On optimal redistributive capital taxation

Leslie J. Reinhorn
Economics Department
Business School
University of Durham
Mill Hill Lane
Durham DH1 3LB
United Kingdom

phone +44 191 334 6365
reinhorn@hotmail.com

6 December 2018

Forthcoming, *Journal of Public Economic Theory*
DOI: 10.1111/jpet.12354

I thank Luca Spataro for introducing me to Straub and Werning (2018). I thank the editors, a referee, and Iván Werning for their constructive input.

Abstract

This paper addresses conflicting results regarding the optimal taxation of capital income. Judd (1985) proves that in steady state there should be no taxation of capital income. Lansing (1999) studies a logarithmic example of one of Judd's models and finds that the optimal steady state tax on capital income is not always zero — it is positive in some specifications, negative in some others. There appears to be a contradiction. However, I show that Lansing derives his result by relaxing the convergence hypotheses of Judd's theorem. With less restrictive hypotheses, a wider range of primitives (parameter values, initial condition, etc) satisfy the hypotheses and since each specification of primitives generates its own optimal time path(s) for the model's variables, it follows that a wider range of time paths with a wider range of steady state properties is possible. This raises a question. What happens if the convergence hypotheses are weakened further so that they are satisfied by a wider yet range of primitives? I find that at any interior steady state for the model's optimal tax equilibrium, either the capital tax is zero or else the elasticity of marginal utility is unitary which is satisfied identically in Lansing's log example. In effect, Lansing's example illustrates the only way in which an interior steady state can violate the zero tax result.

JEL Classification: H2

Keywords: dynamic optimal taxation

1 Introduction

Chamley (1986) and Judd (1985) prove that capital income should not be taxed in a steady state. Lansing (1999) provides a counterexample to this result. The example is particularly intriguing since it is a special case of one of Judd's models. There appears to be a contradiction. Lansing offers explanations to reconcile the differences. He also considers extensions of the model that revive the zero tax result. However, one is still left wondering what goes wrong in the counterexample. Lansing states on page 449, "Future research should be directed at developing a solution method that gives the right answer in all cases." Judd's solution method is optimal control theory (as is Lansing's). It would be very troubling indeed if optimal control theory failed to give the right answer. Fortunately, the contradiction can be resolved: Judd and Lansing have proved two different theorems with two different sets of hypotheses. For the special case considered by Lansing in which the "capitalist" has logarithmic utility, his theorem's hypotheses are less restrictive than Judd's so the range of primitives (parameter values, initial condition, "worker's" utility function, production function) that satisfy his hypotheses is wider. Since each specification of the primitives generates its own optimal time path(s) for the model's variables,[1] Lansing's less restrictive hypotheses allow for a wider range of time paths with a wider range of steady state characteristics. In particular, convergence to a zero capital tax is one possible characteristic, but not the only one.

The hypotheses in question deal with the convergence properties of various co-state variables (Lagrange multipliers). Kemp, Long, and Shimomura (1993) have also observed that the convergence hypotheses of Judd's theorem might not be satisfied. Among the possibilities is that the steady state of the dynamical system could be completely unstable in which case the zero capital tax result may not apply. In Lansing's log example it turns out there is a somewhat different reason why Judd's convergence hypotheses might not be satisfied. The issue is not the local dynamics about the dynamical system's steady state, but rather the very existence of a steady state. Further work regarding the convergence properties has been done by Straub and Werning (2018). They state, "Reinhorn ... correctly clarified that in the logarithmic case the Lagrange multipliers explode, explaining the difference in results" between Judd (1985) and Lansing (1999). Straub and Werning (2018) also state, "[W]e believe the issue can be framed exactly as Reinhorn ... did, emphasizing the non convergence of multipliers."

A main reason why it is difficult to characterize steady states in Judd's optimal tax model is because an explicit solution for the economy's optimal time path(s) is generally not possible. If we had an explicit solution for every possible specification of the primitives, it would be relatively straightforward to identify which primitives generate paths that satisfy the convergence hypotheses and in these cases to identify the limiting tax rate. But this is generally not possible, so the few cases that do admit an explicit solution

[1] For given primitives, the solution to the optimal tax problem might not be unique hence there might be multiple optimal time paths.

are valuable study tools. Example 3.5 presents such a case, one in which the optimal tax equilibrium is time invariant: the tax rate, capital stock, and consumption levels stay forever at their time 0 values. The example has a non-zero optimal capital tax. It satisfies all hypotheses of Lansing's theorem and in particular in Lansing's formulation of the problem the co-state/multiplier is time invariant. However, in Judd's formulation of the problem (which is the exact same optimal tax problem but a different formulation with different co-states), the co-states explode linearly to $\pm\infty$ in the infinite future, a violation of Judd's convergence hypotheses. So Judd's theorem does not apply and this explains how it is possible to have an optimal tax equilibrium that violates Judd's conclusion of a zero capital tax in the limit. It also suggests the need for theorems that do not make assumptions about the behavior of co-state variables.

More specifically, since the co-state variables are unobservable shadow prices that are not part of the economy's equilibrium (and they do not appear in the statement of the optimal taxation problem), one would rather not make assumptions about their behavior. On the other hand, it is quite reasonable to assume that observable macroeconomic variables have stable long run behavior since this is consistent with most developed economies. (E.g., page 304 of Lucas 1990 for the US.) In the case of Judd's model, which abstracts from demographics and technological change, stability boils down to convergence to an interior steady state. Thus, in theorem 3.6 I study the behavior of the optimal tax on capital income, assuming only that the observable macro variables converge to positive limits, with no assumptions about co-states.[2] I find that the following must hold: either the modified golden rule is satisfied in the limit or else savings are insensitive to the after-tax interest rate in the limit. In the former case we get Judd's zero tax result. In the latter case, the income and substitution effects of an interest rate change just cancel, and this is what occurs in Lansing's example with logarithmic utility. If interest does not affect savings, this undermines the benefit from a zero tax on interest/capital income and suggests why Judd's result does not necessarily hold in this case.

Straub and Werning (2018) raise serious concerns about Judd's convergence hypotheses in the case where the capitalist in the model has utility with a constant elasticity of intertemporal substitution (EIS). In particular, when the capitalist's EIS is less than one, the solution to the optimal tax problem cannot converge to an interior steady state. If, in addition, the social welfare function places zero weight on the capitalist and all weight on the worker, then the solution to the optimal tax problem does converge, but to a non-interior steady state with a positive tax rate on capital income. Straub and Werning (2018) conclude that Judd's model cannot be used to unequivocally justify a zero long run tax on capital income. I agree with Straub and Werning (2018). But since the constant EIS case with elasticity less than one leads to a

[2] Throughout the paper, the theorems' hypotheses will be stated in terms of the convergence of endogenous variables. The theorems do not characterize the primitives that satisfy convergence. Some primitives will satisfy Judd's hypotheses, some will satisfy Lansing's, and some neither. However, as discussed above, it seems reasonable to focus on those primitives that lead to stable long run behavior and to reject those that lead to unstable long run behavior (or that lead to corners) since this is inconsistent with the macroeconomic facts.

non-interior steady state, and since this is inconsistent with stable long run behavior, I prefer to exclude these utility functions from consideration and instead focus on utility functions (and other primitives) that do lead to stable long run behavior.

In hindsight, it may seem clear that if Judd finds that the optimal tax rate must converge to zero, while Lansing finds that it might converge to zero (part (ii) of Lansing's proposition 2) but might instead converge to a non-zero limit (parts (i) and (iii)), then Lansing must have relaxed Judd's hypotheses. However, there is another possible explanation, one that is incorrect: in the log case that Lansing considers, Judd's analysis might break down and give the wrong solution to the optimal tax problem. On page 423 Lansing states, "I show that the standard approach [Judd's] to solving the dynamic optimal tax problem yields the wrong answer in this (knife-edge) case [log] because it fails to properly enforce the constraints associated with the competitive equilibrium." On page 438 he adds, "When applying the standard approach in Judd's model, the allocations for k [capital] and c [capitalist's consumption] are assumed to be independent for all [values of the capitalist's constant EIS]. However, as [the EIS converges to 1], this assumption breaks down because the competitive equilibrium requires $c = \rho k$ [where ρ is a preference parameter]. By continuing to treat k and c as independent, the standard approach actually lets in an additional policy instrument through the back door." The error here is the statement that the standard approach treats k and c as independent in the log case; in fact, the standard/Judd approach actually fully accounts for the restriction $c = \rho k$. In particular, Judd's approach constrains the optimal tax problem to satisfy two differential equations, the capitalist's consumption Euler equation and the equilibrium capital accumulation equation, as well as two boundary conditions. I show below that under log utility, these constraints necessarily imply the restriction $c = \rho k$; hence, when these constraints are imposed on the problem then so too is the restriction $c = \rho k$. Pontryagin's theorem then ensures that the necessary conditions for optimality fully reflect this restriction. Appendix D provides a careful derivation of the necessary conditions and confirms the validity of Judd's approach. Furthermore, and crucially, lemma 3.2 below proves that in the log case Judd's approach is completely equivalent to Lansing's — their first order conditions are mathematically equivalent: any solution to Judd's is also a solution to Lansing's and vice versa.[3] Judd's analysis is correct and so is Lansing's. The differences in their conclusions follow from the differences in their hypotheses.

Section 2 presents the model. Section 3 presents the theorems of Judd and Lansing, explains the rela-

[3] With regard to Lansing's "back door" statement, he correctly shows that in the log case Judd's steady state equations imply equality of the welfare-weighted marginal utilities of the model's two consumers (Lansing's equation (25) on page 438). This is a first best condition but it does not imply the implicit existence of some additional policy instrument. If it did, then Lansing's own back door is just as open as Judd's: this exact same condition (25) appears in part (ii) of Lansing's proposition 2. If Judd's result in the log case implicitly uses an additional policy instrument then so too does Lansing's result in part (ii). But in fact neither of them does so. Rather, in the log case Judd's convergence hypotheses are so demanding that generically they cannot be satisfied. See the discussion following equation (4) below. So Lansing's (25) does not indicate the implicit presence of an additional policy instrument; it indicates a fluke. Analogously, under particular conditions the entire first-best time path that maximizes the government's social welfare can be decentralized as an equilibrium, and without any additional policy instruments. See appendix B below. But this too would be a fluke.

tionship between these two theorems, and also provides the general result described above. Section 4 offers a concluding comment.

2 Model

The model has four economic actors: capitalist, worker, firm, government. The capitalist has access to the capital market but does no work. The worker supplies labor inelastically but has no access to the capital market. The firm is a price taking profit maximizer that uses capital and labor to produce output. The government chooses a time path for the tax rate on capital income and uses the proceeds to provide lump sum transfers to the worker. There is no government debt. Hence the transfers must equal the taxes at each point in time. We now proceed to describe the model in detail.

The capitalist has an infinite horizon and maximizes discounted utility, $\int_0^\infty e^{-\rho t}u(c_t^c)dt$, where $\rho > 0$ is the subjective discount rate and $c_t^c \geq 0$ is instantaneous consumption. The superscript identifies the capitalist; c_t^w will be the worker's consumption. The instantaneous utility function u is smooth, strictly increasing, strictly concave, and satisfies Inada conditions. At the beginning of time the capitalist's wealth consists of the economy's entire stock of capital, $k_0 > 0$. This stock of wealth/capital evolves through time according to the capital accumulation equation: $\dot{k}_t = (1 - \tau_{kt})(r_t - \delta)k_t - c_t^c$ where τ_{kt} is the tax rate on net capital income (subsidy rate if negative), r_t is the pre-tax interest rate gross of depreciation, and δ is the depreciation rate. Note the lack of wage income which reflects the assumption that the capitalist supplies no labor. For ease of notation, let $\bar{r}_t := (1 - \tau_{kt})(r_t - \delta)$ denote the after tax, net of depreciation, interest rate. Then the capital accumulation equation is $\dot{k}_t = \bar{r}_t k_t - c_t^c$. Let $\bar{R}_t := \int_0^t \bar{r}_s ds$ be the cumulative interest factor. With this definition we can integrate the capital accumulation equation to get $e^{-\bar{R}_T}k_T - k_0 = -\int_0^T e^{-\bar{R}_t}c_t^c dt$. When $T \to \infty$ this equation gives the capitalist's lifetime budget. In order to prevent Ponzi schemes we will require that the present value of wealth be non-negative in the limit: $\lim_{T\to\infty} e^{-\bar{R}_T}k_T \geq 0$. Then the lifetime present value budget constraint is $\int_0^\infty e^{-\bar{R}_t}c_t^c dt \leq k_0$. The capitalist maximizes lifetime utility subject to this budget.[4] At the solution, the intertemporal marginal rate of substitution must equal the ratio of present value prices, and the budget must hold with equality:

$$e^{-\rho t}u'(c_t^c)/u'(c_0^c) = e^{-\bar{R}_t} \quad \text{and} \quad \int_0^\infty e^{-\bar{R}_t}c_t^c dt = k_0. \tag{1}$$

Equivalently, the first of these conditions can be log differentiated to give the capitalist's consumption Euler equation $\dot{c}_t^c u''(c_t^c)/u'(c_t^c) = \rho - \bar{r}_t$. The second equation in (1) can be expressed in its no-Ponzi form as $\lim_{t\to\infty} e^{-\bar{R}_t}k_t = 0$, or, by the first equation in (1), $\lim_{t\to\infty} e^{-\rho t}u'(c_t^c)k_t = 0$.

The worker inelastically supplies a flow of one unit of labor and immediately consumes all wages and transfers due to the lack of access to the capital market. So the worker is a passive actor who makes

[4] Throughout, control variables in optimization problems are required to be piecewise continuous functions of t. This includes $\bar{r}_t$ since it is the control for the optimal taxation problem in section 3 below.

no decisions. The instantaneous utility function is $v(c_t^w)$. The worker's consumption (and income) is $c_t^w = w_t + TR_t$ where w_t is the wage and TR_t is the transfer. The assumptions that were imposed on the capitalist's utility function u are also imposed on v.

The firm is a price taking profit maximizer with constant returns to scale in labor and capital. The production function in intensive form is $f(k_t)$. The capital to labor ratio coincides with the capital stock since the labor supply is always one unit. We assume that $f(0) = 0$ and that f also satisfies the same conditions as the utility functions u and v. At an interior optimum for the firm, $f'(k_t) = r_t$ and $f(k_t) - k_t f'(k_t) = w_t$.

Given the restriction against government debt, tax revenue must equal the transfer at each instant: $\tau_{kt}(r_t - \delta)k_t = TR_t$. Hence, from the definition of $\bar{r}_t$ and the firm's profit maximization condition, $TR_t = -\bar{r}_t k_t + [f'(k_t) - \delta]k_t$. Then the worker's consumption is

$$c_t^w = w_t + TR_t = [f(k_t) - k_t f'(k_t)] - \bar{r}_t k_t + [f'(k_t) - \delta]k_t = f(k_t) - \delta k_t - \bar{r}_t k_t. \tag{2}$$

In equilibrium, consumption plus investment must equal output: $c_t^c + c_t^w + \delta k_t + \dot{k}_t = f(k_t)$. Substitute for c_t^w to get $\dot{k}_t = \bar{r}_t k_t - c_t^c$, which is satisfied by the capitalist's flow budget constraint (Walras' Law).

3 Optimal taxation

The government maximizes social welfare $\int_0^\infty e^{-\rho t}[\gamma v(c_t^w) + (1-\gamma)u(c_t^c)]dt$ subject to the equilibrium conditions: the capitalist maximizes lifetime utility, the worker consumes all available income, firms maximize profits, the government's budget is in balance at every instant so the worker's income is as described in (2), and markets clear. Note that the government applies the capitalist's discount factor to both consumers, and the welfare weight γ is time invariant. There is one further constraint: $\bar{r}_t \geq 0$. This is a policy restriction that prevents the government from imposing a tax rate in excess of 100 percent.[5] And there are two further assumptions implicit in the analysis of Judd (1985):

- The initial stock of capital satisfies $f(k_0) - \delta k_0 > 0$. Without this, the worker's initial consumption in (2) would not be positive.

- The policy $\bar{r}_t \equiv 0$ does not solve the optimal taxation problem. This requires some background. In nonlinear programming the Fritz John necessary conditions allow for the possibility that the Lagrange multiplier of the objective function equals zero. But if a constraint qualification is satisfied this Lagrange multiplier can be set equal to one and we get the Kuhn–Tucker necessary conditions. For

[5] Judd (1985) and Lansing (1999) both impose $\bar{r}_t \geq 0$. If it were not imposed, the government could set an arbitrarily high tax rate on capital income for an arbitrarily short period of time and thereby raise tax revenue with almost no distortion. However, since the government must balance its budget at each instant of time, the benefits from this brief tax haul would be fleeting. By contrast, if the government were allowed to run a budget surplus then the tax revenue from the haul could be saved and gradually drawn down to pay for transfers to the worker. So the degree of importance of the constraint $\bar{r}_t \geq 0$ depends on the nature of the government's budget constraint. For the sake of consistency with Judd (1985) and Lansing (1999), $\bar{r}_t \geq 0$ is imposed here.

optimal control we follow Seierstad and Sydsæter (1987, p. 86) and say that a solution to the Pontryagin necessary conditions is abnormal if the multiplier of the objective function equals zero. For the optimal taxation problem here, appendix D shows that the only abnormal solution is $\bar{r}_t \equiv 0$. Under the assumption that this does not solve the optimal taxation problem, then any time path that is optimal must be a normal solution. I.e., the multiplier of the objective function is not zero, and it can be set equal to one by normalization, as we do below.

One may feel uncomfortable with the assumption that $\bar{r}_t \equiv 0$ is not optimal. It would be better not to impose an assumption on an endogenous policy variable. Appendix E provides two assumptions on the model's primitives under which we can prove that $\bar{r}_t \equiv 0$ is not optimal. Unfortunately the derivation is quite tedious.

Substitute for c_t^w from (2) to get the following problem:

$$\begin{aligned} \text{maximize} \quad & \int_0^\infty e^{-\rho t}[\gamma v(f(k_t) - \delta k_t - \bar{r}_t k_t) + (1-\gamma)u(c_t^c)]dt \\ \text{subject to} \quad & \dot{k}_t = \bar{r}_t k_t - c_t^c \\ & \dot{c}_t^c = (\rho - \bar{r}_t)u'(c_t^c)/u''(c_t^c) \\ & \bar{r}_t \geq 0 \end{aligned}$$

with $k_0 > 0$ given and $\lim_{t\to\infty} e^{-\rho t}u'(c_t^c)k_t = 0$. The optimal time path for the tax rate can be recovered from the definition of $\bar{r}_t := (1-\tau_{kt})(r_t - \delta)$ with $r_t = f'(k_t)$. The current value Hamiltonian is

$$H(k, c^c, \bar{r}, q_1, q_2, \eta) = \gamma v\big(f(k) - \delta k - \bar{r}k\big) + (1-\gamma)u(c^c) + q_1(\bar{r}k - c^c) + q_2(\rho - \bar{r})u'(c^c)/u''(c^c) + \eta\bar{r}\,.$$

The state variables are k_t (with co-state q_{1t}) and c_t^c (with co-state q_{2t}), $\bar{r}_t$ is the control, and η_t is the Lagrange multiplier for the constraint $\bar{r}_t \geq 0$. The following conditions are necessary for optimality:[6]

$$\partial H/\partial k = \gamma v'(c_t^w)[f'(k_t) - \delta - \bar{r}_t] + q_{1t}\bar{r}_t = \rho q_{1t} - \dot{q}_{1t} \tag{3a}$$

$$\partial H/\partial c^c = (1-\gamma)u'(c_t^c) - q_{1t} + q_{2t}(\rho - \bar{r}_t)\left\{1 - [u''(c_t^c)]^{-2}u'(c_t^c)u'''(c_t^c)\right\} = \rho q_{2t} - \dot{q}_{2t} \tag{3b}$$

$$\partial H/\partial \bar{r} = -\gamma v'(c_t^w)k_t + q_{1t}k_t - q_{2t}u'(c_t^c)/u''(c_t^c) + \eta_t = 0 \tag{3c}$$

$$\partial H/\partial q_1 = \bar{r}_t k_t - c_t^c = \dot{k}_t \tag{3d}$$

$$\partial H/\partial q_2 = (\rho - \bar{r}_t)u'(c_t^c)/u''(c_t^c) = \dot{c}_t^c \tag{3e}$$

$$\eta_t \geq 0, \quad \bar{r}_t \geq 0, \quad \eta_t\bar{r}_t = 0, \quad q_{20} = 0 \tag{3f}$$

[6] Note that the co-states do not appear in the equilibrium conditions nor in the statement of the optimal taxation problem. They are introduced in the Hamiltonian to allow for the use of optimal control theory and thereby derive necessary conditions for optimality. This should be kept in mind with regard to Judd's (1985) theorem and Lansing's (1999) theorem.

together with the problem's two boundary conditions. The last line includes the complementary slackness and transversality conditions.[7]

3.1 Theorem (Judd)[8] *Suppose a solution to* (3) *has the property that* k_t, c_t^c, $\bar{r}_t$, *and* q_{1t} *converge as* t *tends to infinity, with strictly positive limits for* k_t, c_t^c, *and* c_t^w. *Then* $\lim_{t\to\infty} \tau_{kt} = 0$.

Proof Drop the time subscripts to denote limiting values. From (3e),[9] $\bar{r} = \rho$. Therefore (3a) yields $f'(k) - \delta - \bar{r} = 0$. The theorem now follows from the definition $\bar{r}_t = (1 - \tau_{kt})[f'(k_t) - \delta]$. ■

In Lansing's example, $u = \log$. Then (3e) yields $c_t^c = c_0^c e^{\bar{R}_t - \rho t}$, and (3d) yields $d[e^{-\bar{R}_t} k_t]/dt = -e^{-\bar{R}_t} c_t^c$. Substitute the former into the latter to get $d[e^{-\bar{R}_t} k_t]/dt = d[c_0^c e^{-\rho t}/\rho]/dt$ and hence $e^{-\bar{R}_t} k_t = \text{constant} + c_0^c e^{-\rho t}/\rho$ or equivalently $k_t = \text{constant} \cdot e^{\bar{R}_t} + c_0^c e^{\bar{R}_t - \rho t}/\rho$. Now substitute this expression for k_t and the previous expression for c_t^c into the boundary condition $\lim_{t\to\infty} e^{-\rho t} u'(c_t^c) k_t = 0$ with $u = \log$. This determines that the "constant" must be zero, so $k_t = c_0^c e^{\bar{R}_t - \rho t}/\rho$. Evaluate at $t = 0$ and use $\bar{R}_0 = 0$ to get the capitalist's initial consumption, $c_0^c = \rho k_0$. Thus, $c_t^c = \rho k_0 e^{\bar{R}_t - \rho t}$ and $k_t = k_0 e^{\bar{R}_t - \rho t}$ from which we conclude $c_t^c = \rho k_t$. Substitute this and $u = \log$ into (3) to get:

$$\partial H/\partial k = \gamma v'(c_t^w)[f'(k_t) - \delta - \bar{r}_t] + q_{1t}\bar{r}_t = \rho q_{1t} - \dot{q}_{1t} \tag{4a}$$

$$\partial H/\partial c^c = (1-\gamma)/(\rho k_t) - q_{1t} - q_{2t}(\rho - \bar{r}_t) = \rho q_{2t} - \dot{q}_{2t} \tag{4b}$$

$$\partial H/\partial \bar{r} = -\gamma v'(c_t^w) k_t + q_{1t} k_t + q_{2t}\rho k_t + \eta_t = 0 \tag{4c}$$

$$\partial H/\partial q_1 = \bar{r}_t k_t - \rho k_t = \dot{k}_t \tag{4d}$$

$$\partial H/\partial q_2 = -(\rho - \bar{r}_t)\rho k_t = \rho \dot{k}_t \tag{4e}$$

$$\eta_t \geq 0, \quad \bar{r}_t \geq 0, \quad \eta_t \bar{r}_t = 0, \quad q_{20} = 0. \tag{4f}$$

This system characterizes the solution to the optimal tax problem when $u = \log$. One of the properties of (4) is that generically $\lim_{t\to\infty}(k_t, c_t^c, \bar{r}_t, q_{1t})$ does not exist. I.e., it may be that some of these variables

[7] Appendix D provides a derivation of these necessary conditions. For a finite time horizon T, the transversality conditions would be $q_{20} = q_{2T} = 0$. With an infinite time horizon, $q_{20} = 0$ continues to be necessary for optimality and this has implications for the time (in)consistency of the solution. Regarding the necessity of the transversality condition at infinity (TVC∞) for continuous time models, see Halkin (1974) for an early treatment. Kamihigashi (2001) generalizes much of the previous literature on this topic. However, Kamihigashi's (2001) results are not applicable to the optimal taxation problem here. In particular, if we express the problem here in reduced form, the constraint set for $(k_t, c_t^c, \dot{k}_t, \dot{c}_t^c)$ has an empty interior, and this violates assumption 3.1 of Kamihigashi (2001). The TVC∞ may still be necessary for optimality, but we cannot use Kamihigashi's (2001) theorem to reach this conclusion. Fortunately this has no bearing on the main results here. Judd's theorem, Lansing's theorem, and theorem 3.6 below remain true whether or not the TVC∞ is included among the necessary conditions.

[8] See theorem 2 and equations (24) on page 72 of Judd (1985).

[9] The assumption that $\lim_{t\to\infty} \xi_t$ converges does not always imply $\lim_{t\to\infty} \dot{\xi}_t = 0$ (e.g., $t^{-1}\sin t^2$). However, this is not a problem here. Equations (3a, (3d), (3e) are of the form $\dot{\xi}_t = G(k_t, c_t^c, \bar{r}_t, q_{1t})$ with G continuous, where $\dot{\xi}_t$ represents $\dot{q}_{1t}$, $\dot{k}_t$, or $\dot{c}_t^c$. Therefore, under stated assumptions, $\dot{\xi}_t$ has a limit as t tends to infinity. That limit must be zero; otherwise ξ_t (no dot) would fail to converge as t tends to infinity. A similar argument can be applied to Lansing's theorem, and to parts of theorem 3.6, below.

converge, but in general they cannot all converge. Thus, for this special utility function the hypotheses of Judd's theorem generically cannot be satisfied. The reason is as follows. If all these variables were to converge, the proof of Judd's theorem would apply so in the limit $\bar{r} = \rho$ (hence $\eta = 0$) and $f'(k) = \delta + \rho$. The latter condition would uniquely determine k (modified golden rule). Then, from (4c), q_{2t} would converge and its limit would satisfy $\gamma v'(c^w) = q_1 + \rho q_2$. Also, in the limit, (4b) would yield $(1-\gamma)/(\rho k) = q_1 + \rho q_2$. Hence $(1-\gamma)/(\rho k) = \gamma v'(c^w) = \gamma v'\big(f(k) - \delta k - \rho k\big)$, where the last equality uses (2). This would impose a second condition on k, in addition to $f'(k) = \delta + \rho$. Only in exceptional cases will the same value of k satisfy both these conditions. Generically there will be no k that satisfies both. Nonetheless, (4) is still valid — it still characterizes the solution to the optimal tax problem when $u = \log$. The fact that (generically) its variables do not all converge is neither here nor there.

Given the simplifications associated with $u = \log$, Lansing states directly the optimal tax problem for this special case:

$$\begin{aligned} \text{maximize} \quad & \int_0^\infty e^{-\rho t}[\gamma v\big(f(k_t) - \delta k_t - \bar{r}_t k_t\big) + (1-\gamma)\log(\rho k_t)]dt \\ \text{subject to} \quad & \dot{k}_t = (\bar{r}_t - \rho)k_t \\ & \bar{r}_t \geq 0 \end{aligned}$$

with $k_0 > 0$ given. The $\dot{c}_t^c$ equation is dropped because it is redundant. Thus the $\dot{k}_t$ equation has a dual role. Not only is it the capital accumulation equation; it is also the consumption Euler equation for the capitalist. The current value Hamiltonian is $H(k, \bar{r}, q_3, \eta) = \gamma v\big(f(k) - \delta k - \bar{r}k\big) + (1-\gamma)\log(\rho k) + q_3(\bar{r} - \rho)k + \eta\bar{r}$. The following conditions are necessary for optimality:

$$\partial H/\partial k = \gamma v'(c_t^w)[f'(k_t) - \delta - \bar{r}_t] + (1-\gamma)/k_t + q_{3t}(\bar{r}_t - \rho) = \rho q_{3t} - \dot{q}_{3t} \tag{5a}$$

$$\partial H/\partial \bar{r} = -\gamma v'(c_t^w)k_t + q_{3t}k_t + \eta_t = 0 \tag{5b}$$

$$\partial H/\partial q_3 = (\bar{r}_t - \rho)k_t = \dot{k}_t \tag{5c}$$

$$\eta_t \geq 0, \quad \bar{r}_t \geq 0, \quad \eta_t \bar{r}_t = 0 \tag{5d}$$

with $k_0 > 0$ given. In Lansing (1999), this appears as (21) on page 435. Note the new notation q_{3t} for the co-state here in (5). Since the $\dot{k}_t$ equation has a dual role here so does its co-state.[10] Indeed q_{3t} is distinct from both of the co-states in (4), q_{1t} (for capital) and q_{2t} (for the capitalist's consumption). However, they are related to one another.

[10] Cf Lansing (1999) where the same notation q_{1t} is used for the dual role co-state in (21) on page 435 and also for capital's co-state in (17) on page 432 where utility is not restricted to be logarithmic.

3.2 Lemma *Let $u = \log$. Equations* (4) *and* (5) *are equivalent, with*

$$q_{3t} = q_{1t} + \rho q_{2t} \tag{6}$$

$$k_t q_{1t} = (1-\gamma)t + k_t q_{3t} - \rho \int_0^t k_s q_{3s} ds \tag{7}$$

$$k_t q_{2t} = -(1-\gamma)t/\rho + \int_0^t k_s q_{3s} ds. \tag{8}$$

Proof First, given a solution to (4), verify that (5) is satisfied when q_{3t} is defined by (6). From (4a),

$$\begin{aligned} \gamma v'(c_t^w)[f'(k_t) - \delta - \bar{r}_t] &= (\rho - \bar{r}_t)q_{1t} - \dot{q}_{1t} \\ &= (\rho - \bar{r}_t)(q_{3t} - \rho q_{2t}) - (\dot{q}_{3t} - \rho \dot{q}_{2t}) \quad \text{by definition of } q_{3t} \\ &= (\rho - \bar{r}_t)q_{3t} - \dot{q}_{3t} - (1-\gamma)/k_t + \rho q_{1t} + \rho^2 q_{2t} \quad \text{by (4b)} \\ &= (\rho - \bar{r}_t)q_{3t} - \dot{q}_{3t} - (1-\gamma)/k_t + \rho q_{3t} \quad \text{by definition of } q_{3t}. \end{aligned}$$

So (5a) is satisfied. Clearly (5b) follows from (4c), (5c) follows from (4d), and (5d) follows from (4f). This completes the verification of (5).

Next, given a solution to (5), verify that (4) is satisfied when q_{1t} and q_{2t} are defined by (7) and (8). Take the time derivative of (7):

$$\dot{k}_t q_{1t} + k_t \dot{q}_{1t} = 1 - \gamma + \dot{k}_t q_{3t} + k_t \dot{q}_{3t} - \rho k_t q_{3t}.$$

Substitute for $\dot{k}_t$ from (5c) and substitute for $\dot{q}_{3t}$ from (5a):

$$\begin{aligned} &(\bar{r}_t - \rho)k_t q_{1t} + k_t \dot{q}_{1t} \\ &\quad = 1 - \gamma + (\bar{r}_t - \rho)k_t q_{3t} + k_t \{-\gamma v'(c_t^w)[f'(k_t) - \delta - \bar{r}_t] - (1-\gamma)/k_t - q_{3t}(\bar{r}_t - \rho) + \rho q_{3t}\} - \rho k_t q_{3t}. \end{aligned}$$

Simplify and divide by $k_t > 0$ to get (4a). Take the time derivative of (8):

$$\dot{k}_t q_{2t} + k_t \dot{q}_{2t} = -(1-\gamma)/\rho + k_t q_{3t}.$$

Substitute for $\dot{k}_t$ from (5c). We can also substitute for $k_t q_{3t}$: take (7) and add to it ρ times (8) to get $k_t q_{1t} + \rho k_t q_{2t} = k_t q_{3t}$. After these substitutions we have

$$(\bar{r}_t - \rho)k_t q_{2t} + k_t \dot{q}_{2t} = -(1-\gamma)/\rho + k_t q_{1t} + \rho k_t q_{2t}.$$

Divide by $k_t > 0$ to get (4b). Since we have just shown that (7) and (8) yield $k_t q_{1t} + \rho k_t q_{2t} = k_t q_{3t}$, (4c) follows from (5b). Clearly, (4d) and (4e) follow from (5c). Finally, (4f) follows from (5d) and (8). In particular, (8) yields $q_{20} = 0$.[11] ■

[11] With $u = \log$, the lemma's equivalence result has the following consequence. In (4), q_{1t} and q_{2t} affect the real allocation only through the value of $q_{1t} + \rho q_{2t}$. Hence, at any time $t > 0$ we can reset the value of q_{2t} to zero and reset the value of q_{1t} to

3.3 Theorem (Lansing)[12] *Let $u = \log$. Suppose a solution to (5) has the property that k_t, $\bar{r}_t$, and q_{3t} converge as t tends to infinity, with strictly positive limits for k_t, c_t^w, and $f'(k_t) - \delta$. Then, dropping the time subscripts to denote limiting values,* $\operatorname{sgn}(\tau_k) = \operatorname{sgn}\bigl(\rho\gamma v'(c^w)k - 1 + \gamma\bigr)$.

Proof From (5c), $\bar{r} = \rho$. From (5b), $q_3 = \gamma v'(c^w)$ since $\eta = 0$ ($\bar{r} > 0$) and $k > 0$. Therefore, (5a) yields $\gamma v'(c^w)[f'(k) - \delta - \bar{r}] = \rho\gamma v'(c^w) - (1-\gamma)/k$. The theorem now follows from $\bar{r}_t = (1 - \tau_{kt})[f'(k_t) - \delta]$. ∎

When $u = \log$, Judd's hypotheses are more restrictive than Lansing's. That is, in (4) Judd's hypotheses are that k_t, c_t^c, $\bar{r}_t$, and q_{1t} all converge. Recall that generically this does not happen, but when it does, $(1-\gamma)/(\rho k) = \gamma v'(c^w)$. So in this special case Lansing's theorem yields $\tau_k = 0$ in the limit, just like Judd's theorem: When $u = \log$, Judd's theorem is a special (and exceptional) case of Lansing's.

Furthermore, when Judd's hypotheses are satisfied, q_{2t} also converges by (4c). Hence, by (6), q_{3t} converges in (5). So Lansing's hypotheses are satisfied. I.e., when $u = \log$ Judd's hypotheses imply Lansing's hypotheses. The converse does not necessarily hold. It is possible for q_{3t} to converge while q_{1t} and q_{2t} diverge. The following corollary states this formally.

3.4 Corollary *Let $u = \log$. Suppose a solution to (5) has the property that k_t, $\bar{r}_t$, and q_{3t} converge as t tends to infinity, with strictly positive limits for k_t, c_t^w, and $f'(k_t) - \delta$. Then, in (4),*

$$\begin{aligned} \lim_{t\to\infty} q_{1t}/t &= (1 - \gamma - \rho k q_3)/k = (1-\gamma)/k - \rho\gamma v'(c^w) \\ \lim_{t\to\infty} q_{2t}/t &= \bigl(-(1-\gamma)/\rho + k q_3\bigr)/k = -(1-\gamma)/(\rho k) + \gamma v'(c^w) \end{aligned}$$

where $k = \lim_{t\to\infty} k_t$, etc. So if $\rho\gamma v'(c^w)k \neq 1 - \gamma$ then both q_{1t} and q_{2t} fail to converge.

Remark With $u = \log$, theorem 3.3 yields $\operatorname{sgn}(\tau_k) = \operatorname{sgn}\bigl(\rho\gamma v'(c^w)k - 1 + \gamma\bigr)$. It follows that if $\tau_k \neq 0$ then $\rho\gamma v'(c^w)k \neq 1 - \gamma$ in which case the corollary tells us that q_{1t} fails to converge and hence the hypotheses of Judd's theorem 3.1 are not satisfied. That is, if the capital tax result from Judd's theorem is violated, then the hypotheses from Judd's theorem must have been violated.

Proof In (7) and (8), apply l'Hopital's rule to the integrals divided by t, and use $q_3 = \gamma v'(c^w)$ from the proof of theorem 3.3. ∎

The following example rigs the initial condition and parameter values to illustrate the corollary.

$\lim_{s\uparrow t}(q_{1s} + \rho q_{2s})$. Thereafter, the future evolution of q_1 follows (4a) and q_2 follows (4b) and so the future values of $q_1 + \rho q_2$ are exactly as they were before the change. This has no effect on the real allocation. The ability to reset q_{2t} to zero at any point in time, without real consequence, tells us the optimal taxation problem is dynamically consistent when $u = \log$.

[12] See proposition 2 on page 435 of Lansing (1999). In the statement of theorem 3.3 here, the hypothesis that q_{3t} converges is actually a consequence of the other hypotheses, but it is stated explicitly for consistency; see the first line of the proof on page 435 of Lansing (1999). Appendix C below shows that Lansing's (1999) theorem can be derived directly from Judd's (1985) optimality conditions, i.e., directly from (3) with $u = \log$. For some intuition regarding the theorem's result, since $u = \log$ and $c^c = \rho k$, the result can be expressed as $\operatorname{sgn}(\tau_k) = \operatorname{sgn}\bigl(\gamma v'(c^w) - (1-\gamma)u'(c^c)\bigr)$: redistribution goes in favor of the consumer with the larger welfare-weighted marginal utility of consumption.

3.5 Example Let $u = \log$. Suppose

$$\gamma v'(c_0^w)[f'(k_0)-\delta-2\rho]+(1-\gamma)/k_0 = 0 \quad \text{where} \quad c_0^w := f(k_0)-\delta k_0-\rho k_0 > 0 \quad \& \quad 0 < f'(k_0)-\delta \neq \rho. \tag{9}$$

Then the following solves (5): $k_t \equiv k_0$, $\bar{r}_t \equiv \rho$ (hence $\eta_t \equiv 0$), $c_t^w \equiv c_0^w$, and $q_{3t} \equiv \gamma v'(c_0^w)$. So Lansing's hypotheses are satisfied. From (7) and (8), $q_{1t} = q_{30} + \big((1-\gamma)/k_0 - \rho q_{30}\big)t$ and $q_{2t} = -\big((1-\gamma)/k_0 - \rho q_{30}\big)t/\rho$, with $q_{30} = \gamma v'(c_0^w)$. So, from (9), q_{1t} and q_{2t} do not converge. The tax rate on capital income is not zero: $\tau_{kt}[f'(k_0)-\delta] \equiv f'(k_0)-\delta-\rho \neq 0$. ■

Return now to the general case (3) when the capitalist's utility is not necessarily $u = \log$. As stated in the introduction, the focus of attention is time paths for which the observables k_t, c_t^c, $\bar{r}_t$ converge to positive limits as t tends to infinity. Thus, for t sufficiently large the observables are approximately time invariant. To gain some insight we will temporarily take this approximation to the extreme: suppose that for all $t \geq T$, $(k_t, c_t^c, \bar{r}_t) \equiv (k, c^c, \bar{r})$. Although a time invariant path does not in general solve the optimal taxation problem, we will use this approximate solution to derive some implications. This will shed light on the limiting behavior as $t \to \infty$ for the true optimum which we will then analyze rigorously in theorem 3.6.

If a solution to (3) were to satisfy $(k_t, c_t^c, \bar{r}_t) \equiv (k, c^c, \bar{r})$ for all $t \geq T$ with $c^c > 0$ and $k > 0$, then $\bar{r} = \rho$ from (3e) and $c^c = \rho k$ from (3d). Also for all $t \geq T$, $c_t^w \equiv c^w = f(k) - \delta k - \rho k$ from (2); assume this is positive. From (3a), $q_{1t} = q_{1T} - \gamma v'(c^w)[f'(k)-\delta-\rho](t-T)$ for all $t \geq T$. Then from (3c) (with $\eta_t = 0$ for all $t \geq T$ since $\bar{r}_t = \rho$), $q_{2t} = ku''(\rho k)[u'(\rho k)]^{-1}\{q_{1T} - \gamma v'(c^w) - \gamma v'(c^w)[f'(k)-\delta-\rho](t-T)\}$ for all $t \geq T$. All that remains is (3b), which reduces to $(1-\gamma)u'(\rho k) - q_{1t} = \rho q_{2t} - \dot{q}_{2t}$ for all $t \geq T$. With the above solutions for q_{1t} and q_{2t} this requires that the coefficients of t match up:

$$\gamma v'(c^w)[f'(k)-\delta-\rho] = -\rho k u''(\rho k)[u'(\rho k)]^{-1}\gamma v'(c^w)[f'(k)-\delta-\rho]$$

hence

$$\gamma v'(c^w)[f'(k)-\delta-\rho][1+\rho k u''(\rho k)/u'(\rho k)] = 0. \tag{10a}$$

It also requires $(1-\gamma)u'(\rho k) - q_{1T} = \rho q_{2T} - \dot{q}_{2T}$:

$$(1-\gamma)u'(\rho k) - q_{1T} = \rho k u''(\rho k)[u'(\rho k)]^{-1}[q_{1T} - \gamma v'(c^w)] + k u''(\rho k)[u'(\rho k)]^{-1}\gamma v'(c^w)[f'(k)-\delta-\rho]$$

hence

$$q_{1T}[1+\rho k u''(\rho k)/u'(\rho k)] = (1-\gamma)u'(\rho k) - k u''(\rho k)[u'(\rho k)]^{-1}\gamma v'(c^w)[f'(k)-\delta-2\rho]. \tag{10b}$$

The solution to (10a) and (10b) requires one of the following alternatives:

(i) $f'(k) = \rho + \delta$ and $\rho k u''(\rho k)/u'(\rho k) \neq -1$;

(ii) $\rho k u''(\rho k)/u'(\rho k) = -1$ and $\rho(1-\gamma)u'(\rho k) = -\gamma v'(c^w)[f'(k)-\delta-2\rho]$.

In each of these, the first condition ensures that (10a) is satisfied, while the second ensures that (10b) is satisfied. In particular, in (i) the second condition allows us to find a unique value for q_{1T} that satisfies (10b). In (i), the capital tax is zero whereas in (ii), the capital tax is not restricted to be zero. Lansing's example with $u = \log$ is an instance of alternative (ii): the first condition in (ii) is satisfied identically and the second condition determines the value of k. (I.e., it determines the value of k that would lead to a time invariant path.) When $u \neq \log$, alternative (ii) would impose two distinct restrictions on k making it unlikely to have any solution. Thus, other than $u = \log$, alternative (ii) can be effectively dismissed and this leaves us with alternative (i) — zero tax on capital income.

These results for the time invariant approximate solution lead to the following theorem for the limiting behavior of the optimality conditions (3). As stated in the introduction, this theorem assumes that the observable macro variables converge to positive limits but it makes no assumptions about co-states.

3.6 Theorem *Suppose a solution to* (3) *has the property that* k_t, c_t^c, *and* $\bar{r}_t$ *converge as* t *tends to infinity, with strictly positive limits for* k_t, c_t^c, *and* c_t^w. *Then* $\lim_{t\to\infty} \tau_{kt} = 0$ *or* $\lim_{t\to\infty}[c_t^c + u'(c_t^c)/u''(c_t^c)] = 0$ *or both.*

Proof Use (3a) to substitute for $\dot{q}_{1t}$ and use (3e) to substitute for $\dot{c}_t^c$ to get the following:

$$\frac{d}{dt}\left[\frac{q_{1t}}{u'(c_t^c)}\right] = \frac{\dot{q}_{1t}}{u'(c_t^c)} - \frac{q_{1t}u''(c_t^c)\dot{c}_t^c}{[u'(c_t^c)]^2} = -\frac{\gamma v'(c_t^w)[f'(k_t) - \delta - \bar{r}_t]}{u'(c_t^c)}. \tag{11}$$

Then from the mean value theorem, for all $t \geq T$ there exists $s \in [T, t]$ such that

$$q_{1t}/u'(c_t^c) = q_{1T}/u'(c_T^c) - (t - T)\gamma[u'(c_s^c)]^{-1}v'(c_s^w)[f'(k_s) - \delta - \bar{r}_s]. \tag{12}$$

Under the convergence hypotheses, we can choose T sufficiently large so that $[u'(c_s^c)]^{-1}v'(c_s^w)[f'(k_s) - \delta - \bar{r}_s]$ is arbitrarily close to $\lim_{t\to\infty}\left([u'(c_t^c)]^{-1} v'(c_t^w)[f'(k_t) - \delta - \bar{r}_t]\right)$. Then (12) yields $\lim_{t\to\infty}[t^{-1}q_{1t}/u'(c_t^c)] = -\gamma \lim_{t\to\infty}\left([u'(c_t^c)]^{-1} v'(c_t^w)[f'(k_t) - \delta - \bar{r}_t]\right)$. Now consider the limiting behavior of q_{2t}. Since $\lim_{t\to\infty} \bar{r}_t = \rho$ from (3e), we have $\eta_t = 0$ for all t sufficiently large from (3f). Then from (3c), $\lim_{t\to\infty}[t^{-1}q_{2t}u'(c_t^c)/u''(c_t^c)] = \lim_{t\to\infty}[t^{-1}q_{1t}k_t]$. Since the limit of a product is the product of the limits, we can summarize our results thus far:

$$\lim_{t\to\infty}\frac{q_{1t}}{t} = -\gamma \lim_{t\to\infty}\left(v'(c_t^w)[f'(k_t) - \delta - \bar{r}_t]\right) \quad \& \quad \lim_{t\to\infty}\frac{q_{2t}}{t} = -\gamma \lim_{t\to\infty}\left(\frac{k_t u''(c_t^c)}{u'(c_t^c)} v'(c_t^w)[f'(k_t) - \delta - \bar{r}_t]\right).$$

Use (3b) to substitute for $\dot{q}_{2t}$ and use (3e) to substitute for $\dot{c}_t^c$ to get the following:

$$\begin{aligned}\frac{d}{dt}\left[\frac{e^{-\rho t}q_{2t}u'(c_t^c)}{u''(c_t^c)}\right] &= \frac{-\rho e^{-\rho t}q_{2t}u'(c_t^c)}{u''(c_t^c)} + \frac{e^{-\rho t}\dot{q}_{2t}u'(c_t^c)}{u''(c_t^c)} + e^{-\rho t}q_{2t}\left[1 - \frac{u'(c_t^c)u'''(c_t^c)}{[u''(c_t^c)]^2}\right]\dot{c}_t^c \\ &= e^{-\rho t}\frac{[u'(c_t^c)]^2}{u''(c_t^c)}[q_{1t}/u'(c_t^c) - 1 + \gamma].\end{aligned}$$

Integrate over $[t, \infty)$ and use the previous result that $q_{2t} = \mathcal{O}(t)$ as $t \to \infty$:

$$-e^{-\rho t} q_{2t} u'(c_t^c)/u''(c_t^c) = \int_t^\infty e^{-\rho s} \frac{[u'(c_s^c)]^2}{u''(c_s^c)} [q_{1s}/u'(c_s^c) - 1 + \gamma] ds. \tag{13}$$

In preparation for applying integration by parts to (13), let $z_t := \int_t^\infty e^{-\rho s}[u'(c_s^c)]^2[u''(c_s^c)]^{-1}ds$. From l'Hopital's rule, $\lim_{t\to\infty}[z_t/e^{-\rho t}] = \lim_{t\to\infty}[u'(c_t^c)]^2[\rho u''(c_t^c)]^{-1}$. This will be useful later. From (11),

$$\frac{d}{dt}[q_{1t}/u'(c_t^c) - 1 + \gamma] = -\gamma[u'(c_t^c)]^{-1} v'(c_t^w)[f'(k_t) - \delta - \bar{r}_t].$$

We can now express (13) as follows after applying integration by parts to the right side:

$$\begin{aligned} -e^{-\rho t} q_{2t} u'(c_t^c)/u''(c_t^c) &= \Big[-z_s[q_{1s}/u'(c_s^c) - 1 + \gamma]\Big]_t^\infty - \gamma \int_t^\infty z_s[u'(c_s^c)]^{-1} v'(c_s^w)[f'(k_s) - \delta - \bar{r}_s] ds \\ &= z_t[q_{1t}/u'(c_t^c) - 1 + \gamma] - \gamma \int_t^\infty z_s[u'(c_s^c)]^{-1} v'(c_s^w)[f'(k_s) - \delta - \bar{r}_s] ds. \end{aligned}$$

The second line follows from the limiting behavior of z_t and from $q_{1t} = \mathcal{O}(t)$ as $t \to \infty$. Use this equation to substitute for $q_{2t} u'(c_t^c)/u''(c_t^c)$ in (3c):

$$q_{1t}[k_t + e^{\rho t} z_t/u'(c_t^c)] = \gamma v'(c_t^w) k_t + (1-\gamma) e^{\rho t} z_t + \gamma e^{\rho t} \int_t^\infty z_s[u'(c_s^c)]^{-1} v'(c_s^w)[f'(k_s) - \delta - \bar{r}_s] ds - \eta_t. \tag{14}$$

As t tends to infinity, all terms on the right side of this equation converge. In particular, l'Hopital's rule can be applied to the integral divided by $e^{-\rho t}$, while as shown previously $\eta_t = 0$ for all t sufficiently large. Furthermore, the term in square brackets on the left side converges. There are two possible cases: (i) $\lim_{t\to\infty}[k_t + e^{\rho t} z_t/u'(c_t^c)] \neq 0$, or (ii) $\lim_{t\to\infty}[k_t + e^{\rho t} z_t/u'(c_t^c)] = 0$. In case (i), (14) reveals that q_{1t} converges as $t \to \infty$ so Judd's theorem applies and $\lim_{t\to\infty} \tau_{kt} = 0$. In case (ii),

$$0 = \lim_{t\to\infty}[k_t + e^{\rho t} z_t/u'(c_t^c)] = \rho^{-1} \lim_{t\to\infty}[c_t^c + u'(c_t^c)/u''(c_t^c)]$$

where the second equality uses $\lim_{t\to\infty} c_t^c = \rho \lim_{t\to\infty} k_t$ from (3d), (3e), and it also uses the earlier result regarding the limiting behavior of z_t. ∎

3.7 Remark Consider the following cases of the theorem. If $u = \log$, then in the proof $z_t = -e^{-\rho t}/\rho$ and $c_t^c = \rho k_t$, so the left side of (14) is identically zero. Then dropping time subscripts to denote limiting values, the limit of (14) multiplied by $-\rho$ is $0 = \gamma v'(c^w) k[f'(k) - \delta - 2\rho] + 1 - \gamma$ (apply l'Hopital's rule to the integral divided by $e^{-\rho t}$) with $c^w = f(k) - \delta k - \rho k$. This determines the steady state value(s) of k, and hence, by $\rho = \lim_{t\to\infty} \bar{r}_t := \lim_{t\to\infty}\big[(1-\tau_{kt})\big(f'(k_t)-\delta\big)\big]$, it also determines τ_k. If u is any other constant EIS function for which the convergence hypotheses are satisfied, then $c^c + u'(c^c)/u''(c^c) \neq 0$, so the theorem yields $\tau_k = 0$. In this case, q_{1t} converges and (14) determines its limiting value, while the equation $f'(k) = \delta + \rho$ determines k. For general u, if τ_{kt} fails to converge to zero, k must satisfy $\rho k + u'(\rho k)/u''(\rho k) = 0$, and (14) determines the limiting behavior of the indeterminate form $\lim_{t\to\infty} q_{1t}[k_t + e^{\rho t} z_t/u'(c_t^c)]$. ∎

4 Conclusion

This paper has clarified the relationship between the results of Judd (1985) and Lansing (1999). Judd's theorem states that in the long run, the optimal tax rate on capital income converges to zero.[13] Lansing identifies a logarithmic example of one of Judd's models in which this tax rate can converge in the long run to any number, zero or otherwise — the value depends on the model's primitives (parameter values, initial condition, worker's utility function, production function). It seems odd that the same model can generate two different results. The apparent contradiction is resolved by observing that Lansing has relaxed the convergence hypotheses of Judd's theorem. As a consequence, Lansing's hypotheses are satisfied by a wider range of primitives, each of which generates its own optimal time path(s) for the model's variables, and hence a wider range of properties is possible in steady state. One would like to know if a further relaxation of the hypotheses, satisfied by a yet wider range of primitives, will allow for yet more possibilities for steady state properties. Theorem 3.6 addresses this issue and finds that any interior steady state for the optimal tax equilibrium of section 2's model must satisfy one (or both) of the following: (i) the capital tax converges to zero in the long run, or (ii) the elasticity of the capitalist's marginal utility of consumption converges to one in the long run. In (ii), the income and substitution effects of an interest rate change just cancel. This is satisfied identically with logarithmic utility, which was the case considered by Lansing.

[13] Judd (1999, 2002) has returned to this issue, but not with the worker-capitalist model. The range of views on capital income taxation is illustrated by Atkeson, Chari, and Kehoe (1999) on the one hand who say that it is a bad idea to tax capital income, and on the other hand Conesa, Kitao, and Krueger (2009) who say that taxing capital is not a bad idea. The range of models is also vast: infinitely lived agents with uninsurable idiosyncratic shocks (e.g., Aiyagari 1995), overlapping generations of agents with uninsurable idiosyncratic shocks (e.g., Conesa, Kitao, and Krueger 2009), government expenditure that depends on the size of the economy (e.g., Ben-Gad 2017; Lu and Chen 2015), housing capital in addition to business capital (e.g., Eerola and Määttänen 2013), among many others.

References

[1] S. Rao Aiyagari, "Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting," *Journal of Political Economy* 103(6), December 1995, 1158–1175.

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

[3] Michael Ben-Gad, "The optimal taxation of asset income when government consumption is endogenous: Theory, estimation and welfare," *Economic Inquiry* 55(4), October 2017, 1689–1711.

[4] Juan Carlos Conesa, Sagiri Kitao, and Dirk Krueger, "Taxing capital? Not a bad idea after all!" *American Economic Review* 99(1), March 2009, 25–48.

[5] Christophe Chamley, "Optimal taxation of capital income in general equilibrium with infinite lives," *Econometrica* 54(3), May 1986, 607–622.

[6] Essi Eerola and Niku Määttänen, "The optimal tax treatment of housing capital in the neoclassical growth model," *Journal of Public Economic Theory* 15(6), December 2013, 912–938.

[7] Hubert Halkin, "Necessary conditions for optimal control problems with infinite horizons," *Econometrica* 42(2), March 1974, 267–272.

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

[9] Kenneth L. Judd, "Optimal taxation and spending in general competitive growth models," *Journal of Public Economics* 71(1), January 1999, 1–26.

[10] Kenneth L. Judd, "Capital-income taxation with imperfect competition," *American Economic Review Papers and Proceedings* 92(2), May 2002, 417–421.

[11] Takashi Kamihigashi, "Necessity of transversality conditions for infinite horizon problems," *Econometrica* 69(4), July 2001, 995–1012.

[12] Murray C. Kemp, Ngo Van Long, and Koji Shimomura, "Cyclical and noncyclical redistributive taxation," *International Economic Review* 34(2), May 1993, 415–429.

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

[14] Chia-Hui Lu and Been-Lon Chen, "Optimal capital taxation in a neoclassical growth model," *Journal of Public Economic Theory* 17(2), April 2015, 257–269.

[15] Robert E. Lucas, Jr., "Supply-side economics: An analytical review," *Oxford Economic Papers* 42(2), April 1990, 293–316.

[16] Atle Seierstad and Knut Sydsæter, *Optimal Control Theory with Economic Applications*, Amsterdam: North-Holland, 1987.

[17] Ludwig Straub and Iván Werning, "Positive long run capital taxation: Chamley–Judd revisited," working paper, June 2018.

Appendix

This appendix contains material on the following: the effect of interest rates on savings; when can the first best be decentralized as an equilibrium; derivation of Lansing's (1999) theorem from Judd's (1985) optimality conditions; necessary conditions for the infinite horizon optimal taxation problem; and could $\bar{r}_t \equiv 0$ be a solution to the infinite horizon optimal taxation problem.

A: The effect of interest rates on savings

Consider a two period model in which a consumer chooses first period consumption c_1 and second period consumption c_2 to maximize $u_1(c_1) + \beta u_2(c_2)$ subject to the present value budget constraint $c_1 + c_2/R \leq y_1 + y_2/R$. The per-period utility functions u_1 and u_2 are strictly increasing and strictly concave; β is a subjective discount factor (which could have been subsumed in u_2); R is the gross after-tax interest rate; and y_t is exogenous income in period t. Let $y = y_1 + y_2/R$ be present value income. The first order conditions are $u_1'(c_1)/[\beta u_2'(c_2)] = R$ together with the budget with equality. Substitution and re-arrangement yields

$$u_1'(c_1) = \beta R u_2'\big(R(y - c_1)\big).$$

This determines c_1 implicitly as a function of R and y. Differentiation yields

$$u_1''(c_1)\frac{\partial c_1}{\partial R} = \beta u_2'(c_2) + \beta R u_2''(c_2)\left[y - c_1 - R\frac{\partial c_1}{\partial R}\right]$$

and hence

$$\begin{aligned}\left[u_1''(c_1) + \beta R^2 u_2''(c_2)\right]\frac{\partial c_1}{\partial R} &= \beta u_2'(c_2) + \beta c_2 u_2''(c_2) \\ &= \beta u_2''(c_2)\left(\frac{u_2'(c_2)}{u_2''(c_2)} + c_2\right).\end{aligned}$$

Since savings are $s = y_1 - c_1$,

$$\operatorname{sgn}\left(\frac{\partial s}{\partial R}\right) = -\operatorname{sgn}\left(c_2 + \frac{u_2'(c_2)}{u_2''(c_2)}\right).$$

Therefore, in theorem 3.6 the condition $\lim_{t\to\infty}[c_t^c + u'(c_t^c)/u''(c_t^c)] = 0$ has an interpretation that savings are insensitive to the interest rate in the long run: the income effect and the substitution effect cancel each other out.

B: When can the first best be decentralized as an equilibrium?

The first best problem for the model in section 2 is to choose $\{c_t^c, c_t^w, k_t\}_{t\geq 0}$ to

$$\begin{aligned} \text{maximize} \quad & \int_0^\infty e^{-\rho t}[\gamma v(c_t^w) + (1-\gamma)u(c_t^c)]dt \\ \text{subject to} \quad & c_t^c + c_t^w + \dot{k}_t + \delta k_t = f(k_t) \\ & k_0 > 0 \text{ given.} \end{aligned}$$

The current value Hamiltonian is $H(k, c^c, c^w, \lambda) = \gamma v(c^w) + (1-\gamma)u(c^c) + \lambda[f(k) - \delta k - c^c - c^w]$ where λ is the co-state for k. The optimality conditions are

$$\begin{aligned} \partial H/\partial k &= \lambda_t[f'(k_t) - \delta] = \rho\lambda_t - \dot{\lambda}_t \\ \partial H/\partial c^c &= (1-\gamma)u'(c_t^c) - \lambda_t = 0 \\ \partial H/\partial c^w &= \gamma v'(c_t^w) - \lambda_t = 0 \\ \partial H/\partial \lambda &= f(k_t) - \delta k_t - c_t^c - c_t^w = \dot{k}_t \\ & \lim_{t\to\infty} e^{-\rho t}\lambda_t k_t = 0, \quad k_0 > 0 \text{ given.} \end{aligned}$$

We can use the $\partial H/\partial c^c$ equation to eliminate λ and get the following equivalent conditions:

$$\dot{c}_t^c u''(c_t^c)/u'(c_t^c) = \rho + \delta - f'(k_t) \tag{15}$$

$$(1-\gamma)u'(c_t^c) = \gamma v'(c_t^w) \tag{16}$$

$$\dot{k}_t = f(k_t) - \delta k_t - c_t^c - c_t^w \tag{17}$$

$$\lim_{t\to\infty} e^{-\rho t}u'(c_t^c)k_t = 0, \quad k_0 > 0 \text{ given.} \tag{18}$$

Suppose we have a solution to these first best conditions, denoted by asterisks, $\{c_t^{c*}, c_t^{w*}, k_t^*\}_{t\geq 0}$. Our task is to determine when this solution can be decentralized as an equilibrium. I.e., when can we find $\{\bar{r}_t\}_{t\geq 0}$ such that $\{c_t^{c*}, c_t^{w*}, k_t^*, \bar{r}_t\}_{t\geq 0}$ is a solution to

$$c_t^w = f(k_t) - \delta k_t - \bar{r}_t k_t \tag{19}$$

$$\dot{k}_t = \bar{r}_t k_t - c_t^c \tag{20}$$

$$\dot{c}_t^c u''(c_t^c)/u'(c_t^c) = \rho - \bar{r}_t \tag{21}$$

$$\lim_{t\to\infty} e^{-\rho t}u'(c_t^c)k_t = 0, \quad k_0 > 0 \text{ given.} \tag{22}$$

We have omitted the constraint $\bar{r}_t \geq 0$ which was imposed on the government's optimal taxation problem. Since the first best satisfies (15), we will satisfy (21) if and only if

$$\bar{r}_t := f'(k_t^*) - \delta. \tag{23}$$

Since the first best satisfies (17), and with $\bar{r}_t$ as just defined, we will satisfy (20) if and only if

$$c_t^{w*} = f(k_t^*) - k_t^* f'(k_t^*). \tag{24}$$

If the first best does indeed satisfy (24), and with $\bar{r}_t$ defined by (23), then the final equilibrium condition, (19), is also satisfied.

We conclude that the first best can be decentralized as an equilibrium if and only if it satisfies (24) for all $t \geq 0$. There is no reason to expect it to satisfy this condition, so there is no reason to expect it to be decentralizable. However, when this does occur, the equilibrium after-tax interest rate in (23) equals the before-tax rate. Hence, the capital income tax rate is identically zero through all time.

C: Derivation of Lansing's (1999) theorem from Judd's (1985) optimality conditions

C.1 Theorem (Lansing) *Let* $u = \log$*. Consider a solution to* (3) *with* $\lim_{t\to\infty} e^{-\rho t} u'(c_t^c) k_t = 0$ *and with* $k_0 > 0$ *given. Suppose* k_t *and* $\bar{r}_t$ *converge as* t *tends to infinity, with strictly positive limits for* k_t*,* c_t^w*, and* $f'(k_t) - \delta$*. Then, dropping the time subscripts to denote limiting values,* $\operatorname{sgn}(\tau_k) = \operatorname{sgn}\big(\rho\gamma v'(c^w)k - 1 + \gamma\big)$*.*

Remark The hypotheses do not impose the convergence of any co-states. Nonetheless, $q_{1t} + \rho q_{2t}$ will converge to a finite limit. See footnote 12 above.

Proof On page 9, just after the proof of theorem 3.1, we showed that when $u = \log$ the equations (3), together with the boundary condition $\lim_{t\to\infty} e^{-\rho t} u'(c_t^c) k_t = 0$, yield both $c_t^c = \rho k_t$ and also the equations (4). Under the present convergence hypotheses, (4d) and (4e) imply $\bar{r} = \rho$. Then (4f) implies that for all t sufficiently large, $\eta_t = 0$. Then since $k > 0$, (4c) yields that for all t sufficiently large,

$$-\gamma v'(c_t^w) + q_{1t} + \rho q_{2t} = 0. \tag{25}$$

Since c_t^w has a strictly positive limit, this tells us that $q_{1t} + \rho q_{2t}$ has a finite limit as $t \to \infty$. Now take ρ times (4b) and add it to (4a) to get

$$\gamma v'(c_t^w)[f'(k_t) - \delta - \bar{r}_t] + (1-\gamma)/k_t + (\bar{r}_t - \rho)(q_{1t} + \rho q_{2t}) = \rho(q_{1t} + \rho q_{2t}) - (\dot{q}_{1t} + \rho \dot{q}_{2t}).$$

Substitute for $q_{1t} + \rho q_{2t}$ from (25):

$$\gamma v'(c_t^w)[f'(k_t) - \delta - \rho] + (1-\gamma)/k_t = \rho\gamma v'(c_t^w) - \frac{d}{dt}\Big[\gamma v'(c_t^w)\Big] \tag{26}$$

for all t sufficiently large. Loosely speaking, this equation, in which all co-states have been eliminated, is the Euler equation for the optimal taxation problem with $u = \log$. Since the left side of (26) has a finite limit as $t \to \infty$, and since the first term on the right side also has a finite limit, it follows that $\frac{d}{dt}[\gamma v'(c_t^w)]$ has a finite limit, and furthermore this limit must be zero. (See footnote 9.) Let $t \to \infty$ in (26):

$$\gamma v'(c^w)[f'(k) - \delta - \rho] = \rho\gamma v'(c^w) - (1-\gamma)/k.$$

The rest of the proof is the same as theorem 3.3. ■

D: Necessary conditions for the infinite horizon optimal taxation problem[14]

As in section 3, the optimal taxation problem is as follows:

$$\begin{aligned} \text{maximize} \quad & \int_0^\infty e^{-\rho t}[\gamma v(f(k_t) - \delta k_t - \bar{r}_t k_t) + (1-\gamma)u(c_t^c)]dt \\ \text{subject to} \quad & \dot{k}_t = \bar{r}_t k_t - c_t^c \\ & \dot{c}_t^c = (\rho - \bar{r}_t)u'(c_t^c)/u''(c_t^c) \\ & \bar{r}_t \geq 0 \\ & k_0 > 0 \text{ given}, \ \lim_{t\to\infty} e^{-\rho t}u'(c_t^c)k_t = 0. \end{aligned}$$

Let $\{(\bar{r}_t^*, k_t^*, c_t^{c*})\}_{t\geq 0}$ be a solution to this problem. Then following Halkin (1974) we know that for all $T > 0$, $\{(\bar{r}_t^*, k_t^*, c_t^{c*})\}_{0\leq t\leq T}$ is a solution to the following finite horizon problem with clamped terminal state:

$$\begin{aligned} \text{maximize} \quad & \int_0^T e^{-\rho t}[\gamma v(f(k_t) - \delta k_t - \bar{r}_t k_t) + (1-\gamma)u(c_t^c)]dt \\ \text{subject to} \quad & \dot{k}_t = \bar{r}_t k_t - c_t^c \\ & \dot{c}_t^c = (\rho - \bar{r}_t)u'(c_t^c)/u''(c_t^c) \\ & \bar{r}_t \geq 0 \\ & k_0 > 0 \text{ given}, \ k_T = k_T^*, \ c_T^c = c_T^{c*}. \end{aligned}$$

On the last line the terminal values of the state variables are clamped down at the time T values that solve the infinite horizon problem.[15]

We shall proceed to express this clamped terminal state problem in the form that appears in section 3 of chapter 2 of Fleming and Rishel (1975).[16] Unfortunately we need a slight change of notation. Fleming and Rishel (1975) use the symbol u for the control, while the optimal taxation problem already uses u for the capitalist's utility function. So we will replace Fleming and Rishel's (1975) $u(t)$ with $\bar{r}_t$ or $\bar{r}(t)$ which is the control in the main text of the paper. Also, Fleming and Rishel (1975) use the symbol f in the equation of motion for the state of the system, while the optimal taxation problem already uses f for the production function in intensive form. So we will replace Fleming and Rishel's (1975) f with F.

[14] Necessary conditions for the finite horizon (T) optimal taxation problem are similar. Where the derivation differs, this will be indicated with footnotes.

[15] In the finite horizon (T) optimal taxation problem, the capitalist faces the constraint $k_T \geq 0$. This is the finite horizon equivalent to the infinite horizon no-Ponzi condition. The utility maximizing capitalist chooses $k_T = 0$. So in this case the government's optimal taxation problem is almost identical to the clamped terminal state problem except that the boundary conditions are $k_0 > 0$ given and $k_T = 0$, with c_T^c unconstrained.

[16] Wendell H. Fleming and Raymond W. Rishel, *Deterministic and Stochastic Optimal Control*, Berlin: Springer-Verlag, 1975.

The state of the system at time t is

$$\begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} k_t \\ c_t^c \\ x_3(t) \end{pmatrix}.$$

I.e., we are introducing a new component of the state, $x_3(t)$. The equation of motion is $\dot{x}(t) = F\big(t, x(t), \bar{r}(t)\big)$ where

$$F(t, x_1, x_2, x_3, \bar{r}) = \begin{pmatrix} \bar{r}x_1 - x_2 \\ (\rho - \bar{r})u'(x_2)/u''(x_2) \\ e^{-\rho t}[\gamma v(f(x_1) - \delta x_1 - \bar{r}x_1) + (1-\gamma)u(x_2)] \end{pmatrix}.$$

So $\dot{x}_3$ is equal to the third component of F, and we can integrate to get $x_3(t_1) - x_3(t_0) = \int_{t_0}^{t_1} e^{-\rho t}\big[\gamma v\big(f(x_1(t)) - \delta x_1(t) - \bar{r}(t)x_1(t)\big) + (1-\gamma)u\big(x_2(t)\big)\big]\,dt$. Compare this with the welfare objective in the optimal taxation problem with clamped terminal state. If we start the system at the fixed time $t_0 = 0$ and end it at the fixed time $t_1 = T$, then the performance index which we seek to *minimize* is the negative of welfare:

$$\phi_1\big(t_0, t_1, x(t_0), x(t_1)\big) \;=\; x_3(t_0) - x_3(t_1).$$

The end conditions are $\phi_2(\cdot) = \phi_3(\cdot) = \phi_4(\cdot) = \phi_5(\cdot) = \phi_6(\cdot) = 0$ where

$$\begin{aligned}
\phi_2\big(t_0, t_1, x(t_0), x(t_1)\big) &= t_0 \\
\phi_3\big(t_0, t_1, x(t_0), x(t_1)\big) &= t_1 - T \\
\phi_4\big(t_0, t_1, x(t_0), x(t_1)\big) &= x_1(t_0) - k_0 \\
\phi_5\big(t_0, t_1, x(t_0), x(t_1)\big) &= x_1(t_1) - k_T^* \\
\phi_6\big(t_0, t_1, x(t_0), x(t_1)\big) &= x_2(t_1) - c_T^{c*}.
\end{aligned}$$

Since $x(t_0)$ and $x(t_1)$ are 3-tuples, ϕ is a function from $\mathbf{R}^8$ to $\mathbf{R}^6$. The closed control set U, introduced on the last line of page 23 of Fleming and Rishel (1975), is taken to be $U = [0, \infty)$. This captures the constraint $\bar{r}_t \geq 0$.[17]

The Pontryagin necessary conditions for optimality of $(x^*, \bar{r}^*)$ are (5.1) through (5.6) on page 27 of Fleming and Rishel (1975). These conditions are that there exists a non-zero vector $(\lambda_1, \ldots, \lambda_6)$ with $\lambda_1 \leq 0$ and there exists a function $P : [t_0, t_1] \to \mathbf{R}^3$ such that

$$\dot{P}(t)' = -P(t)'F_x\big(t, x^*(t), \bar{r}^*(t)\big) \quad \forall\, t \in [t_0, t_1] \qquad \text{(adjoint equations)}$$

$$P(t)'\big[F(t, x^*(t), \bar{r}^*(t))\big] = \max_{\bar{r} \in U} P(t)'\big[F(t, x^*(t), \bar{r})\big] \quad \forall\, t \in (t_0, t_1) \qquad \text{(maximum principle)}$$

[17] In the finite horizon (T) optimal taxation problem, we modify ϕ as follows: $\phi_5\big(t_0, t_1, x(t_0), x(t_1)\big) = x_1(t_1)$ which captures the constraint $k_T = 0$, and we delete ϕ_6. See footnote 15. Then ϕ is a function from $\mathbf{R}^8$ to $\mathbf{R}^5$.

$$\left.\begin{aligned} P(t_1)' &= \lambda' \phi_{x_1}\big(t_0, t_1, x^*(t_0), x^*(t_1)\big) \\ P(t_0)' &= -\lambda' \phi_{x_0}\big(t_0, t_1, x^*(t_0), x^*(t_1)\big) \\ P(t_1)' F\big(t_1, x^*(t_1), \bar{r}^*(t_1)\big) &= -\lambda' \phi_{t_1}\big(t_0, t_1, x^*(t_0), x^*(t_1)\big) \\ P(t_0)' F\big(t_0, x^*(t_0), \bar{r}^*(t_0)\big) &= \lambda' \phi_{t_0}\big(t_0, t_1, x^*(t_0), x^*(t_1)\big) \end{aligned}\right\} \qquad \text{(transversality conditions)}$$

where ϕ_{x_1} is the partial derivative of ϕ with respect to the arguments of $x(t_1)$, and similarly for ϕ_{x_0}. This notation ϕ_{x_1} is not ideal since x_1 refers to a 3-tuple here whereas x_1 also refers to the first state variable, a scalar. Below, the meaning of x_1 should be clear from the context. In what follows, we drop the asterisks.

We now proceed to re-write these necessary conditions, using the optimal taxation problem's x, F, and ϕ. From the definition of F above we have that $F_x(t, x_1, x_2, x_3, \bar{r})$ is equal to

$$\begin{pmatrix} \bar{r} & -1 & 0 \\ 0 & (\rho - \bar{r})\big\{1 - \big[u''(x_2)\big]^{-2} u'(x_2) u'''(x_2)\big\} & 0 \\ e^{-\rho t} \gamma v'(f(x_1) - \delta x_1 - \bar{r} x_1)\big[f'(x_1) - \delta - \bar{r}\big] & e^{-\rho t}(1-\gamma) u'(x_2) & 0 \end{pmatrix}.$$

Hence the third component of the adjoint equations yields $\dot{P}_3(t) \equiv 0$, so $P_3(t)$ is a constant which we will simply denote P_3. Then the other two adjoint equations can be written as follows, using the definition of x and using the end conditions $t_0 = t_1 - T = 0$:

$$\dot{P}_1(t) = -P_1(t)\bar{r}_t - P_3 e^{-\rho t} \gamma v'(f(k_t) - \delta k_t - \bar{r}_t k_t)\big[f'(k_t) - \delta - \bar{r}_t\big] \quad \forall\, t \in [0, T] \tag{27}$$

$$\dot{P}_2(t) = P_1(t) - P_2(t)(\rho - \bar{r}_t)\big\{1 - \big[u''(c_t^c)\big]^{-2} u'(c_t^c) u'''(c_t^c)\big\} - P_3 e^{-\rho t}(1-\gamma) u'(c_t^c) \quad \forall\, t \in [0, T]. \tag{28}$$

With $U = [0, \infty)$, the maximum principle states that $\bar{r}_t$ must solve

$$\max_{\bar{r} \geq 0} \Big\{ P_1(t)(\bar{r} k_t - c_t^c) + P_2(t)(\rho - \bar{r}) u'(c_t^c)/u''(c_t^c) + P_3 e^{-\rho t}\big[\gamma v\big(f(k_t) - \delta k_t - \bar{r} k_t\big) + (1-\gamma) u(c_t^c)\big] \Big\} \quad \forall\, t \in (0, T).$$

We can disregard the terms that do not involve $\bar{r}$. Then $\bar{r}_t$ must solve

$$\max_{\bar{r} \geq 0} \Big\{ P_1(t)\bar{r} k_t - P_2(t)\bar{r} u'(c_t^c)/u''(c_t^c) + P_3 e^{-\rho t} \gamma v\big(f(k_t) - \delta k_t - \bar{r} k_t\big) \Big\} \quad \forall\, t \in (0, T). \tag{29}$$

The first order necessary condition for this problem is

$$P_1(t) k_t - P_2(t) u'(c_t^c)/u''(c_t^c) - P_3 e^{-\rho t} \gamma k_t v'\big(f(k_t) - \delta k_t - \bar{r}_t k_t\big) = \breve{\eta}_t \leq 0 \ \ \& \ \ \breve{\eta}_t \bar{r}_t = 0 \quad \forall\, t \in (0, T) \tag{30}$$

where $\breve{\eta}_t$ is a Lagrange multiplier.

We now use the definition of ϕ to evaluate its derivative:[18]

$$\phi_{t_0}(t_0,t_1,x_0,x_1) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \phi_{t_1}(t_0,t_1,x_0,x_1) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\phi_{x_0}(t_0,t_1,x_0,x_1) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \phi_{x_1}(t_0,t_1,x_0,x_1) = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Substitute into the transversality conditions to get the following, where we use the result above that P_3 is a constant:

$$\begin{aligned} \begin{pmatrix} P_1(T) & P_2(T) & P_3 \end{pmatrix} &= \begin{pmatrix} \lambda_5 & \lambda_6 & -\lambda_1 \end{pmatrix} \\ \begin{pmatrix} P_1(0) & P_2(0) & P_3 \end{pmatrix} &= \begin{pmatrix} -\lambda_4 & 0 & -\lambda_1 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} P_1(T)(\bar{r}_T k_T - c_T^c) + P_2(T)(\rho - \bar{r}_T)u'(c_T^c)/u''(c_T^c) + P_3 e^{-\rho T}\big[\gamma v\big(f(k_T) - \delta k_T - \bar{r}_T k_T\big) + (1-\gamma)u(c_T^c)\big] &= -\lambda_3 \\ P_1(0)(\bar{r}_0 k_0 - c_0^c) + P_2(0)(\rho - \bar{r}_0)u'(c_0^c)/u''(c_0^c) + P_3\big[\gamma v\big(f(k_0) - \delta k_0 - \bar{r}_0 k_0\big) + (1-\gamma)u(c_0^c)\big] &= \lambda_2. \end{aligned}$$

The last transversality condition is the only place where λ_2 appears so this equation serves as the definition of λ_2 but plays no other role in the solution to the optimal taxation problem. Similarly, the penultimate transversality condition defines λ_3 but plays no other role. The other transversality conditions define $\lambda_4 = -P_1(0)$, $\lambda_5 = P_1(T)$, and $\lambda_6 = P_2(T)$. We also have $\lambda_1 = -P_3$. Recall from the statement of the Pontryagin necessary conditions that $\lambda_1 \leq 0$. Thus, the transversality conditions provide us with the following information:[19]

$$P_2(0) = 0 \quad \& \quad P_3 \geq 0.$$

By way of contradiction, suppose $(P_1(0), P_3) = (0,0)$. Then, together with the transversality condition $P_2(0) = 0$, the unique solution to the differential equations (27) and (28) is $P_1(t) \equiv 0$ and $P_2(t) \equiv 0$. But then the transversality conditions yield $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \lambda_5 = \lambda_6 = 0$ which is a violation of the

[18] In the finite horizon (T) optimal taxation problem, ϕ is a function from $\mathbf{R}^8$ to $\mathbf{R}^5$. See footnote 17. In this case the derivative of ϕ would include only the first 5 rows shown here. Furthermore, the non-zero vector λ would have only 5 components, $(\lambda_1, \ldots, \lambda_5)$.

[19] In the finite horizon (T) optimal taxation problem, there is no sixth component of ϕ and there is no λ_6. See footnote 18. So where λ_6 appears in the transversality conditions in the text, it would be replaced with 0 for the finite horizon optimal taxation problem. In this case, the transversality conditions would provide us with the following information: $P_2(0) = P_2(T) = 0$ & $P_3 \geq 0$. The necessary conditions for optimality would thus be this information together with (27), (28), and (29)/(30).

Pontryagin necessary conditions for optimality. We conclude that the assumption $(P_1(0), P_3) = (0,0)$ leads to a contradiction so, after a normalization, we have[20]

$$P_2(0) = 0 \quad \& \quad P_3 \geq 0 \quad \& \quad \|(P_1(0), P_3)\| = 1. \tag{31}$$

In summary, if $\{(\bar{r}_t, k_t, c_t^c)\}_{t\geq 0}$ solves the infinite horizon optimal taxation problem then it also solves the finite horizon (T) optimal taxation problem with clamped terminal state and hence there exists a function $P : [0,T] \to \mathbf{R}^2$ and there exists a number P_3 such that (27), (28), (29), (30), and (31) are satisfied.

Following Halkin (1974) we consider the finite horizon optimal taxation problem with clamped terminal state for a sequence of time horizons $T^1, T^2, \ldots$ with $\lim_{i\to\infty} T^i = \infty$. As above, for each i there exists a function $P^i : [0, T^i] \to \mathbf{R}^2$ and there exists a number P_3^i such that (27), (28), (29), (30), and (31) are satisfied when the control and state are given by the solution to the infinite horizon optimal taxation problem. Since, by (31), $\|(P_1^i(0), P_3^i)\| = 1$ for all i, there exists a subsequence for which $(P_1^{i_j}(0), P_3^{i_j})$ converges. For ease of notation, and without loss of generality, assume the convergence occurs along the original sequence: $\lim_{i\to\infty}(P_1^i(0), P_3^i) = (P_1(0), P_3)$. We have $P_3 \geq 0$ and $\|(P_1(0), P_3)\| = 1$ since, by (31), these conditions are satisfied for all i. Similarly since $P_2^i(0) = 0$ for all i, if we define $P_2(0) := \lim_{i\to\infty} P_2^i(0)$ then $P_2(0) = 0$.

For $t \in [0, \infty)$ consider the differential equations (27), (28), and $\dot{P}_3(t) \equiv 0$ with initial conditions $(P_1(0), P_2(0), P_3) = \lim_{i\to\infty}(P_1^i(0), P_2^i(0), P_3^i)$ as in the previous paragraph. It should be understood that where the control and state appear in these equations their values are the solution to the infinite horizon optimal taxation problem. Let $P : [0, \infty) \to \mathbf{R}^3$ denote the solution to these differential equations with these initial conditions. Since solutions to differential equations are continuous in initial conditions[21] and since $\lim_{i\to\infty} T^i = \infty$ we have for all $t > 0$, $(P_1(t), P_2(t), P_3(t)) = \lim_{i\to\infty}(P_1^i(t), P_2^i(t), P_3^i)$. Then by continuity, $\{(P_1(t), P_2(t))\}_{t\geq 0}$ and P_3 satisfy (29) and (30).

In summary, we have shown that if $\{(\bar{r}_t, k_t, c_t^c)\}_{t\geq 0}$ solves the infinite horizon optimal taxation problem then there exists a function $P : [0, \infty) \to \mathbf{R}^2$ and there exists a number P_3 such that (27), (28), (29), (30), and (31) are satisfied with $[0,T]$ replaced by $[0, \infty)$ and with $(0,T)$ replaced by $(0, \infty)$. That is:

$$\dot{P}_1(t) \quad = \quad -P_1(t)\bar{r}_t - P_3 e^{-\rho t}\gamma v'(f(k_t) - \delta k_t - \bar{r}_t k_t)\big[f'(k_t) - \delta - \bar{r}_t\big] \quad \forall\, t \geq 0 \tag{32}$$

$$\dot{P}_2(t) \quad = \quad P_1(t) - P_2(t)(\rho - \bar{r}_t)\big\{1 - \big[u''(c_t^c)\big]^{-2} u'(c_t^c) u'''(c_t^c)\big\} - P_3 e^{-\rho t}(1-\gamma)u'(c_t^c) \quad \forall\, t \geq 0 \tag{33}$$

[20]In the finite horizon (T) optimal taxation problem, we have $P_2(0) = P_2(T) = 0$ & $P_3 \geq 0$. See footnote 19. By way of contradiction, if $P_3 = 0$ the first order linear differential equations (27) and (28) with boundary conditions $P_2(0) = P_2(T) = 0$ would yield $P_1(t) \equiv 0$ and $P_2(t) \equiv 0$. But then the transversality conditions would yield $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \lambda_5 = 0$ which would violate the Pontryagin necessary conditions for optimality. Thus, it must be that $P_3 > 0$.

[21]See section 4 of chapter 8 of Morris W. Hirsch and Stephen Smale, *Differential Equations, Dynamical Systems, and Linear Algebra*, San Diego: Academic Press, 1974.

$$\bar{r}_t \text{ solves } \max_{\bar{r}\geq 0}\Big\{P_1(t)\bar{r}k_t - P_2(t)\bar{r}u'(c_t^c)/u''(c_t^c) + P_3 e^{-\rho t}\gamma v\big(f(k_t)-\delta k_t - \bar{r}k_t\big)\Big\} \quad \forall\, t>0 \tag{34}$$

$$P_1(t)k_t - P_2(t)u'(c_t^c)/u''(c_t^c) - P_3 e^{-\rho t}\gamma k_t v'\big(f(k_t)-\delta k_t - \bar{r}_t k_t\big) = \breve{\eta}_t \leq 0 \ \ \& \ \ \breve{\eta}_t \bar{r}_t = 0 \quad \forall\, t>0 \tag{35}$$

$$P_2(0) = 0 \quad \& \quad P_3 \geq 0 \quad \& \quad \|(P_1(0), P_3)\| = 1. \tag{36}$$

We now ask if the system (32) through (36) can have a solution in which $P_3 = 0$. If so, then $P_1(0) = \pm 1$ from (36), while (32) and (33) yield

$$\begin{aligned}
P_1(t) &= P_1(0)e^{-\bar{R}_t} \quad \forall\, t\geq 0\\
\frac{d}{dt}\left[P_2(t)e^{\Omega_t}\right] &= P_1(t)e^{\Omega_t} \quad \forall\, t \geq 0\\
\text{where } \bar{R}_t &:= \int_0^t \bar{r}_s ds\\
\Omega_t &:= \int_0^t \omega_s ds\\
\omega_s &:= (\rho - \bar{r}_s)\big\{1 - \big[u''(c_s^c)\big]^{-2} u'(c_s^c)u'''(c_s^c)\big\}.
\end{aligned}$$

If we integrate the second of these equations and use the boundary condition $P_2(0) = 0$ from (36), we get

$$P_2(t)e^{\Omega_t} \ = \ \int_0^t P_1(s)e^{\Omega_s} ds \ = \ P_1(0)\int_0^t e^{(\Omega_s - \bar{R}_s)} ds$$

where the second equality uses the solution for $P_1(t)$ above. Substitute these solutions for $P_1(t)$ and $P_2(t)$, together with the assumed $P_3 = 0$, into (35):

$$P_1(0)\left\{e^{-\bar{R}_t}k_t - \left[e^{-\Omega_t}\int_0^t e^{(\Omega_s-\bar{R}_s)}ds\right]u'(c_t^c)/u''(c_t^c)\right\} = \breve{\eta}_t \leq 0 \ \ \& \ \ \breve{\eta}_t\bar{r}_t = 0 \quad \forall\, t>0.$$

Since $u'' < 0 < u'$ and since the exponential function is strictly positive, the term in curly braces is strictly positive. Then since $P_1(0) = \pm 1$ from (36), it follows that (35) requires $P_1(0) = -1$ and $\bar{r}_t = 0$ for all $t > 0$. In summary, the assumption $P_3 = 0$ leads to the conclusion that the optimal control for the infinite horizon optimal taxation problem must be $\bar{r}_t = 0$ for all $t > 0$.

The contrapositive of the result from the previous paragraph is the following: if $\bar{r}_t \equiv 0$ does not solve the infinite horizon optimal taxation problem (which is addressed in appendix E), then $P_3 \neq 0$, and hence from (36), $P_3 > 0$.[22] In this case we can let

$$q_{1t} := e^{\rho t}P_1(t)/P_3, \quad q_{2t} := e^{\rho t}P_2(t)/P_3, \quad \eta_t := -e^{\rho t}\breve{\eta}_t/P_3.$$

[22] In the finite horizon (T) optimal taxation problem, we have $P_3 > 0$ and this holds without any need to exclude $\bar{r}_t \equiv 0$ from being optimal. See footnote 20. So in this case we can apply the transformation that appears below in the text and find that (3) is necessary for optimality. But in addition, recall that for the finite horizon optimal taxation problem $P_2(T) = 0$. See footnote 19. Thus, in this case the necessary conditions for optimality include not only (3), but also $q_{2T} = 0$.

Then

$$\begin{aligned} P_1(t) &= P_3 e^{-\rho t} q_{1t} \\ \dot{P}_1(t) &= -\rho P_3 e^{-\rho t} q_{1t} + P_3 e^{-\rho t} \dot{q}_{1t} \\ P_2(t) &= P_3 e^{-\rho t} q_{2t} \\ \dot{P}_2(t) &= -\rho P_3 e^{-\rho t} q_{2t} + P_3 e^{-\rho t} \dot{q}_{2t} \\ \breve{\eta}_t &= -P_3 e^{-\rho t} \eta_t. \end{aligned}$$

Use these to substitute for $P_1(t)$, $\dot{P}_1(t)$, $P_2(t)$, $\dot{P}_2(t)$, and $\breve{\eta}_t$ in (32), (33), (34), (35), and (36). Then divide each of (32) through (36) by $P_3 e^{-\rho t} > 0$ to get

$$\begin{aligned} -\rho q_{1t} + \dot{q}_{1t} &= -q_{1t}\bar{r}_t - \gamma v'(f(k_t) - \delta k_t - \bar{r}_t k_t)\big[f'(k_t) - \delta - \bar{r}_t\big] \quad \forall\, t \geq 0 \\ -\rho q_{2t} + \dot{q}_{2t} &= q_{1t} - q_{2t}(\rho - \bar{r}_t)\big\{1 - \big[u''(c_t^c)\big]^{-2} u'(c_t^c) u'''(c_t^c)\big\} - (1-\gamma)u'(c_t^c) \quad \forall\, t \geq 0 \end{aligned}$$

$$\bar{r}_t \text{ solves } \max_{\bar{r} \geq 0} \Big\{ q_{1t}\bar{r}k_t - q_{2t}\bar{r}u'(c_t^c)/u''(c_t^c) + \gamma v\big(f(k_t) - \delta k_t - \bar{r}k_t\big)\Big\} \quad \forall\, t > 0$$

$$q_{1t}k_t - q_{2t}u'(c_t^c)/u''(c_t^c) - \gamma k_t v'\big(f(k_t) - \delta k_t - \bar{r}_t k_t\big) = -\eta_t \leq 0 \;\;\&\;\; \eta_t \bar{r}_t = 0 \quad \forall\, t > 0$$

$$q_{20} = 0.$$

The first two lines of these expressions coincide with (3a) and (3b) respectively. The penultimate line coincides with (3c) and the first part of (3f). The boundary condition $q_{20} = 0$ on the last line here appears in (3f).

E: Could $\bar{r}_t \equiv 0$ be a solution to the infinite horizon optimal taxation problem?

In section 3 we assumed the optimal time path for the after-tax net of depreciation interest rate is *not* $\bar{r}_t \equiv 0$. By excluding $\bar{r}_t \equiv 0$ from consideration, we were able to use the normal Pontryagin necessary conditions for optimality and exclude the abnormal case. However, proper analysis should not impose an assumption on the time path $\{\bar{r}_t\}_{t\geq 0}$ since it is endogenous to the optimal taxation problem. The purpose of this appendix is to show that under assumptions 1 and 2 below, $\bar{r}_t \equiv 0$ is not optimal. These assumptions are stated in terms of primitives (initial condition, parameter, production function, capitalist's utility function) and not in terms of endogenous variables. Where the analysis brushes over some technical details, this will be pointed out in the presentation.

We begin with some intuition. Recall the government's welfare objective $\int_0^\infty e^{-\rho t}[\gamma v(c_t^w)+(1-\gamma)u(c_t^c)]dt$. Our first result will be that if the worker's welfare weight γ is zero, then $\bar{r}_t \equiv 0$ cannot be optimal. This is obvious. The capitalist is dependent on interest income. So when all welfare weight is on the capitalist it cannot be optimal to tax away all interest income. The more interesting scenario is $\gamma > 0$. The worker's equilibrium consumption is $c_t^w = f(k_t) - \delta k_t - \bar{r}_t k_t$ which is adversely affected in the short run by an increase in $\bar{r}_t$. So if this short run effect is dominant then perhaps it could be optimal to set $\bar{r}_t \equiv 0$. But in the longer run $\bar{r}_t$ affects capital accumulation via $\dot{k}_t = \bar{r}_t k_t - c_t^c$ and capital affects the worker's consumption: $\partial c_t^w/\partial k_t = f'(k_t) - \delta - \bar{r}_t$. In particular, relative to the $\bar{r}_t \equiv 0$ equilibrium, an increase in the capital stock is desirable for the worker if $f'(k_t) - \delta > 0$. Thus, relative to the $\bar{r}_t \equiv 0$ equilibrium, if (i) an increase in $\bar{r}_t$ causes an increase in capital and if (ii) $f'(k_t) - \delta > 0$, then apparently the worker's longer run utility will improve if we increase $\bar{r}_t$ above zero. If this longer run effect is dominant then it would seem $\bar{r}_t \equiv 0$ is not optimal. This is indeed correct when conditions (i) and (ii) are formalized as assumptions 1 and 2 below. We now turn to the analysis.

Let $T \geq 0$ be given and let $\epsilon \geq 0$ be given. Consider the following time path:

$$\bar{r}_t = \begin{cases} 0 & \text{if } 0 \leq t \leq T \\ \epsilon & \text{if } T < t. \end{cases}$$

All results in this appendix are based on this family of variations, parametrized by T and ϵ. Note that $\epsilon = 0$ yields $\bar{r}_t \equiv 0$, the object of study here. If, within this family of variations, $\epsilon = 0$ is not optimal for the optimal taxation problem, then surely $\bar{r}_t \equiv 0$ does not solve the optimal taxation problem more generally. With this parametrized time path for $\{\bar{r}_t\}_{t\geq 0}$, cumulative interest is then

$$\bar{R}_t := \int_0^t \bar{r}_s ds = \begin{cases} 0 & \text{if } 0 \leq t \leq T \\ \epsilon(t-T) & \text{if } T < t. \end{cases}$$

The equilibrium is as follows. The solution to the capitalist's utility maximization problem is given by (1): $e^{-\rho t}u'(c_t^c)/u'(c_0^c) = e^{-\bar{R}_t}$ and $k_0 = \int_0^\infty e^{-\bar{R}_s} c_s^c ds$. The worker's consumption is given by (2):

$c_t^w = f(k_t) - \delta k_t - \bar{r}_t k_t$. The capital accumulation equation is $\dot{k}_t = \bar{r}_t k_t - c_t^c$ which is equivalent to $e^{-\bar{R}_t} k_t - k_0 = -\int_0^t e^{-\bar{R}_s} c_s^c ds$. With $\bar{r}_t$ and $\bar{R}_t$ as above, we have

$$
\begin{aligned}
u'(c_t^c) &= \begin{cases} e^{\rho t} u'(c_0^c) & \text{if } 0 \leq t \leq T \\ e^{\rho t - \epsilon(t-T)} u'(c_0^c) & \text{if } T < t \end{cases} \\
k_t &= k_0 - \textstyle\int_0^t c_s^c ds \quad \text{if } 0 \leq t \leq T \\
e^{-\epsilon(t-T)} k_t &= k_0 - \textstyle\int_0^T c_s^c ds - \int_T^t e^{-\epsilon(s-T)} c_s^c ds \quad \text{if } T < t \\
k_0 &= \textstyle\int_0^T c_s^c ds + \int_T^\infty e^{-\epsilon(s-T)} c_s^c ds \\
c_t^w &= \begin{cases} f(k_t) - \delta k_t & \text{if } 0 \leq t \leq T \\ f(k_t) - \delta k_t - \epsilon k_t & \text{if } T < t. \end{cases}
\end{aligned}
$$

Social welfare is $W(\epsilon, T) := \int_0^\infty e^{-\rho t}[\gamma v(c_t^w) + (1-\gamma) u(c_t^c)] dt$ where the consumption levels are evaluated at the (ϵ, T) equilibrium. The government faces the policy constraint $\bar{r}_t \geq 0$. Thus if there exists T such that $W_\epsilon(0, T) > 0$, then $\bar{r}_t \equiv 0$ cannot solve the optimal taxation problem. Our goal for the remainder of this appendix is to evaluate the partial derivative $W_\epsilon(0, T)$ and show that it is positive for sufficiently large T when assumptions 1 and 2 below are satisfied. We do not use the Laplace transform method of Judd (1985) since the baseline $\epsilon = 0$ equilibrium is not in steady state.

Assuming we can differentiate under the integral sign and all integrals converge, W_ϵ is given by

$$
\begin{aligned}
W_\epsilon(\epsilon, T) &= \int_0^\infty e^{-\rho t} \left[\gamma v'(c_t^w) \frac{\partial c_t^w}{\partial \epsilon} + (1-\gamma) u'(c_t^c) \frac{\partial c_t^c}{\partial \epsilon} \right] dt \\
&= \int_0^\infty e^{-\rho t} \gamma v'(c_t^w) \frac{\partial c_t^w}{\partial \epsilon} dt + (1-\gamma) u'(c_0^c) \int_0^T \frac{\partial c_t^c}{\partial \epsilon} dt + (1-\gamma) u'(c_0^c) \int_T^\infty e^{-\epsilon(t-T)} \frac{\partial c_t^c}{\partial \epsilon} dt \quad (37)
\end{aligned}
$$

where the second line uses the capitalist's first order condition as presented above.

We now proceed to differentiate each of the equilibrium equations with respect to ϵ. We again assume differentiation under the integral sign is justified and all integrals converge:

$$
\begin{aligned}
u''(c_t^c)\,\frac{\partial c_t^c}{\partial \epsilon} &= \begin{cases} e^{\rho t} u''(c_0^c)\frac{\partial c_0^c}{\partial \epsilon} & \text{if } 0 \le t \le T \\ e^{\rho t - \epsilon(t-T)}\left[-(t-T)u'(c_0^c) + u''(c_0^c)\frac{\partial c_0^c}{\partial \epsilon}\right] & \text{if } T < t \end{cases} \\
\frac{\partial k_t}{\partial \epsilon} &= -\int_0^t \frac{\partial c_s^c}{\partial \epsilon} ds \;=\; -u''(c_0^c)\frac{\partial c_0^c}{\partial \epsilon}\int_0^t \frac{e^{\rho s}}{u''(c_s^c)} ds \quad \text{if } 0 \le t \le T \\
e^{-\epsilon(t-T)}\left[-(t-T)k_t + \frac{\partial k_t}{\partial \epsilon}\right] &= -u''(c_0^c)\frac{\partial c_0^c}{\partial \epsilon}\int_0^T \frac{e^{\rho s}}{u''(c_s^c)} ds + \int_T^t e^{-\epsilon(s-T)}(s-T)c_s^c ds \\
&\quad - \int_T^t e^{-\epsilon(s-T)}\frac{e^{\rho s - \epsilon(s-T)}}{u''(c_s^c)}\left[-(s-T)u'(c_0^c) + u''(c_0^c)\frac{\partial c_0^c}{\partial \epsilon}\right] ds \quad \text{if } T < t \\
0 &= u''(c_0^c)\frac{\partial c_0^c}{\partial \epsilon}\int_0^T \frac{e^{\rho s}}{u''(c_s^c)} ds - \int_T^\infty e^{-\epsilon(s-T)}(s-T)c_s^c ds \\
&\quad + \int_T^\infty e^{-\epsilon(s-T)}\frac{e^{\rho s - \epsilon(s-T)}}{u''(c_s^c)}\left[-(s-T)u'(c_0^c) + u''(c_0^c)\frac{\partial c_0^c}{\partial \epsilon}\right] ds \\
\frac{\partial c_t^w}{\partial \epsilon} &= \begin{cases} [f'(k_t) - \delta]\,\frac{\partial k_t}{\partial \epsilon} & \text{if } 0 \le t \le T \\ -k_t + [f'(k_t) - \delta - \epsilon]\,\frac{\partial k_t}{\partial \epsilon} & \text{if } T < t. \end{cases}
\end{aligned}
$$

Throughout, we can use the capitalist's first order condition to replace $e^{\rho t}$ with $u'(c_t^c)/u'(c_0^c)$ if $0 \le t \le T$, and replace $e^{\rho t - \epsilon(t-T)}$ with $u'(c_t^c)/u'(c_0^c)$ if $T < t$. We will also combine the integrals that have $s - T$ as a multiplicative factor in the integrand. And for the sake of completeness, we also repeat the $\partial c_t^w/\partial \epsilon$ equation as is:

$$
\begin{aligned}
u''(c_t^c)\,\frac{\partial c_t^c}{\partial \epsilon} &= \begin{cases} u'(c_t^c)\,\frac{\partial c_0^c}{\partial \epsilon}\,u''(c_0^c)/u'(c_0^c) & \text{if } 0 \le t \le T \\ u'(c_t^c)\left[-(t-T) + \frac{\partial c_0^c}{\partial \epsilon}\,u''(c_0^c)/u'(c_0^c)\right] & \text{if } T < t \end{cases} \\
\frac{\partial k_t}{\partial \epsilon} &= -\left[\frac{\partial c_0^c}{\partial \epsilon}\,u''(c_0^c)/u'(c_0^c)\right]\int_0^t \frac{u'(c_s^c)}{u''(c_s^c)} ds \quad \text{if } 0 \le t \le T \\
e^{-\epsilon(t-T)}\left[-(t-T)k_t + \frac{\partial k_t}{\partial \epsilon}\right] &= -\left[\frac{\partial c_0^c}{\partial \epsilon}\,u''(c_0^c)/u'(c_0^c)\right]\int_0^T \frac{u'(c_s^c)}{u''(c_s^c)} ds \\
&\quad + \int_T^t e^{-\epsilon(s-T)}(s-T)\left[c_s^c + u'(c_s^c)/u''(c_s^c)\right] ds \\
&\quad - \left[\frac{\partial c_0^c}{\partial \epsilon}\,u''(c_0^c)/u'(c_0^c)\right]\int_T^t e^{-\epsilon(s-T)}\frac{u'(c_s^c)}{u''(c_s^c)} ds \quad \text{if } T < t \\
0 &= \left[\frac{\partial c_0^c}{\partial \epsilon}\,u''(c_0^c)/u'(c_0^c)\right]\int_0^T \frac{u'(c_s^c)}{u''(c_s^c)} ds \\
&\quad - \int_T^\infty e^{-\epsilon(s-T)}(s-T)\left[c_s^c + u'(c_s^c)/u''(c_s^c)\right] ds \\
&\quad + \left[\frac{\partial c_0^c}{\partial \epsilon}\,u''(c_0^c)/u'(c_0^c)\right]\int_T^\infty e^{-\epsilon(s-T)}\frac{u'(c_s^c)}{u''(c_s^c)} ds \\
\frac{\partial c_t^w}{\partial \epsilon} &= \begin{cases} [f'(k_t) - \delta]\,\frac{\partial k_t}{\partial \epsilon} & \text{if } 0 \le t \le T \\ -k_t + [f'(k_t) - \delta - \epsilon]\,\frac{\partial k_t}{\partial \epsilon} & \text{if } T < t. \end{cases}
\end{aligned}
$$

Since the goal is to determine the sign of $W_\epsilon(\epsilon, T)$ at $\epsilon = 0$, we now evaluate these equilibrium derivatives at $\epsilon = 0$, in which case $e^{-\epsilon(t-T)} = e^{-\epsilon(s-T)} = 1$. Hence, $\int_0^T \frac{u'(c_s^c)}{u''(c_s^c)} ds + \int_T^t e^{-\epsilon(s-T)}\frac{u'(c_s^c)}{u''(c_s^c)} ds = \int_0^t \frac{u'(c_s^c)}{u''(c_s^c)} ds$, whether t is finite or infinite. Furthermore, with $\epsilon = 0$ the capitalist's consumption Euler equation (3e) can be integrated to yield $\rho \int_0^t \frac{u'(c_s^c)}{u''(c_s^c)} ds = c_t^c - c_0^c$. And for the $t \to \infty$ version of this, note that with $\epsilon = 0$ the capitalist's first order condition is $u'(c_t^c) = e^{\rho t} u'(c_0^c)$, which implies $\lim_{t \to \infty} c_t^c = 0$, hence our integrated Euler equation in the limit is $\rho \int_0^\infty \frac{u'(c_s^c)}{u''(c_s^c)} ds = -c_0^c$. So we will now evaluate the equilibrium derivatives at $\epsilon = 0$ and replace $\int_0^t \frac{u'(c_s^c)}{u''(c_s^c)} ds$ with $(c_t^c - c_0^c)/\rho$ and replace $\int_0^\infty \frac{u'(c_s^c)}{u''(c_s^c)} ds$ with $-c_0^c/\rho$:

$$
\begin{aligned}
u''(c_t^c) \tfrac{\partial c_t^c}{\partial \epsilon}|_{\epsilon=0} &= \begin{cases} u'(c_t^c) \frac{\partial c_0^c}{\partial \epsilon}|_{\epsilon=0} \, u''(c_0^c)/u'(c_0^c) & \text{if } 0 \le t \le T \\ u'(c_t^c) \left[-(t-T) + \frac{\partial c_0^c}{\partial \epsilon}|_{\epsilon=0} \, u''(c_0^c)/u'(c_0^c) \right] & \text{if } T < t \end{cases} \\
\tfrac{\partial k_t}{\partial \epsilon}|_{\epsilon=0} &= -\left[\tfrac{\partial c_0^c}{\partial \epsilon}|_{\epsilon=0} \, u''(c_0^c)/u'(c_0^c) \right] (c_t^c - c_0^c)/\rho \quad \text{if } 0 \le t \le T \\
\tfrac{\partial k_t}{\partial \epsilon}|_{\epsilon=0} &= (t-T)k_t - \left[\tfrac{\partial c_0^c}{\partial \epsilon}|_{\epsilon=0} \, u''(c_0^c)/u'(c_0^c) \right] (c_t^c - c_0^c)/\rho \\
&\quad + \textstyle\int_T^t (s-T) \left[c_s^c + u'(c_s^c)/u''(c_s^c) \right] ds \quad \text{if } T < t \\
0 &= \left[\tfrac{\partial c_0^c}{\partial \epsilon}|_{\epsilon=0} \, u''(c_0^c)/u'(c_0^c) \right] (-c_0^c/\rho) - \textstyle\int_T^\infty (s-T) \left[c_s^c + u'(c_s^c)/u''(c_s^c) \right] ds \\
\tfrac{\partial c_t^w}{\partial \epsilon}|_{\epsilon=0} &= \begin{cases} [f'(k_t) - \delta] \frac{\partial k_t}{\partial \epsilon}|_{\epsilon=0} & \text{if } 0 \le t \le T \\ -k_t + [f'(k_t) - \delta] \frac{\partial k_t}{\partial \epsilon}|_{\epsilon=0} & \text{if } T < t \end{cases}
\end{aligned}
$$

where it is implicit that all economic variables are evaluated at the $\epsilon = 0$ equilibrium. From the penultimate equation:

$$
\left. \frac{\partial c_0^c}{\partial \epsilon} \right|_{\epsilon=0} c_0^c u''(c_0^c)/u'(c_0^c) \;=\; -\rho \int_T^\infty (s-T) \left[c_s^c + u'(c_s^c)/u''(c_s^c) \right] ds. \tag{38}
$$

Note that if $u = \log$ we have $c + u'(c)/u''(c) \equiv 0$, so $\partial c_0^c / \partial \epsilon = 0$ since $u = \log$ implies $c_0^c = \rho k_0$ regardless of the time path for $\{\bar{r}_t\}_{t \ge 0}$.

We now evaluate W_ϵ in (37) at $\epsilon = 0$ and substitute for $(\partial c_t^w / \partial \epsilon)|_{\epsilon=0}$ and for $(\partial c_t^c / \partial \epsilon)|_{\epsilon=0}$ using the results above:

$$
\begin{aligned}
W_\epsilon(0,T) &= \int_0^\infty e^{-\rho t} \gamma v'(c_t^w) \left. \frac{\partial c_t^w}{\partial \epsilon} \right|_{\epsilon=0} dt + (1-\gamma) u'(c_0^c) \int_0^\infty \left. \frac{\partial c_t^c}{\partial \epsilon} \right|_{\epsilon=0} dt \\
&= \int_0^\infty e^{-\rho t} \gamma v'(c_t^w) [f'(k_t) - \delta] \left. \frac{\partial k_t}{\partial \epsilon} \right|_{\epsilon=0} dt - \int_T^\infty e^{-\rho t} \gamma v'(c_t^w) k_t dt \\
&\quad + (1-\gamma) \left. \frac{\partial c_0^c}{\partial \epsilon} \right|_{\epsilon=0} u''(c_0^c) \int_0^\infty \frac{u'(c_t^c)}{u''(c_t^c)} dt - (1-\gamma) u'(c_0^c) \int_T^\infty (t-T) \frac{u'(c_t^c)}{u''(c_t^c)} dt.
\end{aligned}
$$

On the last line we can substitute for $(\partial c_0^c / \partial \epsilon)|_{\epsilon=0}$ from (38), and recall our previous result $\int_0^\infty \frac{u'(c_s^c)}{u''(c_s^c)} ds = -c_0^c/\rho$ when $\epsilon = 0$. For now, we leave the other line as it is:

$$
\begin{aligned}
W_\epsilon(0,T) &= \int_0^\infty e^{-\rho t} \gamma v'(c_t^w) [f'(k_t) - \delta] \left. \frac{\partial k_t}{\partial \epsilon} \right|_{\epsilon=0} dt - \int_T^\infty e^{-\rho t} \gamma v'(c_t^w) k_t dt \\
&\quad + (1-\gamma) u'(c_0^c) \int_T^\infty (t-T) c_t^c dt.
\end{aligned} \tag{39}
$$

This gives a first result which, as mentioned in the introductory remarks to this appendix, is intuitively obvious:

- If the worker's welfare weight is zero, i.e., $\gamma = 0$, then $W_\epsilon(0,T) > 0$ for all T and so $\bar{r}_t \equiv 0$ cannot solve the optimal taxation problem.

We now consider $\gamma > 0$. Use the equilibrium derivatives to substitute for $(\partial k_t/\partial\epsilon)|_{\epsilon=0}$ in formula (39) for $W_\epsilon(0,T)$:

$$\begin{aligned} W_\epsilon(0,T) \quad = \quad & -\left[\left.\frac{\partial c_0^c}{\partial \epsilon}\right|_{\epsilon=0} u''(c_0^c)/u'(c_0^c)\right]\int_0^\infty e^{-\rho t}\gamma v'(c_t^w)[f'(k_t)-\delta]\big((c_t^c-c_0^c)/\rho\big)dt \\ & + \int_T^\infty e^{-\rho t}\gamma v'(c_t^w)[f'(k_t)-\delta](t-T)k_t dt \\ & + \int_T^\infty e^{-\rho t}\gamma v'(c_t^w)[f'(k_t)-\delta]\left[\int_T^t (s-T)\left[c_s^c + u'(c_s^c)/u''(c_s^c)\right]ds\right]dt \\ & - \int_T^\infty e^{-\rho t}\gamma v'(c_t^w)k_t dt \\ & + (1-\gamma)u'(c_0^c)\int_T^\infty (t-T)c_t^c dt. \end{aligned} \tag{40}$$

For the remainder of this appendix we make two assumptions:

- **Assumption 1** Either the capitalist's utility function is $u = \log$, or else $\lim_{c\to 0} cu''(c)/u'(c) \in (-1,0)$.

  For the latter alternative, the assumption is that this limit exists and is strictly between -1 and zero. This is satisfied by, among many others, $u(c) = (c^{1-1/\sigma}-1)/(1-1/\sigma)$ with the elasticity of intertemporal substitution $\sigma > 1$.

  If $u = \log$ we have $c + u'(c)/u''(c) \equiv 0$, so not only does (38) yield $(\partial c_0^c/\partial\epsilon)|_{\epsilon=0} = 0$ thereby causing the first line of (40) to equal zero, but also the third line of (40) equals zero too. Alternatively, if $\lim_{c\to 0} cu''(c)/u'(c) \in (-1,0)$ then $c + u'(c)/u''(c) < 0$ for all c sufficiently small. We will show that this implies that the sum of the first and third lines of (40) is positive if T is sufficiently large. Note, from appendix A, savings (and hence capital) are an increasing function of the interest rate if future consumption satisfies $c + u'(c)/u''(c) < 0$. In effect, we are assuming that an increase in $\bar{r}_t$ causes an increase in capital.

- **Assumption 2** The initial stock of capital satisfies $f'(k_0) - \delta \geq 0$.

  That is, at the beginning of time the pre-tax net of depreciation interest rate is not negative. It then increases monotonically as time evolves since, with $\epsilon = 0$, capital's law of motion is $\dot{k}_t = -c_t^c$ so capital declines monotonically and its marginal product rises monotonically. Assumption 2 is stronger than section 3's $f(k_0) - \delta k_0 > 0$.

We now consider $\lim_{c\to 0} cu''(c)/u'(c) \in (-1,0)$ from assumption 1, and hence $c+u'(c)/u''(c) < 0$ for all c sufficiently small. We shall show that the sum of the first and third lines of (40) is positive if T is sufficiently large. With $\epsilon = 0$ the capitalist's first order condition is $e^{\rho s}u'(c_0^c) = u'(c_s^c)$, which implies $\lim_{s\to\infty} c_s^c = 0$. Thus under assumption 1 there exists T_1 such that if $s \geq T_1$ then $c_s^c + u'(c_s^c)/u''(c_s^c) < 0$. Let $T \geq T_1$ and write the sum of the first and third lines of (40) as follows, where we use (38) to substitute for $(\partial c_0^c/\partial\epsilon)|_{\epsilon=0}$:

$$
\begin{aligned}
&\text{line 1 of (40)} + \text{line 3 of (40)} \\
&= \int_T^\infty (s-T)\,[c_s^c + u'(c_s^c)/u''(c_s^c)]\,ds \times \\
&\qquad \left\{ \int_0^\infty e^{-\rho t}\gamma v'(c_t^w)[f'(k_t)-\delta]\left(\frac{c_t^c - c_0^c}{c_0^c}\right) dt \right. \\
&\qquad \left. + \int_T^\infty e^{-\rho t}\gamma v'(c_t^w)[f'(k_t)-\delta]\left[\frac{\int_T^t (s-T)\,[c_s^c + u'(c_s^c)/u''(c_s^c)]\,ds}{\int_T^\infty (s-T)\,[c_s^c + u'(c_s^c)/u''(c_s^c)]\,ds}\right] dt \right\}.
\end{aligned}
\tag{41}
$$

The symbol $\times$ at the end of the second line of (41) denotes multiplication. Note that the integral on the second line of (41) also appears in the denominator on the fourth line. Since $T \geq T_1$, everywhere that $c_s^c + u'(c_s^c)/u''(c_s^c)$ appears in (41), it is strictly negative. And under assumption 2, everywhere that $f'(k_t)-\delta$ appears in (41), it is strictly positive. With $\epsilon = 0$ the capitalist's first order condition is $u'(c_t^c) = e^{\rho t}u'(c_0^c)$ which implies $c_t^c < c_0^c$ for all $t > 0$ so the third line of (41) is strictly negative and does not depend on T. On the fourth line of (41), the ratio of integrals has a value of zero when the dummy variable t equals T and the value of this ratio increases monotonically as a function of t towards a limit of one as $t \to \infty$, and this holds regardless of the value of T. Therefore, the fourth line of (41) converges to zero as $T \to \infty$. In particular, there exists $T_2 \geq T_1$ such that if $T \geq T_2$ then the entire expression in curly braces in (41) is strictly negative. Since the integral on the second line of (41) is strictly negative, we conclude that the entirety of (41) is strictly positive for all $T \geq T_2$.

Thus far we have shown that under assumptions 1 and 2 the sum of the first, third, and fifth lines of (40) is positive if $T \geq T_2$ (where we can define $T_2 := 0$ under the $u = \log$ alternative in assumption 1). We now address the second and fourth lines.

$$
\begin{aligned}
\text{line 2 of (40)} + \text{line 4 of (40)} &= \int_T^\infty e^{-\rho t}\gamma v'(c_t^w)k_t\left[\big(f'(k_t)-\delta\big)(t-T)-1\right]dt \\
&= \int_0^\infty e^{-\rho(s+T)}\gamma v'(c_{s+T}^w)k_{s+T}\left[\big(f'(k_{s+T})-\delta\big)s-1\right]ds \\
&= e^{-\rho T}\gamma v'(c_{S(T)+T}^w)k_T \times \\
&\qquad \left\{\int_0^{S(T)} e^{-\rho s}\frac{v'(c_{s+T}^w)}{v'(c_{S(T)+T}^w)}\,\frac{k_{s+T}}{k_T}\left[\big(f'(k_{s+T})-\delta\big)s-1\right]ds\right. \\
&\qquad \left.+\int_{S(T)}^\infty e^{-\rho s}\frac{v'(c_{s+T}^w)}{v'(c_{S(T)+T}^w)}\,\frac{k_{s+T}}{k_T}\left[\big(f'(k_{s+T})-\delta\big)s-1\right]ds\right\}. \quad (42)
\end{aligned}
$$

On the fourth and fifth lines of (42) the integrand is the same but the limits of integration differ. We need to define the integration limit $S(T)$. The term in square brackets, namely $\big(f'(k_{s+T})-\delta\big)s-1$, has a value of negative one when $s=0$ and under our assumptions it increases monotonically as a function of s towards a limit of infinity as $s\to\infty$. (Recall that when $\epsilon=0$, $\dot{k}_t=-c_t^c<0$.) Therefore there exists a unique $s>0$, denoted $S(T)$, such that $\big(f'(k_{S(T)+T})-\delta\big)S(T)-1=0$. As a consequence of this definition for $S(T)$, the fourth line of (42) is strictly negative since the dummy variable satisfies $s\le S(T)$ and the fifth line is strictly positive. Furthermore, by implicit differentiation we have

$$
S'(T) \;=\; \frac{-S(T)f''(k_{S(T)+T})\dot{k}_{S(T)+T}}{f'(k_{S(T)+T})-\delta+S(T)f''(k_{S(T)+T})\dot{k}_{S(T)+T}} \;<\; 0.
$$

We can show that $\lim_{T\to\infty}S(T)=0$: Since $S(T)>0$, we have $\lim_{T\to\infty}k_{S(T)+T}=0$ from the capitalist's no-Ponzi condition when $\epsilon=0$, and hence $\lim_{T\to\infty}[f'(k_{S(T)+T})-\delta]=\infty$. The result $\lim_{T\to\infty}S(T)=0$ now follows from the equation $\big(f'(k_{S(T)+T})-\delta\big)S(T)-1=0$ that implicitly defines $S(T)$. Note that

$$
\frac{v'(c_{s+T}^w)}{v'(c_{S(T)+T}^w)} \;=\; \frac{v'\big(f(k_{s+T})-\delta k_{s+T}\big)}{v'\big(f(k_{S(T)+T})-\delta k_{S(T)+T}\big)} \;<\; 1 \quad \text{for all } s\in[0,S(T)) \tag{43}
$$

where we have used the following: $c_t^w=f(k_t)-\delta k_t$ when $\epsilon=0$; $f'(k_t)-\delta>0$ under assumption 2; and $\dot{k}_t<0$ when $\epsilon=0$. Thus the integrand on the fourth line of (42) has an absolute value less than one. (From the definition of $S(T)$, the term in square brackets in the integrand has a value between negative one and zero.) Noting the range of integration, we conclude that the fourth line of (42) is negative with an absolute value less than $S(T)$.

In preparation for analysis of the fifth line of (42), note that with $\epsilon=0$, l'Hopital's rule and the equilibrium laws of motion (3d), (3e) yield

$$
\lim_{T\to\infty}\frac{k_T}{c_T^c} \;=\; \lim_{T\to\infty}\frac{\dot{k}_T}{\dot{c}_T^c} \;=\; \lim_{T\to\infty}\frac{-c_T^c}{\rho u'(c_T^c)/u''(c_T^c)} \;=\; -\frac{1}{\rho}\lim_{c\to 0}\frac{cu''(c)}{u'(c)} \tag{44}
$$

and by assumption this limit exists and is strictly positive. Since we also have $\lim_{T\to\infty} S(T) = 0$, it follows that there exists $T_3 \geq T_2$ such that if $T \geq T_3$ then $k_T/c^c_T > S(T)$. Now consider the fifth line of (42) with $T \geq T_3$:

$$
\begin{aligned}
&\int_{S(T)}^{\infty} e^{-\rho s} \frac{v'(c^w_{s+T})}{v'(c^w_{S(T)+T})} \frac{k_{s+T}}{k_T} \left[\big(f'(k_{s+T}) - \delta\big)s - 1\right] ds \\
\geq\ & \int_{S(T)}^{\infty} e^{-\rho s} \frac{k_{s+T}}{k_T} \left[\big(f'(k_{s+T}) - \delta\big)s - 1\right] ds \\
& \qquad \text{by similar reasoning to (43)} \\
\geq\ & \int_{S(T)}^{\infty} e^{-\rho s} \frac{k_{s+T}}{k_T} \left[\big(f'(k_{S(T)+T}) - \delta\big)s - 1\right] ds \\
& \qquad \text{because with } \epsilon = 0, s \geq S(T) \Rightarrow f'(k_{s+T}) > f'(k_{S(T)+T}) \\
\geq\ & \int_{S(T)}^{k_T/c^c_T} e^{-\rho s} \frac{k_{s+T}}{k_T} \left[\big(f'(k_{S(T)+T}) - \delta\big)s - 1\right] ds \\
& \qquad \text{because } T \geq T_3 \text{ and the integrand is non-negative} \\
\geq\ & \int_{S(T)}^{k_T/c^c_T} e^{-\rho s} \frac{k_T + s\dot{k}_T}{k_T} \left[\big(f'(k_{S(T)+T}) - \delta\big)s - 1\right] ds \\
& \qquad \text{because with } \epsilon = 0,\ t \mapsto k_t \text{ is convex by (3d), (3e)} \\
=\ & \int_{S(T)}^{k_T/c^c_T} e^{-\rho s} (1 - sc^c_T/k_T)\big(s/S(T) - 1\big) ds \\
& \qquad \text{from (3d) for } \dot{k}_T \text{, and from the equation that implicitly defines } S(T) \\
\geq\ & e^{-\rho k_T/c^c_T} \int_{S(T)}^{k_T/c^c_T} (1 - sc^c_T/k_T)\big(s/S(T) - 1\big) ds \\
=\ & e^{-\rho k_T/c^c_T} \frac{\big(k_T/c^c_T - S(T)\big)^3}{6S(T)k_T/c^c_T}
\end{aligned}
$$

where the last line follows from direct calculation of the integral. Since $\lim_{T\to\infty} S(T) = 0$ and since $\lim_{T\to\infty} k_T/c^c_T$ is strictly positive by (44), we conclude that the fifth line of (42) tends to infinity as $T \to \infty$.

Combine this current result with our previous result that the fourth line of (42) is negative with an absolute value less than $S(T)$. Thus there exists $T_4 \geq T_3$ such that if $T \geq T_4$ then the entire expression in (42) is strictly positive. That is, the sum of the second and fourth lines of (40) is strictly positive. Previously we showed that the sum of the first, third, and fifth lines of (40) is positive if $T \geq T_2$. This brings us to the final conclusion of this appendix: Under assumptions 1 and 2, if $T \geq T_4$ then the entire expression in (40) is strictly positive, i.e., $W_\epsilon(0,T) > 0$ and so $\epsilon = 0$ cannot be optimal for the optimal taxation problem, and more generally, $\bar{r}_t \equiv 0$ cannot be optimal for the optimal taxation problem.