ELSEVIER

European Economic Review 46 (2002) 755–769

EUROPEAN ECONOMIC REVIEW

www.elsevier.com/locate/econbase

Macroeconomic Theory

Time-consistent redistribution

Per Krusell[a,b,c,*,1]

[a] Department of Economics, University of Rochester, 109 Harkness Hall, Rochester, NY 14627-0156, USA

[b] Institute for International Economic Studies, Stockholm University, S-106 91 Stockholm, Sweden

[c] Center for Economic Policy and Research, 90–98 Goswell Road, London EC1V 7RR, UK

Abstract

If the government cares more about workers than about capitalists and taxes capital income to finance redistribution to workers, how are inequality and capital accumulation affected in the long run? Assuming that the government cannot commit to future taxes, a *time-consistent* equilibrium – a differentiable subgame-perfect Markov equilibrium – is characterized. In this equilibrium, the current government in part uses the tax, via capital accumulation, to manipulate future governments into setting lower taxes. The equilibrium has substantially lower taxes on capital income than 100%, even though workers do not save and even though the weight on capitalists in government utility is negligible. © 2002 Published by Elsevier Science B.V.

JEL classification: C6; D6; E6; E62; H2; H21

Keywords: Time consistency; Dynamic optimal taxation; Redistribution; Markov-perfect equilibrium; Commitment

1. Introduction

This paper lays out a simple economic model with two goals in mind: (i) to illustrate how heterogeneity can influence the macroeconomy via endogenous policy making; and (ii) to illustrate how recursive, functional-equation methods can be used to characterize time-consistent equilibria. The setup is a neoclassical growth model in which there are two classes – workers and capitalists. There is a proportional tax on production, and the proceeds from this tax are used to fund transfers that go lump-sum to workers; there is no government borrowing or lending. Several other simplifying assumptions

* Corresponding author. Tel.: +1-716-2734903; fax: +1-716-2713900.

E-mail address: pekr@troi.cc.rochester.edu (P. Krusell).

[1] Very helpful comments were received from Marina Azzimonti Renzo, Fernando Leiva Bertran, Torsten Persson, and Harald Uhlig.

0014-2921/02/$ - see front matter
PII: S0014-2921(01)00211-2

are also made: workers (who are all alike) work but do not save, whereas capitalists (who are also all alike) save but do not work. Furthermore, the tax rate in every period is decided upon by a government whose utility is a weighted sum of that of workers and capitalists. That is, the political struggle is summarized by an objective function that gives each group some weight.

An important part of the analysis is the insistence on sequential decision making: the current government only sets current tax rate, and it has no direct influence on how future governments will tax. This lack of commitment is a binding restriction in general. The current government cares about the future, and disagrees with future governments about how to tax. In particular, each future government will view its initial capital stock as inelastically supplied, whereas the current government sees it as elastically supplied. The fact that the lack of commitment is binding significantly complicates the analysis, because it requires any rational policy maker's choices to incorporate views about the decision *rules* of the future policy makers. This paper focuses on solving for differentiable Markov equilibria: equilibria where the decision rules followed by any agent, government or otherwise, is only a function of directly payoff-relevant variables, such as the stock of capital. That is, no reputation mechanisms are considered. Differentiability, moreover, is a key requirement. As in Klein et al. (2001), the approach is to characterize equilibria in terms of first-order conditions. These conditions are derived both for the private sector – these are standard and we focus on their functional-equation versions here – and for the policy maker. The first-order condition for the government, which is also expressed in its functional-equation form, is not standard: its derivation requires the use of recursive methods and it has the unusual property of including derivatives of decision rules. It is generally referred to as a "generalized Euler equation", and it makes clear the inter- and intra-temporal tradeoffs facing a current policy maker. For this equation, differentiability is first of all critical as a selection device: there are potentially a large number of discontinuous Markov equilibria (see Krusell and Smith, 2001), and these equilibria have reputation features, even if they do not use history-dependence explicitly.

The second reason to exploit differentiability is numerical. It namely turns out that a number of standard methods allowing controlled accuracy do not work here, but a perturbation method, which critically relies on taking repeated derivatives, works very well and is easy to implement. Krusell et al. (2001) and Klein et al. (2001) develop and use this method.

The literature on redistribution in a similar context includes Judd (1985), who studies the commitment version of the problem herein for the case where workers and capitalists have the same time preference rates. He finds that taxes have to be zero in the long run.[2] Kemp et al. (1993), in contrast, model the government as choosing taxes without commitment ("closed loop") for more general preferences. Their setup is one of continuous time, and in continuous time the equilibrium definition simplifies in this particular model, because the marginal propensity of saving out of a current tax change has to be one: any changes in current income caused by a tax change will not influence permanent income, and therefore not consumption, unless you are

[2] This result was also derived in Chamley (1986).

constrained in savings/borrowing (in which case the marginal propensity to consume would be zero). A discrete-time model, or a continuous-time model where taxes can be committed to for a significant period of time, seems much more realistic, however. Such a model necessitates dealing with marginal propensities which are not degenerate: it needs to incorporate a description of how the private sector nontrivially responds to any current tax change. Lansing (1999), finally, studies Judd's problem again, but under logarithmic utility. In that case, the commitment solution is time-consistent, and the tax rate is positive in the long run; for any other curvature, long-run taxes are zero. His findings are corroborated here. Moreover, we show that the long-run outcomes for taxes in time-consistent equilibria without commitment vary continuously with the curvature of the utility function; thus, zero taxes do not result. Finally, Bertola (1993) also studies a case without commitment under preferences with constant elasticity of intertemporal substitution.

2. The baseline model

This section contains the description of the economy as well as the analysis of it. Time is discrete and there are two classes of consumers: workers and capitalists. For simplicity, we assume that they are of equal numbers. Agents in the same class are identical: they have identical preferences and face identical economic problems as the economy begins. Because their maximization problems have unique solutions, they also remain identical over time, since they start out with the same initial conditions. There are firms, renting labor and capital inputs from consumers and selling their output to the consumers; these markets are all competitive. Production is taxed at a proportional rate τ_t in period t. The proceeds from the tax in period t are directly given to workers in period t in a lump-sum manner. The government, thus, is restricted to a balanced budget at all points in time; its only choice variable in period t is τ_t.

Worker and capitalist decisions are described next. Thereafter, the behavior of the government will be discussed, leading to an equilibrium definition. The equilibrium definition, in turn, leads onto a path of analysis consisting both of analytical and numerical work.

2.1. The decision problems of consumers and firms

The worker chooses $\{c_{wt}\}_{t=0}^{\infty}$, given an endowment of 1 unit of time, to maximize

$$\sum_{t=0}^{\infty} \beta^t u(c_{wt}) \quad \text{subject to } c_{wt} = w_t \cdot 1 + T_t,$$

where T_t is the transfer from the government. That is, workers by assumption do not save.

The capitalist chooses $\{c_{ct}, k_{t+1}\}_{t=0}^{\infty}$ to maximize

$$\sum_{t=0}^{\infty} \beta^t v(c_{ct}) \quad \text{subject to } c_{ct} + k_{t+1} = (1 - \delta + r_t)k_t$$

given a k_0. That is, the capitalist saves, taking interest rates as given, but does not work.

The functions u and v are assumed to be smooth and have standard properties: they are strictly increasing and strictly concave. The parameter β lies strictly between zero and one. The assumption that workers and capitalists have the same discount rate is used to obtain an exact steady state. The setup with capitalist savers and worker non-savers can be viewed as a simple version of Krusell and Smith (1998). There, a more complicated model with small differences in discount rates between different consumers leads to a "reduced form" where a very small subset of consumers – capitalists – do almost all the saving, leading to a very skewed, and realistic, wealth distribution.

Firms are price takers and choose (k_t, l_t) to maximize

$$(1-\tau_t)F(k_t, l_t) - w_t l_t - k_t r_t.$$

This is a static problem; equilibrium prices for labor and capital services, w_t and r_t, respectively, will be such that firms make zero profits, as F is assumed to have constant returns to scale.

2.2. Aggregate constraints and price determination

The economy is closed:

$$C_{wt} + C_{ct} + K_{t+1} = F(K_t, 1) + (1-\delta)K_t,$$

where capital letters refer to economy-wide averages. In this equation, we have used the equal population shares to obtain unitary weights on the consumption levels of the two classes of consumers.

The government's balanced budget constraint reads

$$T_t = \tau_t F(K_t, 1)$$

and prices are net-of-tax marginal products:

$$w_t = (1-\tau_t)F_l(K_t, 1)$$

and

$$r_t = (1-\tau_t)F_k(K_t, 1).$$

The pricing equations follow as first-order conditions to the firm's problem.

For ease of notation, we will summarize the equilibrium consumption behavior with

$$C_{ct} = C_c(K_t, \tau_t, K_{t+1}) \equiv (1-\tau_t)F_k(K_t, 1)K_t - K_{t+1}, \tag{1}$$

$$C_{wt} = C_w(K_t, \tau_t) \equiv F_l(K_t, 1) + \tau_t F_k(K_t, 1)K_t, \tag{2}$$

where we have substituted in equilibrium expressions for prices and transfers.

2.3. Behavior under commitment: The Ramsey problem

The objective of the government is

$$\sum_{t=0}^{\infty} \beta^t(\lambda v(C_{ct}) + (1-\lambda)u(C_{wt})). \tag{3}$$

When the government can commit to future policy, it chooses $\{\tau_t\}_{t=0}^{\infty}$ to maximize (3) subject to $\{C_{ct}\}_{t=0}^{\infty}$ and $\{C_{wt}\}_{t=0}^{\infty}$ being competitive equilibrium consumption allocations. This is referred to as the "Ramsey problem". In terms of choosing capital and tax sequences, it thus reads

$$\max_{\{K_{t+1},\tau_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t(\lambda v(C_c(K_t,\tau_t,K_{t+1}))) + (1-\lambda)u(C_w(K_t,\tau_t)) \tag{4}$$

subject to

$$v'(C_c(K_t,\tau_t,K_{t+1})) = \beta v'(C_c(K_{t+1},\tau_{t+1},K_{t+2}))[1-\delta+(1-\tau_{t+1})F_k(K_{t+1},1)]. \tag{5}$$

We will characterize steady-state Ramsey allocations in Section 2.5 below. At this point, simply note that this problem leads to a "time-inconsistent" allocation. Why? The variable K_{t+1} appears in one more constraint for every $t>0$ than for $t=0$, and similarly for τ_t. For example, the tax rate at time 1 influences savings in period 0, and this effect is absent in the choice of the tax rate at 0, since savings are predetermined in that period. This means that the first-order conditions will have one more term for all $t>0$, and reoptimization in any later period s would yield a first-order condition for that period violating the one from optimization at 0.

2.4. Behavior when there is no commitment

Without access to a commitment mechanism, the government cannot choose future taxes directly, but it still wants to maximize (3), and it still needs to select an allocation among decentralized equilibria. We are looking for a time-consistent policy equilibrium: the current government sets the current tax correctly foreseeing how the future governments will set taxes. The key here is that the taxes set in any period will, in general, depend on the capital the economy is endowed with as of the beginning of that period. We will look for a stationary subgame-perfect Markov equilibrium, where the tax rate at t does not depend on anything but the capital stock in period t, thus ruling out explicit dependence on any history beyond that summarized in the current capital stock.

By definition, then, we are looking for a government policy that obeys a recursive rule given by the function Ψ:

$$\tau_t = \Psi(K_t).$$

Thus, Ψ is the key endogenous variable: a function. We will look for a Ψ that is differentiable.

2.4.1. Recursive definition of equilibrium

We now apply recursive competitive equilibrium methods to characterize Ψ: we search for a differentiable subgame-perfect Markov equilibrium in government policy. Formally, the game in a given period can be viewed as Stackelberg, with the government leading off with a tax choice and the private sector, in particular the capitalist, making a subsequent consumption/savings choice. Thus, within the period the government has commitment.

In order to specify the government's problem, we need its key inputs: a view of how the private sector responds to its current tax choice. The specification of this response must include what will happen in the future in response to the current tax choice. Given the Markov construction, the idea is to consider a one-period deviation from the rule $\Psi(K)$: the current τ is free of choice, but future taxes are set according to Ψ. For this construction, we will need to find an equilibrium function giving savings as a function of the current tax rate, τ (and of the current state, K):

$$K' = H(K, \tau);$$

we use primes to denote next-period values. The function H thus describes what the private sector will do, for any given capital stock and for any current policy, given that it expects future taxes to be set by future governments according to Ψ. Future savings, moreover, are expected to be given by the function H as well, but evaluated on the equilibrium path, i.e., using $\tau = \Psi(K)$ (where K is any future capital stock).

To see in concrete terms how a given K_t and an arbitrary choice of τ_t affects perceptions about future capital stocks and taxes, we apply the functions H and Ψ repeatedly:

$$K_{t+1} = H(K_t, \tau_t), \quad \tau_{t+1} = \Psi(K_{t+1}), \quad K_{t+2} = H(K_{t+1}, \tau_{t+1}),$$
$$\tau_{t+2} = \Psi(K_{t+2}), \quad \dots .$$

To find the function H, we need a functional-equation version of the standard competitive first-order condition (5) for the capitalist. To derive it, substitute the functions H and Ψ into the appropriate places in this equation to read

$$v'(C_c(K, \tau, H(K, \tau))) = \beta v'(C_c(H(K, \tau), \Psi(H(K, \tau)), H(H(K, \tau), \Psi(H(K, \tau))))) \cdot \\ [1 - \delta + (1 - \Psi(H(K, \tau)))F_k(H(K, \tau), 1)]. \tag{6}$$

Eq. (6) has to hold for all (K, τ). That is, it defines H given Ψ. It is important to recognize that H is a determinant of Ψ. A possibility would be to use the notation $H(K, \tau; \Psi)$ to make this clear; if Ψ were arbitrary, H would depend on it. Similarly, as we shall see, Ψ naturally depends on how the private sector responds, so H is a determinant of Ψ. Rather than making this interdependence between H and Ψ explicit in the notation, however, we will simply require the two functions to simultaneously satisfy two functional equations: thus, we only define these functions "in equilibrium".

Equipped with the knowledge of how the private sector responds to tax rates – H – and how future government sets taxes – Ψ – the government now simply solves

$$\max_{\tau,K'} \lambda v(C_c(K,\tau,K')) + (1-\lambda)u(C_w(K,\tau)) + \beta V(K')$$

subject to

$$K' = H(k,\tau),$$

where V is given by the recursion

$$V(K) = \lambda v(C_c(K,\Psi(K),H(K,\Psi(K)))) + (1-\lambda)u(C_w(K,\Psi(K))) + \beta V(H(K,\Psi(K))).$$

The role of the recursion here is simply to effectively produce the infinite discounted sum of utilities in the objective (3).

It is now apparent what the fixed-point problem is: if, for every K, the solution to the above problem is given by $\Psi(K)$, then the government is rational and we have a subgame-perfect equilibrium.

We can thus describe the government's equilibrium behavior as solving a dynamic-programming equation:

$$V(K) = \max_{\tau,K'} \{\lambda v(C_c(K,\tau,K')) + (1-\lambda)u(C_w(K,\tau)) + \beta V(K')\} \tag{7}$$

subject to

$$K' = H(k,\tau).$$

Recall that whereas H, and the Ψ underlying it, are endogenous in this problem, C_c and C_w are primitive functions.

A *subgame-perfect Markov equilibrium* is now a set of functions Ψ, H, and V such that H solves (6), V solves (7), and Ψ attains the maximum in (7).

2.4.2. Time-consistent equilibrium

There may be many Markov equilibria. Krusell and Smith (2001), who analyze a similar structure, show existence of a continuum of solutions for their policy function – the equivalent of Ψ here. The indeterminacy, however, can only arise with infinite time horizon and the associated policy functions are all discontinuous. The discontinuities play a role similar to that played by histories in trigger-strategy equilibria. Since our focus here is on a "fundamental" equilibrium that is a limit of finite-horizon equilibria, we thus want a time-consistent equilibrium to be one where Ψ is not discontinuous. In particular, we will require it to be differentiable (and for our computations, we will assume that it is smooth). We now proceed toward such a definition. Existence will not be discussed here; below we display a closed-form solution for a specific parametric case.

The sequential version of (7) reads

$$\max_{\{\tau_t,K_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t(\lambda v(C_c(K_t,\tau_t,K_{t+1})) + (1-\lambda)u(C_w(K_t,\tau_t)))$$

subject to

$$K_{t+1} = H(K_t, \tau_t).$$

Whereas the maximization in the dynamic-programming problem in the previous section represents the government's actual choice problem – the choice of current taxes and savings, leaving future taxes and savings to be set by other governments – the sequential formulation here does not. In it, the government can choose any values of future tax rates and capital stocks (subject to the constraints). The point, however, is that even though the sequential problem allows choices that are not part of any government's choice set, its solution will coincide with that of the dynamic-programming problem! This result is an application of Bellman's principle, and it is useful mainly because it allows a simple way of deriving the first-order conditions. Moreover, the sequential problem is still part of a fixed-point problem, since its solution will depend on H and therefore on Ψ.

The first-order conditions are readily derived; τ_t is the control variable and plays the role that consumption plays in a standard optimal growth problem. We will assume that the first-order condition that results from a standard variational argument in this sequential problem is unique, and thus sufficient for a maximum.[3]

Differentiating with respect to τ_t, τ_{t+1}, and K_{t+1} and simplifying yields

$$R_\tau + R_{k'} H_\tau + \beta H_\tau \left\{ R'_k - \frac{H'_k}{H'_\tau} R'_\tau \right\} = 0, \tag{8}$$

where $'$s denote next-period values and we have used the definition

$$R(K, \tau, K') \equiv \lambda v(C_c(K, \tau, K')) + (1 - \lambda) u(C_w(K, \tau))$$

in order to make the condition more compact. Moreover, in (8), all current functions are evaluated at $(K_t, \tau_t) = (K, \Psi(K))$ and all future functions at $(K_{t+1}, \tau_{t+1}) = (H(K, \Psi(K)),\ \Psi(H(K, \Psi(K))))$ – their equilibrium values. This is a functional equation: it has to hold for all K. It is our government's first-order condition: the "generalized Euler equation", or GEE.

We can now define our equilibrium. A *time-consistent equilibrium* is a set of differentiable functions H and Ψ satisfying the functional equations (6) and (8).

2.4.3. Interpretation

We can think of the GEE as a "variation": given values for K and K'', τ and τ' are varied in the best possible way. The effect of an increase in τ can be interpreted as follows. First, there will be an increase in C_w, which provides a direct boost to utility for the worker: $C_{w\tau} = F_k(k, 1)K > 0$ – his marginal propensity to consume out of the new resources is one. As for the capitalist, his behavior is less trivially affected. Under the normal goods assumption, when the capitalist is left with less resources, he chooses to consume less of all goods, current and future: his savings fall, $H_\tau < 0$, as

[3] This assumption, like in most Ramsey problems, is hard to verify in general but can be defended numerically. It is implicit in what follows.

does his consumption, $C_{c\tau} = -C_{w\tau} - H_\tau < 0$. With a λ close to zero, the net current benefit of the tax increase is positive for the government.

The fall in K' has two effects. An immediate effect is that on future incomes; if the production function is such that both capital and labor income are increasing in capital, then the fall in savings will decrease both C_w' and C_c'.[4] Thus, this reduces next-period utility.

The second effect of a fall in K' is to decrease τ': it has to be reduced by $-H_k'/H_\tau'$ times the change in capital next period in order to keep K'' from falling; we obtain this from differentiation of $K'' = H(K', \tau')$ with $\mathrm{d}K'' = 0$. That is, the increase in the current tax rate must lead to a decrease in next period's tax rate, and thus to lower (higher) consumption for the workers (capitalists) that period.

In summary, with a focus on the workers, a tax hike today has an immediate positive benefit in the form of increased consumption, and two negative effects on consumption next period.

2.4.4. A more compact formulation: A comparison with the Ramsey problem

The definition of a time-consistent equilibrium makes clear the distinction between the different players and how they move in sequence: $\Psi(K)$ describes optimal behavior of the government when they freely choose a current τ, and $H(K, \tau)$ describes optimal behavioral response of the private sector to any current τ choice of the government. An alternative, more compact definition is possible as well, and it yields identical outcomes. This alternative does not define the behavioral response of the private sector to any current τ; instead, it uses a function $H(K)$ to describe on-the-equilibrium-path behavior. Now, we would think of the government as simultaneously choosing both τ and K'. The restrictions on these variables, then, would still have to include the first-order condition of the capitalist, in which the private sector's "response" to any current tax rate is incorporated. This restriction on (K', τ) would, thus, read

$$v'(C_c(K, \tau, K')) = \beta v'(C_c(K', \Psi(K'), H(K')))[1 - \delta + (1 - \Psi(K'))F_k(K', 1)]. \tag{9}$$

The alternatively defined *subgame-perfect Markov equilibrium* would then be the set of functions Ψ, H, and V such that V solves

$$V(K) = \max_{K', \tau} \{\lambda v(C_c(K, \tau, K')) + (1 - \lambda)u(C_w(K, \tau)) + \beta V(K')\} \tag{10}$$

subject to (9) and H and Ψ attain the maximum in (10).

It is straightforward to verify that this formulation generates the GEE, either by use of the envelope theorem or by standard differentiation of the corresponding sequential problem

$$\max_{\{\tau_t, K_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t(\lambda v(C_c(K_t, \tau_t, K_{t+1})) + (1 - \lambda)u(C_w(K_t, \tau_t)))$$

[4] Labor income goes up provided that labor and capital are complements in the sense of $F_{kl} > 0$, and $C_{wk} = F_{kl} + \tau(F_{kk}K + F_k)$, which by F being constant returns to scale equals $(1 - \tau)F_{kl} + \tau F_k > 0$. Capital income goes up provided $F_{kl} < F_k$, which is satisfied for a Cobb–Douglas function.

subject to

$$v'(C_c(K_t, \tau_t, K_{t+1})) = \beta v'(C_c(K_{t+1}, \Psi(K_{t+1}), H(K_{t+1})))[1 - \delta + (1 - \Psi(K_{t+1}))$$
$$F_k(K_{t+1}, 1)].$$

We now see that this problem is very close to the Ramsey formulation. The only difference is that the latter allows the savings and tax choices in the future to be subject to choice also in the constraint; here, in contrast, these variables are given by the endogenous functions H and Ψ. Thus, there is still a fixed-point problem here, and the equilibrium is conceptually a very different object than under commitment.

2.5. Steady states

We will now analyze both the time-consistent equilibrium and the Ramsey solution. We will make use of functional-form restrictions for convenience, and we will mainly look at steady states. Throughout, we will assume that F is Cobb–Douglas with capital share α and a TFP normalized to 1, and that $\delta = 1$.

2.5.1. Logarithmic utility

Consider the Ramsey problem first, and suppose that u and v are both logarithmic. Then it is straightforward to verify that the capitalist saves in a "myopic" manner: his current savings decision does not depend on future interest rates or taxes. The reason for this is that income and substitution effects of these variables cancel in the logarithmic case. More specifically, the capitalist's first-order condition can be replaced by the equation

$$K_{t+1} = \beta\alpha(1 - \tau_t)K_t^{\alpha},$$

since it solves that first-order condition for any contemplated sequence of tax rates. Furthermore, this means, since the first-order condition is recursive, delivering next period's capital as a function of current capital and the current control, that the Ramsey problem is a standard recursive problem and therefore time-consistent! This property was also noted in Lansing (1999). It is straightforward to solve for the tax rate; it will be constant over time and equal $1 - \beta - \lambda(1 - \alpha\beta)/\alpha$. Thus, taxes will not be zero in the long run. In order to make taxes go to zero so as to induce first-best accumulation, future taxes would have to be used in order to induce savings to increase, but logarithmic utility does not allow the channel from the future to the present to operate.

Since the Ramsey problem is time-consistent, we have automatically also found the time-consistent equilibrium without commitment. For

$$H(K, \tau) = \beta(1 - \tau)\alpha K^{\alpha} \tag{11}$$

and

$$\Psi(K) = 1 - \beta - \lambda\,\frac{1 - \alpha\beta}{\alpha} \tag{12}$$

can be shown to solve the functional equations in this case by simply substituting into Eqs. (6) and (8).

Since we wish to view the time-consistent equilibrium as the limit of finite-horizon equilibria, consider for a moment a finite-horizon version of this economy. Consider first the special case where λ is zero. In that case, all finite-horizon economies, and thus their limit too, would lead to 100% tax rates and no savings – the economy "dies" in one period. [5] Nevertheless, if λ is arbitrarily close, but not equal, to zero, the situation is actually quite different. Because of continuity, this equilibrium will almost look like the one for $\lambda = 0$ for any finite-horizon economy; in the very last period, tax rates will be (very close to) 100%, and under rational expectations on the part of capitalists, there will therefore be (almost) no saving in the preceding period, and the tax will still be very high. However, working backwards, as the remaining time horizon becomes large, savings increase and taxes fall, and in the limit, we obtain a discretely different equilibrium: we obtain, as the limit of these finite-horizon economies, (11) and (12), which are far from 0 and 1, respectively! That is, we learn that so long as the final period does not have a literally degenerate outcome ("almost" 100% taxes works), the Markov equilibrium can propagate outcomes that are substantially different from the degenerate outcome. [6]

2.5.2. Constant elasticity of intertemporal substitution

Now assume that $u(v)=v(c)=(c^{1-\sigma}-1)/(1-\sigma)$, with $\sigma > 0$. To analyze the Ramsey problem as set up in Section 2.3, derive its first-order conditions and set $(K_t, \tau_t)=(K, \tau)$ for all t. The resulting equations can easily be shown to only be consistent with $\tau = 0$, unless $\sigma = 1$. That is, long-run taxes are discontinuous in the elasticity of intertemporal substitution parameter around $\sigma = 1$: whenever utility is not logarithmic, it is optimal to set the tax to zero in the long run.

The intuition for these results has been discussed elsewhere (see, e.g., Judd, 1985; Lansing, 1999). Most importantly, savings have to be Pareto optimal in the long run, since it is possible to line up the marginal rates of substitution for workers and capitalists. It is then as if workers save directly using the technology.

In contrast, as we shall now see, the long-run properties of the time-consistent equilibrium under lack of commitment do not change discontinuously at $\sigma = 1$. It is no longer possible to solve for the equilibrium in closed form. We therefore use numerical techniques, as discussed in detail in the next section. However, it is possible to obtain intuition from the first-order condition for taxes: the generalized Euler equation, (8). In order to simplify the discussion, we will look at the case where λ is arbitrarily close to zero: the welfare of the capitalists is (almost) irrelevant for the government. [7] In

[5] The equilibrium given by the above equations still exists in the infinite-horizon economy in this case, but as an additional, "sunspot" equilibrium that relies on optimistic expectations.

[6] There is, of course, another smooth equilibrium in the special case of $\lambda=0$: $\tau=\Psi(K)=1$ and $H(K,\tau)=0$ for all K, i.e., complete expropriation of capital (recall that depreciation is 100%) and no saving – the economy "dies" in one period. This equilibrium, in fact, is the limit of the finite-horizon equilibria, and the solution given in the text is just an additional "optimistic" equilibrium which relies on the infinite time horizon. Whenever $\lambda > 0$, the equilibrium in the text is the limit of the finite-horizon equilibria.

[7] Recall footnote 6: the relevant Markov equilibrium for $\lambda = 0$ is 100% taxation.

steady state, the solution for the time-consistent equilibrium satisfies

$$H_\tau \beta \left(\frac{H_k}{H_\tau} - \frac{C_{wk}}{C_{w\tau}} \right) = 1.$$

Since $C_{wk} = F_{kl} + \tau(F_k + F_{kk}K)$ and $C_{w\tau} = F_k K$, with a Cobb–Douglas production function the value of $C_{wk}/C_{w\tau}$ equals $(1 - \alpha + \alpha\tau)/K$. Moreover, the effects on H_k/H_τ can also be calculated: from the first-order condition for savings, Eq. (6), differentiation with respect to K and τ yields $H_k/H_\tau = -(1-\tau)(F_k + F_{kk}K)/(F_k K) = -\alpha(1-\tau)/K$; in this computation, of course, we are using differentiability of Ψ. It is possible to compute this ratio closed form because the effects of K and τ on savings are very similar: they both originate purely through initial income, which is predetermined in the present period. That is, we arrive at a steady-state condition

$$-H_\tau \beta = K$$

in the case of Cobb–Douglas production. To find the effect of changes in the curvature parameter σ on the steady state, we thus need to know how it makes H_τ change. An increased tax rate, which lowers capitalist savings, raises future interest rates. If the curvature parameter is above 1, such interest rate changes will lead to decreased savings, since the income effect dominates in this case; if $\sigma < 1$, the reverse occurs. Thus, if $\sigma > 1$, the income effect reinforces the initial drop in savings: H_τ is higher in absolute value, the higher is σ. This means that capital has to increase. Comparing to (12), thus, taxes would be lower than $1-\beta$ for $\sigma > 1$ and higher for $\sigma < 1$ (recall that λ is (arbitrarily close to) zero in this experiment). The effect is a continuous one: there is no discontinuous drop in taxes around $\sigma = 1$.

Intuitively, the discussion here simply means that a contemplated government tax hike today has a stronger retarding effect on saving with more curvature in the utility function of the capitalists. Since this effect is unwanted – it lowers future consumption, as seen in the GEE – the government chooses to tax less in such an economy.

In our experiments, we use $\lambda = 0$, $\beta = 0.95$, and $\alpha = 0.3$. We use the same utility function for the capitalists as for the workers; the curvature of the utility function of the workers matters here. For $\sigma = 1.5$, taxes go down to 3% from 5% for logarithmic utility. When $\sigma = 0.5$, taxes are up to 9% in steady state. When $\sigma > 1$, it is interesting to note that we find $\Psi'(K) < 0$: an increase in the current tax rate, which lowers savings, thus increases the tax rate next period as well: there would be a permanent effect of any temporary tax hike contemplated by the government.

2.6. Model solution

We now discuss how to solve the model with numerical methods. We have two equations, (6) and (8), in two unknown functions, H and Ψ. The former has to hold for all (K, τ), and it defines a function of two variables, whereas the latter has to hold for all K, since it defines a function of only one variable. Like in the corresponding functional-equation version of the optimal growth model (which contains one equation only in the planners savings function), there is no simple way to find a solution; in general, a closed-form solution does not exist. However, this set of functional equations is a level more complicated still, because (8) contains derivatives of endogenous functions: (H_τ, H_K). Intuitively, this means that even if one is only interested in finding a steady state (H), one needs to simultaneously solve for the dynamics (H_k) around the steady state! Moreover, the dynamics enter because the current government needs to evaluate responses of future decision makers; since these are in disagreement with the current government (as, in particular, they view future capital as less elastically supplied than the current government views it), no envelope argument is available to say that these responses must be of second-order importance.

Can we obtain the derivatives from simply differentiating (6)? First, the derivative cannot in general be solved out explicitly this way. Second, even if it could be, it would depend on the derivative of Ψ, which is not known. The derivative exists, by assumption, but we would be left with a government FOC which is a functional equation containing the derivative of Ψ in addition to Ψ itself. Thus, derivatives of the unknown function(s) are an unavoidable part of the solution, unlike in the standard model.

Can a steady state be computed numerically? It can. Using a "perturbation" method, as in Krusell et al. (2001) and Klein et al. (2001), it is quite straightforward to find a steady state, and one can also solve for dynamics. This method builds on setting the derivatives of Ψ of order n (large) and higher equal to zero, taking n successive derivatives of the first-order conditions, and then solving for all levels and all derivatives in a joint, nonlinear system. Then, n is increased until the low-order derivatives of interest do not change.

For somewhat more detail, and for our initial equilibrium definition, in the first iteration one would make H a linear function, $H^{(1)}(K,\tau) = a_1^{(1)} + a_2^{(1)}K + a_3^{(1)}\tau$, and Ψ a constant, $\Psi^{(1)}(K) = a_4^{(1)}$, thus leaving four parameters to be determined. A fifth unknown, the steady state, $\bar{K}^{(1)}$, can also be defined by $H^{(1)}(\bar{K}^{(1)}, a_4^{(1)}) = \bar{K}^{(1)}$. We then need four equations to determine $a^{(1)} = (a_1^{(1)}, a_2^{(1)}, a_3^{(1)}, a_4^{(1)})$. For this, we use (6) and (8) as well as the two equations obtained by differentiation of (6) with respect to K and τ, respectively, all evaluated at steady state. In taking the derivatives, the derivative of Ψ is needed, but at this stage of the iteration it is zero. Thus, we have four equations in four unknowns in a nonlinear system. As a result, we obtain a steady state.

In the second iteration, $H^{(2)}(K,\tau)$ is a quadratic, thus containing a vector $a^{(2)}$ of six unknown parameters; similarly, $\Psi^{(2)}(K)$ is linear, with two unknown parameters. We have the same four equations as before, evaluated at a new steady state, $\bar{K}^{(2)}$, plus four new equations. Three of the new equations are obtained by second-order differentiation of (6), and the remaining one comes from first-order differentiation of (8). Thus, we have a system of eight equations in eight unknowns, delivering a steady state. This steady state can be compared to the one of the previous step; if they are closer than some prespecified tolerance level, stop. If not, proceed to the next iteration.[8]

The iterative procedure just outlined leaves open three questions: (i) What if the nonlinear equation system has more than one solution?; (ii) Even if it does not, what

[8] For high-order versions of this procedure, it is recommended to use the more compact equilibrium definition, since it saves on unknown coefficients: there, H is a function only of one variable.

if it is hard to solve?; and (iii) How can the successive differentiation of the first-order equations be done efficiently? In response to (i), in fact, in the very first iteration there are two solutions, delivering a stable and an unstable steady state (as given by whether $a_1^{(1)}$ is greater than 1 in absolute value). Here, select the stable one. In our applications so far, further multiplicity has not been found. As for (ii), there is no general answer, but with starting guesses from the previous stage, the problem may not be so difficult, and the first stage is typically easy. And for (iii), the most efficient route is to take numerical derivatives, given that all functional forms are known: they are the primitive functions plus our polynomials. Numerical derivatives can be obtained easily with centered finite differences.

The reader might ask why it is necessary to give detail about numerical solutions – could not standard linearization, or standard nonlinear (global) methods, be used? Linearization around the steady state, as in Krusell and Ríos-Rull (1999), amounts to the first round of iteration above. However, it does not deliver a steady state with controlled accuracy: at the subsequent iterations, the steady states are different. The key value of the perturbation method is its focus on the first-order conditions and its repeated differentiation of these conditions: it critically exploits the differentiability of the time-consistent equilibrium. Methods that do not rely on differentiability may not work, since a continuum of (discontinuous) solutions likely exist in contexts such as the present one (see Krusell and Smith, 2001).

3. Concluding comments

In a democratic world with selfish voting, workers would be an important group behind taxation decisions. In this paper, this idea is represented by assuming that workers are influential in setting income taxes and transfers. In the setup considered in this paper, taxation of income will persist in the long run, provided that taxes cannot be committed to: there will be a positive tax on income, leading to lower long-run output. In a world with commitment, in contrast, taxes will be zero in the long run, and output is at its first-best steady state.

Only special cases were considered here, both in terms of parameter values and assumptions more generally; for example, it was assumed that workers cannot save. It is straightforward to use the techniques herein for more general problems. An important purpose in this paper has been to communicate a method allowing characterization of time-consistent equilibria more broadly. The set of models for which commitment is not available and this constraint binds is large.

References

Bertola, G., 1993. Factor shares and savings in endogenous growth. American Economic Review 83 (5), 1185–1198.

Chamley, C., 1986. Optimal taxation of capital income in general equilibrium with infinite lives. Econometrica 54 (3), 607–622.

Judd, K.L., 1985. Redistributive taxation in a simple perfect foresight model. Journal of Public Economics 28, 59–83.

Kemp, M.C., Long, N.V., Shimomura, K., 1993. Cyclical and noncyclical redistributive taxation. International Economic Review 34 (2), 415–429.

Klein, P., Krusell, P., Ríos-Rull, J.-V., 2001. Time-consistent policy. Manuscript.

Krusell, P., Ríos-Rull, J.-V., 1999. On the size of U.S. government: Political economy in the neoclassical growth model. American Economic Review 89, 1156–1181.

Krusell, P., Smith Jr., A.A., 1998. Income and wealth heterogeneity in the macroeconomy. Journal of Political Economy 106, 867–896.

Krusell, P., Smith, Jr. A.A., 2001. Consumption–savings decisions with quasi-geometric discounting. Manuscript.

Krusell, P., Kuruşçu, B., Smith, Jr. A.A., 2001. Equilibrium welfare and government policy with quasi-geometric discounting. Journal of Economic Theory, forthcoming.

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