

Positive Long-Run Capital Taxation: Chamley-Judd Revisited[†]

By LUDWIG STRAUB AND IVÁN WERNING*

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 nonconstant (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. (JEL H21, H25)

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 conclusions 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.

*Straub: Department of Economics, Harvard University (email: ludwigstraub@fas.harvard.edu); Werning: Department of Economics, Massachusetts Institute of Technology (email: iwerning@mit.edu). John Leahy was the coeditor for this article. This paper benefited from comments by Fernando Alvarez, V. V. Chari, Peter Diamond, Mikhail Golosov, Tom Phelan, Stefanie Stantcheva, 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. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

[†]Go to <https://doi.org/10.1257/aer.20150210> to visit the article page for additional materials and author disclosure statements.

¹To quote from a few examples, Judd (2002, p. 418): “[...] 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, Chari, and Kehoe (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.”

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, Tsyvinski, 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 1. 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

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

²Judd (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, Sections 4–5) allows workers to save, capitalists to work, and considers nonconstant discounting à 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 similar conclusions apply for these variant models in Judd (1985, Sections 4–5).

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

⁴Because 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).

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, $U(C) = C^{1-\sigma}/(1-\sigma)$, we establish that when the intertemporal elasticity of substitution (IES) is below 1, $\sigma > 1$, taxes rise and converge toward a positive limit tax, instead of declining toward 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 percent. 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 1, 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 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 toward zero.

When the IES is above 1, $\sigma < 1$, 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 percent. 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

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 (for surveys, see Brunnermeier, Eisenbach, and Sannikov 2013 and Gertler and Kiyotaki 2010).

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

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.

Returning to the case with additive utility, our second set of results shows 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 1 and debt is high enough. Importantly, the debt level required is below the peak of the 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 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 1, 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 obtain positive

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

⁸For results on capital taxation in overlapping generations (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, Kitao, and Krueger (2009).

long-run taxation by modifying the environment, moving away from the setups in Chamley (1986) and Judd (1985).

One exception is Lansing (1999), which considered a special case of the setup in Judd (1985, Section 3) with $\sigma = 1$ and found that positive long-run capital taxes are possible (see our discussion in Section I); Reinhorn (2019) further clarified the nature of this discrepancy with Judd (1985, Section 3). Bassetto and Benhabib (2006) studies capital taxation in a political 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 linear AK technology and $\sigma > 1$, they also provide a condition for the median voter to prefer indefinite capital taxation. The example in Lansing (1999) was viewed as a knife-edged case, applying only to $\sigma = 1$, while the example in Bassetto and Benhabib (2006) was obtained for a hybrid model that is not a special case of any economy explicitly treated in Judd (1985) or Chamley (1986).⁹ However, our results show that these previous examples were indicative of an unnoticed and more general problem with the zero long-run capital taxation prediction in the precise models of Judd (1985) and Chamley (1986).

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, Christiano, and Kehoe (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.

I. 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 $t = 0, 1, 2, \dots$ ¹⁰ 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

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

⁹Unlike Chamley (1986), their model features heterogeneity and inelastic labor supply; unlike Judd (1985), their model features no financial frictions, so there is no hand-to-mouth worker and the government can issue bonds.

¹⁰Judd (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.

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 $\lim_{c \rightarrow 0} u'(c) = \infty$.

Technology.—Output is obtained from capital and labor using a neoclassical constant returns production function $F(k_t, n_t)$ satisfying standard conditions.¹¹ 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 $g > 0$. 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

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

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

$$\max_{\{C_t, a_{t+1}\}} \sum_{t=0}^{\infty} \beta^t U(C_t),$$

subject to

$$C_t + a_{t+1} = R_t a_t \quad \text{and} \quad a_{t+1} \geq 0,$$

¹¹ We assume that F is increasing and strictly concave in each argument, continuously differentiable, and satisfying the standard Inada conditions $F_k(k, 1) \rightarrow \infty$ as $k \rightarrow 0$ and $F_k(k, 1) \rightarrow 0$ as $k \rightarrow \infty$. Moreover, we assume that capital is essential for production, that is, $F(0, n) = 0$ for all n .

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

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

are necessary and sufficient for optimality.

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.¹²

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 as¹³

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

subject to

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

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

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

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 toward 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.

¹²Equivalently, 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.

¹³Judd (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

$$(2a) \quad \mu_0 = 0,$$

$$(2b) \quad \lambda_t = u'(c_t),$$

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

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

where $\kappa_t \equiv k_t/C_{t-1}$, $v_t \equiv U'(C_t)/u'(c_t)$ and the multipliers on constraints (1b) and (1c) are $\beta^t \lambda_t$ and $\beta^t \mu_t$, respectively.¹⁴ Here, (2a) follows from the first-order condition with respect to C_{-1} .

A. Previous Steady-State Results

Judd (1985, p. 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: $T_t = 1 - R_t/R_t^* \rightarrow 0$.*

The proof is immediate: from equation (2d) we obtain $R_t^* \rightarrow 1/\beta$, while the capitalists' Euler equation requires that $R_t \rightarrow 1/\beta$. 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 IIC.

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

¹⁴We 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 $\mu_t(\beta U'(C_t)(C_t + k_{t+1}) - U'(C_{t-1})k_t)$.

THEOREM 2 (Lansing 1999, Reinhorn 2019): *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,*

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

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 toward 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 $R_t = 1/\beta$ at a steady state. Lansing (1999) first presented the logarithmic case as a counterexample to Judd (1985). Reinhorn (2019) correctly clarified that in the logarithmic case the Lagrange multipliers explode, explaining the difference in results.¹⁵

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 [...]” (p. 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” (p. 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.

B. 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 $s > 0$,

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

Although $s = \beta$ with logarithmic preferences, nothing we will derive depends on this fact, so we can interpret s as a free parameter that is potentially divorced from β .¹⁶

¹⁵Lansing (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 (2a)–(2d) are still satisfied by the optimal allocation for some sequence of multipliers, we believe the issue can be framed exactly as Reinhorn (2019) did, emphasizing the nonconvergence of multipliers.

¹⁶This 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 IC.

The planning problem becomes

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

subject to

$$c_t + \frac{1}{s} k_{t+1} + g = f(k_t) + (1 - \delta) k_t,$$

with k_0 given. This amounts to an optimal neoclassical growth problem, where the price of capital equals $1/s > 1$ 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.¹⁷

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

¹⁷This 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.

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

Toward a contradiction, suppose there existed an optimal allocation that converges to an interior steady state $k_t \rightarrow k$, $C_t \rightarrow C$, $c_t \rightarrow c$ with $k, C, c > 0$. This implies that κ_t and v_t also converge to positive values, κ and v . Moreover, the entire path $\{k_t, C_{t-1}, c_t\}$ must also be interior, such that the first-order conditions (2a)–(2d) necessarily hold at the optimum.¹⁸ 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

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

Now consider whether $\mu_t \rightarrow \mu < 0$ 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 $t = 0, 1, \dots$, a contradiction to $\mu_t \rightarrow \mu < 0$.¹⁹ 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 $k_t \rightarrow 0$, $C_t \rightarrow 0$, or $c_t \rightarrow 0$. With positive spending $g > 0$, $k_t \rightarrow 0$ is not feasible; this also rules out $C_t \rightarrow 0$, since capitalists cannot be starved while owning positive wealth.

Thus, provided the solution converges, $c_t \rightarrow 0$. This in turn implies that either $k_t \rightarrow k_g$ or $k_t \rightarrow k^s$ where $k_g < k^s$ are the two solutions to $(1/\beta)k + g = f(k) + (1 - \delta)k$, using the fact that (1c) implies $C = ((1 - \beta)/\beta)k$ at any steady state.²⁰ We next show that the solution does indeed converge, and that it does so toward 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 $\mu_t \rightarrow \infty$ and $c_t \rightarrow 0$, as argued above, but requires many other steps detailed in the online Appendix.²¹

¹⁸If at any date t one of k_t, C_{t-1} or c_t were zero, then that same variable must remain equal to zero thereafter: for k , see (1b); for C , see (1c); for c , see (2d). This contradicts the assumed convergence to an interior steady state.

¹⁹This argument did not require convexity of the planning problem (1a). It relied, instead, on the fact that the first-order conditions (2a)–(2d) are necessary for an interior allocation $\{k_t, C_{t-1}, c_t\}$.

²⁰Here we assume that government spending g is feasible, that is, $g < \max_k \{f(k) + (1 - \delta)k - (1/\beta)k\}$.

²¹This result also does not rely on convexity of the planning problem (1a). In the online Appendix, after dealing with boundaries explicitly, we only rely on the necessity of the first-order conditions (2a)–(2d).

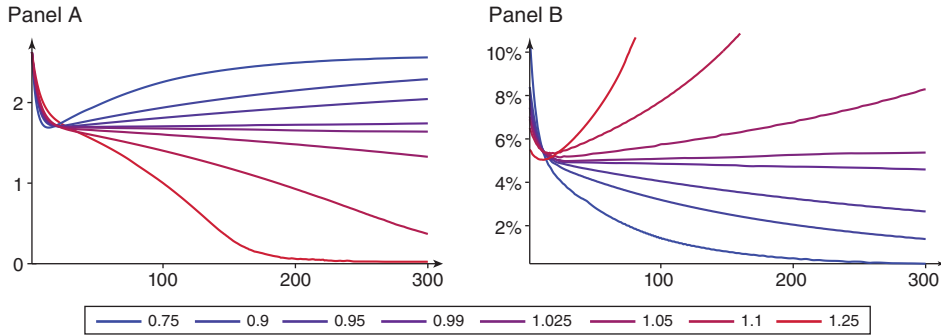


FIGURE 1. OPTIMAL TIME PATHS FOR CAPITAL (PANEL A) AND WEALTH TAXES (PANEL B)

Note: This figure shows the optimal time paths of capital k_t (panel A) and wealth taxes T_t (panel B) for various values of the inverse IES σ .

PROPOSITION 3: *If $\sigma > 1$ and $\gamma = 0$, any solution to the planning problem converges to $c_t \rightarrow 0$, $k_t \rightarrow k_g$, $C_t \rightarrow ((1 - \beta)/\beta)k_g$, with a positive limit tax on wealth: $T_t = 1 - R_t/R_t^* \rightarrow T^g > 0$. The limit tax T^g is decreasing in spending g , with $T^g \rightarrow 1$ as $g \rightarrow 0$.*

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, 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 (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, $\lim_{t \rightarrow \infty} c_t > 0$.

Solution for IES Near 1.—Figure 1 displays the time path for the capital stock and the tax rate on wealth, $T_t = 1 - R_t/R_t^*$, 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 $g/f(k) = 20\%$ 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 online Appendix.

To clarify the magnitudes of the tax on wealth, T_t , consider an example: if $R^* = 1.04$ so that the before-tax net return is 4 percent, then a tax on wealth of 1 percent represents a 25 percent tax on the net return; a wealth tax of 4 percent represents a tax rate of 100 percent on net returns, and so on.

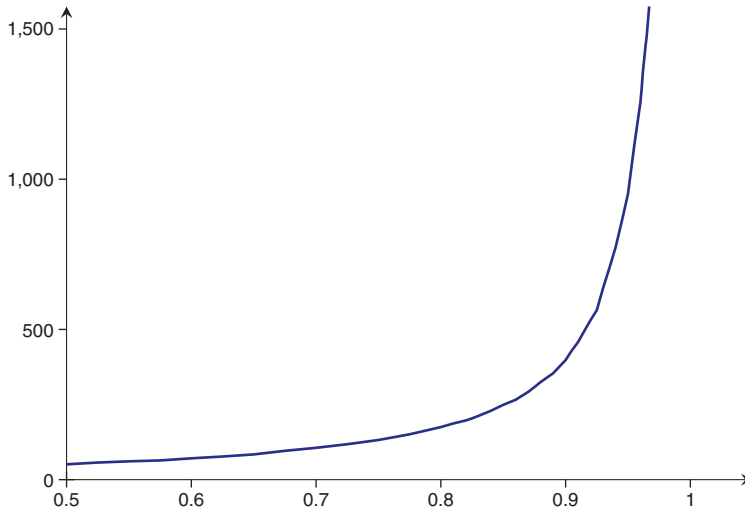


FIGURE 2. SLOW SPEED OF CONVERGENCE TO 0 TAXES FOR σ CLOSE TO, BUT BELOW 1

Note: This plot shows the time it takes until the wealth tax T_t falls below 1 percent for an inverse IES $\sigma \in (1/2, 1)$.

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 monotone²², taking around 200 years for $\sigma = 1.25$. The asymptotic tax rate is very high, approximately $T^s = 1 - R/R^* = 85\%$, lying outside the figure's 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 nonmonotonic²³ 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 $\sigma \rightarrow 1$, 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 percent as a function of $\sigma \in (1/2, 1)$. 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 $T = 1 - \beta = 5\%$. 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

²²This depends on the level of initial capital. For lower levels of capital the path first rises then falls.

²³This 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.

much more slowly either upward or downward, 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 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.²⁴ 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 $\sigma < 1$. 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.

²⁴It 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.

C. 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

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

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 $\beta RS_I \neq 1$, then the limit tax rate is positive and $\beta RS_I < 1$.*

This generalizes Proposition 2, since the case with isoelastic utility and IES less than 1 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 $\beta RS_I < 1$ 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, $RS_I < 1$ is natural, as it ensures local stability for capital given a fixed steady-state return, i.e., the dynamics implied by the recursion $k_{t+1} = S(Rk_t, R, R, \dots)$ 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,

$$(4) \quad \mathcal{T} = 1 - \frac{R}{R^*} = \frac{1 - \beta RS_I}{1 + \sum_{t=1}^{\infty} \beta^{-t+1} \epsilon_{S,t}},$$

where $\epsilon_{S,t} \equiv (R_t/S)(\partial S/\partial R_t)(R_0 k_0; R_1, R_2, \dots)$ 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, Section 3.3, equation (16)).²⁵

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 $T \rightarrow \infty$. For any finite T , the 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 $T \rightarrow \infty$. 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 1.

Second, the derivation we provide in the online 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.²⁶ 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 1: *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

²⁵Their 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.

²⁶Unfortunately, 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.

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 isoelastic 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 $\beta^{-t} \epsilon_{S,t} = -((\sigma - 1)/\sigma)((1 - \beta)/\beta)$ at a steady state and the sum of elasticities diverges. It equals $+\infty$ if the IES is greater than 1, or $-\infty$ if the IES is less than 1.²⁷ 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 1 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 savings function. 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.

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 percent, 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.

D. Redistribution toward Capitalists

In the present model, a desire to redistribute toward 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 toward 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 toward workers or capitalists at the zero-tax steady state (k^*, c^*, C^*) .²⁸ When $\sigma > 1$ and $\gamma > \gamma^*$ the solution converges to the highest sustainable

²⁷ Proposition 12 in the online Appendix shows that the infinite sum $\sum_{t=1}^{\infty} \beta^{-t+1} \epsilon_{S,t}$ also diverges for general recursive, non-additive preferences.

²⁸ 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 percent reduction in the consumption of capitalists, relative to the steady-state value $C = ((1 - \beta)/\beta)k$. The case $\gamma = \gamma^*$ corresponds to 0 in this transformation.

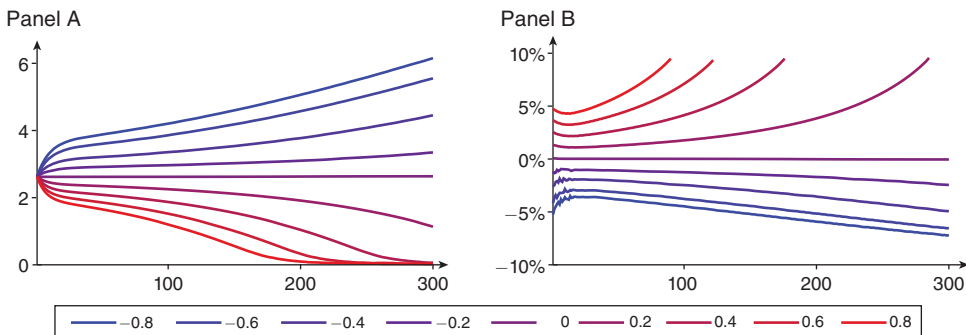


FIGURE 3. WEALTH TAXES DIVERGE AS LONG AS THE PLANNER HAS A DESIRE FOR REDISTRIBUTION

Note: This figure shows the optimal time paths over 300 years for the capital stock (panel A) and wealth taxes (panel B), for various redistribution preferences (zero represents no desire for redistribution; see footnote 28).

capital k^g , the highest solution to $(1/\beta)k + g = f(k) + (1 - \delta)k$, 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 (2a)–(2d). We do so for a continuous-time version of the model, to ensure that our results are comparable to Kemp, Long, and Shimomura (1993). The details are contained in the online Appendix. We obtain the following characterization.

PROPOSITION 5: For a continuous-time version of the model,

- (i) if $\sigma > 1$, the zero-tax steady state is locally saddle-path stable;
- (ii) if $\sigma < 1$ and $\gamma \leq \gamma^*$, the zero-tax steady state is locally stable;
- (iii) 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 toward 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, Long, and Shimomura (1993), which 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 toward 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, Long, and Shimomura (1993).

II. 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 IIA). We then prove, for additively separable preferences, that these bounds may, in fact, bind indefinitely (Section IIB). Readers mainly interested in the latter result may skip Section IIA.

A. 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.”²⁹ 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 $\mathcal{V}(U_0, U_1, \dots)$ 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

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

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

²⁹At 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.

$$(5c) \quad U_t = U(c_t, n_t).$$

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,

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

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 .³⁰

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

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

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

$$c_0 + a_1 \leq w_0 n_0 + R_0 k_0 + R_0^b b_0,$$

$$c_t + a_{t+1} \leq w_t n_t + R_t a_t, \quad t = 1, 2, \dots,$$

and the No Ponzi condition $a_{t+1}/(R_1 R_2 \cdots R_t) \rightarrow 0$. 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 in equilibrium for all $t = 1, 2, \dots$, 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

³⁰ A positive marginal change dU in the constant per-period utility stream increases steady-state utility by some constant dV . By virtue of (5a) this implies $dV = W_U dU + W_V dV$, which yields a contradiction unless $W_V < 1$.

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 $\tau_t \leq 1$.³¹ 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

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

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, Christiano, and Kehoe (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.³²

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

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

$$(8b) \quad R_t \geq 1.$$

The planning problem maximizes $\mathcal{V}(U_0, U_1, \dots)$ 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.

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 $\Lambda_t \rightarrow \Lambda > 0$. Then the tax on capital converges to zero, $R_t/R_t^* \rightarrow 1$.*

³¹ When $R^* - 1$ is negative, however, an upper bound directly imposed on taxes τ_t allows arbitrarily low after-tax interest rates R_t .

³² The argument is identical to that in Lucas and Stokey (1983) and Chari, Christiano, and Kehoe (1994).

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 $\tilde{\Lambda}_t = \Lambda_{t+1} R_{t+1}^*$. 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).³³

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

- (i) *private wealth converges to zero, $a_t \rightarrow 0$; or*
- (ii) *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

³³Note 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 Section IIC.

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 additively 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-edge situation.

B. 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, Chari, and Kehoe (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,³⁴

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

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

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

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

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

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

³⁴Continuous 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.

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, $a_0 > 0$.³⁵ To enforce bounds on capital taxation we follow Chamley (1986) and impose

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

$$(12b) \quad r_t \geq 0,$$

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, p. 615) formulated the following claim regarding the path for capital tax rates.³⁶

CLAIM 1: *There exists a time T with the following three properties:*

- (i) *For $t < T$, the constraint (12b) is binding, that is, $r_t = 0$ and $\tau_t = 1$;*
- (ii) *For $t > T$ capital income is untaxed, that is, $r_t = r_t^*$ and $\tau_t = 0$;*
- (iii) *$T < \infty$.*

At a crucial juncture in the proof of this claim, Chamley (1986, p. 616) states in support of part (iii) that “The constraint $r_t \geq 0$ cannot be binding forever (the marginal utility of private consumption [...] would grow to infinity [...] which is absurd).”³⁷ Our next result shows that there is nothing absurd about this within the logic of the model and that, quite to the contrary, part (iii) 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 though not doing so is feasible.

³⁵ 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.

³⁶ Similar claims are made in Atkeson, Chari, and Kehoe (1999), Judd (1999), and many other papers.

³⁷ It is worth pointing out, however, that although Chamley (1986, p. 619) 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.”

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:

- (i) *For $\sigma > 1$, there exists a lower bound on debt $\underline{b} < \bar{b}$, such that:*
 - *If $b_0 = \bar{b}$, the unique optimum has $T = \infty$ and there is no feasible allocation with $T < \infty$.*
 - *If $b_0 \in [\underline{b}, \bar{b})$, the unique optimum has $T = \infty$ but there exist feasible allocations with $T < \infty$.*
 - *If $b_0 < \underline{b}$, any optimum has $T < \infty$.*
- (ii) *For $\sigma = 1$:*
 - *If $b_0 = \bar{b}$, the unique optimum has $T = \infty$ and there is no feasible allocation with $T < \infty$.*
 - *If $b_0 < \bar{b}$, any optimum has $T < \infty$.*
- (iii) *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, Chari, and Kehoe 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 (i) 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 nonnegative. 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

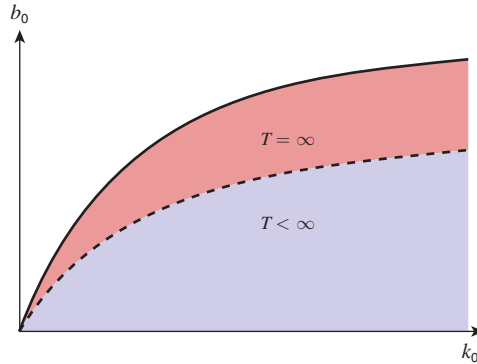


FIGURE 4. GRAPHICAL REPRESENTATION OF THE CASE $\sigma > 1$ IN PROPOSITION 7

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 $c_t, k_t, n_t \rightarrow 0$ 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 $r_t \geq r_t^*(1 - \bar{\tau})$.

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.

C. 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, Manuelli, and Rossi (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 IIB. 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 costate 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 $r_t - r_t^* \rightarrow 0$. Thus, the Chamley (1986) steady-state result actually follows by postulating the convergence of Λ_t , without assuming convergence of the allocation. Judd (1999, p. 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) costate 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,

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

In particular, under the conditions of this theorem, the optimum cannot converge to a steady state with a positive tax on capital.³⁸ 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 .³⁹ 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,

$$MRS_{t,t+s}^{Social} = e^{-\rho s} \frac{\hat{\Lambda}_{t+s}}{\hat{\Lambda}_t} = e^{-\int_0^s r_{t+\bar{s}}^* d\bar{s}}$$

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

$$MRS_{t,t+s}^{Private} = e^{-\rho s} \frac{\theta_{t+s}}{\theta_t} = e^{-\int_0^s r_{t+\bar{s}} d\bar{s}},$$

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

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

This expression 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 $\Lambda_{t+s}/\Lambda_t = e^{\int_0^s (r_{t+\bar{s}} - r_{t+\bar{s}}^*) d\bar{s}}$ is the only source of nonzero taxes. Whenever Λ_t is constant, social and private marginal rate of substitution (MRS) values 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.

³⁸The result is somewhat sensitive to the assumption that $\bar{\tau} = 1$; when $\bar{\tau} \neq 1$ and technology is nonlinear, the costate equation acquires other terms, associated with the bounds on capital taxation.

³⁹In 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.

COROLLARY 2: *At the optimum described in Proposition 7 we have that $\Lambda_t \rightarrow 0$ as $t \rightarrow \infty$. Thus, in 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 $T \rightarrow \infty$ 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 $T \rightarrow \infty$. 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 the following static problem:*

$$(13) \quad \max_{T, c, n} u(c) - v(n),$$

subject to

$$\begin{aligned} (1 + \psi(T))c + G &= k_0 + \omega n, \\ u'(c)c - v'(n)n &= (1 - \tau(T))u'(c)a_0, \end{aligned}$$

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 nondistortionary 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 trade-off between costs and benefits is bounded, $d\psi/d\tau < \infty$, even as $T \rightarrow \infty$. 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.

III. 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 I 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 $\beta^t U'(C_t)(k_{t+1} + b_{t+1}) \rightarrow 0$. 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

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

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 $t = 0, 1, \dots$, 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 $C_t \rightarrow 0$. 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 percent. 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 I 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 $\mathcal{T}_t \rightarrow 100\%$ in some period t and $C_t \rightarrow 0$ for all $t = 0, 1, \dots$.*

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 IIB.

PROPOSITION 11: *Consider the two-class model from Section I 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. We have also maintained the assumption from Judd (1985) that workers do not save. In a political economy context, Bassetto and Benhabib (2006) studies a situation where all agents save (in our context, both workers and capitalists) and are taxed linearly at the same rate. Indeed, they report the possibility that indefinite taxation is optimal for the median voter.

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.

IV. 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 1, the long-run tax on capital can actually be positive. Empirically, an IES below 1 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.⁴⁰ In this paper, we refrain from 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. 2007. "Optimal Capital Income Taxation." NBER Working Paper 13354.
- Aiyagari, S. Rao. 1995. "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting." *Journal of Political Economy* 103 (6): 1158–75.
- Atkeson, Andrew, V. V. Chari, and Patrick J. Kehoe. 1999. "Taxing Capital Income: A Bad Idea." *Federal Reserve Bank of Minneapolis Quarterly Review* 23 (3): 3–18.
- Atkinson, A. B., and J. E. Stiglitz. 1972. "The Structure of Indirect Taxation and Economic Efficiency." *Journal of Public Economics* 1 (1): 97–119.
- Banks, James, and Peter Diamond. 2010. "The Base for Direct Taxation." In *Dimensions of Tax Design: The Mirrlees Review*, edited by Stuart Adam et al., 548–648. Oxford, UK: Oxford University Press.
- Bassetto, Marco, and Jess Benhabib. 2006. "Redistribution, Taxes, and the Median Voter." *Review of Economic Dynamics* 9 (2): 211–23.
- Brunnermeier, Markus K., Thomas M. Eisenbach, and Yuliy Sannikov. 2013. "Macroeconomics with Financial Frictions: A Survey." In *Advances in Economics and Econometrics: Tenth World Con-*

⁴⁰ 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 never taxing capital versus always taxing capital at a fixed rate.

- gress, edited by Daron Acemoglu, Manuel Arellano, and Eddie Dekel, 3–94. Cambridge, UK: Cambridge University Press.
- Chamley, Christophe.** 1980. “Optimal Intertemporal Taxation and the Public Debt.” Yale University Cowles Foundation for Research in Economics Discussion Paper 554.
- Chamley, Christophe.** 1986. “Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives.” *Econometrica* 54 (3): 607–22.
- Chari, V. V., Lawrence J. Christiano, and Patrick J. Kehoe.** 1994. “Optimal Fiscal Policy in a Business Cycle Model.” *Journal of Political Economy* 102 (4): 617–52.
- Chari, V. V., and Patrick J. Kehoe.** 1990. “Sustainable Plans.” *Journal of Political Economy* 98 (4): 783–802.
- Chari, V. V., and Patrick J. Kehoe.** 1999. “Optimal Fiscal and Monetary Policy.” In *Handbook of Macroeconomics*, Vol. 1C, edited by John B. Taylor and Michael Woodford, 1671–1745. Amsterdam: North-Holland.
- Coleman, Wilbur John, II.** 2000. “Welfare and Optimum Dynamic Taxation of Consumption and Income.” *Journal of Public Economics* 76 (1): 1–39.
- Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger.** 2009. “Taxing Capital? Not a Bad Idea after All!” *American Economic Review* 99 (1): 25–48.
- Diamond, Peter A., and James A. Mirrlees.** 1971. “Optimal Taxation and Public Production: I: Production Efficiency.” *American Economic Review* 61 (1): 8–27.
- Erosa, Andrés, and Martin Gervais.** 2002. “Optimal Taxation in Life-Cycle Economies.” *Journal of Economic Theory* 105 (2): 338–69.
- Farhi, Emmanuel, Christopher Sleet, Iván Werning, and Sevin Yeltekin.** 2012. “Nonlinear Capital Taxation without Commitment.” *Review of Economic Studies* 79 (4): 1469–93.
- Farhi, Emmanuel, and Iván Werning.** 2010. “Progressive Estate Taxation.” *Quarterly Journal of Economics* 125 (2): 635–73.
- Farhi, Emmanuel, and Iván Werning.** 2013. “Estate Taxation with Altruism Heterogeneity.” *American Economic Review* 103 (3): 489–95.
- Gertler, Mark, and Nobuhiro Kiyotaki.** 2010. “Financial Intermediation and Credit Policy in Business Cycle Analysis.” In *Handbook of Monetary Economics*. Vol. 3, edited by Benjamin M. Friedman and Michael Woodford, 547–99. Amsterdam: Elsevier.
- Golosov, Mikhail, Aleh Tsyvinski, and Iván Werning.** 2006. “New Dynamic Public Finance: A User’s Guide.” In *NBER Macroeconomics Annual*, Vol. 21, edited by Daron Acemoglu, Kenneth Rogoff, and Michael Woodford, 317–88. Cambridge, MA: MIT Press.
- Greulich, Katharina, Sarolta Laczó, and Albert Marcet.** 2016. “Pareto-Improving Optimal Capital and Labor Taxes.” Unpublished.
- Jones, Larry E., Rodolfo E. Manuelli, and Peter E. Rossi.** 1993. “Optimal Taxation in Models of Endogenous Growth.” *Journal of Political Economy* 101 (3): 485–517.
- Jones, Larry E., Rodolfo E. Manuelli, and Peter E. Rossi.** 1997. “On the Optimal Taxation of Capital Income.” *Journal of Economic Theory* 73 (1): 93–117.
- Judd, Kenneth L.** 1985. “Redistributive Taxation in a Simple Perfect Foresight Model.” *Journal of Public Economics* 28 (1): 59–83.
- Judd, Kenneth L.** 1993. “Optimal Taxation in Dynamic Stochastic Economies: Theory and Evidence.” Unpublished.
- Judd, Kenneth L.** 1999. “Optimal Taxation and Spending in General Competitive Growth Models.” *Journal of Public Economics* 71 (1): 1–26.
- Judd, Kenneth L.** 2002. “Capital-Income Taxation with Imperfect Competition.” *American Economic Review* 92 (2): 417–21.
- Kemp, Murray C., Ngo Van Long, and Koji Shimomura.** 1993. “Cyclical and Noncyclical Redistributive Taxation.” *International Economic Review* 34 (2): 415–29.
- Koopmans, Tjalling C.** 1960. “Stationary Ordinal Utility and Impatience.” *Econometrica* 28 (2): 287–309.
- Lansing, Kevin J.** 1999. “Optimal Redistributive Capital Taxation in a Neoclassical Growth Model.” *Journal of Public Economics* 73 (3): 423–53.
- Lucas, Robert E., Jr.** 1990. “Supply-Side Economics: An Analytical Review.” *Oxford Economic Papers* 42 (2): 293–316.
- Lucas, Robert E., Jr., and Nancy L. Stokey.** 1983. “Optimal Fiscal and Monetary Policy in an Economy without Capital.” *Journal of Monetary Economics* 12 (1): 55–93.
- Lucas, Robert E., Jr., and Nancy L. Stokey.** 1984. “Optimal Growth with Many Consumers.” *Journal of Economic Theory* 32 (1): 139–71.
- Phelan, Christopher, and Ennio Stacchetti.** 2001. “Sequential Equilibria in a Ramsey Tax Model.” *Econometrica* 69 (6): 1491–1518.

- Piketty, Thomas, and Emmanuel Saez.** 2013. "A Theory of Optimal Inheritance Taxation." *Econometrica* 81 (5): 1851–86.
- Reinhorn, Leslie J.** 2019. "On Optimal Redistributive Capital Taxation." *Journal of Public Economic Theory* 21 (3): 460–87.
- Saez, Emmanuel.** 2013. "Optimal Progressive Capital Income Taxes in the Infinite Horizon Model." *Journal of Public Economics* 97: 61–74.
- Sargent, Thomas J., and Lars Ljungqvist.** 2004. *Recursive Macroeconomic Theory*, Vol. 2. Cambridge, MA: MIT Press.
- Stokey, Nancy L.** 1991. "Credible Public Policy." *Journal of Economic Dynamics and Control* 15 (4): 627–56.
- Uzawa, Hirofumi.** 1968. "Time Preference, the Consumption Function, and Optimum Asset Holdings." *Value, Capital, and Growth: Papers in Honour of Sir John Hicks*, edited by James N. Wolfe, 485–504. Edinburgh: University of Edinburgh Press.
- Werning, Iván.** 2007. "Optimal Fiscal Policy with Redistribution." *Quarterly Journal of Economics* 122 (3): 925–67.
- Zhu, Xiaodong.** 1992. "Optimal Fiscal Policy in a Stochastic Growth Model." *Journal of Economic Theory* 58 (2): 250–89.

This article has been cited by:

1. David R. Agrawal, William H. Hoyt, John D. Wilson. 2022. Local Policy Choice: Theory and Empirics. *Journal of Economic Literature* **60**:4, 1378-1455. [[Abstract](#)] [[View PDF article](#)] [[PDF with links](#)]
2. Louis Kaplow. 2021. Market Power and Income Taxation. *American Economic Journal: Economic Policy* **13**:4, 329-354. [[Abstract](#)] [[View PDF article](#)] [[PDF with links](#)]
3. Jess Benhabib, Bálint Szoóke. 2021. Optimal Positive Capital Taxes at Interior Steady States. *American Economic Journal: Macroeconomics* **13**:1, 114-150. [[Abstract](#)] [[View PDF article](#)] [[PDF with links](#)]