condition for the inverse elasticity rule is not satisfied or the allocation does not converge to an interior steady state.

F Infinite Sum of Elasticities with Recursive Utility

In this section, we prove the result that the infinite sum $\sum_{\tau=1}^{\infty} \beta^{-\tau} \epsilon_{S,\tau}$ does not converge for any recursive utility function that is locally non-additive.

More specifically, we consider the capitalist's optimization problem as in Section 2.3, just with recursive preferences as in Section 3.1, with $U = c$. In particular, the capitalist's utility is characterized by the recursion $V_t = W(C_t, V_{t+1})$, assuming W is twice continuously differentiable and strictly increasing in both arguments. Analogous to our analysis in Section 3.1, we define $\overline{\beta}(c) \equiv W_V(c, V(c))$ as the steady state discount factor along a constant consumption stream yielding steady state utility $V(c) = W(c, V(c))$.

Any such recursive utility function naturally yields an optimal savings function $a_{t+1} = S(R_t a_t, R_{t+1}, \ldots)$. Fix now constant interest rates R and a steady state of the capitalist's optimization problem (a, c, V). Let $\beta = W_V(c, V(c)) = \overline{\beta}(c)$ the discount factor in that specific steady state. Define $\epsilon_{S,\tau} = \frac{1}{a} R \frac{\partial \log S}{\partial \log R_\tau}$. The following proposition characterizes the behavior of the infinite sum $\sum_{\tau=1}^{\infty} \beta^{-\tau} \epsilon_{S,\tau}$.

Proposition 12. *Suppose capitalists have recursive preferences represented by (5a) (see Section 3.1, with $U = c$). Then, if the discount factor is locally non-constant, $\overline{\beta}'(c) \neq 0$, the series $\sum_{\tau=1}^{T} \beta^{-\tau} \epsilon_{S,\tau}$ does not have a finite limit as $T \to \infty$.*

Proof. We first compute the elasticities $\epsilon_{S,\tau}$ and then prove that the infinite sum does not have a finite limit. To compute $\epsilon_{S,\tau}$, we consider an agent with the recursive preferences introduced above, who is at a steady state (a, c, V) given a constant interest rates R. Note that because utility is strictly increasing in a permanent increase in consumption at the steady state, we have $\beta = W_V \in (0, 1)$.

The conditions for optimality are then,

$$V_t = W\left(R_t a_t - a_{t+1}, V_{t+1}\right)$$
$$W_C\left(R_t a_t - a_{t+1}, V_{t+1}\right) = R_{t+1} W_V\left(R_t a_t - a_{t+1}, V_{t+1}\right) W_C\left(R_{t+1} a_{t+1} - a_{t+2}, V_{t+2}\right).$$

The first equation is the recursion for utility V_t and the second equation is the Euler equation. In particular, note that the latter implies that $\beta R = 1$ at the steady state. Linearizing these equations around the steady state (denoted without time subscripts) yields,

$$W_V \, dV_{t+1} = -W_C R \, da_t + W_C \, da_{t+1} + dV_t - W_C a \, dR_t \tag{37}$$

and

$$
\begin{aligned}
(RW_C W_{VC} - RW_{CC} - W_{CC})\, da_{t+1} + W_{CC}\, da_{t+2} - (W_V W_C + W_{CC} a)\, dR_{t+1} \\
+ (W_{CV} - RW_C W_{VV})\, dV_{t+1} - W_{CV}\, dV_{t+2} \\
= (R^2 W_C W_{VC} - W_{CC} R)\, da_t + (RW_C W_{VC} a - W_{CC} a)\, dR_t,
\end{aligned} \tag{38}
$$

where all derivatives are evaluated at the steady state $((R-1)\,a, V)$. To save on notation, we define $\omega \equiv W_{VC} - \beta W_{CC}/W_C \in \mathbb{R}$, which is a term that will appear multiple times below. We solve (37) and (38) by the method of undetermined coefficients, guessing

$$
da_{t+1} = \omega\lambda\, da_t + a \sum_{\tau=0}^{\infty} \beta^{\tau} \theta_{\tau}\, dR_{t+\tau} \tag{39a}
$$

$$
dV_t = W_C R\, da_t + (W_C a) \sum_{\tau=0}^{\infty} \beta^{\tau}\, dR_{t+\tau}. \tag{39b}
$$

The form of equation (39b) is what is required by the Envelope condition. We are left to find λ and the sequence $\{\theta_\tau\}$, where $\theta_\tau = \beta^{-\tau}\epsilon_{S,\tau}$, for $\tau \geq 1$, is exactly the sequence of elasticities were are looking for. Substituting the guesses (39a) and (39b) into (38), we obtain an expression featuring da_t, da_{t+1}, da_{t+2} and $dR_{t+\tau}$ for $\tau = 0, 1, \ldots$. Setting the coefficient on da_t to zero gives a quadratic for λ,

$$
\omega^2\lambda^2 + \left(-(1+R)\omega + (R-1)\overline{\beta}'(c)\right)\lambda + R = 0. \tag{40}
$$

Note that the solution of this equation can never be zero, i.e. $\lambda \neq 0$. Also, if $\overline{\beta}'(c) = 0$, the term in parentheses simplifies to $-(1+R)\omega$ and the solutions are just $\lambda = \omega^{-1}$ and $\lambda = \omega^{-1}R$.

Setting the coefficient on dR_t to zero gives

$$
\theta_0 = \beta\omega\lambda.
$$

Similarly for dR_{t+1} we find after various simplifications,

$$
\theta_1 = \omega\lambda\,(\theta_0 - 1) + \lambda\left(\beta^2 a^{-1} + (1-\beta)\overline{\beta}'(c)\right)
$$

and for $dR_{t+\tau}$ after some more simplifications

$$
\theta_\tau = \omega\lambda\theta_{\tau-1} + \lambda\,(1-\beta)\,\overline{\beta}'(c), \tag{41}
$$

for $\tau = 2, 3, \ldots$. The result then follows from this expression: If $\overline{\beta}'(c) \neq 0$, the sum $\sum_{\tau=1}^{T} \beta^{-\tau}\epsilon_{S,\tau} = \sum_{\tau=1}^{T} \theta_\tau$ cannot converge. To see this, consider

$$
\sum_{\tau=1}^{T} \theta_\tau = \theta_1 + \sum_{\tau=2}^{T} \theta_\tau = \theta_1 + \sum_{\tau=1}^{T-1} \omega\lambda\theta_\tau + \sum_{\tau=2}^{T} \lambda(1-\beta)\overline{\beta}'(c).
$$

If the left hand side of this equation converged to some limit $\Theta \in \mathbb{R}$, the right hand side of this equation would diverge since the last sum diverges (while all other terms would remain finite). Therefore, $\sum_{\tau=1}^{T} \beta^{-\tau} \epsilon_{S,\tau}$ cannot converge to a finite limit. $\square$

G Linearized Dynamics and Proof of Proposition 5

A natural way to prove Proposition 5 would be to linearize our first order conditions in (2), and to solve forward for the multipliers μ_t and λ_t using transversality conditions, arriving at an approximate law of motion of the form

$$\begin{pmatrix} k_{t+1} \\ C_t \end{pmatrix} - \begin{pmatrix} k_t \\ C_{t-1} \end{pmatrix} = \hat{J} \begin{pmatrix} k_t - k^* \\ C_{t-1} - C^* \end{pmatrix}.$$

To maximize similarity with Kemp et al. (1993), however, we do not take that route; rather we start with the continuous time problem, derive its first order conditions and linearize them around the zero tax steady state. The problem in continuous time is

$$\begin{aligned} \max \int_0^\infty e^{-\rho t} \left(u(c_t) + \gamma U(C_t) \right) dt \\ \text{s.t. } c_t + C_t + g + \dot{k}_t = f(k_t) - \delta k_t \\ \dot{C}_t = \frac{C_t}{\sigma} \left(\frac{f(k_t)}{k_t} - \delta - \frac{c_t}{k_t} - \rho \right). \end{aligned}$$

Let p_t and q_t denote the costates corresponding respectively to the states k_t and C_t. The FOCs are,

$$\begin{aligned} u_t'(c_t) &= p_t c_t + q_t \frac{1}{\sigma} \frac{C_t}{k_t} \\ \dot{p}_t &= \rho p_t - p_t (f'(k_t) - \delta) + q_t \frac{\dot{C}_t}{k_t} - q_t \frac{C_t}{k_t} (f'(k_t) - \delta) \\ \dot{q}_t &= \rho q_t - \gamma U'(C_t) - q_t \frac{1}{\sigma} \left(\frac{f(k_t)}{k_t} - \delta - \frac{c_t}{k_t} - \rho \right). \end{aligned}$$

Just like Kemp et al. (1993), we require the two transversality conditions to hold,

$$\lim_{t \to \infty} e^{-\rho t} q_t C_t = 0 \quad \text{and} \quad \lim_{t \to \infty} e^{-\rho t} p_t k_t = 0. \tag{42}$$

Denote the 4-dimensional state of this dynamic system by $x_t = (k_t, C_t, p_t, q_t)$ and its unique positive steady state (the zero-tax steady state) by $x^* = (k^*, C^*, p^*, q^*)$. The linearized system is,

$$\dot{x}_t = J(x_t - x^*), \tag{43}$$

where J is a 4×4 matrix with determinant

$$\det J = (1-\sigma)\underbrace{\frac{f''(k^*)u'(c^*)}{u''(c^*)}}_{>0}\frac{\rho^2}{\sigma^2}. \tag{44}$$

Its eigenvalues can be written as,

$$\lambda_{1-4} = \frac{\rho}{2} \pm \left[\left(\frac{\rho}{2}\right)^2 - \frac{\chi}{2} \pm \frac{1}{2}\left(\chi^2 - 4\det J\right)^{1/2}\right]^{1/2}, \tag{45}$$

with

$$\chi = \frac{\rho}{\sigma}\frac{u'(c^*) - \gamma U'(C^*)}{u''(c^*)k^*} - \frac{f''(k^*)u'(c^*)}{u''(c^*)}. \tag{46}$$

There are two "$\pm$" signs in (45). In the remainder, we number eigenvalues according to those two signs in (45): λ_1 has $++$, λ_2 has $+-$, λ_3 has $-+$, and λ_4 has $--$. For convenience, define γ^* by $\gamma^* = u'(c^*)/U'(C^*)$.

In general, a solution x_t to the linearized FOCs (43) can load on all four eigenvalues. However, taking the two transversality conditions into account restricts the system to only load on eigenvalues with $\mathrm{Re}(\lambda_i) \leq \rho/2$. In Lemma 13 below, we show that this means the solution loads on eigenvalues λ_3 and λ_4.

Lemma 13. *The eigenvalues in* (45) *can be shown to satisfy the following properties.*

(a) It is always the case that

$$Re\lambda_1 \geq Re\lambda_2 \geq \rho/2 \geq Re\lambda_4 \geq Re\lambda_3.$$

(b) If $\sigma > 1$, then $\det J < 0$, *implying that*

$$Re\lambda_1 = \lambda_1 > \rho > Re\lambda_2 \geq \rho/2 \geq Re\lambda_4 > 0 > \lambda_3 = Re\lambda_3. \tag{47}$$

In particular, there is a exactly one negative eigenvalue. The system is saddle-path stable.

(c) If $\sigma < 1$ and $\gamma \leq \gamma^*$, then $\det J > 0$ *and $\delta < 0$, implying that*

$$Re\lambda_1, Re\lambda_2 > \rho > 0 > Re\lambda_4, Re\lambda_3. \tag{48}$$

In particular, there exist exactly two eigenvalues with negative real part. The system is locally stable.

(d) If $\sigma < 1$ and $\gamma > \gamma^*$, the system may either be locally stable, or locally unstable (all eigenvalues having positive real parts).

Proof. We follow the convention that the square root of a complex number a is defined as the *unique* number b that satisfies $b^2 = a$ *and has nonnegative real part* (if $\mathrm{Re}(b) = 0$ we also require $Im(b) \geq 0$). Hence, the set of all square roots of a is given by $\{\pm b\}$. We prove the results in turn.

(a) First, observe the following fact: Given a real number x and a complex number b with nonnegative real part, it holds that $\mathrm{Re}\left(\sqrt{x+b}\right) \geq \mathrm{Re}\left(\sqrt{x-b}\right)$.[60] From there, it is straightforward to see that $\mathrm{Re}\lambda_1 \geq \mathrm{Re}\lambda_2$ and $\mathrm{Re}\lambda_4 \geq \mathrm{Re}\lambda_3$. Finally $\mathrm{Re}\lambda_2 \geq \rho/2 \geq \mathrm{Re}\lambda_4$ holds according to our convention of square roots having nonnegative real parts.

(b) The negativity of $\det J$ follows immediately from (44). This implies

$$-\frac{\delta}{2}+\frac{1}{2}\left(\delta^2-4\det J\right)^{1/2}>0>-\frac{\delta}{2}-\frac{1}{2}\left(\delta^2-4\det J\right)^{1/2},$$

and so (47) holds, using monotonicity of $\mathrm{Re}\sqrt{x}$ for real numbers x.

(c) The signs of $\det J$ and δ follow immediately from (44) and (46). In this case, $-\delta/2 \pm 1/2\mathrm{Re}\left(\delta^2-4\det J\right)^{1/2}>0$ proving (48).

(d) This is a simple consequence of the fact that if $\det J>0$, then either $-\delta/2 \pm 1/2\mathrm{Re}\left(\delta^2-4\det J\right)^{1/2}>0$, or $-\delta/2 \pm 1/2\mathrm{Re}\left(\delta^2-4\det J\right)^{1/2}<0$, where under the latter condition the system is locally unstable. $\square$

H Proof of Proposition 6

In this proof, we first exploit the recursiveness of the utility $\mathcal{V}$ to recast the IC constraint (7) entirely in terms of V_t and $W(U,V')$. Then, using the first order conditions, we are able to characterize the long-run steady state. Throughout this section, we denote by X_{zt} the derivative of quantity X with respect to z, evaluated at time t. To save on notation, we define $f(k,n)\equiv F(k,n)+(1-\delta)k$.

Let $\beta_t \equiv \prod_{s=0}^{t-1} W_{Vs}$. Using the definition of the aggregator in (3) this implies that $\mathcal{V}_{ct}=\beta_t W_{Ut}U_{ct}$ and $\mathcal{V}_{nt}=\beta_t W_{Ut}U_{nt}$. Thus the IC constraint (7) can be rewritten as

$$\sum_{t=0}^{\infty}\beta_t W_{Ut}(U_{ct}c_t+U_{nt}n_t)=W_{U0}U_{c0}\left(R_0k_0+R_0^b b_0\right), \tag{49}$$

and the planning problem becomes

$$\begin{aligned} &\max_{\{V_t,c_t,n_t,R_0\}} V_0 \\ &\quad \text{s.t. } V_t = W(U(c_t,n_t),V_{t+1}) \\ &\qquad\quad \text{RC (6), IC (49)}, R_t \geq 1. \end{aligned} \tag{50}$$

[60] To prove this, let $\bar{b}$ denote the complex conjugate of b and note that $\mathrm{Re}\left(\sqrt{x+b}\right)$ is monotonic in the real number x. Then, $\mathrm{Re}\left(\sqrt{x+b}\right)=\mathrm{Re}\left(\sqrt{x+\bar{b}}\right)=\mathrm{Re}\left(\sqrt{x-b+(\bar{b}+b)}\right)\geq \mathrm{Re}\left(\sqrt{x-b}\right)$ where $\bar{b}+b=2\mathrm{Re}(b)\geq 0$ and monotonicity are used.

To state the first order conditions, define for each $t \geq 0$, $A_{t+1} \equiv \frac{1}{\beta_{t+1}} \frac{\partial}{\partial V_{t+1}} \sum_{s=0}^{\infty} \beta_s W_{Us}(U_{cs} c_s + U_{ns} n_s)$ and $B_t \equiv \frac{1}{\beta_t} \sum_{s=0}^{\infty} \frac{\partial(\beta_s W_{Us})}{\partial U_t}(U_{cs} c_s + U_{ns} n_s)$. Let $\beta_t \nu_t$ be the present value multiplier on the Koopmans constraint (50), $\beta_t \lambda_t$ the present value multiplier on the resource constraint (6), and μ the multiplier on the IC constraint (49). As stated in the proposition, we assume that the capital tax bound $R_t \geq 1$ is not binding eventually, say from period T onwards. The first order conditions for V_{t+1}, c_t, n_t, and k_{t+1} (in that order) are for each $t \geq T$ given by

$$\begin{aligned}
-\nu_t + \nu_{t+1} + \mu A_{t+1} &= 0 \\
-\nu_t W_{Ut} U_{ct} + \mu W_{Ut} \left(U_{ct} + U_{cc,t} c_t + U_{nc,t} n_t\right) + \mu B_t U_{ct} &= \lambda_t \\
\nu_t W_{Ut} U_{nt} - \mu W_{Ut} \left(U_{nt} + U_{cn,t} c_t + U_{nn,t} n_t\right) - \mu B_t U_{nt} &= \lambda_t f_{nt} \\
-\lambda_t + \lambda_{t+1} W_{Vt} f_{kt+1} &= 0.
\end{aligned}$$

Suppose the allocation converges to an interior steady state in c, k, and n. Then U_t and V_t converge, as well as their first and second derivatives (when evaluated at c_t, k_t, and n_t). Similarly, the representative agent's assets a_t converge to a value a, which can be characterized using a time $t+1$ version of the IC constraint,

$$\begin{aligned}
a &= \lim_{t \to \infty} a_{t+1} = \lim_{t \to \infty} \left(W_{Ut+1} U_{ct+1} \beta_{t+1} R_{t+1}\right)^{-1} \sum_{s=t+1}^{\infty} \beta_s W_{Us}(U_{cs} c_s + U_{ns} n_s) \\
&= \left((1-\beta) U_c R\right)^{-1} \left(U_c c + U_n n\right),
\end{aligned}$$

where $\beta \equiv \bar{\beta}(V) = W_V \in (0,1)$ (see footnote 27). Using this expression, we see that A_{t+1}, which can be written as,

$$A_{t+1} = \frac{W_{UV,t}}{W_{Vt}}(U_{ct} c_t + U_{nt} n_t) + \frac{W_{VV,t}}{W_{Vt}} \underbrace{\beta_{t+1}^{-1} \sum_{s=t+1}^{\infty} \beta_s W_{Us}(U_{cs} c_s + U_{ns} n_s)}_{W_{Ut+1} U_{ct+1} R_{t+1} a_{t+1} \to W_U U_c R a},$$

converges as well, to some limit A,

$$\begin{aligned}
A_{t+1} &\to \frac{\beta_U}{\beta}(U_c c + U_n n) + \frac{\beta_V}{\beta} W_U U_c R a \\
&= \left(\frac{1-\beta}{W_U} \beta_U + \beta_V\right) \frac{1}{\beta} W_U U_c R a = \frac{\bar{\beta}'(V)}{\beta} W_U U_c R a \equiv A. \qquad (51)
\end{aligned}$$

where we defined $\beta_X \equiv W_{VX}$ and $X = U, V$. Similarly, we can show that B_t converges to some finite value B. Taking the limits of quantities in the first order conditions above, we thus find a system of equations for multipliers ν_t, μ, λ_t,

$$-\nu_t + \nu_{t+1} + \mu A = 0 \tag{52a}$$

$$-\nu_t + \mu\left(1 + \frac{U_{cc}c}{U_c} + \frac{U_{nc}n}{U_c}\right) + \mu\frac{B}{W_U} = \lambda_t\frac{1}{W_U U_c} \tag{52b}$$

$$-\nu_t + \mu\left(1 + \frac{U_{cn}c}{U_n} + \frac{U_{nn}n}{U_n}\right) + \mu\frac{B}{W_U} = -\lambda_t\frac{f_n}{W_U U_n} \tag{52c}$$

$$-\lambda_t + \lambda_{t+1}\beta f_k = 0. \tag{52d}$$

Substituting out λ_t from (52d) using (52a) and (52b), we find

$$\beta f_k - 1 = \frac{\lambda_t}{\lambda_{t+1}} - 1 = -\frac{W_U U_c}{\lambda_{t+1}}\mu A. \tag{53}$$

We now move to the two main results of this section. First, we show that steady state capital taxes are zero. Second, we show that steady state labor taxes are also zero, unless $\bar{\beta}'(V) = 0$, when preferences are locally additive separable.

Lemma 14. *At an interior steady state, capital taxes are zero, i.e.* $\beta f_k = 1$.

Proof. If $A = 0$ or $\mu = 0$ the result is immediate from (53). Suppose instead that $A \neq 0$ and $\mu \neq 0$. Then, (52a) implies that ν_t and hence λ_t diverges to $+\infty$ or $-\infty$. Then again, $\beta f_k = 1$ follows from (53).[61] $\square$

We move to our second result.

Lemma 15. *At an interior steady state, labor taxes are zero, i.e.* $\tau^n \equiv 1 + \frac{U_n}{U_c f_n} = 0$ *if* $\bar{\beta}'(V) \neq 0$ *and* $a > 0$.

Proof. By combining equations (52b) and (52c) we find an expression for τ^n,

$$\lambda_t \tau^n = \mu\frac{W_U U_n}{f_n}\left(\frac{U_{cc}c}{U_c} + \frac{U_{nc}n}{U_c} - \frac{U_{cn}c}{U_n} - \frac{U_{nn}n}{U_n}\right), \tag{54}$$

Note that by normality of consumption and labor the term in brackets is negative, $\frac{U_{cc}c}{U_c} + \frac{U_{nc}n}{U_c} - \frac{U_{cn}c}{U_n} - \frac{U_{nn}n}{U_n} < 0$. It is immediate from (54) that $\tau^n = 0$ if λ_t diverges to either $+\infty$ or $-\infty$.[62] Suppose $\lambda_t \to \lambda \in \mathbb{R}$. We distinguish whether $\mu = 0$ or $\mu \neq 0$. If $\mu = 0$, the economy was first best to start with, and the labor tax must be zero at any date, including at the steady state. If $\mu \neq 0$, convergence of λ_t (equivalent to convergence of ν_t) necessitates that $A = 0$, using (52a). But then (51) contradicts our assumptions that preferences are not locally additively separable, $\bar{\beta}'(V) \neq 0$, and steady state assets are positive $a > 0$. $\square$

[61] Notice that $\lambda_t \to 0$ requires $\mu = 0$ by (54), so the optimal allocation is first best to begin with, implying $\beta f_k = 1$.

[62] Since $A_t \to A \neq 0$ and μ is constant over time, ν_t and thus also λ_t have a well-defined limit in $[-\infty, \infty]$.

I Proof of Proposition 7

In this section, we prove Proposition 7. The proof is organized as follows. In Section I.1 we introduce the planning problem, derive and discuss the first order conditions, and define the largest feasible level of initial government debt $\bar{b}$. Section I.2 then focuses on parts A and B (i) of Proposition 7. Finally, Section I.3 proves the bang-bang property and parts B (ii) and C of Proposition 7.

I.1 Planning problem and first order conditions

As in the statement of the proposition, we fix some positive initial level of capital $k_0 > 0$. The problem under scrutiny is

$$V(b_0) \equiv \max_{\{c_t, n_t, k_t, r_t\}} \int_0^\infty e^{-\rho t} \left(u(c_t) - v(n_t)\right) dt \tag{55a}$$

$$\dot{c}_t = c_t \frac{1}{\sigma} (r_t - \rho) \tag{55b}$$

$$c_t + g + \dot{k}_t = f(k_t, n_t) - \delta k_t \tag{55c}$$

$$\int_0^\infty e^{-\rho t} \left(u'(c_t)c_t - v'(n_t)n_t\right) dt \geq u'(c_0)(k_0 + b_0) \tag{55d}$$

$$c_t > 0, n_t \geq 0, k_t \geq 0, r_t \geq 0$$

where recall that $u(c) = c^{1-\sigma}/(1-\sigma)$ and $v(n) = n^{1+\zeta}/(1+\zeta)$, $\zeta > 0$. In the entire analysis in this section, we write value functions such as $V(b_0)$ without explicit reference to k_0 since we treat k_0 as fixed. The current-value Hamiltonian of this optimal control problem with subsidiary condition (55d) (see, e.g., Gelfand and Fomin, 2000) can be written as

$$H(c, k; n, r; \lambda, \eta, \mu) = \Phi_u u(c) - \Phi_v v(n) + \eta c \frac{1}{\sigma} (r - \rho) + \lambda \left(f(k, n) - \delta k - c - g\right), \tag{56}$$

where we defined $\Phi_v \equiv 1 + \mu(1+\zeta)$ and $\Phi_u \equiv 1 + \mu(1-\sigma)$ with μ being the multiplier on the IC constraint; and where we denoted the costates of consumption and capital by η_t and λ_t, respectively. Notice that $\eta_t \leq 0$ or else $r_t = \infty$ were optimal, violating the resource constraint. Problem (55a) implies the following first order conditions for the controls $\{n_t, r_t\}$,

$$\Phi_v v'(n_t) = \lambda_t f_n(k_t, n_t) \tag{57a}$$

$$r_t \begin{cases} = 0 & \text{if } \eta_t < 0 \\ \in [0, \infty) & \text{if } \eta_t = 0, \end{cases} \tag{57b}$$

the following laws of motion for the costates,

$$\dot{\eta}_t - \rho\eta_t = \eta_t\frac{\rho}{\sigma} + \lambda_t - \Phi_u u'(c_t) \tag{57c}$$

$$\dot{\lambda}_t = (\rho - r_t^*)\lambda_t \tag{57d}$$

and the following optimality condition for the initial state of consumption c_0,

$$\eta_0 = -\mu\sigma c_0^{-\sigma-1}(k_0 + b_0). \tag{57e}$$

In equation (57d) we defined the before-tax return on capital as $r_t^* = f_k(k_t, n_t) - \delta$. The conditions (57a)–(57e), together with the constraints (55b)–(55d) and the two transversality conditions

$$\lim_{t\to 0} e^{-\rho t}\lambda_t k_t = 0 \tag{57f}$$

$$\lim_{t\to 0} e^{-\rho t}\eta_t c_t = 0 \tag{57g}$$

are sufficient for an optimum *if* we are able to establish that the planning problem (55a) is a concave maximization problem, or can be transformed into one using variable transformations.

The first order conditions (57a)–(57e) (though not the transversality conditions (57f) and (57g)) are necessary at an optimum since interiority is ensured by the imposition of Inada conditions; that is, with the exception when that optimum is also maximizing the subsidiary constraint, which is the IC constraint in our case (see Gelfand and Fomin, 2000). More specifically, the above first order conditions are not necessary when the optimum to (55a) achieves the supremum in

$$\bar{b} \equiv \sup_{\{c_t, n_t, k_t\}} \frac{1}{u'(c_0)} \int_0^\infty e^{-\rho t}\left(u'(c_t)c_t - v'(n_t)n_t\right) dt - k_0 \tag{58}$$

subject to the two other constraints (55b) and (55c). We deliberately formulated (58) in a way to define $\bar{b}$ as the highest level of b_0 for which there can possibly exist a feasible allocation. Notice that $\bar{b} \in [-\infty, \infty]$, allowing for $\bar{b} = -\infty$ if no feasible allocation exists at all (which might happen if g is very large), and $\bar{b} = \infty$ if there exists a feasible allocation for any value of b_0.

Since in the case that $b_0 = \bar{b}$ the supremum in (58) is attained, there are still necessary first order conditions the allocation satisfies, namely the ones corresponding to (58). These are exactly the same as (57a)–(57e) after substituting $\mu\eta_t$ for η_t and $\mu\lambda_t$ for λ_t, and then dividing by μ and setting $\mu = \infty$. This replaces Φ_u by $(1 - \sigma)$ and Φ_v by $(1 + \zeta)$ in (57a)–(57c), leaves (57d) unchanged and alters (57e) to $\eta_0 = -\sigma c_0^{-\sigma-1}(k_0 + b_0)$.

One additional remark about the setup in (55a) is in place. We stated an inequality IC constraint (55d), corresponding to a non-negative multiplier μ. This is without loss of generality in our setup, since at any optimum, μ will indeed be non-negative: From the first order condition (57e), we see that our assumption of positive initial private wealth,

$k_0 + b_0 > 0$, together with the non-positivity of η_0 means that $\mu \geq 0$.

I.2 Proof of parts A and B (i)

Our proof in this subsection proceeds in three steps. First, we characterize the space of solutions to a *restricted* planning problem, in which the length T of capital taxation is restricted to be infinity. Then we use these insights to prove that $T = \infty$ is in fact optimal in the *unrestricted* planning problem for levels of initial debt $b_0 \in [\underline{b}, \overline{b}]$ (with non-empty interior if $\sigma > 1$). Finally, we show that for all $b_0 < \overline{b}$ there are feasible policies with $T < \infty$. Throughout, we assume that $\sigma \geq 1$, as is assumed in parts A and B (i) of Proposition 7.

1^{st} **step: The restricted problem.** We start by studying a restricted planning problem, where we restrict ourselves to the case of indefinite capital taxation (at its upper bound). Effectively, this implies that $r_t = 0$ for all t and the path of c_t is entirely characterized by c_0 and (55b). To characterize this restricted problem, it will prove useful to define the minimum discounted sum of labor disutilities, henceforth *effective disutility from labor*, needed to sustain this path $\{c_t\}$ as

$$\tilde{v}(c_0) \equiv \min_{\{n_t, k_t\}} \int_0^\infty e^{-\rho t} v(n_t) dt \tag{59a}$$

$$\text{s.t. } c_0 e^{-\rho/\sigma t} + g + \dot{k}_t = f(k_t, n_t) - \delta k_t. \tag{59b}$$

We prove important properties of the effective disutility $\tilde{v}$ and the optimal control problem (59a) in Lemma 16.

Lemma 16. *The function $\tilde{v} : \mathbb{R}_+ \to \mathbb{R}_+$ is strictly convex, strictly increasing and continuously differentiable at any $c_0 \in \mathbb{R}_{++}$. It satisfies $\tilde{v}(0) > 0$. Moreover, for any $c_0 \geq 0$, there exists a unique solution $\{n_t^\infty, k_t^\infty\}$ and a costate of capital $\{\lambda_t^\infty\}$. Upon defining $c_t^\infty = c_0 e^{-\rho/\sigma t}$ it holds that, $\{c_t^\infty, n_t^\infty, k_t^\infty, \lambda_t^\infty\} \to (c^\infty, n^\infty, k^\infty, \lambda^\infty)$, where*

$$c^\infty = 0$$

$$f(k^\infty, n^\infty) = \delta k^\infty + g \tag{60a}$$

$$f_k(k^\infty, n^\infty) = \rho + \delta \tag{60b}$$

$$v'(n^\infty) = \lambda^\infty f_n(k^\infty, n^\infty). \tag{60c}$$

In particular, the transversality condition $\lim_{t\to\infty} e^{-\rho t} \lambda_t^\infty k_t^\infty = 0$ holds and

$$\tilde{v}'(c_0) = \int_0^\infty e^{-(\rho+\rho/\sigma)t} \lambda_t dt. \tag{60d}$$

Proof. The proof has 4 steps: First, we prove existence and uniqueness of the solution to a "bounded" version of the optimal control problem (59a) with bounds on n_t and k_t. Second, we characterize the optimal paths (k_t^∞, n_t^∞) of this problem. Third, we show that increasing the bounds on k_t and n_t makes the bounded problem equivalent to (59a). Finally, we establish that the claimed properties of $\tilde{v}$.

First step. For this step, relax the constraint (59b) to be an inequality "$\le$" and introduce upper bounds $\bar{k} > 0$ and $\bar{n} > 0$ on k and n. Using the definition of k^∞ and n^∞ in (60a)–(60b), pick $\bar{k} > \max\{k_0, k^\infty\}$ and pick $\bar{n}$ large enough so that $\bar{n} > n^\infty$ and so that $\dot{k}_0 > 0$ is feasible at time $t = 0$.[63] This means the problem is given by

$$\tilde{v}(c_0) \equiv \min_{\{n_t,k_t\}} \int_0^\infty e^{-\rho t} v(n_t) dt \tag{61a}$$
$$\text{s.t. } c_0 e^{-\rho/\sigma t} + g + \dot{k}_t \le f(k_t, n_t) - \delta k_t$$
$$k_t \in [0, \bar{k}],\ n_t \in [0, \bar{n}].$$

This problem is clearly a strictly convex minimization problem (strictly convex objective and a convex constraint), even without bounds on k and n, and therefore at most admits a single solution. A straightforward application of Seierstad and Sydsaeter (1987, Section 3.7, Theorem 15) to the optimal control problem (61a) reveals that there always exist paths $\{n_t^\infty, k_t^\infty\}$ that attain the minimum in (61a).[64]

Second step. We now study the long-run properties of the solution to the problem (61a). Before we dive into the details, we note that $k^\infty > 0$ and $n^\infty > 0$ are uniquely determined by (60a) and (60b) due to the Inada properties of $f_k(\cdot, n)$ and the fact that $f/k \ge f_k$. λ^∞ follows from (60c). At each point where $k_t < \bar{k}$ and $n_t < \bar{n}$, the necessary first order conditions corresponding to (61a) are given by

$$v'(n_t) = \lambda_t f_n(k_t, n_t) \tag{62a}$$
$$\dot{\lambda}_t = \lambda_t(\rho - r_t^*), \tag{62b}$$

where λ_t denotes the costate of k_t. Notice that n_t is continuous, as an immediate consequence of (62a) and of the fact that both k_t and λ_t are continuous. Also note that (62a) implies $\lambda_t \ge 0$, meaning our relaxation of the resource constraint (59b) to an inequality was without loss of generality. Using the resource constraint (59b) and (62a)–(62b), we can derive an ODE system entirely in terms of n_t and k_t, consisting of the resource constraint (59b) itself and of

$$(\zeta + \alpha_t)\frac{\dot{n}_t}{n_t} = \rho + (1 - \alpha_t)\delta - \alpha_t \frac{g + c_t}{k_t},$$

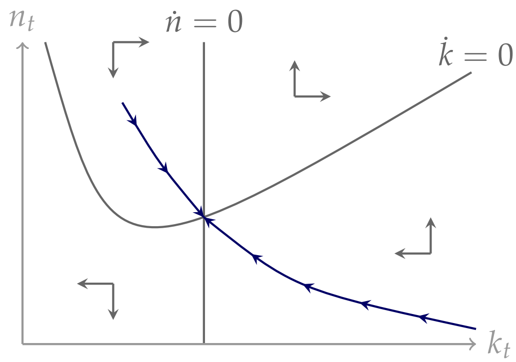
where $\alpha_t = \alpha(k_t/n_t) \equiv \frac{\partial \log f_n}{\partial \log(k_t/n_t)}$. We can also abbreviate the ODEs as $\dot{k} = \dot{k}(k, n, c_t)$ and $\dot{n} = \dot{n}(k, n, c_t)$. Define the two sets

$$\mathcal{A}_t \equiv \{(k, n) \mid \dot{n}(k, n, c_t) > 0,\ \dot{k}(k, n, c_t) > 0\}$$
$$\mathcal{B}_t \equiv \{(k, n) \mid \dot{n}(k, n, c_t) < 0,\ \dot{k}(k, n, c_t) < 0\}.$$

[63] $\dot{k}_0 > 0$ iff $f(k_0, \bar{n}) - \delta k_0 - g - c_0 > 0$.

[64] This relies on our choice of $\bar{n}$ which ensures that $\dot{k}_0 > 0$, so even for low values of k_0 there exist admissible paths $\{n_t, k_t\}$.

Figure 6: Phase diagram characterizing the solution to the restricted problem (59a).



To illustrate these sets, note that for large t, $c_t \approx 0$, we can draw the phase diagram that corresponds to the ODE system. This is done in Figure 6 for the Cobb-Douglas case where $\alpha_t = const$. In that figure, $\mathcal{A}_t$ is the top right area, while $\mathcal{B}_t$ is the bottom left area. We now argue that the state (k_t, n_t) can never be in $\mathcal{A}_t$ for any t, and never be in $\mathcal{B}_t$ for large t. If for any t, $(k_t, n_t) \in \mathcal{A}_t$, n_t can be lowered to achieve $\dot{k}_t = 0$ at all times, clearly improving the objective. If there does not exist a time s such that $(k_t, n_t) \notin \mathcal{B}_t$ for $t > s$, then it must be that asymptotically $(k_t, n_t) \in \mathcal{B}_t$ for all sufficiently large t. But in that case, $k_t \to 0$, contradicting feasibility (since government spending is positive, $g > 0$). Therefore, it must be that $(k_t^\infty, n_t^\infty) \to (k^\infty, n^\infty)$.

Note that the optimal costate λ_t^∞ can be computed using the first order condition for labor, (62a). Due to the steady state convergence of the system, the transversality condition $\lim_{t\to\infty} e^{-\rho t}\lambda_t^\infty k_t^\infty = 0$ naturally holds.

Third step. We now show that there exists a sufficiently large bound $\overline{n}$ such that the solutions of the problem without bounds, (59a) and the problem with bounds (61a) coincide. This is the case if there exists a $\overline{n}$ such that $n_t^\infty < \overline{n}$ at the optimum at all times t. Assume the contrary held, that is, no matter how large $\overline{n}$ is, at the corresponding optimal path, which we denote by $(k_t^\infty(\overline{n}), n_t^\infty(\overline{n}))$ to emphasize the dependence on $\overline{n}$, there exist times t where $n_t^\infty(\overline{n}) = \overline{n}$. Since $n_t^\infty(\overline{n})$ can never approach $\overline{n}$ from below (this would require $(k_t, n_t) \in \mathcal{A}_t$), it must be that there exists a time $s > 0$ such that $n_t^\infty(\overline{n}) = \overline{n}$ for any $t \in [0, s]$ and any arbitrarily large $\overline{n}$. It is straightforward to see that this lets $k_s^\infty(\overline{n})$ grow unboundedly large, in particular leading to $(k_s^\infty(\overline{n}), n_s^\infty(\overline{n})) \in \mathcal{A}_s$—a contradiction. This completes our proof that problem (59a) admits a unique solution, which approaches the steady state (k^∞, n^∞) asymptotically.

Fourth step. In our final step, we derive the claimed properties of $\tilde{v}$. First, since the objective is strictly convex, $\tilde{v}$ is strictly convex. It is also strictly increasing since the constraint tightens with larger c_0. $\tilde{v}(0) > 0$ follows directly from $g > 0$. For differentiability, pick any $\hat{c}_0 \in \mathbb{R}_{++}$ and denote the associated optimal path for capital by $\{\hat{k}_t^\infty\}$. Following the logic in Benveniste and Scheinkman (1979) we can define a strictly convex and differentiable function $w(c_0) = \int_0^\infty e^{-\rho t}\frac{1}{1+\zeta}N\left(\hat{k}_t^\infty, c_0 e^{-\rho/\sigma t} + g + \dot{\hat{k}}_t^\infty + \delta\hat{k}_t^\infty\right)^{1+\zeta} dt$ where $N(k, y) \equiv f(k, \cdot)^{-1}(y)$ is the level of labor needed to fund output $y \geq 0$ given capital

$k > 0$. By contsruction, $w(\hat{c}_0) = v(\hat{c}_0)$ and $w(c_0) \geq v(c_0)$ locally around $\hat{c}_0$.[65] This implies that $\tilde{v}$ is differentiable at any $c_0 \in \mathbb{R}_{++}$ with derivative[66]

$$\tilde{v}'(c_0) = \int_0^\infty e^{-(\rho+\rho/\sigma)t} \frac{v'(n_t^\infty)}{f_n(k_t^\infty, n_t^\infty)} dt. \tag{63}$$

This formula for the derivative of $\tilde{v}$ is equivalent to (60d) after expressing the former in terms of λ_t using the first order condition for labor (62a). This concludes our proof of Lemma 16. □

The effective disutility $\tilde{v}(c_0)$ is convenient since in the original planning problem (55a), labor disutility appears in present value terms both in the objective as well as in the IC constraint (55d). Moreover, due to the assumption of power disutility, both present values are essentially $\tilde{v}(c_0)$ up to a constant factor. The restricted version of the original planning problem (55a) can now be simply written as restricted problem

$$V_\infty(b_0) \equiv \max_{c_0 > 0} u(c_0)\frac{\sigma}{\rho} - \tilde{v}(c_0) \tag{64a}$$

$$c_0^{1-\sigma}\frac{\sigma}{\rho} - (1+\zeta)\tilde{v}(c_0) \geq c_0^{-\sigma}(k_0 + b_0). \tag{64b}$$

We obtained (64a) from the original problem (55a) by requiring that $T = \infty$ and using the definition of $\tilde{v}$. We characterize the restricted problem (64a) in the following lemma.

Lemma 17. *There exists a level of initial debt $\overline{b}^r \in \mathbb{R}$ such that a solution to the restricted planner's problem (64a) exists if and only if $b_0 \leq \overline{b}^r$. For each $b_0 \leq \overline{b}^r$, there is a unique optimum $c_0^\infty(b_0) \in \mathbb{R}_{++}$ and for each $b_0 < \overline{b}^r$ there is a unique multiplier $\mu^\infty(b_0) \in [0, \infty)$ on the IC constraint (64b) such that*

$$\Phi_u u'(c_0)\frac{\sigma}{\rho} - \Phi_v \tilde{v}'(c_0) = -\sigma\mu c_0^{-\sigma-1}(k_0 + b_0), \tag{65}$$

for $c_0 = c_0^\infty(b_0)$, $\mu = \mu^\infty(b_0)$. Finally, there exists some $b^* < \overline{b}^r$ such that $\mu^\infty : [b^*, \overline{b}^r) \to [0, \infty)$ is a continuous and strictly increasing bijection.

Proof. First, notice that the IC constraint of the restricted planning problem, (64b), can be rewritten as

$$c_0\frac{\sigma}{\rho} - (1+\zeta)c_0^\sigma \tilde{v}(c_0) \geq k_0 + b_0. \tag{66}$$

Observe that this is a convex constraint, as its left hand side is strictly concave. It is also strictly increasing at $c_0 = 0$ and diverges to $-\infty$ for large c_0.[67] Therefore, there exists an

[65] The expression for w is obtained by substituting the resource constraint (59b) into the objective (59a).

[66] Notice that the derivative must be finite since $\tilde{v}$ is strictly convex and finite-valued for any $c_0 \in \mathbb{R}_+$.

[67] Note that for $\sigma = 1$, (66) reads $c_0(\frac{\sigma}{\rho} - (1+\zeta)\tilde{v}(c_0)) \geq k_0 + b_0$ and by positivity of $k_0 + b_0$ and monotonicity of $\tilde{v}$, this means that $\frac{\sigma}{\rho} - (1+\zeta)\tilde{v}(0) > 0$ (which is exactly equal to the derivative of the left hand side of (66) at $c_0 = 0$).

interior maximum at some $\bar{c} > 0$. By definition, $c_0 = \bar{c}$ is the only value that is compatible with the IC constraint if $b_0 = \bar{b}^r$, where we defined

$$\bar{b}^r \equiv \max_{c_0 > 0} c_0 \frac{\sigma}{\rho} - (1+\zeta) c_0^{\sigma} \tilde{v}(c_0) - k_0. \tag{67}$$

The maximizer $\bar{c}$ is then characterized by the first order conditions

$$\frac{\sigma}{\rho} \bar{c}^{-\sigma} = (1+\zeta)\sigma \bar{c}^{-1} \tilde{v}(\bar{c}) + (1+\zeta)\tilde{v}'(\bar{c}). \tag{68}$$

For any $b_0 > \bar{b}^r$ the set of feasible c_0 compatible with the IC constraint (64b) is empty, so the restricted planning problem (64a) has a solution precisely when $b_0 \leq \bar{b}^r$.

An advantage of writing the IC constraint as in (66) is that it allows us to see that the restricted problem (64a) has a strictly concave objective with a convex and bounded constraint set. The objective attains its unconstrained maximum at some $c^* \in (0, \infty)$ satisfying $u'(c^*)\frac{\sigma}{\rho} = \tilde{v}'(c^*)$. We can show that $c^* > \bar{c}$ since the objective is increasing at $\bar{c}$,

$$u'(\bar{c})\frac{\sigma}{\rho} - \tilde{v}'(\bar{c}) = (1+\zeta)\sigma \bar{c}^{-1} \tilde{v}(\bar{c}) + \zeta \tilde{v}'(\bar{c}) > 0,$$

where we used the first order condition for $\bar{c}$, (68). Define $b^* \equiv c^* \frac{\sigma}{\rho} - (1+\zeta) c^* \tilde{v}(c^*) - k_0$, so that c^* lies in the constraint set (64b) if and only if $b_0 \leq b^*$—or in other words, the constraint holds with equality for any $b_0 \geq b^*$. We next show that there exists (a) a strictly decreasing (and hence continuous) bijection $c^\infty : [b^*, \bar{b}^r) \to (\bar{c}, c^*]$ and (b) a strictly increasing (and hence continuous) bijection $\mu^\infty : [b^*, \bar{b}^r) \to [0, \infty)$ such that $c^\infty(b_0)$ is the unique solution to the strictly concave problem (64a), and constraint (64b) has Lagrange multiplier $\mu^\infty(b_0)$, for any $b_0 \in [b^*, \bar{b}^r)$.

Take any $c_0 \in (\bar{c}, c^*]$. Clearly, c_0 is optimal with Lagrange multiplier μ when initial debt is b_0 if the three objects c_0, μ, b_0 satisfy the first order condition of the problem—which can easily be seen to be given by (65)—and the constraint (64b). By substituting out b_0 from (65) using the constraint, the first order condition can be expressed as function of μ,

$$\mu = \frac{\frac{\sigma}{\rho} - c_0^{\sigma} \tilde{v}'(c_0)}{(1+\zeta)\sigma c_0^{\sigma-1} \tilde{v} + (1+\zeta) c_0^{\sigma} \tilde{v}'(c_0) - \sigma/\rho} \equiv M(c_0).$$

For $c_0 \in (\bar{c}, c^*]$, the denominator is positive and strictly increasing in c_0, approaching 0 for $c_0 \searrow \bar{c}$; while the numerator is strictly decreasing and non-negative, with a zero at $c_0 = c^*$. This defines a strictly decreasing bijection $M : (\bar{c}, c^*] \to [0, \infty)$. From the constraint (64b), we see that

$$b_0 = c_0 \frac{\sigma}{\rho} - (1+\zeta) c_0^{\sigma} \tilde{v}(c_0) - k_0 \equiv B(c_0)$$

which, by definition of $\bar{b}^r$ and $\bar{c}$, defines a strictly decreasing bijection $B : (\bar{c}, c^*] \to [\tilde{b}, \bar{b}^r)$.

It follows that for any $b_0 \in [\tilde{b}, \overline{b}^r)$, the unique solution to (64a) is given by $c^\infty(b_0) = B^{-1}(b_0)$, with associated multiplier $\mu^\infty(b_0) = M(B^{-1}(b_0))$. This concludes the proof. $\square$

We finished our characterization of the restricted planning problem and are now ready for the second and main part of the proof of Proposition 7.

2^{nd} **step: Optimality of $T = \infty$ in the unrestricted problem.** Before we proceed to prove the optimality of $T = \infty$ in the unrestricted problem, we establish that $\overline{b}^r$ is not just the upper bound of possible initial debt in the restricted planning problem, but equal to $\overline{b}$, the one in the unrestricted planning problem (55a).

Lemma 18. *Let $b_0 \in \mathbb{R}$ and $\sigma \geq 1$. The constraints (55b), (55c), (55d) define a non-empty set for $\{c_t, n_t, k_t, r_t\}$ if and only if $b_0 \leq \overline{b}^r$. In particular, $\overline{b} = \overline{b}^r$. Moreover, if $b_0 = \overline{b}^r$ then capital is necessarily taxed at the maximum, $T = \infty$.*

Proof. It suffices to show that the constraint set in the original problem is empty for $b_0 > \overline{b}^r$, and that $T = \infty$ is necessary for $b_0 = \overline{b}^r$. We show both by proving that any $b_0 \geq \overline{b}^r$ is infeasible with if capital is not taxed at its upper bound in all periods.

Hence fix some $b_0 \geq \overline{b}^r$ and assume it was achievable without $T = \infty$ by $\{c_t, n_t, k_t, r_t\}$. Then, it must be that $r_t > 0$ on some non-trivial interval, and the path of consumption is described by the Euler equation (55b), as always. Let the initial consumption value be c_0 and denote by $\hat{c}_t$ the path which starts at the same initial consumption $\hat{c}_0 = c_0$ but keeps falling at the fastest possible rate $-\rho/\sigma$ forever, corresponding to $T = \infty$. Similarly, define by $\hat{n}_t$ the path for labor which keeps k_t fixed but satisfies the resource constraint with consumption equal to $\hat{c}_t$. Clearly, $\hat{n}_t \leq n_t$ for all t and $\hat{n}_t < n_t$ on a positive-measure set of times t. Because the left hand side of (55d) is weakly decreasing in c_t and strictly decreasing in n_t, this strictly relaxes the IC constraint. Hence,

$$\int_0^\infty e^{-\rho t} \hat{c}_t^{1-\sigma} dt - \int_0^\infty e^{-\rho t} v(\hat{n}_t) dt > \hat{c}_0^{-\sigma}(k_0 + b_0).$$

Notice, however, that for $T = \infty$, we can do even better by optimizing over labor (not necessarily keeping capital constant, see (59a)), leading to

$$\hat{c}_0^{1-\sigma} \frac{\sigma}{\rho} - (1 + \zeta)\tilde{v}(\hat{c}_0) > \hat{c}_0^{-\sigma}(k_0 + b_0).$$

By definition of $\overline{b}^r$ in (67) this is a contradiction to $b_0 \geq \overline{b}^r$. Therefore, $\overline{b}^r$ is equal to the highest sustainable debt level in the original problem, $\overline{b}$, and can only be achieved with $T = \infty$. $\square$

Our next lemma establishes that the unrestricted problem (55a) is a strictly concave maximization problem with convex constraints. This will be helpful when proving uniqueness in Lemma 20 below.

Lemma 19. *Suppose $\sigma \geq 1$. The unrestricted problem (55a) can be transformed into a strictly concave maximization problem with convex constraints, using variable substitution. Therefore, any optimum of (55a) is unique when $\sigma \geq 1$.*

Proof. We rewrite (55a) in terms of the two variables $u_t \equiv u(c_t) \in (-\infty, 0)$ and $v_t \equiv v(n_t) \in [0, \infty)$ instead of c_t and n_t. We only consider the case $\sigma > 1$; the case $\sigma = 1$ is analogous. For $\sigma > 1$, the substitution yields

$$
\begin{aligned}
V(b_0) \equiv \max_{\{u_t, v_t, k_t\}} & \int_0^\infty e^{-\rho t} (u_t - v_t)\, dt \\
& \dot{u}_t \geq (\sigma - 1)\frac{\rho}{\sigma} u_t \\
& ((1-\sigma)u_t)^{-1/(\sigma-1)} + g + \dot{k}_t \leq f\left(k_t, ((1+\zeta)v_t)^{1/(1+\zeta)}\right) - \delta k_t \\
& \int_0^\infty e^{-\rho t} ((1-\sigma)u_t - (1+\zeta)v_t)\, dt \geq ((1-\sigma)u_0)^{\sigma/(\sigma-1)} (k_0 + b_0) \\
& u_t < 0, v_t \geq 0, k_t > 0.
\end{aligned}
\tag{69}
$$

We made two additional simplifications in (69): We incorporated the inequality for the control $r_t \geq 0$ in the Euler equation constraint (55b); and the (strictly convex) resource constraint was relaxed to be an inequality, which is without loss of generality since by (57a) we know that its Lagrange multiplier, the costate of capital λ_t, is necessarily positive at any optimum. Since the resource constraint binds and is strictly convex, all other constraints in (69) are also convex and the objective is linear, this planning problem can at most have a single solution. And, (57a)–(57e), (57f), (57g), (55b)–(55d) are sufficient conditions to find this solution. □

Our next lemma finally establishes the optimality of $T = \infty$ in the unrestricted problem (55a).

Lemma 20. *Suppose $\sigma > 1$ and define $\underline{b} \equiv (\mu^\infty)^{-1}\left(\frac{1}{\sigma-1}\right)$ with μ^∞ as in Lemma 17. Indefinite capital taxation is optimal in the Chamley problem (55a) if and only if $b_0 \in [\underline{b}, \overline{b}]$.*

Proof. As a consequence of Lemma 19, the unrestricted planning problem (55a) can be transformed into a strictly concave maximization problem with convex constraints. This implies that the first order conditions (57a)–(57e), together with transversality conditions (57f), (57g), and constraints (55b)–(55d) are in fact sufficient to characterize the unique optimum of the unrestricted planning problem (55a). In this proof we guess a solution and verify the sufficient conditions in a first step. In a second step, we prove that any $b_0 < \underline{b}$ does not imply positive long run capital taxation, where $T < \infty$. Throughout the proof, we focus on $b_0 < \overline{b}$ since we know from Lemma 18 that initial debt of $\overline{b}$ requires indefinite capital taxation.

First step: Let $b_0 \in [\underline{b}, \overline{b})$. We now construct an allocation $\{c_t, n_t, k_t, r_t\}$ and multipliers $\{\lambda_t, \eta_t\}, \mu$ that satisfy all the sufficient conditions. We define $c_0 \equiv c^\infty(b_0)$ as in Lemma 17; given c_0, $\{c_t, n_t, k_t\} \equiv \{c_t^\infty, n_t^\infty, k_t^\infty\}$ and $\lambda_t \equiv \Phi_v \cdot \lambda_t^\infty$ with notation as in Lemma 16; $\mu \equiv \mu^\infty(b_0)$ as in Lemma 17; $\eta_0 \equiv \Phi_u u'(c_0)\frac{\sigma}{\rho} - \Phi_v \tilde{v}'(c_0)$ (which is negative since $\Phi_u \leq 0$ by construction of μ) and η_t as solution to the ODE (57c) with initial condition η_0. The first order conditions (57a)–(57d) are satisfied by construction and by the fact that the allocation $\{n_t^\infty, k_t^\infty, \lambda_t^\infty\}$ satisfies (62a) and (62b). The first order condition for initial consumption (57e) is equivalent to (65) in Lemma 17. The Euler equation constraint (55b) is trivially satisfied by construction of $\{c_t\}$. The resource constraint holds for $\{c_t^\infty, n_t^\infty, k_t^\infty\}$ (see (59b) and Lemma 16) and therefore also for $\{c_t, n_t, k_t\}$. Due to the fact that $\{n_t^\infty, k_t^\infty\}$ solves (59a) and $c_t = c_0 e^{-\rho/\sigma t}$, the IC constraint (55d) can be seen to be equivalent to (64b) and hence is satisfied since c_0 was chosen to be $c^\infty(b_0)$. Finally, Lemma 16 implies that the transversality condition for capital, (57f), holds. And, concluding the second step, the transversality condition for consumption, (57g), holds since

$$e^{-\rho t}\eta_t c_t = c_0 e^{-(\rho+\rho/\sigma)t}\eta_t = -c_0\int_t^\infty e^{-(\rho+\frac{\rho}{\sigma})s}\lambda_t dt + c_0\Phi_u u'(c_0)\frac{\sigma}{\rho}e^{-\frac{\rho}{\sigma}t} \to 0. \tag{70}$$

and by this expression it also follows that $\eta_t < 0$ at all times t. The second equality in (70) builds on an integral version of the law of motion of η_t, which we obtained by combining (57c) with our definition of η_0 as $\Phi_u u'(c_0)\frac{\sigma}{\rho} - \Phi_v \tilde{v}'(c_0)$ and the expression for $\tilde{v}'(c_0)$ in (60d) from Lemma 16. It will become important in the second step below that (70) also reveals the limiting behavior of η_t itself: $\lim_{t\to\infty}\eta_t = -\infty$ but $\lim_{t\to\infty} e^{-\rho t}\eta_t = \Phi_u u'(c_0)\frac{\sigma}{\rho}$.

Second step: We proceed by contradiction. Suppose $b_0 < \underline{b}$ gave rise to indefinite capital taxation (at the maximum rate). Then, reversing the logic of the first step, it must be the case that the allocation $\{c_t, n_t, k_t\}$ is also optimal in the labor disutility minimization problem (59a) with multipliers $\lambda_t^\infty = \frac{1}{\Phi_v}\lambda_t$, given c_0; and c_0 and μ must be optimal given b_0 in the restricted planning problem (64a), that is, $c_0 = c^\infty(b_0)$ and $\mu = \mu^\infty(b_0) < \frac{1}{\sigma-1}$. Since the first order condition (57e) is necessary, it must then be the case that $\eta_0 = \Phi_u u'(c_0)\frac{\sigma}{\rho} - \Phi_v\tilde{v}'(c_0)$ by comparing it to (65). Equation (70) thus holds as in the second step, implying $\lim_{t\to\infty} e^{-\rho t}\eta_t = \Phi_u u'(c_0)\frac{\sigma}{\rho}$ which now is positive since $\Phi_u > 0$, a contradiction to the optimality of capital taxes. □

3^{rd} **step: Feasibility of finite capital taxation for all** $b_0 < \bar{b}$**.** We now move to the third and last part of this section. Here, we establish:

Lemma 21. *For any initial government debt level $b_0 < \bar{b}$, there are implementable allocations with nonzero capital taxation for only a finite time, $T < \infty$.*

Proof. Fix $b_0 \le \bar{b}$ and fix the allocation $\{c_t^\infty, n_t^\infty, k_t^\infty\}$ that is optimal among all allocations with indefinite capital tax. By construction, this allocation satisfies the restricted problem (64a). We now explicitly construct an allocation $\{\tilde{c}_t, \tilde{n}_t, \tilde{k}_t\}$ for which there is no capital tax, $\dot{\tilde{c}} = \frac{1}{\sigma}(r_t^* - \rho)\tilde{c}_t$, after time some time $T < \infty$ but that is feasible—satisfying constraints (55b)–(55d)—with initial debt $b_0 - \epsilon$, for $\epsilon > 0$ arbitrarily small. First, we describe the allocation for all times $t \ge T$. Consider

$$\begin{aligned} V^{\text{zero tax}}(\hat{k}) \equiv \max_{\{c_t, n_t, k_t\}_{t\ge T}} & \int_T^\infty e^{-\rho(t-T)}\left(u(c_t) - v(n_t)\right)dt \\ \text{s.t. } & c_t + g + \dot{k}_t = f(k_t, n_t) - \delta k_t \\ & k_T = \hat{k} \\ & c_t > 0, n_t \ge 0, k_t \ge 0 \end{aligned}$$

which is the social planning problem of a standard neoclassical growth model with power utilities in consumption and labor, and a Cobb-Douglas technology (i.e. zero labor and zero capital taxes). It is known that such a model has optimal paths $\{c_t^*, n_t^*, k_t^*\}$ that monotonically converge to a unique positive steady state (c^*, n^*, k^*). This implies that $\{n_t^*\}$ is bounded from above by $\overline{n}(\hat{k}) = \max\{n^*, n(\hat{k})\}$ where $n(\cdot)$ denotes the (continuous) policy function for labor supply. Moreover, the undistorted Euler condition holds along the path for consumption $\{c_t^*\}$. Also, it is well known that the consumption policy function $c(k)$ of this problem is continuous and strictly increasing, with $c(k) > 0$ for any $k > 0$. Fix $\hat{k} \equiv k_T^\infty$. Since k_t^∞ converges to a positive limit $k^\infty > 0$ but $c_t^\infty \to 0$ (see Lemma 16), it is the case for sufficiently large T that $c(\hat{k}) > c_T^\infty$. Focus on such T. Also let $\overline{n} \equiv \sup_t \overline{n}(k_t^\infty) < \infty$ be an upper bound for labor (which by construction is uniform in T). Notice that $\overline{n} < \infty$ since k_t^∞ converges to some $k^\infty > 0$.

Now construct the paths $\{\tilde{c}_t, \tilde{n}_t, \tilde{k}_t\}$ by piecing together $\{c_t^\infty, n_t^\infty, k_t^\infty\}$ for $t < T$ and a zero-tax path $\{c_t^*, n_t^*, k_t^*\}$, starting with $k_T^* = k_T^\infty$, for $t \geq T$. By design, the capital stock is continuous at $t = T$ and consumption jumps upwards at $t = T$.[68] Using this construction, the allocation satisfies the resource constraint at all periods, and the Euler equation with equality for $t > T$. Also,

$$\int_0^\infty e^{-\rho t} \left(u'(c_t^\infty)c_t^\infty - v'(n_t^\infty)n_t^\infty\right) dt - \int_0^\infty e^{-\rho t} \left(u'(\tilde{c}_t)\tilde{c}_t - v'(\tilde{n}_t)\tilde{n}_t\right) dt =$$

$$\underbrace{\int_T^\infty e^{-\rho t} \left(u'(c_t^\infty)c_t^\infty - u'(c_t^*)c_t^*\right) dt}_{\leq e^{-\rho T} u'(c_T^\infty) c_T^\infty \frac{\sigma}{\rho}} + \underbrace{\int_T^\infty e^{-\rho t} \left(v'(n_t^*)n_t^* - v'(n_t^\infty)n_t^\infty\right) dt}_{\leq e^{-\rho T} \frac{1}{\rho} \overline{n}^{1+\zeta}}. \tag{71}$$

As $e^{-\rho T} u'(c_T^\infty)c_T^\infty \to 0$ both terms in (71) approach zero. This is why for T sufficiently large, $\int_0^\infty e^{-\rho t} \left(u'(\tilde{c}_t)\tilde{c}_t - v'(\tilde{n}_t)\tilde{n}_t\right) dt$ approaches $u'(\tilde{c}_0)(k_0 + b_0)$. Thus, for any $\epsilon > 0$, there exists a T such that the allocation $\{\tilde{c}_t, \tilde{n}_t, \tilde{k}_t\}$ is implementable without capital taxes after time T, for initial debt $b_0 - \epsilon$,

$$\int_0^\infty e^{-\rho t} \left(u'(\tilde{c}_t)\tilde{c}_t - v'(\tilde{n}_t)\tilde{n}_t\right) dt \geq u'(\tilde{c}_0)(k_0 + b_0 - \epsilon)$$

which is what we set out to show. This proves that for any $b_0 < \overline{b}$, there exists a feasible (but not necessarily optimal) path with only a finite period of positive capital taxation. □

Summary This concludes the proof of parts A and B (i) of Proposition 7. For part A, we proved (i) in Lemma 18 , (ii) in Lemma 20 and (iii) in Lemma 21. Part B (i) was shown in Lemma 18.

[68] We think of this as a very high capital subsidy for a very short amount of time (which would definitely not be violating any capital tax constraints). If one prefers to avoid this simple limit case, one could easily smooth out this jump over some very small interval. This makes no difference whatsoever for the argument that follows.

I.3 Proof of the bang-bang property and parts B (ii) and C

We proceed in three steps. We first establish a transversality condition that is necessary at any optimum (in general, transversality conditions are not necessary). Then, using this transversality condition, we derive the "bang-bang" property of capital taxes. Notice that previous proofs of this property relied on the assumption that indefinite capital taxation is not optimal, which we showed is not the case. The bang-bang property lets us summarize an optimal capital tax plan by the date $T \in [0, \infty]$ at which capital taxes jump from the upper bound $\bar{\tau}$ to zero. In the final step, we prove parts B (ii) and C, that is, $T < \infty$ if either $\sigma < 1$ or $\sigma = 1$ and $b_0 = \overline{b}$.

1^{st} **step: A necessary transversality condition.**

Lemma 22. *Let $\{c_t, n_t, k_t, r_t\}$ be a solution to problem (55a), with multipliers $\{\lambda_t, \eta_t, \mu\}$. If $\exists\, s \geq 0$ such that $c_t = c_s e^{-\rho(t-s)/\sigma}$ for all $t \geq s$, then the transversality condition for consumption (57g) holds.*

Proof. We first establish that under the conditions of the lemma, $\{k_t, n_t\}$ converges to a positive steady state. If $c_t = c_s e^{-\rho(t-s)/\sigma}$, then $\{k_t, n_t\}_{t \geq s}$ must be minimizing the stream of labor disutilities (61a) given initial capital k_s and initial consumption c_s. Therefore, $\{k_t, n_t\} \to \{k^\infty, n^\infty\}$, using the notation from Lemma 16.

Thus, there exists some large enough $\overline{n} > 0$ such that

$$f(k_t, \overline{n}) - \delta k_t - c_t - g > 0 \tag{72}$$

for all t. Since the time t controls maximize the time t Hamiltonian H_t (see (56)), we then have for any $\overline{n}$

$$e^{-\rho t}(\Phi_u u(\hat{c}_t) - \Phi_v v(\overline{n})) + e^{-\rho t}\eta_t \hat{c}_t \frac{1}{\sigma}(-\rho) + e^{-\rho t}\lambda(f(\hat{k}_t, \overline{n}) - \delta \hat{k}_t - \hat{c}_t - g) \leq e^{-\rho t} H_t \to 0 \tag{73}$$

where the left hand side is the present value Hamiltonian with controls $r_t = 0$ and $n_t = \overline{n}$, and the right hand side is the present value Hamiltonian with optimal controls r_t, n_t (both along the optimal path for c_t, k_t). The right hand side converges to zero following Michel (1982). Notice that in (73), $e^{-\rho t}(\Phi_u u(c_t) - \Phi_v v(\overline{n})) \to 0$. Suppose $\liminf_{t\to\infty} e^{-\rho t}\eta_t c_t$ were negative. Then, according to (73) it would have to be that

$$\limsup_{t\to\infty} e^{-\rho t}\lambda(f(k_t, \overline{n}) - \delta k_t - c_t - g) \leq \liminf_{t\to\infty} e^{-\rho t}\eta_t c_t \frac{1}{\sigma}\rho < 0$$

contradicting (72). This means the transversality condition for consumption (57g) holds. $\square$

2^{nd} **step: The bang bang property.** We move to the first main result of this subsection.

Lemma 23. *A solution to problem (55a) is of the form that the capital tax τ_t binds at the upper bound for some time $T \in [0, \infty]$ and is equal to zero thereafter.*

Proof. Let $\{c_t, n_t, k_t, r_t\}$ be an optimal allocation solving (55a), for some initial debt $b_0 \in \mathbb{R}$. Let $\{\lambda_t, \eta_t, \mu\}$ be a set of multipliers such that allocation and multipliers satisfy the necessary first order conditions for the case $b < \overline{b}$, (57a)–(57e) and constraints (55b)–(55d). Our proof is analogous if $b = \overline{b}$. We first show that if $\tau_t < \bar{\tau}$ on some non-trivial interval, then $\tau_t = 0$ on that interval. Second, we prove that $\tau_t = 0$ at all times after that interval as well. The proof utilizes the fact that once $\tau_t = 0$, it must not only be that $r_t^* \geq 0$ at that time (or else the $r_t \geq 0$ constraint would be binding); but also that $r_t^* > 0$ for all future times. We formally prove this fact in Lemma 24 below.

First, suppose $\tau_t < \bar{\tau}$ for some non-trivial interval $t \in [s_0, s_1]$. Then, by (57b), $r_t > 0$ and $\eta_t = 0$ on that interval. Hence, by (57c), $\lambda_t = \Phi_u u'(c_t)$. Taking logs and differentiating implies an undistorted Euler equation of the agent. Therefore, $\tau_t = 0$ for $t \in [s_0, s_1]$. Second, suppose there is a later time where capital taxes are positive, that is $s' \equiv \inf\{t > s_1 \,|\, \eta_t < 0\} < \infty$. Observe that, between $t = s_1$ and $t = s'$, both $u'(c_t)$ and λ_t grow at the common rate $\rho - r_t^*$, so $\lambda_{s'} = \Phi_u u'(c_{s'})$. For any $t > s'$, $u'(c_t)$ still grows at least as fast as λ_t, and, by definition of s', for a positive-measure set of times t after s', $u'(c_t)$ grows at the faster rate $\rho > \rho - r_t^*$ since $\eta_t < 0$ and $\tau_t = \bar{\tau}$ for those t. Therefore, for any $t > s'$, $\Phi_u u'(c_t) > \lambda_t$. By (57c), this means $\dot{\eta}_t < \eta_t \left(\rho + \frac{\rho}{\sigma}\right)$, or in other words, $\eta_t < 0$ and $c_t = c_{s'} e^{-\rho(t-s')/\sigma}$ for $t > s'$. Moreover, $\limsup_{t\to\infty} e^{-\rho t}\eta_t c_t < e^{-\rho s'}\eta_{s'} c_{s'} < 0$, contradicting Lemma 22. This concludes our proof of Lemma 23. $\square$

Lemma 24. *If $\tau_s = 0$ for $s \geq 0$, then $r_s^* \geq 0$ and $r_{s'}^* > 0$ for all $s' > s$.*

Proof. For convenience we introduce $R_t \equiv f_k(k_t, n_t)$. R_t has the following law of motion,

$$(\zeta + \alpha_t)\beta_t^{-1}\frac{\dot{R}_t}{R_t} = \rho + (1+\zeta)\delta + \zeta\frac{g + c_t}{k_t} - \zeta f(1, h(R_t)) - R_t,$$

which was obtained by log-differentiating the first order condition of labor (57a) and combining it with the resource constraint (55c). Here, $\alpha_t \equiv \frac{\partial \log f_n}{\partial(k_t/n_t)}$ as before, $\beta_t \equiv \frac{\partial \log f_k}{\partial(n_t/k_t)}$, and $h(x) \equiv f_k(1,\cdot)^{-1}(x)$. Observe that $h : \mathbb{R}_+ \to \mathbb{R}_+$ is strictly increasing and bijective. Since $\dot{R}$ depends implicitly (through α and β) and explicitly on k_t, R_t, and c_t, we also write $\dot{R}_t = \dot{R}(k, R, c)$.

Our proof of this lemma proceeds in two steps. First, we show an auxiliary result, namely that whenever

$$(k_t, R_t, c_t) \in \mathcal{A} \equiv \{(k, R, c) \,|\, R \leq \delta,\, \dot{R}(k, R, c) \leq 0\},$$

for some time $t = t_0$, then $(k_t, R_t, c_t) \in \mathcal{A}$ for all later times $t > t_0$ too. Second, we establish the result stated in the lemma.

First step. To prove the auxiliary result, it suffices to consider points (k_t, R_t, c_t) at the boundary of $\mathcal{A}$ and study whether the flows induced by the differential equation point to the inside of $\mathcal{A}$. There are two kinds of boundary points. If $R_t = \delta$, it trivially holds that $\frac{d}{dt}R_t \leq \frac{d}{dt}\delta = 0$. Suppose now that $\dot{R}(k_t, R_t, c_t) = 0$ and ask whether $\frac{d}{dt}\dot{R}(k_t, R_t, c_t) \leq 0$.

Generally, whenever $(k_t, R_t, c_t) \in \mathcal{A}$, it is straightforward to see that

$$\frac{\dot{k}}{k} = \frac{f(k,n)}{k} - \delta - \frac{g + c_t}{k_t} \geq \frac{\rho}{\zeta} > 0. \tag{74}$$

Moreover, $\dot{c}_t = -\frac{\rho}{\sigma} c_t$ since $r_t^* = R_t - \delta \leq 0$. The fact that k_t is increasing and c_t is decreasing mean that

$$(\zeta + \alpha_t)\beta_t^{-1} \frac{\frac{d}{dt}\dot{R}_t}{R_t} = \frac{d}{dt}(\zeta + \alpha_t)\beta_t^{-1} \frac{\dot{R}}{R} = \underbrace{\frac{d}{dt}\frac{g + c_t}{k_t}}_{<0} - \underbrace{\frac{d}{dt}\left(\zeta f(1, h(R)) + R\right)}_{=0} < 0$$

establishing the auxiliary result.

Second step. Suppose $\tau_s = 0$ for some $s \geq 0$. The fact that $r_s^* \geq 0$ follows directly from the constraint $(1 - \tau_s) r_s^* = r_s \geq 0$. Let $s' \equiv \inf\{t > s | r_t^* \leq 0\}$ and suppose $s' < \infty$. Since r_t^* is continuous and differentiable, this means that $r_{s'}^* = 0$ and $\frac{d}{dt} r_t^* |_{t=s'} \leq 0$, or in terms of R_t, $R_{s'} = \delta$ and $\dot{R}_{s'} \leq 0$. Applying the auxiliary result, $(k_t, R_t, c_t) \in \mathcal{A}$ for any $t > s'$. Moreover, $k_t \to \infty$ due to (74) at all times $t > s'$. This is in sharp contradiction to Lemma 16 (which applies here using $k_{s'}$ as initial capital stock since $\dot{c}_t = -\frac{\rho}{\sigma} c_t$ for all $t \geq s'$. Therefore $r_t^* > 0$ for all $t > s$. □

3^{rd} **step: Finite capital taxation $T < \infty$ in parts B (ii) and C.**

Lemma 25. *If either $\sigma < 1$ or $\sigma = 1$ and $b_0 < \bar{b}$, then $T < \infty$.*

Proof. If either $\sigma < 1$ or $\sigma = 1$ and $b_0 < \bar{b}$, then $\Phi_u > 0$ for any $\mu \geq 0$.[69] In the following, we prove that this is incompatible with $T = \infty$. By contradiction, suppose it were the case that there exists an optimal allocation $\{c_t, n_t, k_t, r_t\}$ with $T = \infty$, i.e. $c_t = c_0 e^{-\rho t/\sigma}$. Applying Lemma 22, $(k_t, n_t) \to (k^\infty, n^\infty)$. In particular, $r_t^* \to \rho > 0$ following the definition of (k^∞, n^∞) in Lemma 16. Now, $\Phi_u u'(c_t)$ grows at rate ρ while λ_t only grows at rate $\rho - r_t^* < \rho$. Therefore, there exists some finite time s such that $\lambda_t < \Phi_u u'(c_t)$ for all $t > s$. Using law of motion of η_t, (57c), this means $\dot{\eta}_t < \eta_t \left(\rho + \frac{\rho}{\sigma}\right)$ for all $t > s$ and so $\limsup_{t\to\infty} e^{-\rho t} \eta_t c_t < e^{-\rho s} \eta_s c_s < 0$, contradicting Lemma 22. □

Summary. This concludes our proofs of the bang bang property (Lemma 23) and parts B (ii) and C (Lemma 25).

J Proof of Proposition 8

We proceed by providing an explicit solution to the first order and transversality conditions to problem (55a) with zero government spending and certain combinations of k_0, b_0. We do so in two steps. First, taking k_0 as given, we find paths $\{c_t, n_t, k_t, r_t\}$, $\{\lambda_t, \eta_t\}$, μ and

[69] If $b_0 = \bar{b}$ and $\sigma < 1$, then as we explain in Section I, Φ_u can be taken to be $(1 - \sigma)$, and thus is positive here.

a level of initial debt b_0 which together satisfy all first order conditions, transversality conditions and constraints, with the one exception that η_t need not necessarily be negative. In a second step, we choose k_0 such that $\mu \geq 1/(\sigma-1)$ which will ensure that $\eta_t < 0$ at all times t.

The reason this construction is analytically tractable is that along the optimum, c_t, n_t, k_t will all fall to zero at the exact same growth rate, which needs to equal $\frac{\rho}{\sigma}$ by the Euler equation (55b). At the same time, $r_t = 0$ (since $T = \infty$). Taken together, to find the solution for a given k_0, it is necessary to find $c_0, n_0, \{\lambda_t, \eta_t\}, \mu, b_0$. Again, we use the previous notation $\Phi_u = 1 + \mu(1-\sigma)$ and $\Phi_v = 1 + \mu(1+\zeta)$.

First step. We conjecture that $c_t = c_0 e^{-\rho/\sigma t}$, $n_t = n_0 e^{-\rho/\sigma t}$, $k_t = k_0 e^{-\rho/\sigma t}$, $r_t = 0$, $\lambda_t = \lambda_0 e^{-\zeta\rho/\sigma t}$. The Euler equation (55b) obviously holds. The resource constraint (55c) is satisfied iff

$$c_0 = f(k_0, n_0) - \delta k_0 + \frac{\rho}{\sigma} k_0. \tag{75}$$

The IC constraint (55d) is satisfied iff

$$b_0 = c_0 \frac{\sigma}{\rho} - \frac{1}{\rho + (1+\zeta)\rho/\sigma} c_0^{\sigma} n_0^{1+\zeta} - k_0. \tag{76}$$

The first order condition for labor (57a) and the costate λ_t (57d) hold iff

$$f_k(k_0, n_0) = \zeta \frac{\rho}{\sigma} + \rho + \delta \tag{77}$$

and

$$\Phi_v n_0^{\zeta} = \lambda_0 f_n(k_0, n_0). \tag{78}$$

Given k_0, (77) pins down n_0, (75) c_0, and (76) b_0. The law of motion of η_t (57c) and the associated transversality condition (57g) are satisfied iff

$$\eta_t = -\frac{\lambda_0}{\rho + (1+\zeta)\rho/\sigma} e^{-\zeta\rho/\sigma t} + \frac{\sigma}{\rho} \Phi_u c_0^{-\sigma} e^{\rho t}. \tag{79}$$

Notice that (57b) holds, i.e.$\eta_t < 0$, as long as $\Phi_u \leq 0$, requiring $\mu \geq \frac{1}{\sigma-1}$. The transversality condition for capital (57f) obviously holds.

It remains to determine λ_0, η_0, and μ subject to (79) (at $t = 0$), $\mu \geq \frac{1}{\sigma-1}$, (78), and the first order condition for c_0 (57e). For expositional reasons, define the initial labor tax as $\tau_0^{\ell} \equiv 1 - n_0^{\zeta} c_0^{\sigma}/w^*$. Then, we can solve for μ as

$$\mu = \frac{\tau_0^{\ell} + \sigma + \zeta}{\sigma\left((1-\tau_0^{\ell})\frac{n_0}{c_0} w^* - 1\right) - \tau_0^{\ell}(1+\zeta)}. \tag{80}$$

Notice that whenever $\mu \in [\frac{1}{\sigma-1}, \infty)$, $\lambda_0 > 0$ is given by (78) and $\eta_0 < 0$ is given by (79) (at $t = 0$). So the last step in our construction is to determine whether there are levels for k_0 for which $\mu \in [\frac{1}{\sigma-1}, \infty)$.

Second step. The only object on the right hand side of (80) that depends on k_0 is τ_0^ℓ, and τ_0^ℓ is a strictly decreasing function of $k_0 \in [0, \infty)$, with $\tau_0^\ell \to 1$ as $k_0 \to 0$ and $\tau_0^\ell \to -\infty$ as $k_0 \to \infty$. Moreover, μ is increasing in $\tau_0^\ell \in (-\infty, 1]$ and has a pole at $\tau_{0,\text{pole}}^\ell = \frac{\sigma w^* n_0/c_0 - \sigma}{\sigma w^* n_0/c_0 + 1 + \zeta} < 1$, where it rises to $+\infty$ from the left. For $\tau_0^\ell = 1$, $\mu = -1 < 0$. We define $\underline{k}$ to be the value of k_0 corresponding to $\tau_{0,\text{pole}}^\ell$. Putting the mapping from k_0 to τ_0^ℓ and the one from τ_0^ℓ to μ together, we find a function $\mu(k_0)$ with the properties that

$$\begin{aligned} \mu(k_0) &< 0 && \text{for} \quad k_0 < \underline{k} \\ \mu(k_0) &\geq 1/(\sigma - 1) && \text{for} \quad k_0 \in (\underline{k}, \bar{k}] \\ \mu(k_0) &< 1/(\sigma - 1) && \text{for} \quad k_0 > \bar{k}, \end{aligned}$$

where $\bar{k} \equiv \inf_{k_0 \geq \underline{k}} \left\{ k_0 \geq \underline{k} \,|\, \mu(k_0) < \frac{1}{\sigma - 1} \right\} \in (\underline{k}, \infty]$. This proves that for $k_0 \in (\underline{k}, \bar{k}]$, there exists a debt level $b_0(k_0)$ for which the quantities c_t, n_t, k_t all fall to zero at equal rate $-\rho/\sigma$ and the sufficient optimality conditions of the problem are satisfied.

K Proof of Proposition 9

First, we show that the planner's problem is equivalent to (13). Then we show that the functions $\psi(T)$ and $\tau(T)$ are increasing, have $\psi(0) = \tau(0) = 0$ and bounded derivatives.

The planner's problem in this linear economy can be written using a present value resource constraint, that is,

$$\begin{aligned} \max \int_0^\infty & e^{-\rho t} \left(u(c_t) - v(n_t) \right) dt \\ \text{s.t. } \dot{c} &\geq c \frac{1}{\sigma} ((1 - \bar{\tau}) r^* - \rho) \\ \int_0^\infty & e^{-r^* t} (c_t - w^* n_t) dt + G = k_0 \\ \int_0^\infty & e^{-\rho t} \left[(1 - \sigma) u(c_t) - (1 + \zeta) v(n_t) \right] dt \geq u'(c_0) a_0, \end{aligned} \tag{81}$$

where $G = \int_0^\infty e^{-r^* t} g dt$ is the present value of government expenses, k_0 is the initial capital stock, a_0 is the representative agent's initial asset position, and per-period utility from consumption and disutility from work are given by $u(c_t) = c_t^{1-\sigma}/(1 - \sigma)$ and $v(n_t) = n_t^{1+\zeta}/(1 + \zeta)$. Note that we assumed $\sigma > 1$. The FOCs for labor imply that given n_0,

$$n_t = n_0 e^{-(r^* - \rho) t/\zeta}. \tag{82}$$

An analogous argument to the bang-bang result in Appendix **??** implies the existence of $T \in [0, \infty]$ such that $\tau_t = \bar{\tau}$ for $t \leq T$ and zero thereafter. In particular, the after-tax (net) interest rate will be $r_t = (1 - \bar{\tau}) r^* \equiv \bar{r}$ for $t \leq T$ and $r_t = r^*$ for $t > T$. Then, by the representative agent's Euler equation, the path for consumption is determined by

$$c_t = c_0 e^{-\frac{\rho-\bar{r}}{\sigma}t + \frac{r^*-\bar{r}}{\sigma}(t-T)^+}. \tag{83}$$

Substituting equations (82) and (83) into (81), the planner's problem simplifies to,

$$\begin{aligned} \max_{T,c_0,\bar{n}} \quad & \psi_1(T)u(c_0) - \psi_3 v(n_0) \\ \text{s.t. } & \psi_2(T)(\chi^*)^{-1}c_0 + G = k_0 + \psi_3 w^* n_0 \\ & \psi_1(T)u'(c_0)c_0 - \psi_3 v'(n_0)n_0 = \chi^* u'(c_0)a_0, \end{aligned} \tag{84}$$

where $\psi_1(T) = \frac{\chi^*}{\chi}\left(1-e^{-\chi T}\right) + e^{-\chi T}$, $\psi_2(T) = \frac{\chi^*}{\hat{\chi}}\left(1-e^{-\hat{\chi} T}\right) + e^{-\hat{\chi} T}$, $\psi_3 = \chi^*\left(r^* + \frac{r^*-\rho}{\zeta}\right)^{-1}$ and $\chi = \frac{\sigma-1}{\sigma}\bar{r} + \frac{\rho}{\sigma}$, $\chi^* = \frac{\sigma-1}{\sigma}r^* + \frac{\rho}{\sigma}$, $\hat{\chi} = r^* + \frac{\rho-\bar{r}}{\sigma}$. Notice that $\hat{\chi} > \chi^* > \chi$.

Now normalize consumption and labor

$$c \equiv \psi_1(T)^{1/(1-\sigma)}c_0/\chi^* \qquad\qquad n \equiv \psi_3^{1/(1+\zeta)} n_0/\left(\chi^*\right)^{(1-\sigma)/(1+\zeta)}$$

and define an efficiency cost $\psi(T) \equiv \psi_2(T)\psi_1(T)^{1/(\sigma-1)} - 1$, a capital levy $\tau(T) \equiv 1 - \psi_1(T)^{-\sigma/(\sigma-1)}$, and the present value of wage income $\omega n \equiv w^*\psi_3^{\zeta/(1+\zeta)} n$. Here, we note that by definition, ψ is bounded away from infinity and τ is bounded away from 1. Then, we can rewrite problem (84) as

$$\begin{aligned} \max_{T,c,n} \quad & u(c) - v(n) \\ \text{s.t. } & (1+\psi(T))c + G = k_0 + \omega n \\ & u'(c)c - v'(n)n = (1-\tau(T))u'(c)a_0, \end{aligned}$$

which is what we set out to show. Notice that $\psi_1(0) = \psi_2(0) = 1$ and so $\psi(0) = \tau(0) = 0$. Further, given our assumption that $\sigma > 1$, $\psi_1(T)$ and $\tau(T)$ are increasing in T. To show that $\psi'(T) \geq 0$, notice that, after some algebra,

$$\frac{d}{dT}\left(\psi_2\psi_1^{1/(\sigma-1)}\right) \geq 0 \quad \Leftrightarrow \quad \hat{\chi}\left(e^{\chi T} - 1\right) \leq \chi\left(e^{\hat{\chi} T} - 1\right),$$

which is true for any $T \geq 0$ because $\hat{\chi} > \chi$. Therefore, $\psi'(T) \geq 0$, with strict inequality for $T > 0$, implying that $\psi(T)$ is strictly increasing in T.

Now consider the ratio of derivatives,

$$\frac{\psi'(T)}{\tau'(T)} = \frac{1}{\sigma}\psi_2\psi_1^{(1+\sigma)/(\sigma-1)}\left((\sigma-1)\frac{\psi_2'}{\psi_2}\frac{\psi_1}{\psi_1'} + 1\right).$$

Notice that $\psi_1(T) \in [1, \chi^*/\chi]$ and $\psi_2(T) \in [\chi^*/\hat{\chi}, 1]$, so both are bounded away from infinity and zero. Further, the ratio ψ_2'/ψ_1' is also bounded away from infinity, $\psi_2'/\psi_1' = -\frac{1}{\sigma-1}e^{-(\hat{\chi}-\chi)T} \in [-1/(\sigma-1), 0]$, implying that $\psi'(T)/\tau'(T)$ is bounded away from ∞.

L Proof of Proposition 10

The planning problem is given by

$$\sup_{\{c_t, C_t, k_{t+1}\}} \sum_{t=0}^{\infty} \beta^t u(c_t) \tag{85}$$

$$c_t + C_t + k_{t+1} = f(k_t) + (1-\delta)k_t \tag{86}$$

$$\sum_{t=0}^{\infty} \beta^t C_t^{1-\sigma} = C_0^{-\sigma} a_0. \tag{87}$$

First, note that a_0 must be positive or else the IC constraint (87) cannot be satisfied (recall that $\sigma > 1$). Second, note that there exists a unique solution $C_0(\varphi)$ to the equation

$$C_0^{\sigma} \varphi_0 \varphi^{1-\sigma} + C_0 = a_0$$

for any $\varphi_0, \varphi > 0$, and that $C_0(\varphi) \to 0$ as $\varphi \to 0$. We now use this to construct a sequence of feasible paths $\{C_t^{(n)}\}_{t=0}^{\infty}$, $n = 0, 1, \ldots$, with $C_t^{(n)}$ uniformly converging to 0 as $n \to \infty$. Take any sequence $\{C_t^{(0)}\}_{t=0}^{\infty}$ that satisfies (87). Define

$$C_t^{(n)} = \begin{cases} C_0(\varphi^n) & t = 0 \\ \varphi^n C_t^{(0)} & t > 0 \end{cases}$$

for some $\varphi \in (0,1)$, $\varphi_0 \equiv C_0^{-\sigma}(a_0 - C_0)$. By construction, $C_t^{(n)} \to 0$ uniformly and the supremum in (85) approaches the maximum of the planning problem of a standard neo-classical growth model,

$$\max_{\{c_t, C_t, k_{t+1}\}} \sum_{t=0}^{\infty} \beta^t u(c_t)$$

$$c_t + k_{t+1} = f(k_t) + (1-\delta)k_t.$$

The way $\{C_t^{(n)}\}$ was constructed in this proof, it suggests an implementation via a wealth tax $\mathcal{T}_1 = R_1/R_1^* \to 100\%$. Analogous to this construction, a wealth tax approaching 100% in any period would implement the same allocation. This also shows that only a single period of unconstrained taxation is necessary to implement the supremum.

M Proof of Proposition 11

As in Section 2, labor supply is inelastic at $n_t = 1$. Denote the capitalist's initial wealth by $a_0 \equiv R_0 k_0 + R_0^b b_0$. The planning problem is then

$$\max_{\{c_t, C_t, k_{t+1}\}} \sum_{t=0}^{\infty} \beta^t u(c_t) \tag{88a}$$

$$C_{t+1} \geq C_t \beta^{1/\sigma} \tag{88b}$$

$$c_t + C_t + k_{t+1} = f(k_t) + (1-\delta)k_t \tag{88c}$$

$$\sum_{t=0}^{\infty} \beta^t U'(C_t) C_t = U'(C_0) a_0. \tag{88d}$$

The necessary first order conditions for C_t and c_t in problem (88a) are

$$\beta^{1/\sigma}\eta_t - \beta^{-1}\eta_{t-1} = \lambda_t - \Phi_u U'(C_t) \tag{89}$$

$$u'(c_t) = \lambda_t \tag{90}$$

$$\beta^{1/\sigma}\eta_0 = \lambda_0 - \Phi_u U'(C_0) - \mu\sigma C_0^{-\sigma-1} a_0 \tag{91}$$

where we defined $\Phi_u \equiv \mu(1-\sigma)$. Here, μ is the multiplier on the IC constraint (88d), λ_t is the multiplier of the resource constraint (88c)—which is positive by (90)—and η_t denotes the costate of capitalists' consumption C_t. If $\eta_t < 0$, constraint (88b) is binding. Also, it follows from (88d) that

$$\sigma C_0^{-\sigma-1} a_0 = \sigma C_0^{-1} \cdot U'(C_0) a_0 = \sigma C_0^{-1} \cdot \sum_{t=0}^{\infty} \beta^t U'(C_t) C_t > (\sigma - 1) U'(C_0),$$

where the inequality is obtained by dropping all terms with $t > 0$ from the infinite sum and observing that $\sigma > \sigma - 1$. Using this inequality, (91) implies that μ must be positive and $\Phi_u < 0$.

Suppose now there existed a period $T \geq 0$ where constraint (88b) is slack. In that case, $\eta_T = 0$ and (89) becomes for $t = T + 1$

$$\Phi_u U'(C_{T+1}) = \lambda_{T+1} - \beta^{1/\sigma}\eta_{T+1} > 0$$

contradicting $\Phi_u < 0$. Therefore, (88b) binds in all periods, or equivalently, $R_t = 1$ for all t.