

Time-Consistent Public Policy

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

^a*University of Western Ontario*

^b*Princeton University, IIES, CAERP, CEPR, and NBER*

^c*University of Pennsylvania, CAERP, CEPR, and NBER*

July 30, 2006

Abstract

In this paper we study how a benevolent government which cannot commit to future policy should trade off the costs and benefits of public expenditure. We characterize and solve for Markov-perfect equilibria of the dynamic game between successive governments. The characterization consists of an intertemporal first-order condition (a “generalized Euler equation”) for the government and we use it both to gain insight into the nature of the equilibrium and as a basis for computations. For a calibrated economy, we find that when the only tax base available to the government is capital income—an inelastic source of funds at any point in time—the government still refrains from taxing at confiscatory rates. We also find that, when the only tax base is labor income, then the Markov equilibrium features less public expenditure and lower tax rates than the Ramsey equilibrium.

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

JEL classification: E62, H21.

*Correspondence to Krusell at the Department of Economics, Princeton University. Krusell thanks the National Science Foundation, Ríos-Rull thanks the National Science Foundation (Grant SES-0079504) and the University of Pennsylvania Research Foundation, and Klein thanks the Social Science and Humanities Research Council of Canada for support. The authors thank Harald Uhlig, the editor, and two anonymous referees for very useful comments.

1 Introduction

In this paper we study how a benevolent government should trade off the costs and benefits of public expenditure. We do this in the context of a neoclassical growth model where consumers value government consumption. Because this framework features dynamic inconsistency, we have to take a stand on whether the government can commit itself to future policy. The view we take here is that it cannot. It follows from this view that the situation is best described as a dynamic game between successive governments. We characterize and solve for Markov-perfect equilibria of this game, focusing on those Markov equilibria that feature differentiable policy functions. This refinement concept is designed to rule out anything resembling a reputation-like mechanism for sustaining “good” equilibria.¹

The characterization involves deriving a first-order condition for the government that captures the trade-off we are interested in. We call this condition the “generalized Euler equation” or GEE. It reveals how the government optimally trades off tax wedges over time, and this trade-off involves, among other things, the future government’s marginal propensity to tax out of their initial capital.

We provide analysis for two sets of economies. The first set is designed to admit analytical solutions. When the horizon is finite, the solution is typically unique, and as the time horizon tends to infinity the solution tends to a well-defined limit that coincides with the differentiable Markov-perfect equilibrium of the corresponding infinite-horizon economy. This equilibrium has a unique steady state, and convergence to it is monotone.² In some examples, however, there are two equilibria with constant tax rates: one with a high rate and one with a low rate. The former is the limit of finite-horizon equilibria; the latter requires an infinite horizon.

The second set of economies is calibrated to match some features of the postwar U.S. economy and solved numerically. Here, the purpose is not to provide an empirical evaluation of the model nor to account for any quantitative facts. Rather, the purpose

¹ In employing this refinement concept, we differ from Phelan and Stacchetti (2001), who have developed methods to find all “sustainable” equilibria in environments similar to ours. For a fuller discussion of how our work is situated in the literature, see Section 5.

² This contrasts with the nonmonotone convergence result in Hassler, Krusell, Storesletten, and Zilibotti (2004). See also Section 5.

is to determine the robustness of the analytically obtained results to changes in parameter values in the direction of those that are reasonable given the available data. The computations also serve to illustrate our solution method. In general, numerical solutions can be difficult to obtain in setups like ours. This is because the nature of the equilibrium is quite different from what is mostly studied in macroeconomics. Even finding a steady state—the lowest-order representation of the government’s policy function—is difficult, because the level of taxes in a steady state is determined by a condition that involves the future marginal propensity to tax, which is a higher-order feature of the same, unknown policy function. This feature is not present either in standard competitive equilibria or in Ramsey (optimal-tax) equilibria. Our method is based on approximating the policy function and its derivatives evaluated at one point only: the steady state. For the economies we consider, the method performs very well: in the case where an analytical solution exists, the numerical method finds it with very small errors, and in other cases, convergence is fast and the solutions are very close to those obtained with global methods.

The results of our analytical and numerical calculations can be summarized as follows. When capital income is the only tax base, then our Markov equilibrium, perhaps surprisingly, since capital income is inelastic ex post, does not result in equating the marginal rate of substitution between private and government-provided goods to the marginal rate of transformation. Rather, public goods are underprovided in order to mitigate underinvestment. The steady state represents a compromise between the distortion of the public/private good margin and the consumption/investment margin.³ When, on the other hand, labor income is the only tax base and labor supply is elastic, the Markov equilibrium steady-state ratio of public to private goods consumption is smaller than the corresponding Ramsey equilibrium long-run ratio. The reason is that a higher tax rate in period t encourages labor supply in previous periods; this beneficial effect (it is beneficial since labor supply is suboptimal in equilibrium) is ignored ex post by a government without commitment. Finally, if the government is able to tax labor and capital income separately, but faces the constraint that labor cannot be subsidized, then the optimal choice is to set the labor tax equal to zero, as we show in Appendix

³ If the government were allowed to encourage investment in some other way than by refraining from confiscatory taxation of capital income, say by providing a lump-sum transfer or by subsidizing labor, it would. This is discussed in Martin (2005). See also Appendix B.

B. Thus, this case collapses to the case where only capital income is taxed.

An important conceptual insight coming out of our work concerns the nature of the interaction between successive governments. It turns out that in our framework, the current government does not want to manipulate its successor via the state variable, in the spirit of Persson and Svensson (1989), or as in the case of savers with time-inconsistent preferences. The reason is that, in our environment, it is the constraints that are time-inconsistent and not the preferences. 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—it is constrained by private agents forming rational expectations based on equilibrium prices—but its policy choice will influence private capital accumulation which has an equilibrium influence on future policy, and thus on current expectations. This effect *is* taken into account by the government.

The 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 Markov equilibrium (Section 2.3) step by step. The section presents closed-form solutions for specific versions of the model. Section 3 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 4 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). Section 5 concludes and discusses the relation of the present paper with the literature. The Appendix includes the description of the computational procedures we use and a treatment of the case with separate tax rates for capital and labor income.

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 lives for T periods 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:

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

and

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

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 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 $\{g_t\}_{t=0}^T$.⁴ A simple way to describe this problem formally is to 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

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

⁴Formally, letting the government choose g instead of τ can be a problem if more than one tax rates are 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.

for all $t \geq 0$. We refer to the solution of this problem as the Ramsey allocation. As is well known, this solution is typically not time-consistent.

2.3 No commitment: Markov equilibrium

The equilibrium concept we consider here can be stated in different ways; the particular formulation we use is practical in that it allows for compact statements.⁵ We assume an infinite time horizon, but the definition can very easily be modified to apply to a finite horizon.

An equilibrium consists of a value function v , a government policy function Ψ and a savings function h such that, for all k , $g = \Psi(k)$ and $k' = h(k)$ solve

$$\begin{aligned} & \max_{k',g} \{u(\mathcal{C}(k, k', g), g) + \beta v(k')\} \\ & \text{subject to} \\ & u_c(\mathcal{C}(k, k', g), g) = \\ & \beta u_c(\mathcal{C}(k', h(k'), \Psi(k')), \Psi(k')) \cdot \{1 + [1 - \mathcal{T}(k', \Psi(k'))][f_k(k') - \delta]\} \end{aligned} \quad (2)$$

and

$$v(k) \equiv u(\mathcal{C}(k, h(k), \Psi(k)), \Psi(k)) + \beta v(h(k)).$$

Here, for the recursive construction to be useful, one needs assumptions on primitives (such as on u , the domain for capital, etc.) such that this v also solves the consumer's problem when stated sequentially. Our examples below satisfy these conditions.

The following definition will prove useful later on.

$$\begin{aligned} \eta(k, g, k') & := u_c(\mathcal{C}(k, k', g), g) - \\ & \beta u_c(\mathcal{C}(k', h(k'), \Psi(k')), \Psi(k')) \cdot \{1 + [1 - \mathcal{T}(k', \Psi(k'))][f_k(k') - \delta]\}. \end{aligned} \quad (3)$$

Given an equilibrium (v, Ψ, h) we can define the function \mathcal{H} via $\eta(k, g, \mathcal{H}(k, g)) \equiv 0$. This function describes the response of the private sector to a “one-shot deviation” on

⁵ It was suggested to us by Harald Uhlig.

the part of the government in the following sense.⁶ If the current capital stock is k , current government consumption is g and the private sector expects that future government policy will be determined by the equilibrium policy function Ψ , then savings will be given by $\mathcal{H}(k, g)$.

2.3.1 The GEE

The first-order condition of the government—the generalized Euler equation—can be derived by using the envelope condition to substitute out the derivative of the value function.⁷ With use of the function η (defined above), the GEE simplifies to

$$0 = (-u_c + u_g)\eta_{k'} - \eta_g \left[-u_c + \beta u'_c(f'_k + 1 - \delta) - \beta(-u'_c + u'_g) \frac{\eta'_k}{\eta'_g} \right]. \quad (4)$$

Meanwhile, the \mathcal{H} is defined so that $\mathcal{H}_g = -\frac{\eta_g}{\eta_k}$ and $\mathcal{H}_k = -\frac{\eta_k}{\eta_{k'}}$, so that we may write

$$\beta \mathcal{H}_g \left\{ u'_c [f'_k + 1 - \delta - \mathcal{H}'_k] + \frac{\mathcal{H}'_k}{\mathcal{H}'_g} (u'_c(\mathcal{H}'_g + 1) - u'_g) \right\} = 0, \quad (5)$$

This is a *functional* equation, a fact that is obscured by the suppression of the functional arguments.

2.3.1.1 A “wedge” interpretation The GEE can be rewritten as a linear combination of wedges, the public/private wedge $\gamma := u_g - u_c$ and the savings/consumption wedge $\nu := -u_c + \beta u'_c(1 + f'_k - \delta)$. Thus, rearranging terms we obtain

$$\gamma + \mathcal{H}_g \nu + \beta \mathcal{H}_g \left(-\frac{\mathcal{H}'_k}{\mathcal{H}'_g} \right) \gamma' = 0. \quad (6)$$

⁶ The function \mathcal{H} can be used more directly in the equilibrium definition; see, e.g., Krusell and Ríos-Rull (1999).

⁷ It can also be derived by using Bellman’s principle to identify a Markov equilibrium with the solution to an appropriate sequence problem. (This is of course not the problem that the policy-maker actually solves.) The Euler equation of this sequence problem is the GEE.

Three terms appear: these are the three different “wedges” that are affected by the change in the current tax rate: the current public/private good wedge, the consumption/saving wedge and the next period’s public/private good wedge. The weight on the final term can be thought of as the derivative $\frac{dg'}{dg}$ when the household’s first-order condition $\eta(k', g', k'') = 0$ is satisfied and k'' is held constant.

2.3.2 Comparison with the commitment case

In the commitment case, the functions h and Ψ are not applicable, so we need to modify the definition of η that we gave in (3) as follows.⁸

$$\begin{aligned} \tilde{\eta}(k, k', k'', g, g') &:= u_c(\mathcal{C}(k, k', g), g) - \\ &\beta u_c(\mathcal{C}(k', k'', g'), g') \cdot \{1 + [1 - \mathcal{T}(k', g')][f_k(k') - \delta]\}. \end{aligned} \quad (7)$$

With this new definition, the no-commitment GEE becomes, in sequential terms,

$$\gamma_t - \frac{\tilde{\eta}_{4,t}}{\tilde{\eta}_{2,t} + \tilde{\eta}_{3,t}h_{k,t+1} + \tilde{\eta}_{5,t}\Psi_{k,t+1}} \left(\nu_t - \frac{\tilde{\eta}_{1,t+1}}{\tilde{\eta}_{4,t+1}}\beta\gamma_{t+1} \right) = 0. \quad (8)$$

By contrast, the first-order condition for optimal government spending under commitment, which follows from straightforward calculus, is

$$\gamma_t - \tilde{\eta}_{4,t}D_t \left(\nu_t - \frac{\tilde{\eta}_{1,t+1}}{\tilde{\eta}_{4,t+1}}\beta\gamma_{t+1} - \frac{\tilde{\eta}_{3,t-1}}{\tilde{\eta}_{5,t-1}}\gamma_t \right) - \frac{\tilde{\eta}_{5,t-1}D_{t-1}}{\beta} \left(\nu_{t-1} - \frac{\tilde{\eta}_{1,t}}{\tilde{\eta}_{4,t}}\beta\gamma_t - \frac{\tilde{\eta}_{3,t-2}}{\tilde{\eta}_{5,t-2}}\gamma_{t-1} \right) = 0 \quad (9)$$

where $\tilde{\eta}_{4,t}D_t$ captures a “ $\frac{dk_{t+1}}{dg_t} |_{\tilde{\eta}_t=0}$ ” effect (the implied change in k_{t+1} that keeps $\tilde{\eta}_t = 0$):

$$\frac{1}{D_t} := \tilde{\eta}_{2t} - \tilde{\eta}_{4t} \frac{\tilde{\eta}_{3t-1}}{\tilde{\eta}_{5t-1}} - \tilde{\eta}_{5t} \frac{\tilde{\eta}_{1t+1}}{\tilde{\eta}_{4t+1}}.$$

The first three terms of (9) have direct counterparts in (8).⁹ The remaining terms

⁸ By definition of h and Ψ , we obviously have $\eta(k, g, k') \equiv \tilde{\eta}(k, k', h(k'), g, \Psi(k'))$.

⁹ Note only that under commitment, future control variables can be chosen directly, so $-\tilde{\eta}_{4t} \frac{\tilde{\eta}_{3t-1}}{\tilde{\eta}_{5t-1}}$ and $-\frac{\tilde{\eta}_{1t+1}}{\tilde{\eta}_{4t+1}}$ here replace $\tilde{\eta}_{3,t}h_{k,t+1}$ and $\Psi_{k,t+1}$, respectively, in the no-commitment GEE.

involve, in one way or another, period $t - 1$; they represent the impact of current decisions on past expectations, and are hence ignored when there is no commitment. Specifically, the fourth term captures how the induced change in k_{t+1} alters last period's expectations, thus altering g_t indirectly: it captures " $\frac{dg_t}{dk_{t+1}}|_{\eta_{t-1}=0}$ ". Finally, terms five through seven are a replica of terms two through four: the latter represent the effects of the change in g_t on k_{t+1} keeping $\eta_t = 0$, whereas the former involve the same kinds of effects, but via how the change in g_t changes k_t keeping $\eta_{t-1} = 0$, i.e., via the household's first-order condition last period.

We have stated the commitment GEE in its full dynamic form here; in steady state, of course, it simplifies greatly and the commitment and no-commitment GEEs are easier to compare. In the steady-state context, the additional terms appearing in the commitment GEE are given some quantitative content in Section 4 below.

2.3.3 An economy that allows a closed-form solution

We will consider an example that illustrates the nature of a Markov-perfect equilibrium and how it relates to the Ramsey solution. So suppose that $u(c, g) = \log c + \gamma \log g$, $f(k) = k^\theta$, and $\delta = 1$. We will assume that the only tax base is capital income (without any deduction for depreciation), because the two other cases (a tax on labor income and a tax on total income) turn out not to be very interesting: in the former case, the solution coincides with the Pareto optimum, since labor taxes are non-distortionary; in the latter case, the Ramsey optimal solution is time-consistent and hence it is also a Markov-perfect equilibrium.

In this case we have $g_t = \tau_t \theta k_t^\theta$ and $\mathcal{T}(k, g) = \frac{g}{\theta k^\theta}$. Savings rules can be written as

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

where savings rates satisfy the recursion

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

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 decrease 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 a time-consistent equilibrium. Optimal government policy can be represented via

$$g_t = \tau_t \theta k_t^\theta$$

and this leads to a value function which can be written as

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

where

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

and B_t 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

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

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.

2.3.4 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 derivative 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 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.

3 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,

$$\frac{u_\ell(\mathcal{C}(k, \mathcal{H}(k, g), g), 1 - \mathcal{L}(k, g), g)}{u_c(\mathcal{C}(k, \mathcal{H}(k, g), g), 1 - \mathcal{L}(k, g), g)} = f_\ell(k, \mathcal{L}(k, g)) (1 - \mathcal{T}(k, g)) \quad (10)$$

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

$$\begin{aligned} & \mathcal{L}_g \left[u_c f_\ell - u_\ell \right] + \left[u_g - u_c \right] + \mathcal{H}_g \left[-u_c + \beta u'_c (1 + f'_k - \delta) \right] + \\ & \beta \mathcal{H}_g \left\{ \mathcal{L}'_k \left[u'_c f'_\ell - u'_\ell \right] - \frac{\mathcal{H}'_k}{\mathcal{H}'_g} \left(\mathcal{L}'_g \left[u'_c f'_\ell - u'_\ell \right] + \left[u'_g - u'_c \right] \right) \right\} = 0 \end{aligned} \quad (11)$$

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 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 2.3.3, it is straightforward to see that $\Psi(k) = \frac{\gamma}{\alpha + \gamma} (1 - \beta\theta) k^\theta \ell^{1-\theta}$, with $\mathcal{L}(k, g) = \frac{\alpha(1-\theta)}{\alpha(1-\theta) + (1-\alpha)(1-\beta\theta)}$ 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 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

$$\mathcal{L}_g \left[u_c f_\ell - u_\ell \right] + \left[u_g - u_c \right] + \beta \mathcal{H}_g \left\{ \left(\mathcal{L}'_k - \frac{\mathcal{H}'_k}{\mathcal{H}'_g} \mathcal{L}'_g \right) \left[u'_c f'_\ell - u'_\ell \right] - \frac{\mathcal{H}'_k}{\mathcal{H}'_g} \left[u'_g - u'_c \right] \right\} = 0. \quad (12)$$

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

$$\left[u_g - u_c \right] + \mathcal{H}_g \left[-u_c + \beta u'_c (1 + f'_k - \delta) \right] + \beta \mathcal{H}_g \left(-\frac{\mathcal{H}'_k}{\mathcal{H}'_g} \right) \left[u'_g - u'_c \right] = 0. \quad (13)$$

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.

Finally, if the government can tax labor and capital income at separate rates, but faces the constraint that labor cannot be subsidized, then the optimal choice is to set the labor tax equal to zero, as we show in Appendix B. Thus, this case collapses to the case where only capital income is taxed. This case is also discussed in Martin (2005).

4 Equilibrium policy for an economy calibrated to postwar U.S. data

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.¹⁰

¹⁰Related insights are also obtained in Stockman (2001). For earlier analysis of a setup without commitment, see Klein and Ríos-Rull (2003) 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.

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

$$u(c, \ell, g) = \frac{\left[(1 - \alpha_p) (\alpha_c c^\rho + (1 - \alpha_c) \ell^\rho)^{\psi/\rho} + \alpha_p g^\psi \right]^{\frac{1-\sigma}{\psi}} - 1}{1 - \sigma}. \quad (14)$$

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

$$u(c, \ell, g) = (1 - \alpha_p) \alpha_c \ln c + (1 - \alpha_p) (1 - \alpha_c) \ln \ell + \alpha_p \ln g \quad (15)$$

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 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). The parameter values of the baseline economy are given by $\theta = 0.36$, $\alpha_c = 0.30$, $\alpha_p = 0.13$, $\beta = 0.96$, $\delta = 0.08$, $\rho = 0$, $\psi = 0$, and $\sigma = 1.0$.

As indicated, we look for the smooth Markov-perfect equilibrium, and our belief is that it is the limit of finite-horizon equilibria. This is not substantiated with proofs, but we base our belief both on the fact that our numerical algorithms always found only one solution as well as on a theoretical parallel with the cases discussed above and where analytical proofs were possible. There, a “spurious” equilibrium could exist if the limit equilibrium were one with zero savings, and zero production, which was shown to sometimes occur using backwards-induction logic. Here, such a limit equilibrium could not occur, because the taxation in the last period will not lead to zero savings in preceding periods: the stock of capital is not taxed, and even if it were, it would more

than suffice for covering the public expenditure needs, given our calibration.

4.1 The baseline economy with different tax regimes

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 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 distortive taxation, both one with access to commitment (Ramsey) and one without such access (Markov, because we look at the Markov equilibrium). Table 1 reports the steady-state allocations of these three economies under three different assumptions about the tax base.

	Labor income taxes			Capital income taxes			Total income taxes		
Steady state	Type of government			Type of government			Type of government		
statistic	Pareto	Ramsey	Markov	Pareto	Ramsey	Markov	Pareto	Ramsey	Markov
y	1.000	0.700	0.719	1.000	0.588	0.478	1.000	0.669	0.693
k/y	2.959	2.959	2.959	2.959	1.734	1.149	2.959	2.527	2.649
c/y	0.509	0.509	0.573	0.509	0.712	0.688	0.509	0.532	0.587
g/y	0.254	0.254	0.190	0.254	0.149	0.220	0.254	0.265	0.201
c/g	2.005	2.005	3.017	2.005	4.779	3.123	2.005	2.005	2.928
l	0.350	0.245	0.252	0.350	0.278	0.285	0.350	0.256	0.258
τ	–	0.397	0.297	–	0.673	0.821	–	0.334	0.255

Table 1: Baseline model economy

4.1.1 Taxes on labor income

The absence of capital income taxes 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 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.¹¹

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 (12) 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 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 (12) positive and the first one negative. The difference with the Ramsey case can perhaps be best described by the fact that the Ramsey policymaker takes into account the fact that a tax hike at t not only lowers labor supply at t but raises it at $t - 1$.¹² In contrast, a Markov policy-maker treats the latter as a bygone and hence chooses lower tax rates.

4.1.2 Taxes on capital income

The tax on capital income 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

¹¹This is a simple implication of the first-order conditions of the Ramsey problem when written in primal form.

¹² For the case of inelastic labor supply, the technical nature of this kind of consideration is discussed in Section 2.3.2.

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 labor. 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 would be chosen if all income were labor income—it equals $1 - \alpha_c$ exactly, independently of the wage (in this case 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.

4.1.3 Taxes on total income

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

First, the Ramsey government sets 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.

4.1.4 Comparisons across tax regimes

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 income 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.¹³ 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 actually vary very little across environments. Third, perhaps the most surprising feature 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 the private-to-public consumption ratio closest to the first best is that of the total income-tax economy.

¹³Note that because the tax base excludes depreciation, the tax base of a capital income tax is not a constant fraction of GDP.

4.2 Sensitivity analysis

In this Section we report results from changing the parameters of the utility function given in Equation (14). We consider four cases; in Case 1, the elasticity of substitution between consumption and leisure is 1/2 and in Case 2 it is 2. In Case 3, the elasticity of substitution between the private consumption/leisure composite good and public consumption is 1/2 and, finally, in Case 4, that elasticity is 2. As we can see in Tables 2-3, an increase in the substitutability between consumption and leisure leads to a rise in leisure in equilibrium. On the other hand, an increase in the substitutability between the private consumption/leisure composite leads to less government spending relative to private consumption. This is true for any equilibrium concept: Pareto, Ramsey or Markov. Another notable feature of these results is that the underprovision of public goods in the Markov equilibrium as compared to the Ramsey equilibrium is particularly pronounced when $\rho = 1/2$; this is when labor supply is particularly elastic and the effect discussed in Section 1 is particularly strong.

Steady state statistic	Total income taxes, $\rho = -1, \psi = 0$			Total income taxes, $\rho = 1/2, \psi = 0$		
	Type of government			Type of government		
	Pareto	Ramsey	Markov	Pareto	Ramsey	Markov
y	1.000	0.816	0.813	1.000	0.220	0.507
k/y	2.959	2.608	2.652	2.959	2.029	2.662
c/y	0.528	0.568	0.589	0.489	0.358	0.594
g/y	0.236	0.223	0.199	0.274	0.479	0.193
c/g	2.239	2.548	2.957	1.784	0.747	3.070
l	0.379	0.333	0.328	0.323	0.088	0.174
τ	–	0.282	0.253	–	0.572	0.246

Table 2: Sensitivity analysis: cases 1 and 2

5 Concluding remarks: summary and literature discussion

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-

Steady state statistic	Total income taxes, $\rho = 0, \psi = -1$			Total income taxes, $\rho = 0, \psi = 1/2$		
	Type of government			Type of government		
	Pareto	Ramsey	Markov	Pareto	Ramsey	Markov
y	1.000	0.528	0.588	1.000	0.819	0.827
k/y	2.959	2.145	2.484	2.959	2.771	2.807
c/y	0.424	0.393	0.514	0.620	0.650	0.670
g/y	0.340	0.435	0.287	0.143	0.128	0.106
c/g	1.248	0.902	1.790	4.344	5.058	6.315
l	0.393	0.249	0.255	0.307	0.261	0.261
τ	–	0.526	0.358	–	0.165	0.137

Table 3: Sensitivity analysis, Cases 3 and 4

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.

The solution concept we have used is Markov-perfect equilibrium, as defined and defended by Maskin and Tirole (2001). Specifically, we focus on a refinement of Markov-perfect equilibrium, focusing only on solutions featuring differentiable policy functions. 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 model economies parameterized to generate U.S.-like aggregate statistics, the level of (capital income) taxation without commitment is surprisingly low given that the tax base is total income. This result is due in part to the fact 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.

What is the significance of the assumption of no commitment? Our hunch is that governments are much more than machines implementing past decisions. Whether or not our hunch is correct, however, what we hope to accomplish here is simply to take another step toward deriving implications that one may eventually be able to test against alternatives, such as that based on believing that governments have full commitment. Since Kydland and Prescott first pointed to the time-inconsistency problem, most of the attempts to deal with it have been attempts to fully overcome the problem. In short, the idea has been to introduce (full or partial) commitment through other mechanisms: “rules” (e.g., Kydland and Prescott (1977)), delegation (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 use reputational equilibria to argue—see Chari and Kehoe (1990)—that good outcomes are feasible without explicit commitment, assuming that the time horizon is infinite and that agents are sufficiently patient. Here we simply wonder what might occur if reputation mechanisms fail. In addition, 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. Hopefully, the methods we use in this paper help in this respect.

There is earlier work on analyzing no-commitment outcomes in economies with long horizon. 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, and in some recent papers on taxation and political economy that take the same approach (see footnote 16 above). 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 appearing 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. Special functional-form assumptions of other sorts as well can be useful for characterizing Markov-perfect equilibria; see Bassetto and Sargent (2005) for an example involving quasilinear utility.

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 (2003)) 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.¹⁴ In contrast, the methods proposed and used here do. An alternative approach to ours can be found in Phelan and Stacchetti (2001), who have looked at environments like those studied in this paper and who have developed methods, based on those in Abreu, Pearce, and Stacchetti (1990), to find all “sustainable” equilibria. Their methods, however, do not allow Markov equilibria to be identified and explicitly interpreted, which is what we do here.

The numerical algorithm that we propose, though it works very well for the purposes of the present paper, is not intended to be the last word. Rather, we hope that our contribution will encourage more work in the development of numerical methods for solving models in political economy and optimal policy without commitment. Indeed, this should be a fertile area for future research, both because we think it is important and because so little is known about it, especially in contrast to what is known about efficient methods of solving versions of the standard growth model. The presence of derivatives of unknown functions in what we call the GEE (equation 4) means that the insights from the large (and still growing) literature on numerical solutions to the growth model are not sufficient for solving problems similar to those of this paper.

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

¹⁴ For their economies, Sarte, Azzimonti-Renzo, and Soares (2003) compare linear-quadratic approximation with the method used here. They find significant differences for some cases.

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. Moreover, in related work—Krusell, Martin, and Ríos-Rull (2004)—it is shown that Markov-perfect equilibria in models where debt is allowed have a peculiarity that may provide a partial justification for abstracting from bonds. The reason for this is that there are many—a countable number—of steady states, and these steady states can be indexed by the level of outstanding government bonds. That is, if the multiple-steady-states result in that work, which uses a deterministic version of the economy without capital studied by Lucas and Stokey (1983), can be extended to a neoclassical framework like the present one, it would state that at least for steady-state analysis, an assumption of a balanced budget is not restrictive.¹⁵ However, these are speculations the assessment of which are beyond the scope of the present paper.

Most earlier studies in the literature to which this paper belongs focus on public finance; i.e., they take public expenditures as given. Alternatively, they examine how various paths of public expenditures influence growth and business cycles but do not explicitly discuss the full optimal determination of these paths. Exceptions include some papers using endogenous-growth frameworks, such as Barro (1990), where the commitment solution is time-consistent, and a recent paper by Sarte, Azzimonti-Renzo, and Soares (2003), which studies development from the perspective of productive public capital. Another exception is Hassler, Krusell, Storesletten, and Zilibotti (2004), which studies a linear-quadratic economy where utility is linear in private and public consumption but where there are quadratic investment costs.¹⁶ There, the key finding is one of dynamics: paths under commitment (and possibly those without commitment as well) feature oscillating taxes and public-goods provision. Here, in contrast, convergence is monotone. The key reason for the difference is that the present paper uses a standard geometric depreciation structure for capital; Hassler et al. focus on human capital and the kinds of physical capital which are more of the one-hoss-shay nature. Moreover,

¹⁵For insights on the dynamics of optimal debt where commitment is assumed, see Aiyagari, Marcet, Sargent, and Seppala (2002), Abraham, Marcet, and Scott (2003), and Gorostiaga (2003).

¹⁶Related papers in the political-economy literature also discuss government expenditures, which do not have to take the form of redistribution but more generally can be public goods which are not perfect substitutes with private goods; see, e.g., Hassler, Storesletten, and Zilibotti (2003) and Hassler, Storesletten, and Zilibotti (2005).

their paper does not allow standard calibration since the *AK* production technology used there abstracts from labor and labor taxation. In sum, the Hassler et al. and Azzimonti et al. papers are the only ones we know of in the literature that compare optimal paths under commitment to those under no commitment. We also know of no other quantitative assessment of the optimal role of public expenditures in the context of a standard neoclassical growth framework, and we are not aware of any attempts to approach data on public expenditures using flexible dynamic theory based on microeconomic underpinnings—which we advocate here—as an organizing tool.

Another closely related literature upon which the present work builds quite directly is that analyzing dynamic games between successive selves, as outlined in the economics and psychology literature by Strotz (1956), Phelps and Pollak (1968), Laibson (1997), and others. This literature contains the derivation of a GEE, and Krusell, Kuruşçu, and Smith (2002) show how to solve it numerically for a smooth decision rule equilibrium. As is elaborated on in Appendix A, the smooth rule can be difficult to find with standard methods, and Krusell, Kuruşçu, and Smith (2002) resort to a perturbation method of sorts, which we also use here. This method relies on successive differentiation of the GEE. Thus, we view the approach taken here as an adaptation to optimal-policy environments of the tools suggested in these one-agent problems. Finally, there is a recent paper on monetary theory that studies Markov-perfect outcomes using numerical techniques: Díaz-Giménez, Giovannetti, Marimón, and Teles (2006). We hope that the methods we employ here and those of the mentioned papers will prove to be of general applicability; they could possibly be used 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

- Abraham, A., A. Marcet, and A. Scott (2003). Debt and deficit fluctuations and the structure of bond markets. Economics Working Papers 558, Department of Economics and Business, Universitat Pompeu Fabra.
- Abreu, D., D. Pearce, and E. Stacchetti (1990). Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica* 58, 1041–1063.

- Aiyagari, S. R., A. Marcet, T. Sargent, and J. Seppala (2002). Optimal taxation without state-contingent debt. *Journal of Political Economy* 110, 1220–1254.
- Barro, R. J. (1990). Government spending in a simple model of endogenous growth. *Journal of Political Economy* 98, S103–S125.
- Bassetto, M. and T. Sargent (2005). Politics and efficiency of separating capital and ordinary government budgets. NBER Working Paper No. 11030.
- 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.
- Díaz-Giménez, J., G. Giovannetti, R. Marimón, and P. Teles (2006). Nominal debt as a burden to monetary policy. CREA Barcelona Economics WP #8.
- Gorostiaga, A. (2003). Should fiscal policy be different in a non-competitive framework? *Journal of Monetary Economics* 50(6), 1311–31.
- Hassler, J., P. Krusell, K. Storesletten, and F. Zilibotti (2004). On the optimal timing of capital taxes. Mimeo.
- Hassler, J., K. Storesletten, and F. Zilibotti (2003). The survival of the welfare state. *American Economic Review* 93(1), 87–112.
- Hassler, J., K. Storesletten, and F. Zilibotti (2005). Democratic public good provision. Forthcoming in the *Journal of Economic Theory*.
- Klein, P. and J.-V. Ríos-Rull (2003). Time-consistent optimal fiscal policy. *International Economic Review* 44(4), 1217–1246.
- Krusell, P., B. Kuruşçu, and A. Smith (2002). Equilibrium welfare and government policy with quasi-geometric discounting. *Journal of Economic Theory* 105, 42–72.
- Krusell, P., F. Martin, and J.-V. Ríos-Rull (2004). Time-consistent debt. Mimeo.
- 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.
- 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.
- 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.
- Martin, F. (2005). Optimal taxation without commitment. Mimeo, Simon Fraser University.
- Maskin, E. and J. Tirole (2001, October). Markov perfect equilibrium. *Journal of Economic Theory* 100(2), 191–219.
- 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).
- Rogoff, K. (1985). The optimal degree of commitment to an intermediate monetary target. *Quarterly Journal of Economics* 100(4), 1169–1190.
- Sarte, P.-D. G., M. Azzimonti-Renzo, and J. Soares (2003). Optimal public investment with and without government commitment. FRB Richmond Working Paper 03-10.
- Stockman, D. R. (2001). Balanced-budget rules: Welfare loss and optimal policies vol. 4, no. 2, april, 2001. *Review of Economic Dynamics* 4(2), 438–459.
- Strotz, R. H. (1956). Myopia and inconsistency in dynamic utility maximization. *res* 23, 165–80.

Appendix A : Numerical algorithm

A.1 The functional equations

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

$$0 = u_c - \beta u'_c [1 + (1 - \mathcal{T}') (f'_k - \delta)], \quad (16)$$

and the one for the government,

$$0 = (-u_c + u_g) \eta_{k'} - \eta_g \left[-u_c + \beta u'_c (f'_k + 1 - \delta) - \beta (-u'_c + u'_g) \frac{\eta'_k}{\eta_g} \right]. \quad (17)$$

These are functional equations: they hold for all k . The derivatives of η are derived from the definition of η in Section 2.3; they are

$$\eta_k = u_{cc} \mathcal{C}_k \quad (18)$$

$$\eta_g = u_{cc} \mathcal{C}_g + u_{cg} \quad (19)$$

$$\begin{aligned} \eta_{k'} = & u_{cc} \mathcal{C}_{k'} + u'_c [1 - \mathcal{T}'] f'_{kk} - u'_c [f'_k - \delta] [\mathcal{T}'_k + \mathcal{T}'_g \Psi'_k] + \\ & \{1 + [1 - \mathcal{T}'] [f'_k - \delta]\} \{u_{cc} [\mathcal{C}'_k + \mathcal{C}'_{k'} h'_k + \mathcal{C}'_g \Psi'_k] + u_{cg} \Psi'_k\} \end{aligned} \quad (20)$$

If we substitute equations (18-20) into equations (16) and (17) we obtain a system of two equations that we can write compactly as

$$0 = \xi^p \{k, h(k), \Psi(k), h[h(k)], \Psi[h(k)]\} \quad (21)$$

$$0 = \xi^g \{k, h(k), \Psi(k), h[h(k)], \Psi[h(k)], h'_k(k), \Psi'_k(k)\}. \quad (22)$$

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

$$0 = \xi^p(k^*, k^*, g^*, k^*, g^*) \equiv \xi^{p*}(k^*, g^*) \quad (23)$$

$$0 = \xi^g(k^*, k^*, g^*, k^*, g^*, h_k^*, \Psi_k^*) \equiv \xi^{g*}(k^*, g^*, h_k^*, \Psi_k^*). \quad (24)$$

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 (2002). In short, it relies on a successive set of approximations to the decision rules that are polynomial functions and that only use steady-state information.

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 (23) and (24)—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

$$0 = \xi^p \tag{25}$$

$$0 = \xi^g \tag{26}$$

$$0 = \xi_k^p + \varphi_1^{k,1} \xi_{k'}^p + \varphi_1^{g,1} \xi_p^p + \left(\varphi_1^{k,1}\right)^2 \xi_{k''}^p + \varphi_1^{g,1} \varphi_1^{k,1} \xi_{g'}^p \tag{27}$$

$$0 = \xi_k^g + \varphi_1^{k,1} \xi_{k'}^g + \varphi_1^{g,1} \xi_g^g + \left(\varphi_1^{k,1}\right)^2 \xi_{k''}^g + \varphi_1^{g,1} \varphi_1^{k,1} \xi_{g'}^g, \tag{28}$$

where equations (27) and (28) 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 simultaneously solve more than two equations in two unknowns with this recursive method.

The solutions reported in our quantitative section have been compared to solutions obtained with global methods for solving for the fixed-point decision rules. In particular, when Chebyshev polynomials were used, we obtained steady states and derivatives of decision rules at steady state that are very close (up to the third decimal point) to those obtained using our steady-state-based method. Moreover, global plots show that these rules are also close on a much larger domain for the state variable; thus, the steady-state-based methods are not just efficient but, at least for this environment, appear to deliver reliable global decision rule features as well. This is not a surprise, perhaps, since if decision rules are functions which are analytic (as are those in our closed-form examples), information at one point can, with standard polynomial expansions, be used to provide accurate approximations of the functions far from this point. Details of these comparisons are available upon request from the authors.

Appendix B: Allowing taxation of both labor and capital

If the government can tax both labor and capital income, it will want to subsidize labor income in order to encourage saving. If we impose the constraint that the labor income tax τ_t is not negative, then the Kuhn-Tucker multiplier associated with that constraint

is strictly positive. To see this, denote the capital income tax by θ_t and associate μ_t with the constraint $k_{t+1} = \mathcal{H}(k_t, \theta_t, \tau_t)$ and λ_t with the constraint $-\tau_t \leq 0$.

The FOCs are, using the usual shorthand notation,

$$u_c \mathcal{C}_\theta + u_\ell \mathcal{L}_\theta + u_g \mathcal{G}_\theta + \mu \mathcal{H}_\theta = 0,$$

$$u_c \mathcal{C}_\tau + u_\ell \mathcal{L}_\tau + u_g \mathcal{G}_\tau + \mu \mathcal{H}_\tau + \lambda = 0$$

and the optimal investment condition which does not concern us right now.

Solving for λ , we have

$$\lambda = u_c \left[\frac{\mathcal{H}_\tau}{\mathcal{H}_\theta} \mathcal{C}_\theta - \mathcal{C}_\tau \right] + u_\ell \left[\frac{\mathcal{H}_\tau}{\mathcal{H}_\theta} \mathcal{L}_\theta - \mathcal{L}_\tau \right] + u_g \left[\frac{\mathcal{H}_\tau}{\mathcal{H}_\theta} \mathcal{G}_\theta - \mathcal{G}_\tau \right].$$

But

$$\mathcal{C}_\theta = f_\ell \mathcal{L}_\theta - \mathcal{H}_\theta - \mathcal{G}_\theta$$

and similarly with \mathcal{C}_τ , so we obtain

$$\lambda = (u_c f_\ell + u_\ell) \cdot \left[\frac{\mathcal{H}_\tau}{\mathcal{H}_\theta} \mathcal{L}_\theta - \mathcal{L}_\tau \right] + (u_g - u_c) \cdot \left[\frac{\mathcal{H}_\tau}{\mathcal{H}_\theta} \mathcal{G}_\theta - \mathcal{G}_\tau \right] - u_c \left[\frac{\mathcal{H}_\tau}{\mathcal{H}_\theta} \mathcal{H}_\theta - \mathcal{H}_\tau \right].$$

The first term vanishes at $\tau = 0$ and the third term vanishes everywhere. So we are left with

$$\lambda = (u_g - u_c) \cdot \left[\frac{\mathcal{H}_\tau}{\mathcal{H}_\theta} \mathcal{G}_\theta - \mathcal{G}_\tau \right].$$

Although we have not been able to establish assumptions on primitives that guarantee this expression to be strictly positive, we have verified numerically that it is in all the applications we have considered.