

Centro Atlántico de Estudios Ríos Pérez

Time-Consistent Public
Expenditures

Paul Klein, Per Krusell, José-Víctor Ríos-Rull

CAERP

Documento de Trabajo #15

Working Paper #15

Time–Consistent Public Expenditures

Paul Klein,^a Per Krusell,^b José-Víctor Ríos-Rull^{c, *}

^a*University of Western Ontario*

^b*University of Rochester, IIES, CAERP*

^c*University of Pennsylvania, CAERP*

First draft: 10 August 2001

This version: January 9, 2004

Abstract

In this paper, we study the optimal choice of public expenditures when there is no way of committing to future policy and “reputational” mechanisms are not operative. This amounts to confining our attention to Markov equilibria. The environment is a neoclassical growth model where consumers derive utility from a public good. This environment gives rise to a dynamic game between successive governments and the private sector and this game is made interesting by the presence of a state variable: the capital stock. We characterize equilibria in terms of an intertemporal first–order condition (a “generalized Euler equation”, GEE) for the government and we use this condition both to gain insight into the nature of the equilibrium and as a basis for computation.

The GEE reveals how the government optimally trades off tax wedges over time. It also allows us to discuss in what sense a current government may be strategically influencing future governments in their taxation decisions. For a calibrated economy, we find that when the tax base available to the government is capital income—an inelastic source of funds at any moment in time—the government still refrains from taxing at high rates in order to smooth distortions over time. As a result, the economy is far from the mix of public and private goods that would be optimal in a static context; in return, the savings distortion is not very severe and the capital stock is quite high.

Keywords: Time–consistency, Markov–perfect equilibrium, optimal taxation.

JEL classification: E62, H21.

*Correspondence to Krusell at the Department of Economics, University of Rochester. Krusell thanks the National Science Foundation for support and Ríos-Rull thanks the National Science Foundation (Grant SES–0079504) and the University of Pennsylvania Research Foundation for their support.

1 Introduction

Governments are not machines implementing past decisions. Instead, at each point in time they can choose whether or not to implement a plan made in the past. In this paper we attempt to understand the qualitative and quantitative tradeoffs that a benevolent government faces when it is unable to commit to its future policy choices. We analyze a very simple setup in which the solution under commitment exhibits dynamic inconsistency. There is a choice between private and public consumption at each point in time, and because the financing of public consumption must involve distortionary taxation (e.g., capital income is taxed), as time evolves the government would in general want to modify any plan made in the past. We assume period-by-period budget balance but focus on the case where there are physical implications of current policy on future feasible sets, through the effect of policy on capital accumulation. Our main contributions here are to develop analytical tools for the study of time-consistent Markov-perfect equilibria and to evaluate the quantitative significance—using a calibration based on U.S. data—of the lack of commitment in the public spending choice.

An understanding of what guides government decisions regarding the provision of public goods when there is no commitment is not only a challenging intellectual exercise, but we also believe it is of quantitative importance. Public goods are, in all developed economies, a sizeable fraction of GDP, and the taxes necessary to pay for these expenditures are arguably an important determinant of economic performance, especially if taxes on the income of accumulated factors such as physical and human capital are considered. In order to understand these policy outcomes—both the cross-country and time-series variation of public spending on goods and services and the associated tax choices—we need to understand how governments make their decisions. Of course, in contrast with what we assume, governments may not be benevolent, or they may be torn between constituencies with conflicting goals, and the political process itself, as well as markets, may be less than perfect. However, before proceeding to such arguably more realistic setups, one needs to understand the underlying basics of policy choice over time when commitment is lacking even under benevolence and no frictions other than those implied by taxation itself. Our goal here is to develop those basics.

The lack of commitment makes the policy choice of the government into more than a

decision problem: the current government needs to predict how future governments—which it cannot directly influence—respond to current policy. We embed the choice of public goods into a framework with capital accumulation: we use the neoclassical growth model. This is done for two reasons. First, with capital as a state variable, the implications of current policy for the future capital stock become an important channel through which current and future policy are linked. In other words, we are in a world with fundamental dynamic links, as opposed to a repetition of static economies. Second, because taxation in practice involves distortions to factor accumulation, which are believed to be important for welfare as well as for long-run output, a quantitatively aimed analysis needs to consider accumulation; the neoclassical growth model is the standard workhorse for such analysis. Thus, the setup we study boils down to a sequential game between successive governments, with the restriction that allocations need to be competitive equilibrium allocations. The bulk of the analysis here is devoted to understanding how such an economy evolves over time both qualitatively and quantitatively.

The easiest way to illustrate some basics of our setup is to suppose that the economy has a finite time horizon, say, two periods, and that the government can tax total income at a proportional rate subject to a balanced budget period by period. In the second period, the government makes a policy choice given the resources available in the economy, as implied by the amount of capital at that time. It therefore solves the static problem of balancing the tax wedge in the period—two labor—leisure choice with the wedge between the marginal utility of public and private goods at that time. The resulting decision is to tax income quite heavily, as at least one part of that income—capital income—is provided inelastically *ex post*. Predicting such tax rates, the private sector will save little in the first period. What is the resulting behavior of the government in the first period? It too would like to tax current (period—one) income at a high rate, but it also realizes that if it holds back on taxation a bit, thus sacrificing public goods and increasing the inefficiency in their current provision, it leaves resources in the hands of the private sector, and part of these resources will be saved, assuming that consumption at different points in time are normal goods and that the tax rate in the last period is not 100%. Thus, having somewhat lower taxation in the first period than in the second period is beneficial, as it will constitute a way of counteracting the future government's high *ex-post* taxation of (capital) income:

it will increase savings. One realizes that this line of reasoning becomes more potent with more periods: tax rates early on may be quite a bit lower than those in the last period. As a quantitative matter, one would like to know how important this channel is in counteracting the negative aspects of a lack of commitment.

Our analysis in this paper amounts to analyzing economies like the one just described when the time horizon is long. If the time horizon is literally infinite, one could appeal to reputation mechanisms in order to support “better” equilibria, such as even the allocation that a government with access to commitment could attain. We do not. We instead focus on the limit of finite-horizon equilibria, because we take the position that this case is an equally important benchmark. Reputation mechanisms may not work, either because agents are not patient enough, or because the equilibrium selection mechanism may not result in a good equilibrium. Thus, a loose way of stating what we do here is: we answer the question “What if reputation cannot be relied upon—what happens then?”. We provide the answer by means of two tools: (i) an interpretable first-order condition for the government’s choice of policy, which dictates how the government will balance present and future wedges due to taxes and non-optimal provision of public goods; and (ii) quantitative, numerical solutions of a model calibrated to U.S. data, allowing us to assess the importance of the different wedges. By our analysis in (ii), we are thus able to assess the quantitative relevance of the lack of commitment in our simple economy.

Formally, we develop tools for the analysis of differentiable Markov-perfect equilibria. The “Markov-perfect” part is just the requirement that outcomes cannot be history-dependent other than through the physical state variable capital. The “differentiable” part, i.e. the requirement that decision rules be differentiable, is introduced for two reasons. First, differentiability is a convenient requirement that rules out the kind of multiplicity of Markov equilibria that has been shown to be present in similar environments. In particular, Krusell and Smith (2003) show that there is an indeterminacy of Markov-perfect equilibria in an infinite-horizon game between “selves” with different time orientation. The many equilibria in that setup, neither of which can occur in a finite economy, rely on discontinuous decision rules, and differentiability would rule these equilibria out.

Second, and more substantially, we use differentiability to derive interpretable first-

order conditions for the choice problem of the government: current policy choices will, through an impact on capital, influence future policy choices, and we analyze environments in which marginal conditions make sense to use. Thus, the first-order condition for the government will generally contain a derivative of the function, used by the next government, that maps capital into policy choice. Although this first-order condition can be given an intuitive form, it is a significantly more complex object mathematically than the standard first-order condition for the consumption-savings choice. One implication of this complication is that it is necessary to jointly solve for steady states and dynamics in this class of models. Another implication is that, in order to solve the model numerically, even if one is only interested in the steady state, one needs to go beyond linearization techniques. The methods we use can, however, be viewed as a straightforward extension of linearization.

We use recursive language to define and characterize equilibria. The main output of our analytical results are a first-order necessary condition for government choice. We label this equation the “GEE” to represent the “Government”, or “Generalized”, Euler Equation.¹ We state this condition in two ways. The first version of the GEE parallels the typical intertemporal first-order condition in macroeconomics where the current contemporaneous effects of a tax hike are compared with those that arise in the future. We make appeal to an envelope theorem: because the current government is benevolent, it would make the same decisions as any future government given the same starting condition, i.e., given the same amount of capital. This means that “the future” can be summarized by effects on the following period; any further effects are of second order. Thus, the first-order condition weighs current effects on marginal utility against effects on marginal utility one period later. The alternative specification—which is merely a reordering of terms—presents the first-order condition as a weighted average of wedges. In this case the government can be seen as choosing the optimal relative size of deviations from the first best, with the GEE providing a formula for how to weigh those wedges: if a certain wedge is narrowed today, some other wedge has to increase, either today or in the future, and the GEE tells us which and by how much. This approach follows a tradition in the public finance literature by describing optimal taxes as a compromise between different wedges.

¹The term “generalized Euler equation” was, as far as we know, first used in the literature on time-inconsistent preferences; see Laibson (1994). Our GEE is similar to the GEEs of that literature.

We also discuss whether the current government wants to manipulate its successor via the state variable, in the spirit of Persson and Svensson (1989). We find that it does not. The reason is that, in our environment, it is the constraints that are time-inconsistent and not the preferences. More precisely, the current government will not manipulate tomorrow's capital stock so as to make its successor behave more in line with its own preferences via the effect of tomorrow's capital stock on the successor's policy. It will not because both governments share the same objective function; hence the envelope theorem referred to above. What the government however would like to do, if it could, is to alter the private sector's expectations of future policy so as to increase savings. It cannot do this directly, but its policy choice will influence private capital accumulation which has an equilibrium influence on future policy, and thus on itself. This effect is taken into account by the government.

In our quantitative experiments it turns out that the properties of taxes and allocations in the time-consistent and Ramsey (commitment) equilibria can, but do not necessarily, differ markedly. We find that even though reputation is by definition ruled out, the mechanisms that are left—which involve the effects of current taxation on the capital stock bequeathed to the next decision maker—can be quite powerful. In an economy where labor supply is exogenous and the government taxes current capital income alone to finance the current provision of public goods, it does not produce a mix of private and public goods that equates the marginal rate of substitution to the marginal rate of transformation, even though the capital income tax is equivalent to a lump-sum tax. This has to do with the wealth effect outlined in the two-period example above, and this effect has quantitative significance: long-run Ramsey tax rates on capital are not so much lower than time-consistent tax rates as one might suspect. When labor supply is elastic, Ramsey equilibria lead to higher income taxes than do time-consistent equilibria, since higher taxes at time t lead to higher work effort at earlier times, which is a desirable effect of taxation for a Ramsey government; the effects on past labor effort are always ignored by a time-consistent planner.

Since Kydland and Prescott stated the time-inconsistency problem, most of the attempts to deal with time inconsistency of public policy have been attempts to fully resolve the problem. In short, the idea has been to introduce (full or partial) commitment through other mechanisms: “rules” (e.g., Kydland and Prescott (1977)), delega-

tion (e.g., Rogoff (1985)), a richer set of policy instruments with built-in irreversibilities (such as long-maturity bonds which by assumption cannot be defaulted upon; see, e.g., Lucas and Stokey (1983)), and so on. To us, it is not clear that these alternatives are feasible. Finally, it is possible to argue—see Chari and Kehoe (1990), following Abreu, Pearce, and Stacchetti (1990)—that good outcomes are feasible without explicit commitment, assuming that the time horizon is infinite and that agents are sufficiently patient.

The tools developed in this paper are, we think, quite general and applicable to a wide variety of contexts. However, there is earlier work in this direction. First, Markov equilibria of the type that we are interested in have been studied in Cohen and Michel (1988) and Currie and Levine (1993), who explore linear-quadratic economies.² In such economies, Markov equilibria can be characterized and computed rather easily, since the first-order conditions become linear in the state variable. The problem of unknown derivatives cropping up in equilibrium conditions is not nearly as severe as in environments that are not linear-quadratic, since second and higher-order derivatives of decision rules vanish. The drawback, of course, of linear-quadratic settings is that they only apply in extremely special settings. Thus, either one has to give up on quantitative analysis to apply them, or accept reduced-form objective functions and/or reduced-form private decision rules.

There is also a literature both in political economy (Krusell, Quadrini, and Ríos-Rull (1997), Krusell and Ríos-Rull (1999)) and in optimal policy with a benevolent government (Klein and Ríos-Rull (1999)) that has used computational methods to find quantitative implications of Markov equilibria for a variety of questions. This work is closely related to the present one, but it has two drawbacks. First, the methods used—essentially, numerical solution of value functions based on linear-quadratic approximations—are of the “black-box” type: they do not deliver interpretable conditions, such as first-order conditions for the key decision maker. The present paper fills this gap. Secondly, the numerical methods do not deliver controlled accuracy. In contrast, the methods proposed and used here do.

In a related paper, Phelan and Stacchetti (2001) have looked at environments like those

²More recently, Ambler (2001) has extended the Cohen–Michel approach to the non-linear-quadratic case.

studied in this paper and have developed methods to find all equilibria. Their methods, however, do not allow Markov equilibria to be identified and explicitly interpreted. The only other closely related literature is that upon which the present work builds quite directly: the analysis of dynamic games between successive selves, as outlined in the economics and psychology literature by Strotz (1956), Phelps and Pollak (1968), Pollak (1971), Laibson (1997), and others. This literature contains the derivation of a GEE, and Krusell, Kuruşçu, and Smith (2000) show how to solve it numerically for a smooth decision rule equilibrium. As will be elaborated on below, the smooth rule can be difficult to find with standard methods, and Krusell, Kuruşçu, and Smith (2000) resort to a perturbation method, which we also use here. This method relies on successive differentiation of the GEE.

The broad outline of the paper is as follows. In Section 2 we describe our baseline environment, in which the only private economic decision is the consumption-savings choice (Section 2.1), define a Ramsey equilibrium (Section 2.2), and then define and discuss our time-consistent, Markov equilibrium (Section 2.3) step by step. The section ends with some illustrative examples where closed-form solutions are available. The analysis of our equilibrium—which involves interpretations of the GEE and of government behavior as well as comparisons with alternative ways of stating the government problem/defining equilibrium—is contained in Section 3. Section 4 then discusses an extension to our baseline setup where leisure is valued and where there are different possibilities for what tax base might be used. Section 5 discusses the properties of the policies that arise in an environment calibrated to U.S. data where governments do not have access to a commitment technology (Markov policies) and compares them to those that arise both in environments with commitment (Ramsey policies) and in environments where the government has access to lump sum taxation (Pareto policies).

We argue that the canonical structure we study captures central components of any more complex story of government decision making. Often, for example, successive governments may disagree, because they may represent different voter groups, or because governments may simply not be benevolent. However, in these contexts as well, the mechanisms we explore here are present. That is, our setup is a useful benchmark for any further explorations into positive analysis of government decision making over time. Section 6 concludes with a discussion along these lines. The Appendix includes

some auxiliary formal definitions and the description of the computational procedures we use.

2 The model

In this section, we describe the specific setup. We then define a benchmark “Ramsey equilibrium”—the solution to an optimal–policy problem where the government can commit to future policies. After that, we proceed toward a definition of a time–consistent equilibrium where the government does not have the ability to commit.

2.1 The environment

Our setup is a canonical model of public-goods provision embedded in a neoclassical growth framework. The representative consumer is infinitely lived and there is a benevolent government with a period-by-period balanced budget and proportional taxation. To begin with, the tax base is total income and leisure is not valued.

In a competitive equilibrium, households maximize

$$\sum_{t=0}^T \beta^t u(c_t, g_t),$$

where T is either a natural number or $+\infty$, subject to

$$c_t + k_{t+1} = k_t + (1 - \tau_t) [w_t + (r_t - \delta)k_t]$$

taking the price and tax sequences as given. Firms maximize profits; using a constant-returns-to-scale production function $f(k, l)$, where f is concave, they employ inputs so that w_t and r_t are the marginal products of labor and capital, respectively. The resource constraint in this economy reads

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

It follows that the government's balanced-budget constraint is

$$g_t = \tau_t [f(k_t, 1) - \delta k_t].$$

We will make use of the following functions:

$$\mathcal{T}(k, g) \equiv g / [f(k, 1) - \delta k]$$

and

$$\mathcal{C}(k, k', g) \equiv f(k, 1) + (1 - \delta)k - k' - g,$$

where 's denote next-period values. These functions— \mathcal{C} representing consumption as a function of current and next-period capital and the current public expenditure and \mathcal{T} representing the balanced-budget tax rate as a function of current capital and the current public expenditure—are exogenous and will economize on notation significantly.

2.2 Commitment: the Ramsey problem

If lump-sum taxes were available, the optimal allocation in this economy would involve two conditions: $u_c(c_t, g_t) = \beta(1 + f_k(k_{t+1}, 1) - \delta)u_c(c_{t+1}, g_{t+1})$ (optimal savings) and $u_c(c_t, g_t) = u_g(c_t, g_t)$ (optimal public expenditures). In our economy lump-sum taxes are assumed not to be available, and the the optimal allocation using a proportional income tax is more involved.

We will first assume that the government has the ability to commit to all its future policy choices at the beginning of time. The government's decision problem is therefore to choose a sequence of tax rates $\{\tau_t\}_{t=0}^T$ in order to maximize utility, taking into account how the private sector will respond to these taxes. In order to simplify notation from here and on, we will assume that the government chooses a sequence of expenditures instead: it chooses $\{g_t\}_{t=0}^T$.³ A simple way to describe this problem formally is to

³Formally, letting the government choose g instead of τ can be a problem if there are more than one tax rate associated with a given g (which can occur). One could then specify a selection rule, or alternatively simply state the choice in terms of the tax rate directly. None of our conclusions depend on this notational simplification.

choose $\{g_t, k_{t+1}\}_{t=0}^T$ to maximize

$$\sum_{t=0}^T \beta^t u(\mathcal{C}(k_t, k_{t+1}, g_t), g_t)$$

subject to the private sector's first-order condition for savings

$$u_c(\mathcal{C}(k_t, k_{t+1}, g_t), g_t) = \beta u_c(\mathcal{C}(k_{t+1}, k_{t+2}, g_{t+1}), g_{t+1}) [1 + (1 - \mathcal{T}(k_{t+1}, g_{t+1}))(f_k(k_{t+1}, 1) - \delta)] \quad (1)$$

for all $t \geq 0$. We refer to the solution of this problem as the Ramsey allocation.

This problem has a noteworthy feature: its solution will, in general, not be time-consistent. That is, the optimal sequence of taxes and capital stocks will not be optimal ex post: if the government could reoptimize at a $t > 0$, they would choose to not follow the original sequence. For this reason, the assumption that the government can commit to future taxes is a binding one.

Technically, the source of the time-inconsistency is the special status of period 0. For any other period s , two versions of (1) apply, one with $t = s$ and the other with $t = s - 1$. For $s = 0$ only one constraint applies. This means that resetting the clock matters to zero changes the optimal solution.

Intuitively, the tax rate chosen by the government for time $t > 0$ does influence—distort—the savings choice in period $t - 1$ (and therefore also in any earlier periods), but if it were to reoptimize at time t , it would not recognize this distortion. The characterization of the solution to this Ramsey problem will briefly be discussed in our section with quantitative analysis below; the full analysis is contained in another paper (Klein, Krusell, and Ríos-Rull (2003)).

2.3 No commitment: time-consistent equilibrium

The development of the time-consistent equilibrium concept will take place in three steps. First we define equilibrium in a finite-horizon environment via backward induction. Then we illustrate this solution concept in some environments where a closed-form

solution is available. In those environments, we can easily analyze what happens when the horizon tends to infinity. We then go on to define equilibrium in a more general environment with an infinite horizon.

2.3.1 Definition of equilibrium when the horizon is finite

Consider again the environment described in Section 2.1. We will define equilibrium with recursive methods, thus describing equilibrium choices of private savings and public policy in terms of *functions* of state variables. The optimal policy choice by the government period t will be given by a function Ψ_t mapping any relevant history—which is summarized by our only state variable, the capital stock—into a choice for public expenditures: $g_t = \Psi_t(k_t)$.⁴ Denote the value of being born as a representative agent into the economy at t when the capital stock is k_t by $v_t(k_t)$.

Private savings will also be given by a function of the state variable. However, in order to define the government's choice problem at any point in time, we need to define private savings as a function not only of current capital but also of any current choice for government expenditures (and hence the tax rate). Let privately optimal saving in period t when the capital stock is k_t and the current policy is g_t be given by the function \mathcal{H}_t so that $k_{t+1} = \mathcal{H}_t(k_t, g_t)$.

Having established this notation, a Markov-perfect equilibrium is a sequence of functions $\{\mathcal{H}_t, v_t, \Psi_t\}_{t=0}^T$ such that

1. the government maximizes consumer utility, i.e.,

$$\Psi_t(k) \in \underset{g}{\operatorname{argmax}} \{u(\mathcal{C}(k, \mathcal{H}_t(k, g), g), g) + \beta v_{t+1}(\mathcal{H}_t(k, g))\}$$

for $k \geq 0$ and $t = 0, 1, \dots, T$,

$$v_t(k) = u(\mathcal{C}(k, \mathcal{H}_t(k, \Psi_t(k)), \Psi_t(k)), \Psi_t(k)) + \beta v_{t+1}(\mathcal{H}_t(k, \Psi_t(k)))$$

⁴We will presume that the equilibrium in any finite-horizon economy is unique. In case it is not, the history might include more information than the capital stock.

for $k \geq 0$ and $t = 0, 1, \dots, T$,

$$v_{T+1}(k) = 0,$$

for $k \geq 0$; and

2. the private sector saves optimally, i.e.,

$$u_c(\mathcal{C}(k, k', g), g) = \beta u_c(\mathcal{C}(k', k'', g'), g') [1 + (1 - \mathcal{T}(k', g'))(f_k(k', 1) - \delta)] \quad (2)$$

for $k \geq 0$ and $g \geq 0$ where, in (2), $k' = \mathcal{H}_t(k, g)$, $g' = \Psi(k')$ and $k'' = \mathcal{H}_{t+1}(k', g')$ for $t = 0, 1, \dots, T - 1$ and

$$\mathcal{H}_T(k, g) = 0.$$

Our finite-horizon equilibrium can be solved for backwards. We will look at some examples below where functional form assumptions on preferences and technology allow closed-form solution. Among cases which do not admit closed-form solutions, we will restrict attention to those which deliver differentiable policy rules for the private and the public sectors.⁵ Our main tool for analyzing our equilibria will rely on a first-order condition for the government's choice. To save on space, we will postpone the derivation of such a first-order condition to the section below which uses a finite time horizon.

To illustrate our solution concept, we now briefly consider environments where the Markov equilibrium can be solved for explicitly.

2.3.2 Some examples with a closed-form solution

We will consider two examples that illustrate the nature of the Markov-perfect equilibrium and how it relates to the Ramsey solution.

⁵Differentiability follows when decision problems yield unique interior solutions assuming sufficient differentiability of the primitive functions. The purpose here is not to state precise conditions under which such properties obtain; finding general conditions on primitives guaranteeing differentiability is difficult. It does not, however, seem that the applications to policy choice in the neoclassical growth model we consider lead to multiple solutions for government choice.

2.3.2.1 A proportional tax on total income

Here we maintain the case where the tax rate on capital income equals that on wage income. Assume that the period utility function is $\ln c + \gamma \ln g$, that the production function is k^θ , and that there is full depreciation. Then it is straightforward to solve for a competitive equilibrium for *any* sequence of government expenditures/tax rates. It is given by

$$k_{t+1} = s_t(1 - \mathcal{T}(k_t, g_t))k_t^\theta$$

with

$$s_t = \beta\theta \frac{1 + \beta\theta + \dots + (\beta\theta)^{t-1}}{1 + \beta\theta + \dots + (\beta\theta)^{t-1} + (\beta\theta)^t}$$

and

$$\mathcal{T}(k, g) = g/k^\theta.$$

This competitive equilibrium has the feature that future government spending patterns/tax rates do not matter for current savings. This result depends crucially on the functional forms we employ. We know, first, that with logarithmic utility, changes in future returns to capital—which directly depend on tax rates on income—do not affect current savings because income and substitution effects cancel for an isoelastic utility function. Second, a tax on future labor income is a negative wealth effect and should cause current savings to increase, since consumption at different points in time are normal goods. However, the net present value of future labor income net of taxes is also affected by discounting, and if the tax rate on capital income goes up too, the net interest rate will go down, thus increasing the current value of any future income. If the rate of depreciation is 100%, or if the tax is on gross capital income, the decreased net interest rate exactly offsets the decreased future net labor income; hence, current savings will be unaffected by future taxes. The Cobb-Douglas form of the production function is helpful because it allows closed-form solution of the savings function.

Given that current savings do not depend on future levels of government expenditures/tax rates, the commitment solution will be time-consistent: there is no incentive to change plans for future policy once current savings decisions are made. To find the optimal tax rate, notice that there are direct and indirect effects on utility. The direct effects occur on current private consumption—a higher g lowers c —and on public consumption. The indirect effects occur through savings. These effects cannot be ignored

here, because the consumer does not internalize the effects of savings on future public-goods provision that occur because capital income is taxed. Thus, we need to find out how savings influence utility. In our finite-horizon setup, it is easy to see that in the last period government expenditures will be a constant fraction $\tau_T = \gamma/(1 + \gamma)$ of output and, indeed, the same property will also hold at all earlier dates. Thus “guessing” that $\Psi_t(k) = \tau_t k^\theta$, the value function at time t can be seen to equal

$$v_t(k) = A_t \ln k + B_t,$$

where

$$A_t = \theta(1 + \gamma) \frac{1 - (\beta\theta)^{T-t+1}}{1 - \beta\theta},$$

and

$$B_t = \ln(1 - \tau_t)(1 - s_t) + \gamma \ln \tau_t + \beta A_{t+1} \ln s_t(1 - \tau_t) + \beta B_{t+1},$$

with $A_{T+1} = B_{T+1} = 0$. The optimal choice of g at t can now be found by maximizing over τ_t :

$$\tau_t \in \underset{\tau}{\operatorname{argmax}} \{ \ln(1 - \tau - s_t) + \gamma \ln \tau + \beta A_{t+1} \ln(1 - \tau) \}$$

and the unique solution is

$$\tau_t = \frac{\gamma}{1 + \gamma} \frac{1 - \beta\theta}{1 - (\beta\theta)^{T-t+1}}.$$

We see that tax rates are between zero and one, and that they are increasing over time for any given time horizon: the longer the remaining time horizon, the higher are savings so the higher is the positive impact on future expenditures by taxing less now. In other words, tax rates that are lower than the statically optimal level will be used in order to increase current savings, which are suboptimally low due to tax distortions. In particular, it is possible that the statically optimal tax rate is very close to 100% and yet that taxes early on are far from 100%; the time-zero tax rate is decreasing in the time horizon, approaching $(1 - \beta\theta)\gamma/(1 + \gamma)$ as T goes to infinity.

2.3.2.2 A proportional tax on capital income

When the tax base is capital income only— $g_t = \tau_t \theta k_t^\theta$ —we obtain savings rules of a

very similar form:

$$k_{t+1} = s_t(1 - \theta\mathcal{T}(k_t, g_t))k_t^\theta$$

the same form as before, but savings rates now satisfy the recursion

$$s_t = \beta\theta \frac{(1 - \beta\theta)(1 - \tau_{t+1})}{1 - \theta\tau_{t+1} - s_{t+1} + \beta(1 - \tau_{t+1})\theta}$$

with $s_T = 0$. Here, savings at t depend on all future tax rates: they decrease in all future tax rates. The reason is that future tax rates decreases discounting, thus making any future income more worth in present terms; this positive wealth effect will increase current consumption and decrease savings. As a result, the commitment solution will not be time-consistent in this case. We will not solve for the commitment solution here—it does not admit closed-form expressions—but we will solve for the time-consistent equilibrium. As in the case with a tax on total income, optimal government policy at t will be a constant fraction τ_t of output at t , leading to the same indirect utility as above with the difference that B_t now satisfies

$$B_t = \ln(1 - \theta\tau_t)(1 - s_t) + \gamma \ln \theta\tau_t + \beta A_{t+1} \ln s_t(1 - \theta\tau_t) + \beta B_{t+1}.$$

Inspecting this expression, one sees that the optimal choice of τ_t does not interact with future taxes, delivering

$$\tau_t = \frac{1}{\theta} \frac{\gamma}{1 + \gamma} \frac{1 - \beta\theta}{1 - (\beta\theta)^{T-t+1}}.$$

Thus, $\theta\tau_t$ here obeys the same form as τ_t did before. However, with a high enough γ in this case, capital income will not suffice to provide for ex-post optimal public consumption levels in the last period, leading to a tax rate above 100%. If this occurs in period T , there will be no savings in any earlier periods and both c and g will be 0 in all periods but the very first one. With a literally infinite time horizon, however, there will also be an “expectations-driven” equilibrium with savings in this case if $(1 - \beta\theta)\gamma/(1 + \gamma)/\theta < 1$: if agents—private and public—believe future capital income will not be taxed at high rates, there will be savings.

2.3.3 Infinite time horizon

The main purpose of our paper is to develop an applied version of the model where the time horizon is not an important determinant of outcomes. Like in most of the macroeconomic literature, we therefore assume that the time horizon is infinite. I.e., the presumption is that T is large enough that small changes in it will not significantly change the model's predictions. The resulting version of our Markov-perfect equilibrium is simply a "stationary" version of our earlier definition: it consists of time-invariant value and policy functions. Thus, a Markov-perfect equilibrium consists of three functions \mathcal{H} , v , and Ψ such that

1. Given \mathcal{H} and Ψ , v is the value function, i.e.

$$v(k) = u(\mathcal{C}(k, \mathcal{H}(k, \Psi(k)), \Psi(k)), \Psi(k)) + \beta v(\mathcal{H}(k, \Psi(k))) \quad (3)$$

for all $k \geq 0$;

2. given \mathcal{H} and v , Ψ delivers an optimal choice for the government, i.e.

$$\Psi(k) \in \operatorname{argmax}_g \{u(\mathcal{C}(k, \mathcal{H}(k, g), g), g) + \beta v(\mathcal{H}(k, g))\} \quad (4)$$

for all $k \geq 0$; and

3. given Ψ , \mathcal{H} is an optimal savings function, i.e.

$$u_c(\mathcal{C}(k, k', g), g) = \beta u_c(\mathcal{C}(k', k'', g'), g') [1 + (1 - \mathcal{T}(k', g'))(f_k(k', 1) - \delta)] \quad (5)$$

for all $k \geq 0$ and $g \geq 0$ where, in (5), $k' = \mathcal{H}(k, g)$, $g' = \Psi(k')$ and $k'' = \mathcal{H}(k', g')$.

With an infinite horizon, the Bellman equation expresses the time-consistency of government behavior: its stationarity guarantees that consecutive governments agree, given initial conditions. Agreement across governments may be surprising, and in Section 3.2.1 we will discuss the (limited) sense in which consecutive governments disagree on taxation issues.

How many Markov-perfect equilibria exist in general is a question that we do not address in this paper. We do not have a proof of existence or uniqueness.

Because we wish to encapsulate a strong sense of lack of commitment, we will focus on equilibria that are limits of the finite-horizon equilibria defined in the previous section. In order to make such equilibria operationally useful, however, we will employ a more precise definition of Markov-perfect equilibria: we will refer to a time-consistent equilibrium as a Markov-perfect equilibrium whose policy function Ψ is *differentiable*.

The differentiability requirement is both useful computationally and a tool for eliminating potential indeterminacy of Markov-perfect equilibria. Computationally, first, it allows the analysis to build on first-order conditions—one standard condition characterizing private sector behavior and a new condition summarizing government behavior—rather than on an iterative procedure whereby equilibria are solved backwards. We will discuss our computational algorithm in more detail below. Second, indeterminacy of Markov-perfect equilibria has been demonstrated in similar contexts. In particular, the setup in Krusell and Smith (2003) is one where the consecutive decision makers have different objectives, and the indeterminacy there builds on decision rules—those equivalent to Ψ here—that are discontinuous (step functions). The differentiability requirement therefore rules out such equilibria, though it is an open question whether they would exist here.

2.3.4 Deriving the GEE: the government’s first-order condition

We now assume that our equilibrium policy functions Ψ and \mathcal{H} are differentiable and proceed to derive a functional-equation first-order condition for the government’s choice. This equation, which we will refer to as the GEE (the Government’s, or Generalized, Euler Equation), will be in focus in the analysis below.

There are several ways to derive the GEE. The most straightforward way at this point is to first derive a first-order condition from the government’s recursive problem, which will contain the unknown value-function derivative, v_k , and then use an envelope condition to eliminate v_k . This procedure is nonstandard only in that it is somewhat more roundabout to eliminate v_k here than in the standard growth model.

The first-order condition for the government produces

$$u_c(-\mathcal{H}_g - 1) + u_g + \beta v'_k \mathcal{H}_g = 0.$$

Here, we are economizing on notation by suppressing the arguments of the functions. This equation, thus, pins down g . The notation v'_k stands for the derivative of v with respect to k (subscript denote derivatives) evaluated at $\mathcal{H}(k, g)$ (primes denote forward lags). To obtain an expression for it, we differentiate Equation (3) with respect to k . We obtain, noting that $\mathcal{C}_g = \mathcal{C}_{k'} = -1$,

$$v_k = u_c(\mathcal{C}_k - (\mathcal{H}_k + \mathcal{H}_g \Psi_k) - \Psi_k) + u_g \Psi_k + \beta v'_k (\mathcal{H}_k + \mathcal{H}_g \Psi_k) = 0.$$

This equation contains *indirect effects*, via Ψ_k . Notice, too, that a grouping of the Ψ_k terms leads to $\Psi_k \cdot 0$, where the “0” results from use of the first-order condition above: this is the envelope theorem. However, unlike in the setting of the standard growth model, the use of the envelope theorem does not suffice to make v'_k disappear here: v_k still depends on v'_k :

$$v_k = u_c(\mathcal{C}_k - \mathcal{H}_k) + \beta v'_k \mathcal{H}_k = 0.$$

This may appear like a problem, but it is not: v'_k can be expressed in terms of primitives and decision rules from the first-order condition above. This delivers

$$\beta v'_k = \frac{1}{\mathcal{H}_g} (u_c(\mathcal{H}_g + 1) - u_g).$$

Thus, the expression for v_k in terms of primitives and decision rules reads

$$v_k = u_c(\mathcal{C}_k - \mathcal{H}_k) + \frac{\mathcal{H}_k}{\mathcal{H}_g} (u_c(\mathcal{H}_g + 1) - u_g).$$

We can now update this expression one period and substitute back into the original first-order condition to obtain our GEE:

$$-u_c[\mathcal{H}_g + 1] + u_g + \beta \mathcal{H}_g \left\{ u'_c [f'_k + 1 - \delta - \mathcal{H}'_k] + \frac{\mathcal{H}'_k}{\mathcal{H}'_g} (u'_c(\mathcal{H}'_g + 1) - u'_g) \right\} = 0, \quad (6)$$

where we have also used the definition of \mathcal{C} in terms of primitives. Equation (6), where arguments are still suppressed for readability and primes on functions indicate that the function is evaluated in the next period, holds for all k . It is our fundamental functional equation determining $\Psi(k)$ given $\mathcal{H}(k, g)$. I.e., it defines the government policy rule Ψ as the optimal policy determination under the assumption that the private sector behaves according to an arbitrary \mathcal{H} . In the next section, we will discuss the government's behavior from the perspective of this equation.

3 Properties of government behavior

We now move to our discussion of the GEE—the government's first-order condition. After providing various interpretations of this condition, we present two reformulations of our problem. The first one of these is more compact than the one above, although perhaps less transparent. It is particularly useful in the numerical work later. The second reformulation casts the government's problem as a sequential one. This problem is useful in delivering an alternative, and perhaps easier, method for deriving the GEE than the one above.

3.1 Interpretation of the government's Euler equation

We begin with two alternative interpretations of the costs and benefits of raising current taxes. Thereafter, we address the question of whether there is a sense in which the current government, through the effect its taxation decision has on capital accumulation, manipulates its successors.

3.1.1 The macroeconomist's version

A first notable property of the GEE is that it has a finite number of terms. That is, even though the current tax rate choice in general has repercussions into the infinite future—recall that the present government cannot “keep future variables constant” because it cannot commit future governments—the marginal costs and benefits at an optimum can be summarized with terms involving only two consecutive periods. This,

of course, is due to the recursive structure and the use of the envelope theorem. The envelope theorem in this context means that, *when k' is viewed as given*, the current government agrees with the next government on how to set g' ; these two governments have identical ways of evaluating utility from tomorrow and on. That is, disagreement is only present if the effects of g' on k' are taken into account.

The recursive structure of the government's problem makes it equivalent to a sequential problem, which we will state and discuss in some detail in Section 3.1.3. Thus, one can also view the GEE as resulting from a *variational* (2-period) problem: keeping the state variables k and k'' fixed, vary k' , through the control variables g and g' , in order to obtain the highest possible utility. Viewed this way, one sees that any change in g' , which is effectuated by a change in g , requires an accompanying change in g' so that k'' remains unchanged. Total differentiation of $k'' = \mathcal{H}(k', g')$ thus states that this change in g' has to be $\frac{dg'}{dk'} = -\frac{\mathcal{H}'_k}{\mathcal{H}'_g}$. Notice that this term appears in the GEE: it thus reflects a partial change in g' coming about due to a change in k' . It is not, however, equal to Ψ' , which is the net change in g' .

Second, the GEE contains both primitive functions, such as marginal utility, and (endogenous) decision rules: \mathcal{H} and Ψ . Moreover, and this is why the term “generalized” Euler equation is appropriate, the equation contains derivatives of decision rules; it contains \mathcal{H}_k and \mathcal{H}_g , both evaluated in the present and in the future. It is possible to eliminate the derivatives of \mathcal{H} by use of the consumer's Euler equation for savings, equation (5). This equation defines $H(k, g)$, and by differentiation with respect to k and g , respectively, these derivatives can be obtained. This differentiation, however, will produce another unknown derivative: Ψ'_k . Since Ψ is present in the forward-looking consumer's Euler equation, its derivative is a determinant of how changes in current policy and capital will induce equilibrium changes in savings. Thus, our equilibrium system of functional equations is actually a *differential equation system*.

We will now interpret the GEE in terms of marginal benefits and marginal costs of changing g . These benefits and costs will involve the unknowns \mathcal{H}_k and \mathcal{H}_g . In our discussion, we will assume that the former of these is positive and the latter negative. These assumptions seem natural: under the normal goods assumption regarding c and c' (i.e., time-additive utility and a concave u), one would expect increased income to increase savings, and both an increase in k and a decrease in g reflect increased income.

However, notice that, to the extent k also changes tax rates (recall that k is not the individual's state variable, but the economy-wide capital stock), which it does in our economy, other effects could be present. In our analytically solved example below as well as in our quantitative section, we confirm the assumed signs of \mathcal{H}_k and \mathcal{H}_g .

We can thus describe our marginal benefits and costs as follows. The effects on today's utility-relevant variables from a marginal increase in current government expenditures are on

1. current consumption, which
 - (a) goes up via lower savings, delivering a utility effect of $-u_c \mathcal{H}_g > 0$, and
 - (b) down via higher government spending, with an effect on utility of $-u_c < 0$; and on
2. current government spending, whose rise leads to a utility change of $u_g > 0$.

The effects of the increase in g on future utility-relevant variables occur via a decrease in savings ($H_g < 0$), leading to

3. effects on next period's consumption which
 - (a) goes down via a direct effect on production and undepreciated capital, affecting utility by $\beta \mathcal{H}_g u'_c (f'_k + 1 - \delta) < 0$;
 - (b) goes up via an indirect negative effect from lowered saving ($\frac{dk'}{dg} \frac{dk''}{dk'} = \mathcal{H}_g \mathcal{H}'_k < 0$), affecting utility by $\beta \mathcal{H}_g u'_c (-\mathcal{H}'_k) > 0$; and
4. two induced effects which occur via the above-mentioned decrease in next period's government expenditures, $\frac{dg'}{dg} = -\mathcal{H}_g \frac{\mathcal{H}'_k}{\mathcal{H}'_g} < 0$; this effect
 - (a) raises next period's consumption, which results in a change in next period's utility by $\beta \mathcal{H}_g (-\frac{\mathcal{H}'_k}{\mathcal{H}'_g}) u'_c [-\mathcal{H}'_g - 1] > 0$ (assuming $\mathcal{H}_g + 1 > 0$), and
 - (b) lowers next period's government spending, which leads to a utility change of the amount $\beta \mathcal{H}_g (-\frac{\mathcal{H}'_k}{\mathcal{H}'_g}) u'_g < 0$.

In our numerical work below, we assign (steady-state) values to these different terms: we learn which of the effects in the GEE are quantitatively important and which are not.

3.1.2 The public economics version

The GEE can be rewritten so that it is a linear combination of wedges. Thus, rearranging terms we obtain the following equation.

$$\left[u_g - u_c \right] + \mathcal{H}_g \left[-u_c + \beta u'_c (1 + f'_k - \delta) \right] + \beta \mathcal{H}_g \left(-\frac{\mathcal{H}'_k}{\mathcal{H}'_g} \right) \left[u'_g - u'_c \right] = 0. \quad (7)$$

Three terms in brackets appear: these are the three different “wedges” that are affected by the change in the current tax rate. Note that only wedges in the current and in the next period appear, even though this intertemporal economy has wedges in every period: again, envelope theorems imply that future wedges are handled optimally and hence can be ignored in comparing marginal costs and benefits of a current tax increase.

How are the different distortions traded off against each other? First, an increase in g influences the gap between u_g and u_c . This gap, which would be zero with lump-sum taxes since private and public goods are perfect substitutes in production, must be positive since it is costlier to provide g than c here. This means that a g increase, which also lowers c , makes this gap smaller.

Second, since the increase in g leads to a decrease in savings, the intertemporal distortion is affected. The second bracket, $-u_c + \beta u'_c (1 + f'_k - \delta)$, actually equals $u'_c (f'_k - \delta) \tau'$ from the consumer’s Euler equation: so long as the tax rate next period is positive, marginal utility of consumption today is too low (because savings are too low). Thus, the decrease in savings resulting from an increase in current taxes will be detrimental: it increases the intertemporal distortion further.

Third, the lowered savings will lead to changes in the provision of public goods next period, and it will thus influence the gap between the marginal utilities of public and private goods in that period. The channel is that lowered saving induces a decrease in next period’s government expenditures: $\frac{dg'}{dg} = \mathcal{H}_g \left(-\frac{\mathcal{H}'_k}{\mathcal{H}'_g} \right) < 0$. Thus, this effect is a

negative: the distortion is increased next period.

In sum, we are weighing one positive effect of increasing the current government expenditures—it decreases the wedge between public and private goods—against two negative ones: it increases the same wedge next period, and it also increases the intertemporal wedge. Not all wedges can be zero, because the optimal provision of public goods— $u_c = u_g$ —demands a positive tax rate, which necessarily makes the intertemporal distortion nonzero. This result is perhaps surprising: the use of the income tax seems nondistortionary in this model from the perspective of the current government: it is like a lump-sum tax. Nevertheless, the government *does not* tax at the high rates that would be necessary to deliver (statically) optimal public-goods provision. This is because the government finds it in their interest to leave more resources than that in the hands of the private sector, because some of those resources will be saved, and this will help alleviate the intertemporal distortion. At an optimum, an optimizing government makes sure that a marginally decreased current public-goods-provision wedge is exactly counterbalanced by increases in the other wedges.

3.1.3 A sequential formulation

By Bellman’s principle, it follows that we can alternatively characterize the problem of the government as one where it chooses a policy sequence, $\{g_t\}_{t=0}^T$, to solve the following sequential problem:

$$\begin{aligned} \max_{\{g_t, k_{t+1}\}_{t=0}^T} \quad & \sum_{t=0}^T \beta^t u(\mathcal{C}(k_t, k_{t+1}, g_t), g_t) \\ \text{subject to} \quad & \\ & k_{t+1} = \mathcal{H}(k_t, g_t). \end{aligned} \tag{8}$$

Problem (8), however, does not correspond to the decision problem the government is actually facing, because it does feature a different feasible set for taxes: it does allow the government full power to choose any $\{g_t\}_{t=0}^T$. In contrast, in our time-consistent equilibrium, the government at time t only has a one-dimensional way of affecting future taxes: by the choice of g_t , which via \mathcal{H} influences k_{t+1} , which in turn via Ψ affects g_{t+1} , and so on. But despite being formally different problems, the equilibrium

sequence chosen by the government in our Markov-perfect equilibrium will also solve this problem, and vice versa!

Problem (8) is not autonomous: it is still part of a fixed-point problem in Ψ . How does Ψ enter in the problem? As just pointed out, it is incorporated into \mathcal{H} : the shape of this function reflects the expectations of future taxes through the private sector's Euler equation.

The usefulness of the sequential formulation is that it provides a direct way of deriving the GEE by using standard variational methods.

3.2 An alternative equilibrium definition

It is possible to provide a more compact definition of equilibrium.⁶ We state this definition because its compactness actually matters for computation (as we will comment on later in detail) and because it is more closely connected to the Ramsey problem of Section 2.2. The compactness is accomplished by not allowing a distinction between the government and the private sector. The government thus chooses both k' and g directly, with associated equilibrium mappings $h(k)$ and $\Psi(k)$.

The problem that the government solves—maximize utility subject to the Euler equation of the households—reads as follows:

$$\begin{aligned}
 v(k) &= \max_{k',g} \{u(\mathcal{C}(k, k', g), g) + \beta v(k')\} \\
 &\text{subject to} \\
 u_c(\mathcal{C}(k, k', g), g) &= \\
 \beta u_c(\mathcal{C}(k', h(k'), \Psi(k')), \Psi(k')) \cdot \{1 + [1 - \mathcal{T}(k', \Psi(k'))][f_k(k') - \delta]\}.
 \end{aligned} \tag{9}$$

Thus, k' is left as a choice variable and the restriction that it be consistent with private-sector behavior is not captured through a function \mathcal{H} but instead by including the consumer's Euler equation explicitly. This equation, moreover, is not a functional equation; it restricts k' and g , while k'' as well as g' , which enter on the right-hand side of the Euler equation, are restricted to equal $h(k')$ and $\Psi(k')$, respectively. This

⁶We owe gratitude to Harald Uhlig, who suggested this very natural alternative definition.

is how the lack of commitment shows up in this formulation.

Problem (9) is, of course, also a fixed-point problem: taking as given expectations about future behavior as captured by h and Ψ , current optimal behavior has to reproduce these functions. This formulation is obviously closer to the Ramsey formulation: the difference is that future savings and tax choices here are restricted by h and Ψ whereas they are free in the Ramsey problem.

It is straightforward to see that our two equilibrium definitions are equivalent: the function \mathcal{H} is *defined* to solve the Euler equation above with the only difference that what appears on the right-hand side is $\mathcal{H}(k', \Psi(k'))$, not $h(k')$, and $\Psi(k')$, not $\Psi(k)$. Thus, if \mathcal{H} and Ψ constitute an equilibrium according to our original definition, then h and Ψ defined by $h(k) \equiv \mathcal{H}(k, \Psi(k))$ and $\Psi \equiv \Psi$ are an equilibrium as defined in this section; and if h and Ψ satisfy the above equilibrium definition, then one can define \mathcal{H} from the Euler equation above (solve for k' as a function of k and g) and set $\Psi \equiv \Psi$ and it is evident that \mathcal{H} and Ψ are an equilibrium according to our original definition.

It is also possible to derive the GEE from (8). However, as perhaps is evident, it will lead to an equation which is very long, because both g and, especially, k' , appear in a large number of places. Among the many terms, both Ψ' and h' —the derivatives of the policy rules—will appear in this first-order condition.

One can simplify matters by summarizing the Euler equation by

$$\eta(k, g, k') = 0, \quad (10)$$

where η is defined as the left-hand side of the restriction in (9) minus the right-hand side. The GEE, then, becomes (after deriving the first-order condition and utilizing the envelope theorem)

$$0 = (-u_c + u_g)\eta_{k'} - \eta_g \left[-u_c + \beta u'_c (f'_k + 1 - \delta) - \beta (-u'_c + u'_g) \frac{\eta'_{k'}}{\eta'_g} \right]. \quad (11)$$

Noting that \mathcal{H} is defined by

$$\eta(k, g, \mathcal{H}(k, g)) = 0$$

for all (k, g) , we can differentiate with respect to k and g and obtain, respectively,

$$\mathcal{H}_k = -\frac{\eta_k}{\eta_{k'}}$$

and

$$\mathcal{H}_g = -\frac{\eta_g}{\eta_{k'}}.$$

Dividing the GEE—equation (11)—by $\eta_{k'}$ and rearranging, we obtain our original GEE—equation (6).

3.2.1 Strategic policy: does the current government manipulate its successors?

The dynamic game played between governments involves a disagreement: the current government would like to see the next government choose a lower tax on income, τ' , than it ends up choosing. Does this mean that the current government attempts to “manipulate” the next government in its tax choice? It could influence τ' through its influence on saving, k' . Suppose, for example, that $g' = \Psi(k')$ is increasing. Then the current government might see a reason to increase g a little extra, so as to decrease k' and thereby decrease g' : it could influence the government expenditure choice next period through savings.

Our GEEs, however, do not directly contain the derivative of the tax policy rule Ψ , as one might think it would. In fact, from our arguments earlier, and the very fact that the government’s problem can be written recursively, the successive governments actually agree in one important dimension: *given the value for current savings*, they agree on how to set next period’s taxes. That is why the derivative of Ψ does not appear directly in the government’s first-order conditions. It appears indirectly, as a determinant of \mathcal{H}_g . But this appearance does not reflect strategic behavior; rather, it simply captures how the effects on private-sector savings of a current change in g depends on how those extra savings will alter next period’s tax rate. That is, \mathcal{H}_g reflects how a current tax change influences the *expectations* of private agents, and therefore their savings. More precisely, if the tax rate today is changed, how much extra (or less) capital is saved— \mathcal{H}_g —depends on how the determination of the expenditure on

g' is perceived by the private sector.

To illustrate the role of Ψ in the determination of the savings response, let us compare a “myopic” government to the kind of government we model: a myopic government does not realize that their current taxation behavior influences future taxes. Suppose that the time-consistent equilibrium has Ψ as an increasing function: the higher the savings today, the higher the government expenditures will be next period. In contrast, the myopic government perceives $\Psi(k)$ to be constant. How, then, would the myopic government’s first-order condition look? The answer is that it would look the same, with the one difference that \mathcal{H}_g would be a different number: in terms of our compact equilibrium definition, we have $\mathcal{H}_g = -\frac{\eta_g}{\eta_{k'}}$, and here the denominator (but not the numerator) depends on the derivate of Ψ . Assuming that $u(c, g)$ is additively separable, that $\eta_g > 0$, and that $\eta_{k'} > 0$, one sees that if a change in future government expenditures is ignored, $\eta_{k'}$ would could be too high—because of the lowered consumption, and therefore increased future marginal utility value of savings, implied by the higher future tax rate—or too low—because of the lower net-of-tax return from future savings. That is, a myopic government would misperceive \mathcal{H}_g , but whether this leads to lower or higher equilibrium taxes is a quantitative question.

4 Extensions: valued leisure and other tax bases

Suppose now that leisure is valued: we assume that utility is given by

$$\sum_{t=0}^{\infty} \beta^t u(c_t, 1 - \ell_t, g_t).$$

We continue assuming that the tax base is total income. Our equilibrium definition works as before, but one more element is needed: we need to describe the equilibrium labor response to (k, g) . The relevant mapping is $\mathcal{L}(k, g)$, which is obtained from the consumer’s first-order condition for the labor-leisure choice. Thus,

$$\frac{u_{\ell}(\mathcal{C}(k, \mathcal{H}(k, g), g), 1 - \mathcal{L}(k, g), g)}{u_c(\mathcal{C}(k, \mathcal{H}(k, g), g), 1 - \mathcal{L}(k, g), g)} = f_{\ell}(k, \mathcal{L}(k, g)) (1 - \mathcal{T}(k, g)) \quad (12)$$

for all (k, g) and the first-order condition for savings (which now contains a leisure argument, but which we will not restate) *jointly* define the functions $\mathcal{H}(k, g)$ and $\mathcal{L}(k, g)$.

The equilibrium conditions now include three functional equations: the private sector's first-order conditions for labor and savings and the government's first-order condition. We go straight to the latter—to the GEE—which can be derived with the same procedure as above. It reads

$$\begin{aligned} & \mathcal{L}_g \left[u_c f_\ell - u_\ell \right] + \left[u_g - u_c \right] + \mathcal{H}_g \left[-u_c + \beta u'_c (1 + f'_k - \delta) \right] + \\ & \beta \mathcal{H}_g \left\{ \mathcal{L}'_k \left[u'_c f'_\ell - u'_\ell \right] - \frac{\mathcal{H}'_k}{\mathcal{H}'_g} \left(\mathcal{L}'_g \left[u'_c f'_\ell - u'_\ell \right] + \left[u'_g - u'_c \right] \right) \right\} = 0 \end{aligned} \quad (13)$$

for all k (again, the arguments of the functions are suppressed for readability). We see a new wedge appearing: $u_c f_\ell - u_\ell$, in the current period as well as in the next. This wedge, which equals $u_c \tau$, must be positive so long as public goods are provided ($\tau > 0$). A current tax increase will now decrease labor supply (presumably) and thus increase this intratemporal distortion. Similarly, there will be repercussions through lowered savings on the same wedge in the future, in parallel with the induced effects on future savings.

In a closed-form application of the economy with leisure and taxation of total income, using $u(c, 1 - \ell, g) = \alpha \ln c + (1 - \alpha) \ln(1 - \ell) + \gamma \ln g$ and the same production technology as used in Section 2.3.2, it is straightforward to see that $\Psi(k) = \frac{\gamma}{\alpha + \gamma} (1 - \beta\theta) k^\theta$, with $\mathcal{L}(k, g) = \frac{\alpha}{\alpha + (1 - \alpha)(1 - \beta\theta)}$ and $\mathcal{H}(k, g) = \beta\theta(1 - \mathcal{T}(k, g)) k^\theta \ell^{1 - \theta}$, solves this functional equation: the tax rate is constant. Here, as above, future taxes do not influence present savings decisions, and present work decisions are not influenced either because wealth effects are not present: future taxes will lower net-present-value income for *given* net interest rates (and should increase work effort) but net interest rates go down to exactly cancel the lowering of future income flows.

Below in the quantitative section, we will look at an economy with less than full depreciation of capital and income taxation (i.e., the stock of capital is not taxed). There, in contrast, an increase in future taxes on total income would decrease present-value income. This is because the net interest rate would fall by less in percentage terms than would the value of the future labor endowment, and hence current work

effort (and savings) would increase.

A similar effect would be present if only labor income were taxed: there, there would be no counteracting decrease in the interest rate at all, and increased future tax rates would lower raise work effort and increase current savings. In this and the latter kinds of economies, hence, this by-product of future taxation—the induced rises in current work and savings efforts—will counteract the distortions present by taxation and thus be desirable. In particular, in an economy with labor taxation only the commitment outcome is expected to lead to *higher* taxation than the outcome without commitment, which does not internalize the positive impact of current taxes on past work efforts. The GEE with labor income taxes only becomes

$$\begin{aligned} & \mathcal{L}_g \left[u_c f_\ell - u_\ell \right] + \left[u_g - u_c \right] + \\ & \beta \mathcal{H}_g \left\{ \left(\mathcal{L}'_k - \frac{\mathcal{H}'_k}{\mathcal{H}'_g} \mathcal{L}'_g \right) \left[u'_c f'_\ell - u'_\ell \right] - \frac{\mathcal{H}'_k}{\mathcal{H}'_g} \left[u'_g - u'_c \right] \right\} = 0. \end{aligned} \quad (14)$$

A wedge version of the GEE that would have arisen in an economy with lump-sum taxes is given by

$$\left[u_g - u_c \right] + \beta \mathcal{H}_g \left(-\frac{\mathcal{H}'_k}{\mathcal{H}'_g} \right) \left[u'_g - u'_c \right] = 0. \quad (15)$$

Here, a policy that sets marginal utility of government expenditures equal to that of private consumption satisfies equation (15) and hence is an equilibrium policy. Meanwhile, to satisfy the private Euler equations with no distortionary taxes, labor supply will be Pareto optimal as well in the Markov equilibrium. Thus the Markov equilibrium is Pareto optimal. When the first best can be achieved, time inconsistency is no longer a problem, and the GEE shows that.

Finally, consider the case where only (net) capital income can be taxed. Then the GEE becomes

$$\left[u_g - u_c \right] + \mathcal{H}_g \left[-u_c + \beta u'_c (1 + f'_k - \delta) \right] + \beta \mathcal{H}_g \left(-\frac{\mathcal{H}'_k}{\mathcal{H}'_g} \right) \left[u'_g - u'_c \right] = 0. \quad (16)$$

Notice that this is *the same* GEE as in the model without leisure. This does not mean that the equilibrium tax rate is the same—the remaining equilibrium equation elements

are different. As in the case without valued leisure, it will not be optimal to go all the way to (statically) optimal public-goods provision.

5 Optimal policy for actual economies

We proceed next to look at numerical solutions for a selected set of economies with some aggregate statistics that resemble those of the United States postwar economy. For the sake of comparison we also provide the optimal policy under the first best (lump-sum taxation) allocation and those implied by a benevolent government that has access to commitment but not to a technology to save resources, that is, the Ramsey equilibrium given a period-by-period balanced budget constraint.⁷

We specify the per-period utility function of the CES class as

$$u(c, \ell, g) = \frac{\left[(1 - \alpha_p) (\alpha_c c^\rho + (1 - \alpha_c) \ell^\rho)^{\Psi/\rho} + \alpha_p g^\Psi \right]^{\frac{1-\sigma}{\Psi}} - 1}{1 - \sigma}. \quad (17)$$

This function reduces to a separable function with constant expenditure shares when $\sigma \rightarrow 1$, $\rho \rightarrow 0$ and $\Psi \rightarrow 0$, yielding

$$u(c, \ell, g) = (1 - \alpha_p) \alpha_c \ln c + (1 - \alpha_p)(1 - \alpha_c) \ln \ell + \alpha_p \ln g \quad (18)$$

Meanwhile, the production function is a standard Cobb-Douglas function with capital share θ : $f(K, L) = A \cdot K^\theta L^{1-\theta}$.

Our parameterization of the baseline economy is also standard. We calibrate the baseline model economy, which is the one with only labor taxes, to have some statistics within the range of U.S. data in the lack-of-commitment economy. So we set the share of GDP that is spent by the government to be slightly under 20%, the capital share to 36%, the investment-to-output ratio to a little over 20%, hours worked to about

⁷The discussion of the intuition underlying the Ramsey equilibrium in this section is kept quite brief; more extensive formal analysis of this case is contained in another paper: Klein, Krusell, and Ríos-Rull (2003). Related insights are also obtained in Stockman (1998). For earlier analysis of a setup without commitment, see Klein and Ríos-Rull (1999) who perform a quantitative analysis of optimal taxation (labor and capital income taxes) for exogenous public expenditures under a period-by-period balanced budget constraint.

one fourth of total time, and the capital-to-output ratio to about 3. These choices are common in the macroeconomic literature.

We choose the baseline economy to have logarithmic utility which makes preference separable (making cross derivatives zero). We report the values of the parameters that implement our choices in Table 1.

Parameter values		
$\theta = 0.36$	$\alpha_c = 0.30$	$\alpha_p = 0.13$
$\beta = 0.96$	$\delta = 0.08$	$\rho = 0$
$\Psi = 0$	$\sigma = 1.0$	

Table 1: Parametrization of the baseline model economy

5.1 Labor income taxation

We now look at the steady states of the baseline economy under three different benevolent governments that we label Pareto, Ramsey, and Markov. These labels, respectively, refer to: a government with commitment and access to lump-sum taxation (Pareto); a government restricted by a period-by-period balanced-budget constraint and to the use of labor income taxation, both one with access to a commitment technology (Ramsey) and one which does not have access to such commitment technology (Markov, because we look at the Markov equilibrium). Table 2 reports the steady-state allocations of these three economies.

The absence of capital income taxes in all economies ensures that the steady-state interest rate is equated to the rate of time preference, yielding an equal capital-to-output ratio in all economies. Comparing the Pareto and the Ramsey economies, we see the effect of distortionary labor taxation. The Pareto economy delivers the optimal allocation while the Ramsey economy has a distortionary tax that discriminates against produced goods and in favor of leisure. As a result, leisure is significantly higher in the

Labor taxes, endogenous g			
steady state statistic	type of government		
	Pareto	Ramsey	Markov
y	1.000	0.700	0.719
k/y	2.959	2.959	2.959
c/y	0.509	0.509	0.573
g/y	0.254	0.254	0.190
c/g	2.005	2.005	3.017
ℓ	0.350	0.245	0.252
τ	–	0.397	0.297

Table 2: Baseline model economy

Ramsey economy than in the Pareto economy, and because of this and the equal rate of return, the steady-state stock of capital and output are much lower in the Ramsey economy. However, the ratio between private and public consumption is the same in both economies given that this margin is undistorted. This latter feature is a special implication of the functional form that we have chosen and it relies on preferences being separable in all three goods and on being of the CRRA class with respect to consumption.⁸

When we look at the behavior of the Markov economy, we see two things: first, qualitatively, the distortion introduced by the tax on labor is also present in this economy, inducing more leisure and less consumption (both private and public) than in the Pareto economy; and second, the ratio between private and public consumption is not the same as in the other economies (where it was equal to the relative share parameter in preferences). Recall that from equation (14) the optimal policy of the Markov case amounted to striking a balance between achieving the first best in terms of equating the marginal utility of the private and public good and the distortion that the labor tax induces on the leisure–private consumption margin. This balance does not imply

⁸This is a simple implication of the first-order conditions of the Ramsey problem when written in the primal form.

setting the margin between the public and the private good to zero. Indeed, the term $u_g - u_c$ is positive in the Markov case, making the second term of equation (14) positive and the first one negative. The difference with the Ramsey case can perhaps be best described by the fact that the Ramsey policy maker takes into account the fact that a tax hike at t not only lowers labor supply at t but raises it at $t - 1$. In contrast, a Markov policy-maker treats the latter as a bygone and hence chooses lower tax rates.

5.2 Capital income taxation

Table 3 shows the steady state when the only available tax is the capital income tax.

Capital income taxes, endogenous g			
steady state statistic	type of tovernment		
	Pareto	Ramsey	Markov
y	1.000	0.588	0.488
k/y	2.959	1.734	1.193
c/y	0.509	0.712	0.690
g/y	0.254	0.149	0.215
c/g	2.005	4.779	3.211
ℓ	0.350	0.278	0.255
τ	–	0.673	0.812

Table 3: Baseline model economy

This tax is in general very distortionary. The Ramsey government understands this and, therefore, reduces future taxes so as to mitigate the distortionary effect. However, since no other tax base is available here, the result is that the ratio of private to public consumption is much lower than in the unconditional first best. The Markov government, however, does not see the current tax as distortionary at all, as capital is already installed when the government chooses the tax rate: capital is inelastically supplied.

The Markov government, however, understands that the government that follows one period later will distort the allocation significantly, and is therefore willing to attempt to transfer resources into the future to increase future consumption. For this reason, it does not tax capital so as to set the private-to-public consumption ratio at the first-best level. The ability of the Markov government to influence the future choices is of course smaller than that of the Ramsey government, and as a result its capital tax rate is higher and capital and output are lower.

Another interesting feature of this case is that leisure is the lowest in the Pareto case, even when there is no tax on leisure. With the preferences of this model economy, in any market implementation, the household's choice of leisure can be decomposed into two parts. One part is what it would choose if all income were labor income—it equals $(1 - \alpha_c)$ exactly, independently of the wage (that in this case is 0.7). The other part comes from the amount of additional income that the household has, so that leisure is increasing in that additional income. In the Pareto economy, the lump-sum tax levied is larger than the amount of capital income, inducing the household to enjoy less leisure than 0.7, while in all the other economies, the after-tax capital income is always positive, which accounts for why workers enjoy leisure of more than 0.7 in those economies.

5.3 General income taxation

With respect to the case of a tax on total income, a couple of points are worth stressing.

First, the Ramsey government can set the ratio of private to public consumption to its unconditionally optimal level. Due partly to the special nature of the preferences used in this model economy, the distortions that affect the intertemporal margin and the consumption leisure margin do not affect the private-to-public-consumption margin. From the point of view of the Markov government, however, this is not the case. An uncommitted policy maker does not take into account that today's taxes increase yesterday's incentives to work, and in addition it wishes to increase savings by taxing less today, and these effects induce a smaller government sector. This result is perhaps surprising because one might have guessed that a Markov government, which views its taxes as less distortionary than does the Ramsey government, would tax more.

Total income taxes, endogenous g			
steady state statistic	type of government		
	Pareto	Ramsey	Markov
y	1.000	0.669	0.693
k/y	2.959	2.527	2.649
c/y	0.509	0.532	0.587
g/y	0.254	0.265	0.201
c/g	2.005	2.005	2.928
ℓ	0.350	0.256	0.258
τ	–	0.334	0.255

Table 4: Baseline model economy

In addition to the comparisons that we have performed between the three taxing technologies that the government may have access to (and that yield the Pareto, Ramsey, and Markov cases), for each of the tax tools, we should also compare the allocations for the Markov case across tax instruments.

From the point of view of the Markov government, taxing capital is not distortionary since it is already installed and hence is like a lump-sum tax. On the other hand, the tax base is quite small, as capital income is much smaller than labor income.⁹ On the other hand, labor taxes are distortionary but its base is larger. Finally, total income taxes have the highest tax base and they are as distortionary as the labor income tax rate for the same tax rate, or less distortionary for the same revenue.

With respect to tax outcomes, first, as should have been expected, the larger is the role of capital income taxes (which implies an ordering with capital income first, followed by total income and last labor income), the lower is the stock of capital, and hence the lower is output. The differences are large. Second, hours worked are actually varying very little across environments. Third, perhaps the most surprising feature

⁹Note that because the tax base excludes depreciation, the tax base of a capital income tax is not a constant fraction of GDP.

that we obtain is that the ratio of private consumption to public consumption is the highest in the capital-tax economy. This is very surprising, since we should expect that the government, since it considers taxes to be non-distortionary, would allocate current resources optimally across these goods, thus equating the marginal utility of public and private consumption (which is what the Pareto government does). The reason why this does not occur is that the government in the capital-income economy understands that the next government will tax capital heavily (more heavily, indeed, than what this government would like), and in an effort to move resources into the future it thus sacrifices current public consumption. Note also that this effect is non-linear in that the private-to-public consumption ratio closest to the first best is that of the total income-tax economy.

6 Conclusion

In this paper we have characterized time-consistent equilibria of an environment where a benevolent government that does not have access to commitment sets tax rates to finance a public good. We have shown how the problem of the government has a recursive structure with only one nonstandard feature: the restriction on the government—the behavior of the private sector—is summarized by an equation that embodies private-sector expectations of future government taxation behavior. Given these expectations, successive governments agree on taxation. The relevance of the lack of commitment is thus not the lack of direct power over next period’s taxation decisions but rather the inability to influence private-sector expectations. Thus, the government only has an imperfect ability to influence current private decisions such as savings and work effort. For example, absent the ability to make the private sector “believe” that capital income will not be taxed next period, the government can only increase current savings by spending less resources on current public goods.

Our analysis leads to a natural characterization of government behavior in terms of a first-order condition—the GEE. We have emphasized the interpretative and computational value of the GEE, and we have explored calibrated versions of it. For U.S.-like parameter values, the level of (capital income) taxation without commitment is surprisingly low presuming that the tax base is total income. This result is due in part to the

idea that holding back on taxation helps saving, thus alleviating the distortion that future capital income taxation has. However, the equilibrium without commitment also has much lower taxes on labor income than does the equilibrium with commitment, because a government with commitment realizes that current taxation is a way of making earlier consumers work harder, thus alleviating the labor-leisure distortion.

We know that our conclusions depend at least quantitatively on the assumption of a balanced budget: the government might otherwise tax extremely heavily early on to build up a large surplus from which future spending can be financed. We do not, however, observe large asset accumulation for governments in present-day economies and it is an open question as to why this is so. The research program to which the present paper is a beginning aims to systematically explore optimal government decision making with and without commitment, with and without the ability to borrow and lend, and with and without other intertemporal “frictions” such as differences in preferences that may be due to the political process or agency problems, with voters being principals and governments agents. We believe that the organization of this analysis around a stochastic version of the GEE, summarizing the tradeoffs governments face, could be as useful as the Euler equation has been to the consumption literature. Before empirical work can begin, however, we need to understand how to derive these optimality conditions for various settings and how to compare them across different economies, and the present paper precisely starts this investigation. An important problem in this context is computational tractability, and our work suggests methods that are a simple generalization of, and very close to as easy to use as, the well-known linearization techniques used in the business cycle literature.

Finally, we believe that the methods we use here are straightforwardly applicable to other areas with dynamic interaction without commitment, e.g., for studying optimal monetary policy, dynamic political economy, dynamic industrial organization issues (e.g., the durable goods monopoly and dynamic oligopoly), models with impure intergenerational altruism, and so on.

References

- Abreu, D., D. Pearce, and E. Stacchetti (1990). Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica* 58, 1041–1063.
- Ambler, S. (2001). Time consistent control in non-linear models. Unpublished working paper, UQAM.
- Chari, V. V. and P. J. Kehoe (1990). Sustainable plans. *Journal of Political Economy* 98(4), 784–802.
- Cohen, D. and P. Michel (1988). How should control theory be used to calculate a time-consistent government policy? *Review of Economic Studies* 55, 263–274.
- Currie, D. A. and P. L. Levine (1993). *Rules, Reputation and Macroeconomic Contracts*. Cambridge University Press, Cambridge.
- Klein, P., P. Krusell, and J.-V. Ríos-Rull (2003). Time-consistent public expenditures. Mimeo, University of Pennsylvania.
- Klein, P. and J.-V. Ríos-Rull (1999). Time-consistent optimal fiscal policy. manuscript.
- Krusell, P., B. Kuruşçu, and A. A. Smith (2000). Tax policy with quasi-geometric discounting. Manuscript.
- Krusell, P., V. Quadrini, and J.-V. Ríos-Rull (1997, January). Politico-economic equilibrium and economic growth. *Journal of Economic Dynamics and Control* 21(1), 243–272.
- Krusell, P. and J.-V. Ríos-Rull (1999). On the size of U.S. government: Political economy in the neoclassical growth model. *American Economic Review* 89(5), 1156–1181.
- Krusell, P. and A. A. Smith (2003). Consumption-savings decisions with quasi-geometric discounting. Forthcoming in *Econometrica*.
- Kydland, F. E. and E. C. Prescott (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy* 85(3), 473–91.
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *qje*, 443–477.
- Laibson, D. I. (1994). *Hyperbolic Discounting and Consumption*. PhD Thesis, Massachusetts Institute of Technology.
- Lucas, R. E. J. and N. L. Stokey (1983). Optimal fiscal and monetary policy in an economy without capital. *Journal of Monetary Economics* 12, 55–93.
- Persson, T. and L. E. O. Svensson (1989). Why a stubborn conservative would run a deficit: Policy with time-inconsistent preferences. *Quarterly Journal of Economics* 104(2), 325–45.

- Phelan, C. and E. Stacchetti (2001). Sequential equilibria in a ramsey tax model. *Economica* 69(6), 1491–1518.
- Phelps, E. S. and R. A. Pollak (1968). On second-best national saving and game-equilibrium growth. *Review of Economic Studies* 35(185-199).
- Pollak, R. (1971). Additive utility functions and linear engel curves. *Review of Economic Studies* 38, 401–413.
- Rogoff, K. (1985). The optimal degree of commitment to an intermediate monetary target. *Quarterly Journal of Economics* 100(4), 1169–1190.
- Stockman, D. R. (1998). Balanced-budget rules: Welfare loss and optimal policies. Mimeo, University of Delaware.
- Strotz, R. H. (1956). Myopia and inconsistency in dynamic utility maximization. *res* 23, 165–80.

Appendix

A Our computational algorithm

The compact definition of equilibrium of Section 3.2 is most convenient at this stage: h is a simpler object to characterize than \mathcal{H} , since it has one less argument.

A.1 The functional equations

Recall the two first-order conditions: the one for the private sector,

$$0 = u_c - \beta u'_c [1 + (1 - \mathcal{T}') (f'_k - \delta)], \quad (19)$$

and the one for the government,

$$0 = (-u_c + u_g) \eta_{k'} - \eta_g \left[-u_c + \beta u'_c (f'_k + 1 - \delta) - \beta (-u'_c + u'_g) \frac{\eta'_k}{\eta'_g} \right]. \quad (20)$$

These are functional equations: they hold for all k . The derivatives of η are derived from the definition of η in Section 3.2; they are

$$\eta_k = u_{cc} \mathcal{C}_k \quad (21)$$

$$\eta_g = u_{cc} \mathcal{C}_g + u_{cg} \quad (22)$$

$$\eta_{k'} = u_{cc} \mathcal{C}_{k'} + u'_c [1 - \mathcal{T}'] f'_{kk} - u_c [f'_k - \delta] [\mathcal{T}'_k + \mathcal{T}'_g \Psi'_k] + \{1 + [1 - \mathcal{T}'] [f'_k - \delta]\} \{u_{cc} [\mathcal{C}'_k + \mathcal{C}'_{k'} h'_k + \mathcal{C}'_g \Psi'_k] + u_{cg} \Psi_k(k')\} \quad (23)$$

If we substitute equations (21-23) into equations (19) and (20) we obtain a system of two equations that we can write compactly as

$$0 = \xi^p \{k, h(k), \Psi(k), h[h(k)], \Psi[h(k)]\} \quad (24)$$

$$0 = \xi^g \{k, h(k), \Psi(k), h[h(k)], \Psi[h(k)], h'_k(k), \Psi'_k(k)\}. \quad (25)$$

Global computation of a solution to the pair of functional equations could be operationalized in a number of ways, including postulating flexible parameterized functional forms for h and Ψ and requiring that the functional equations hold exactly on an appropriately chosen grid, or that the error to these equations be minimized over a large number of grid points. Here, however, we will only solve for steady states, and thus a simple generalization of a linearization method can be used.

A.2 The steady state

A steady state is a pair of values k^* and g^* such that the two functional equations are satisfied when setting $k = k' = k''$ and $g = g'$. Doing this yields

$$0 = \xi^p(k^*, k^*, g^*, k^*, g^*) \equiv \xi^{p*}(k^*, g^*) \quad (26)$$

$$0 = \xi^g(k^*, k^*, g^*, k^*, g^*, h_k^*, \Psi_k^*) \equiv \xi^{g*}(k^*, g^*, h_k^*, \Psi_k^*). \quad (27)$$

Using this compact form, we see two equations and four unknowns: the vector of steady-state values for k and g and the first derivatives of their associated decision rules evaluated at the steady state: $(k^*, g^*, h_k^*, \Psi_k^*)$. This means that levels cannot be solved for without knowing derivatives.

The method we use to solve for a steady state is outlined for a simpler problem in Krusell, Kuruşçu, and Smith (2000).

The algorithm builds on (i) constructing a set of local approximations of order m —here, m -order polynomials—to the functions h and Ψ ; (ii) denoting these approximations $\varphi^{p,m}(k)$ and $\varphi^{g,m}(k)$, respectively, solving for the steady state given m ; (iii) increasing m until the steady state changes by less than some convergence criterion. We now show in more detail how such an algorithm is implemented.

1. When $m = 0$, the functions φ are constants. With two equations—equations (26) and (27)—and two unknowns (using the fact that the derivatives are zero) there is typically a unique solution. Denote the implied steady state $\{k^0, g^0\}$.
2. For $m = 1$, the functions φ are linear, yielding $k' = \varphi_0^{k,1} + \varphi_1^{k,1} k$ and $g = \varphi_0^{g,1} + \varphi_1^{g,1} k$; this means that all derivatives of order 2 and above are zero and that the functions are entirely specified by their levels and derivatives at the steady state. Now the 4 unknowns necessitate 4 equations. We thus keep the equations from the previous step and differentiate each of these with respect to k ; this is valid (assuming differentiability) since the equations have to hold for all k . Thus we have 4 equations and 4 unknowns. Imposing the steady-state condition and substituting k by $\frac{\varphi_0^{k,1}}{1-\varphi_1^{k,1}}$, g by $\frac{\varphi_0^{g,1}}{1-\varphi_1^{g,1}}$, $h_k(k)$ by $\varphi_1^{k,1}$, and $\Psi_k(k)$ by

$\varphi_1^{g,1}$ we have

$$0 = \xi^p \tag{28}$$

$$0 = \xi^g \tag{29}$$

$$0 = \xi_k^p + \varphi_1^{k,1} \xi_{k'}^p + \varphi_1^{g,1} \xi_p^p + \left(\varphi_1^{k,1}\right)^2 \xi_{k''}^p + \varphi_1^{g,1} \varphi_1^{k,1} \xi_{g'}^p \tag{30}$$

$$0 = \xi_k^g + \varphi_1^{k,1} \xi_{k'}^g + \varphi_1^{g,1} \xi_g^g + \left(\varphi_1^{k,1}\right)^2 \xi_{k''}^g + \varphi_1^{g,1} \varphi_1^{k,1} \xi_{g'}^g, \tag{31}$$

where equations (30) and (31) use the fact that $h_{kk}(k)$ and $\Psi_{kk}(k)$ are zero because these functions are assumed to be linear at this stage of the iteration. In this equation system, the ξ functions and their derivatives of course depend on the four unknowns, and a nonlinear solver has to be used to deliver the unknowns, and hence $\{k^1, g^1, h_k^1, \Psi_k^1\}$.

3. Turning to $m = 2$, there are six unknowns which are uniquely determined by the values of h and Ψ and their first two derivatives at a given point. The six equations are the four equations from the previous step plus those that result from differentiating the last two equations once more with respect to k .
4. The procedure is repeated until the steady-state values for k and g (and possibly some low-order derivatives, if local dynamics are also an object of study) change by a small amount.

Two specific additional comments are in order. First, to differentiate the first-order condition (multiple times) one can either use numerical differentiation or use symbolic differentiation using a package like MAPLE. The latter imposes no bound on the number of derivatives that can be computed; numerical derivatives of high order are hard to obtain with precision.

Second, to solve the nonlinear equation system at step m , which involves $2(m + 1)$ equations and unknowns, one can of course use brute force. However, it is also possible to use an inherent recursivity in the system. This recursivity, however, requires computing not the coefficients in the polynomials for h and Ψ but the associated sequence of derivatives. In terms of these derivatives, (i) the first equation always contains two levels and no derivatives; (ii) the next two equations contain the two levels and two first-order derivatives; (iii) the next two equations contain the two levels, the two first-order derivatives, and the two second-order derivatives; and so on until the last equation, which contains no new higher-order derivatives, since these are assumed to be zero. Thus, one guesses on, say, k^m , uses the first equation to solve for g^m , the next two to solve for the two first-order derivatives, the following two to solve for the two

second-order derivatives, etc., until all the non-zero derivatives have been calculated; the last equation remains, and it has to be satisfied, which is ensured by iteration on the initial choice k^m . Thus, at no stage is it necessary to solve more than two equations and two unknowns with this recursive method.