On the role of commitment for tax dynamics PRELIMINARY AND INCOMPLETE

John Hassler,*Per Krusell,†
Kjetil Storesletten,‡and Fabrizio Zilibotti§

March, 2004

Abstract

We argue that in very natural settings, optimal tax dynamics involve fluctuations. We illustrate our point in a simple model of human capital accumulation where the payoffs from current effort involves a stream of future productivity improvements, as opposed to when labor effort only gives static payoffs. Thus, the general principle of smoothing distortions, as opposed to smoothing taxes, can actually give rise to tax fluctuations. Not being convinced that tax fluctuations characterize taxes implemented in real-world economies, we go on to argue that if the government maximizes consumer welfare but cannot commit to future tax rates, a natural dampening, or elimination, of tax fluctuation occurs. The conclusion from this argument is that, if institutions allowing commitment could be set up, the resulting policy changes should move toward higher fluctuations in tax rates (and lower fluctuations in distortions).

*John Hassler: IIES and CEPR; e-mail: John.Hassler@iies.su.se.
†Per Krusell: University of Rochester, IIES, CAERP, CEPR, and NBER; e-mail: pekr@troi.cc.rochester.edu.
‡Kjetil Storesletten: University of Oslo, IIES, and CEPR; e-mail: Kjetil.Storesletten@econ.uio.no.
§Fabrizio Zilibotti: IIES, IFS, and CEPR; e-mail: fabrizio.zilibotti@iies.su.se

1 Introduction

What is the desired path of taxes and of public expenditures over time? We argue in this paper that (i) in very natural settings, fluctuations in taxes and expenditures are to be expected if policy maximizes consumer welfare, and that (ii) the extent of the fluctuations depends on the extent to which the government can commit to how to set future taxes: less commitment leads to a dampening of the fluctuations. What we mean by a "natural setting", first, involves the assumption of capital accumulation of certain forms and, second, involves restrictions on the government's ability to differentiate taxes across different types of capital. We argue that these two assumptions are plausible, although in this paper we do not carry out these arguments in detail.[1] Here, our focus is more on the role of commitment for fluctuations and in making the point how the *absence* of commitment leads to *less* fluctuations.

Our application involves the optimal choice of public expenditures, and accompanying taxes, over time in a model with an infinitely-lived household and a benevolent "Ramsey planner". However, the basic argument can be carried over to a variety of other contexts, such as models of political economy where the government revenue is used for redistribution, not public goods, and where decisions are made under conflicts between the different agents in the population. In particular, democratic rule naturally limits the desirability of commitments to future policy choices, in addition to the general difficulty of implementing commitment. Thus, when applied to a political-economy context, where commitment is an even less natural assumption than here, our model implies a natural tendency for taxes not to fluctuate or at least fluctuate less, i.e., the smaller or no "politically driven cycles".

The framework we study is very stylized in order to just capture the key features behind our argument. We assume here that capital is human capital that is accumulated in period t and that is productive in periods t and $t+1$. To attach words to these assumption, we can think of two "periods of life" of a worker, and of education, or on-the-job training, that occurs early in life and whose benefits remain with the worker until he exits the labor market. More generally, what is essential is that, first, the benefits of effort are persistent, i.e., that they lead to a form of capital, and, second, that these benefits do not depreciate geometrically (such as in our case where the capital dies abruptly with the worker).[2] Although there is an overlapping-generations feature to our capital technology, our results are not driven by overlapping generations per se; the model can be regarded as of a single household that uses the income from the sequence of workers in order to consume in every period. Nor is the two-period horizon of investment crucial. As long as the building up of capacity requires an irreversible investment up-front, our result extends to cases when investment pays off over a longer horizon (although a number of properties of the optimal policy, e.g., the periodicity of the fluctuations, depend on the length of the horizon).

The second key assumption in our framework is that there are restrictions on how the government can raise revenue. In particular, we assume that taxation, because it is assumed to be proportional to income, is distortionary. Moreover, we assume that it cannot distinguish different capital vintages. I.e., income from capital built in periods t and $t-k$ are taxed at the same rate; in the context of our two-period-lived-worker model, "young" and "old" workers are taxed at the same rate. Whether it is feasible in reality to tax capital of different vintages at different rates probably depends on the context, and here we assume no ability to differentiate for illustration

[1] Another paper (Hassler, Krusell, Storesletten, and Zilibotti (2004)) extensively discusses which features of capital accumulation are crucial.

[2] The paper mentioned in footnote 1 characterizes the nature of fluctuations in detail as a function of the form of depreciation, including cases of quasi-geometric depreciation and flat ("lightbulb") depreciation schedules of different lifetimes.

only.

Why does the Ramsey planner choose tax fluctuations in this environment? Consider the following example. Suppose that at a certain time t, a country is subject to an unexpected surge in an external security threat (e.g., "terrorism"), causing an increase in the social value of defense. Suppose, for simplicity, that this increase is perceived as permanent. What is the optimal way to increase tax revenue to finance the larger expenditure? A tax-smoothing argument would suggest that the government should set a higher constant tax rate. Such a response, however, would forego the opportunity to extract more tax revenue from the generation who made their investment before the surge of the threat. This generation sunk its investment under the expectation of lower taxes, and this investment is, at t, an inelastic tax base. So, from a public finance standpoint, a high tax rate is called for at t. This high tax rate can be counteracted by a lower tax rate in $t+1$ so that effort decisions in period t are not too distorted. Then, since the $t+1$ tax rate is low, the government can afford a higher $t+2$ rate, and so on. This oscillating plan features a rather smooth path of distortions, since the present discounted value of taxes on investment for each investor is rather stable over time. At the same time, it allows the planner to exploit the lower elasticity of the tax base at t.

The above example relies on a "surprise". This might suggests that the incentives for the planner to expropriate installed capital in the first period of the planning horizon (similarly to Chamley (1986) and Judd (1985)) plays a crucial role. Indeed, if there were no installed capital in the first period, or if the planner were not allowed to tax it, the optimal tax sequence would be constant. However, we show that optimal tax fluctuations also arise when this incentive is missing in an environment where there are stochastic shock affecting the marginal value of expenditure over time (like in the example above, but without assuming that the change comes as a complete surprise) and the planner sets a state-contingent tax sequence with full commitment. Moreover, allowing the government to save or issue debt does not eliminate fluctuations. There is, however, one scenario in which tax smoothing is optimal: if the government (planner) has access to complete markets, i.e., it can issue state-contingent debt, then fluctuations disappear. This requires, essentially, that the government can commit to raising taxes if the threat does not materialize in order to honor claims to output in case the threat does materialize. We regard this as a very strong assumption, and we conclude therefore that tax fluctuations are rather robust.

Our analysis of the commitment model relates to Barro's (1979) result that tax smoothing is optimal. Barro looked at debt-vs-tax finance of a given stream of expenditures. Our model differs from Barro's in a number of respects. First, in Barro, the distortionary effects of taxation have a static (e.g., labor supply) rather than a dynamic nature. Second, in our specification expenditure is endogenous and determined by the desire to finance a public good. Third, the main focus of our analysis is not on debt-vs-tax finance, and in the benchmark model we impose, for simplicity, that the government budget must balance period-by-period. Of these differences, only the first is essential. Exogenous changes in the stream of public expenditure correspond, in our setting, to an increase in the marginal value of public good inducing the government to increase taxes (e.g., a war).[3] Moreover, we show in an extension that our main results are robust to allowing the government to run budget surpluses or deficits subject to an intertemporal budget constraint.

Our main point in concluding that fluctuations in taxation can be optimal is not to say that Barro's analysis is inaccurate, but merely to emphasize that the smoothing should occur for distortions, not for taxes. In particular, in our model, the distortions to an agent's effort choice can be summarized by the *present value* of extra taxes incurred by the effort choice: whether to become

[3]The analysis in Hassler et al. (2004) is conducted under the assumption that the government must finance an exogenous stream of expenditures; whether the stream is exogenous or not is immaterial.

educated (a higher-earning career, presumably) or not depends on what one thinks will happen with the taxes over the entire course of one's working life. So a fluctuating tax rate on income is not bad per se and, as we show, is desirable in order to implement a higher taxation of already installed (more inelastic) sources of income.

In Barro's model, the commitment solution is time-consistent; however, in related debt-vs-tax analyses, such as the paper by Lucas and Stokey (1983) and the literature that followed, the commitment solution is not. That literature has discussed various ways to achieve commitment, involving institutional design and the introduction of additional instruments. The source of the time-inconsistency is different there than here; there, it is the desire to lower interest rates in order to ease the burden of carry government debt (Barro assumed an exogenous interest rate), whereas here it is the more standard ex-ante vs. ex-post taxation of capital income. Therefore, a comparison of the specifics of the proposed institutions/instrument there and here is not useful. However, more general "mechanisms" for solving commitment problems have been proposed in the Lucas-Stokey (and similar) contexts as well: the appeal to reputation equilibria, along the lines of Abreu, Pearce, and Stacchetti (1990). Here, in contrast, we rule out reputation mechanisms by requiring Markov perfection of the equilibrium; we essentially demand that the equilibrium be a limit of finite-horizon equilibria.[4]

When there is no commitment, the government's tradeoff between costs and benefits changes. Under commitment, it sets the marginal benefit from raising public expenditure at t (which is constant due to our assumption that preferences are linear in this argument) equal to the marginal cost, and the marginal cost has two components. The first component has to do with how effort at t is lowered, and the second with how effort at $t-1$ is lowered. The second of these, clearly, is not present when there is no commitment, just like it is not present in the first period of the commitment problem. Thus, when the commitment problem sets the sum of each future *pair* of marginal costs equal to the first-period marginal cost, and thus leaves it open for each future marginal cost to fluctuate, the problem under no commitment requires the marginal costs to be constant over time, one by one. This allows some movements in tax rates because of the initial sunk effort not being equal to steady-state effort, but the fluctuations are minor. Interestingly, it turns out that the commitment problem, where both taxes and effort fluctuate over time, leads to constant output—i.e., the sum of past and present effort. Under no commitment, output has minor movements.

Our paper is also related to the recent study of public expenditure choice in Klein, Krusell, and Ríos-Rull (2003). Their model is a neoclassical growth setup where the government has no access to debt and has no commitment; like in this paper, they focus on Markov-perfect equilibria. The focus there is on (i) deriving and interpreting first-order conditions for the government and (ii) numerical methods and a quantitative evaluation. The present paper is different not mainly because it derives closed-form solutions but because it emphasizes conditions under which non-monotonic dynamics arise (and which are plausible). The neoclassical framework studied in Klein et al.'s work uses only physical capital, with geometric depreciation, thus ruling out the possibility of oscillations in the solution with commitment.

In Section 2, we describe the basic setup. Then we look at the commitment and no-commitment solutions in the following two sections (Section 3 and Section 4, respectively). The issue of whether the fluctuations in our examples are all memories of the initial period is then dealt with in a very

[4]Using the Lucas-Stokey model, Krusell, Martin, and Ríos-Rull (2004) show that the assumption of lack of commitment under Markov perfection leads to outcomes that are surprisingly similar to the commitment outcomes in terms of observables. However, the results in their paper are of a very different nature than those here, since their equilibrium relies on discontinuous decision rules; here, decision rules are (piece-wise) linear.

simple extension of the basic setup to uncertainty in the second period. There we show that, if government debt is not state-contingent, "new" fluctuations occur as a result of the shock (Section 5.1) whereas if it is state-contingent, no new fluctuations occur: those that are present are indeed a memory of the initial period (Section 5.2). Section 6 concludes.

2 The model

2.1 Population, preferences, technology, and policy

The model economy is populated by a continuum one of dynasties of two-period lived agents. In the first period of their lives, agents undertake an investment in human capital. The cost of investment to each individual is e^2, and the return is spread over two periods. In particular, the individual earn labor earning equal to $e \cdot w$ in the first period of her life and $e \cdot w \cdot z$ in the second period. $z \leq 1$ captures the fact that agents retire within the second period of their life (see, for instance, Matsuyama (2003) for a similar assumption).[5]

Dynasties derive utility from the consumption of a private and a public good. Each period's felicity depends on the total consumption (net of the investment cost) of the dynasty's member, irrespectively of the split of consumption between the old and the young agent. The preferences of the dynasty's cohort that is alive at t are described by the following linear-quadratic utility function

$$U_t = c_t + v(g_t) - e_t^2 + \beta U_{t+1},$$

where $\beta \in [0, 1)$ is the discount factor and g_t denotes the public good available at t. In most of the analysis below, we will assume that $v(g) = Ag$, where A is a parameter describing the marginal utility of the public good to the living cohorts.[6] The marginal cost of the public good is unity and we focus on the case where the social good is valuable; in the linear case, we assume $A \geq 1$, which will imply that the public good is socially valuable. Our quasi-linear utility formulation is chosen so as to be able to ignore savings issues and interest rates that respond to public policy. We thus have that the discount rate, $(1 - \beta)/\beta$, equals the market interest rate. Since the savings decisions can be abstracted from, the welfare of a dynasty is simply given by the present discounted value of their income net of investment costs, i.e.,

$$U_t = \Sigma_{j=0}^{\infty} \beta^j \left((1 - \tau_{t+j}) \, y_{t+j} + v(g_{t+j}) - e_{t+j}^2 \right),$$

where

$$y_{t+j} = (z e_{t+j-1} + e_{t+j}) \, w, \tag{1}$$

i.e., the gross income accruing to the dynasty at $t + j$, given by the sum of the labor incomes generated by the parent born at $t + j - 1$ and her offspring born at $t + j$. The parent's human capital depends on her investment at $t+j-1$ (e_{t+j-1}) while the offspring's human capital depends on her investment at $t+j$ (e_{t+j}). Since agents live for two periods, and the effect of the human capital investment dies with them, y_t only depends on the realization of two subsequent investments.[7]

[5] This effect can be offset by a rising age-earnings profile. We assume here that the effect of retirement is more important, implying that the total earnings are larger in the first than in the second period of an agent's life.

[6] This is equivalent to assuming that the public good provides a marginal utility equal to $A/2$ to each living individual.

[7] The assumption presented so far is consistent, for instance, with agents only consuming when they are old and parents paying for the education of their children.

Note that the environment described so far is equivalent to an economy populated by infinitely lived agents (as opposed to infinitely lived dynasties with overlapping generations), where production is carried out through a sequence of two-period lived projects. In every period t, the representative agent can run a project of size e_t, at the cost e_t^2, without any liquidity constraints. The project generates an income of we_t in period t and of zwe_t in $t+1$ (where $1-z$ would be the first-period depreciation rate of the project), and zero return at $t+j$ for $j>1$ (the depreciation rate is 100% in the second period).

Due to a standard free-riding problem, there is no private provision of the public good. This is instead provided by an agency that will be called "government" that has access to a technology to turn one unit of revenue into one unit of public good. The government revenue is collected by taxing agents' labor income at the flat rate τ, subject to a balanced budget constraint. More formally, the government budget constraint requires that $g_t = \tau_t (ze_{t-1} + e_t) w$, where, at time t, e_{t-1} is predetermined, whereas e_t depends on expectations about the current and future tax rate. In particular, for $\tau_t, \tau_{t+1} \in [0,1]$, the optimal investment of a young agent at t is given by

$$e_t^* = e(\tau_t, \tau_{t+1}) \equiv \frac{1+\beta z - (\tau_t + \beta z \tau_{t+1})}{2} w. \tag{2}$$

Equation (2) shows the distortionary effect of taxation on investments. In particular, τ_{t+j} distorts the investment of two generations: that born at $t+j-1$, as $e_{t+j-1}^* = e(\tau_{t+j-1}, \tau_{t+j})$, and that born at $t+j$, as $e_{t+j}^* = e(\tau_{t+j}, \tau_{t+j+1})$).

The government budget constraint allows us to express the provision of public good at t as a function of current and future (next-period) taxes plus the level of investments sunk at $t-1$. More formally:

$$g_t = \tau_t (ze_{t-1} + e(\tau_t, \tau_{t+1})) w = g(\tau_t, \tau_{t+1}, e_{t-1}). \tag{3}$$

Finally, we restrict $\tau_t \in [0,1] \, \forall t$, which implies that investments, public good provision and private net income (e_t^*, g_t and $(1-\tau_t) y_t$) all are non-negative.

3 The Ramsey allocation with commitment

In this section, we characterize the optimal tax sequence set by a benevolent planner maximizing the utility of the representative dynasty. The choice set of the planner is the set of sequences of taxes, $\{\tau_t\}_{t=0}^{\infty}$, that are feasible for some sequence of public good provision, $\{g_t\}_{t=0}^{\infty}$, and associated private investment choices. We assume that the planner can commit to future taxes; we refer to this problem as the Ramsey problem and to its solution as the Ramsey allocation. More formally, the Ramsey planner chooses the sequence of $\tau_t \; \forall t \geq 0$ in order to maximize

$$W(e_{-1}) = \Sigma_{j=0}^{\infty} \beta^j \left((1-\tau_{t+j}) (ze_{t+j-1} + e_{t+j}) w + v(g_{t+j}) - e_{t+j}^2 \right), \tag{4}$$

subject to

$$g_t = \begin{cases} g(\tau_t, \tau_{t+1}, e_{t-1}) & \text{for } t=0 \\ g(\tau_t, \tau_{t+1}, e(\tau_{t-1}, \tau_t)) & \text{for } t \geq 1. \end{cases}, \tag{5}$$

where the functions $e(.)$ and $g(.)$ are given by (2) and (3), respectively, $\tau_t \in [0,1] \, \forall t$, and e_{-1} is predetermined.

The Ramsey problem does not admit a standard recursive formulation, and the Ramsey allocation is time-inconsistent. Intuitively, when planning at time zero, the planner takes into account the distortion of τ_{t+j} on e_{t+j-1} (where $j \geq 1$), but this investment is sunk when, at $t+j$, the plan is implemented. This relates to the special formulation of the time-zero problem (see the implementability constraint, (5)); while labor tax for all $t > 0$ distort the investments of two cohorts, τ_0 only distorts the investment of agents born in period zero, since the investment of agents born in period minus one is sunk when the tax sequence is set.

We first discuss some general features of the Ramsey solution by deriving first-order conditions. These first-order conditions make clear that monotone dynamics toward a steady state can be ruled out, and argues why this is. Since the first-order conditions are not sufficient, we also provide a complete characterization of the solution to the Ramsey problem by showing how it can be written recursively.

3.1 First-order conditions

As a preliminary, we will first derive the first-order conditions for the planner's choice of taxes. As we will see, they will provide some important hints regarding the optimal solution. Letting $y(\tau_{t-1}, \tau_t, \tau_{t+1}) \equiv (ze(\tau_{t-1}, \tau_t) + e(\tau_t, \tau_{t+1}))w$ denote gross income at t, given privately optimal investments in $t-1$ and t, depending on taxes in t, t and $t+1$, the government's objective can be written directly in terms of tax rates as

$$\sum_{t=0}^{\infty} \beta^t \left(y(\tau_{t-1}, \tau_t, \tau_{t+1})(1 - \tau_t) + v(y(\tau_{t-1}, \tau_t, \tau_{t+1})\tau_t) - e(\tau_t, \tau_{t+1})^2 \right) .$$

What is useful about this objective is that the first-order condition is simpler than one might expect due to most effects of a marginal change in τ_t being "indirect" and thus dropping out. The first-order condition simply reads

$$y_t(v_t' - 1) = -\left(\beta^{-1} v_{t-1}' y_{t-1,3} \tau_{t-1} + v_t' y_{t,2} \tau_t + \beta v_{t+1}' y_{t+1,1} \tau_{t+1} \right) , \tag{6}$$

where $y_t \equiv y(\tau_{t-1}, \tau_t, \tau_{t+1})$, $y_{t,j}$ is the partial of y_t with respect to the j:th argument and v_t' is the marginal value of the public good at period t.

In (6), the left-hand side captures the direct effect of raising taxes. Letting $\gamma_{g,t} \equiv v_t' - 1$ denote the public-expenditures wedge, this term provides the marginal benefit of taxation, given by $\gamma_{g,t}$ times the tax base. Clearly, in the case when public good preferences are linear, this term is always positive. Second, the right-hand side captures the marginal costs: the "budget externalities" that arise because private agents do not take into account how their investment choice influences distortions via the government's budget: as the tax rate at time t increases, this lowers the investment in periods $t-1$ and t and thus lowers the tax bases in three periods: $t-1$, t, and $t+1$, leading to a loss which is proportional to the tax rate which multiplies each of these tax bases.

For the understanding of why oscillations must arise, which we will show formally below, it is instructive first to note in the first-order condition above that (i) the marginal benefit of taxation, $y\gamma_g$, is decreasing in each tax rate, since output is lower with higher tax rates (recall that γ_g is constant with a linear v) and that (ii) the marginal cost of taxation, the sum of the three right-hand side terms, is increasing in each tax rate, because these costs are proportional to the tax rates (recall that output is linear in tax rates so that the output derivatives are constant). Thus, if any one tax rate is increased, since the marginal benefit must fall and the marginal cost must rise, the government must decrease another tax rate in order to still be satisfying its first-order condition! The argument below will be that at time zero, the government will want to have a high tax rate, and that will force the time-1 tax rate down, which from the next first-order condition will force the time-2 tax rate down, and so on.

More precisely, and focusing on the case where v is linear, substituting in the expression for y and its derivatives, and simplifying, (6) can be written

$$z\tau_{t-1} + \left(1 + \beta z^2\right)\tau_t + \beta z \tau_{t+1} = (1+z)(1+\beta z)\frac{A-1}{2A-1}. \tag{7}$$

Turning to the dynamics of taxes implied by this equation, note that the characteristic polynomial of this difference equation, given by $(z+\rho)(\rho\beta z + 1)$, has two roots, $\rho = -z$ and $\rho = -\frac{1}{\beta z}$, both being strictly negative for $z > 0$. Thus, oscillatory behavior is generic. Moreover, as we will see, optimal dynamics will be determined by the root that is smallest in absolute value, $\rho = -z$. In order to show that this type of dynamics is indeed optimal, it is necessary to make sure that appropriate second-order conditions are met, to make sure that none of the tax choices in (7) violates non-negativity, and finally to explicitly incorporate the choice of the period-0 tax rate. The period-0 tax choice also leads to a linear first-order condition, but the key question is whether the tax there should be chosen so that oscillations result in the following periods or so that the future tax rates will be constant: recall that equation (7) is indeed consistent with a constant solution if τ_0 is chosen to the steady state value implied by (7), namely $\tau = \frac{A-1}{2A-1}$. Carrying this analysis out in the sequential formulation of the problem is less convenient than with recursive methods. The purpose of the next section is thus to provide a recursive formulation, where sufficient conditions for optimality are easy to verify. Before doing that, we note that (i) taxes are oscillatory, (ii) effort is oscillatory, but (iii) gross income is constant whenever (7) is satisfied. To see this, note that

$$\begin{aligned} y(\tau_{t-1},\tau_t,\tau_{t+1}) &= \left(ze(\tau_{t-1},\tau_t) + e(\tau_t,\tau_{t+1})\right)w \\ &= \frac{w^2}{2}(1+\beta z)(1+z) - \frac{w^2}{2}\left(z\tau_{t-1} + \left(1+\beta z^2\right)\tau_t + \beta z\tau_{t+1}\right) \end{aligned} \tag{8}$$

where the first term is a constant and the second is constant by (7). Specifically, (7) implies that

$$y(\tau_{t-1},\tau_t,\tau_{t+1}) = \frac{w^2(1+z)(1+\beta z)}{2}\frac{A}{2A-1}.$$

As we see, income is decreasing in A.This is easy to understand. As A increases, it is optimal to increase taxes, despite the fact that higher taxes leads to less investments and thus less income.

3.2 Optimal dynamic taxation, a complete characterization

It is well known that Ramsey problems admit a two-stage formulation whereby future decisions, in stage two, can be described as coming from a recursive problem with an additional state variable whereas the time-zero decisions, in stage one, can be derived from a "static" problem with payoffs that are given by the value function associated with the solution to the recursive problem.[8] In this framework we will show that the second-stage recursive problem is particularly simple in that the choice of τ_{t+1} involves only one state variable: the *current* period's tax rate τ_t is sufficient for determining the optimal τ_{t+1}. This result follows from the fact that since the investment horizon is two periods only, a benevolent planner who can commit to benefits one period ahead would choose the same level of redistribution as would a planner who could commit all future periods. Specifically, if the planner at period t chooses τ_{t+1}, she would have chosen the same τ_{t+1} if she had had the ability to commit at any period $s < t$ (assuming τ_t is held constant). Furthermore,

[8]See, e.g., Marcet and Marimon (1999), where the additional state variable is marginal utility or the Lagrange multiplier associated with the incentive constraint.

although the flow of felicity in period t is affected by both the predetermined variables e_{t-1} and τ_t, the optimal choice of τ_{t+1} is only affected by τ_t. Therefore, the recursive program has only τ_t as a state variable, with τ_{t+1} being the choice variable. As to the choice in the initial period, the planner is not subject to earlier pre-commitments and thus chooses τ_0 and τ_1 simultaneously. The following lemma can thus be proved.

Lemma 1 *The utilitarian planner program (4) is equivalent to the following recursive program:*

$$W(e_{-1}) = \max_{\tau_0 \in [0,1]} \{Y_0(e_{-1}, \tau_0) + V(\tau_0)\} \tag{9}$$

$$V(\tau_t) = \max_{\tau_{t+1} \in [0,1]} \{Y(\tau_t, \tau_{t+1}) + \beta V(\tau_{t+1})\} \quad \textit{for } t \geq 0, \tag{10}$$

where

$$Y_0(e_{-1}, \tau_0) = (1 + (A-1)\tau_0) zwe_{-1}$$

and

$$\begin{aligned} Y(\tau_t, \tau_{t+1}) &= e(\tau_t, \tau_{t+1})((1 - \tau_t + \beta z(1 - \tau_{t+1})) + Aw(\tau_t + \beta z \tau_t) - e(\tau_t, \tau_{t+1})) \\ &= e(\tau_t, \tau_{t+1})(e(\tau_t, \tau_{t+1}) + wA(\tau_t + \beta z \tau_{t+1})). \end{aligned}$$

Moreover, the mapping $\Gamma(v) \equiv \max_{\tau' \in [0,1]} \{Y(\tau, \tau') + \beta v(\tau')\}$ is a contraction mapping with V as the unique fixed point.

Proof. Consider the Ramsey-problem as formulated in (4). Using (3), we can define the planner's period t felicity for $t \geq 1$ as

$$\begin{aligned} F(\tau_{t-1}, \tau_t, \tau_{t+1}) &= wz(1 + (A-1)\tau_t) e(\tau_{t-1}, \tau_t) \\ &\quad + w(1 + (A-1)\tau_t) e(\tau_t, \tau_{t+1}) - e(\tau_t, \tau_{t+1})^2, \end{aligned}$$

and for $t = 0$ as

$$\begin{aligned} F_0(e_{-1}, \tau_0, \tau_1) &\equiv wz(1 + (A-1)\tau_0) e_{-1} \\ &\quad + w(1 + (A-1)\tau_t) e(\tau_0, \tau_1) - e(\tau_0, \tau_1)^2. \end{aligned}$$

Clearly, $F(.)$ can be separated into two additive terms: $F(\tau_{t-1}, \tau_t, \tau_{t+1}) = D(\tau_{t-1}, \tau_t) + H(\tau_t, \tau_{t+1})$. Here,

$$D(\tau_{t-1}, \tau_t) \equiv wz(1 + (A-1)\tau_t) e(\tau_{t-1}, \tau_t)$$

and

$$H(\tau_t, \tau_{t+1}) \equiv w(1 + (A-1)\tau_t) e(\tau_t, \tau_{t+1}) - e(\tau_t, \tau_{t+1})^2.$$

Note that $D(\tau_{t-1}, \tau_t)$ is the felicity in period t that accrues from private consumption of the old and public consumption financed by taxes on the old. Similarly, $H(\tau_t, \tau_{t+1})$ is the felicity in period t that accrues from private consumption of the young and public consumption financed by taxes on the young minus investment costs.

For $t = 0$, we define

$$F_0(e_{-1}, \tau_0, \tau_1) \equiv wz(1 + (A-1)\tau_0) e_{-1} + H(\tau_0, \tau_1).$$

Now define

$$Y(\tau_t, \tau_{t+1}) \equiv H(\tau_t, \tau_{t+1}) + \beta D(\tau_t, \tau_{t+1}).$$

Furthermore, using (2) and $\tau_t \in [0,1]\,\forall t$ we can write

$$Y(\tau_t, \tau_{t+1}) = e(\tau_t, \tau_{t+1})\left(e(\tau_t, \tau_{t+1}) + wA(\tau_t + \beta z \tau_{t+1})\right).$$

The contribution of the period 0 old is

$$\begin{aligned} Y_0(e_{-1}, \tau_0) &\equiv wz\left(1 + (A-1)\tau_0\right)e_{-1} \\ &= F_0(e_{-1}, \tau_0, \tau_1) - H(\tau_0, \tau_1). \end{aligned}$$

The planner problem under commitment (4) can now be expressed as

$$\begin{aligned} W(e_{-1}) &= \max_{\{\tau_t\}_{t=0}^{\infty}} \left\{ F_0(e_{-1}, \tau_0, \tau_1) + \sum_{t=1}^{\infty} \beta^t F(\tau_{t-1}, \tau_t, \tau_{t+1}) \right\} \\ &= \max_{\{\tau_t\}_{t=0}^{\infty}} \left\{ F_0(e_{-1}, \tau_0, \tau_1) + \sum_{t=1}^{\infty} \beta^t \left(D(\tau_{t-1}, \tau_t) + H(\tau_t, \tau_{t+1})\right) \right\} \\ &= \max_{\{\tau_t\}_{t=0}^{\infty}} \left\{ F_0(e_{-1}, \tau_0, \tau_1) - H(\tau_0, \tau_1) + \sum_{t=0}^{\infty} \beta^t \left(\beta D(\tau_t, \tau_{t+1}) + H(\tau_t, \tau_{t+1})\right) \right\} \\ &= \max_{\{\tau_t\}_{t=0}^{\infty}} \left\{ Y_0(e_{-1}, \tau_0) + \sum_{t=0}^{\infty} \beta^t Y(\tau_t, \tau_{t+1}) \right\}. \end{aligned} \tag{11}$$

Defining the value function

$$V(\tau_t) \equiv \max_{\{\tau_{t+s}\}_{s=1}^{\infty}} \sum_{s=0}^{\infty} \beta^s Y(\tau_{t+s}, \tau_{t+s+1}), \tag{12}$$

we can rewrite (11) as

$$\begin{aligned} W(e_{-1}) &= \max_{\tau_0} \left\{ Y_0(e_{-1}, \tau_0) + V(\tau_0) \right\} \qquad (13) \\ &= \max_{\tau_0} \left\{ Y_0(e_{-1}, \tau_0) + \max_{\tau_1} \left\{ Y(e_{-1}, \tau_0) + \beta V(\tau_1) \right\} \right\} \end{aligned}$$

and standard recursion on (12) for $t \geq 1$ yields the functional Bellman equation (10). Since Y is bounded by the fact that $\tau \in [0,1]$ and since $0 \leq \beta < 1$, the Bellman equation (10) is a contraction mapping with a unique solution, which must also be the solution to the sequential continuation problem (Theorem 4.3 in Stokey and Lucas, 1989). ■

$Y(\tau_t, \tau_{t+1})$ is the contribution to the planner's utility (evaluated at time t) of the cohort born in period t. This contribution consists of the cohort's lifetime private consumption, i.e., $1-\tau_t+\beta z(1-\tau_{t+1})$, plus the utility from public consumption financed by the cohort's contribution $(wAe_t(\tau_t + \beta z \tau_{t+1}))$, minus the cost of period t investment (e_t^2). Similarly, $Y_0(e_{-1}, \tau_0)$ is the contribution of the initial old. Clearly, the planner's objective is to maximize the discounted sum of the contribution of all cohorts and this can be done in a recursive fashion when the tax rate is determined one period in advance.

The recursive formulation form facilitates the characterization of the optimal sequence of taxes. This characterization is provided in the following proposition.

Proposition 1 *The optimal solution to the Ramsey program, (4), is*

$$\tau_{t+1} = \max\{0, \tau^* - z(\tau_t - \tau^*)\} < 1, \tag{14}$$

for $t \geq 0$ *and*

$$\tau_0 = \begin{cases} \tau_0 = \left(1 + \frac{2ze_{-1}}{w(1-\beta z)}\right)\tau^* & \text{if } e_{-1} \leq \frac{w(1-\beta z)}{2z^2} \\ \min\left\{1, \left(1 + \beta z + \frac{2ze_{-1}}{w}\right)\tau^*\right\} & \text{else,} \end{cases},$$

where

$$\tau^* = \frac{A-1}{2A-1} \in [0, \frac{1}{2})$$

is the steady-state tax rate. If $z < 1$*, the Ramsey tax sequence converges asymptotically in an oscillatory fashion to* τ^**. If* $z = 1$*, the Ramsey tax sequence is a 2-period cycle such that,*

$$\tau_t = \begin{cases} \tau_0 & \text{if } t \text{ is even} \\ \max\{0, 2\tau^* - \tau_0\} & \text{if } t \text{ is odd.} \end{cases}$$

Proof. We first guess the characterization of value function, $V(\tau_t)$, and then show that this is the unique solution to (10). We guess that:

$$V(\tau_t) = \begin{cases} \frac{w^2}{4}\left(B_0 + B_1\tau_t + B_2\tau_t^2\right) & \tau_t \leq \frac{1+z}{z}\tau^* \\ \frac{w^2}{4}\left(-(2A-1)\tau_t^2 + 2(A-1)(1+\beta z)\tau_t + (1+\beta z)^2 + \beta B_0\right) & \text{else,} \end{cases}, \tag{15}$$

where

$$\begin{aligned} B_2 &\equiv -\left(1-\beta z^2\right)(2A-1) \\ B_1 &\equiv 2\left(1-\beta z^2\right)(A-1) \\ B_0 &\equiv = \frac{(2A-(1-\beta z))(1+\beta z(2A-1)) + \beta(A-1)^2(1-z)^2}{(1-\beta)(2A-1)} \\ \tau^* &= \frac{A-1}{2A-1}. \end{aligned}$$

We also guess that the optimal policy is linear in the state variable, i.e., that

$$\tau_{t+1} = T(\tau_t) = T_0 + T_1\tau_t,$$

where

$$\begin{aligned} T_1 &= -z \\ T_0 &= \frac{(1+z)(A-1)}{2A-1}. \end{aligned}$$

Note that this guess implies that $\tau^* = \frac{A-1}{2A-1}$. When $\tau_t \leq 2\tau^*$, the optimal τ_{t+1} is interior, and otherwise the restriction $\tau_{t+1} \geq 0$ binds.

The first-order condition for τ_{t+1} in the Bellman equation is given by

$$\begin{aligned} 0 &= Y_2(\tau_t, \tau_{t+1}) + \beta V'(\tau_{t+1}), \\ 0 &= -\beta z\frac{1}{2}w^2(2A-1)\left(\tau_t + \beta z\tau_{t+1} - \frac{A-1}{2A-1}(1+\beta z)\right) + \beta\frac{w^2}{4}(B_1 + 2B_2\tau_{t+1}). \end{aligned}$$

The solution to this yields (14). The right-hand side of the Bellman equation is concave in τ_{t+1} and (14) implies that $\tau_{t+1} < 1 \; \forall t > 0$ so the constraint $\tau_{t+1} \leq 1$ never binds. The first-order condition $(Y_2(\tau_t, \tau_{t+1}) + \beta V'(\tau_{t+1}))\, \tau_{t+1} = 0$ is therefore necessary and sufficient for maximization of the right-hand side of the Bellman equation. Substituting the expressions for B_1 and B_2 into the right-hand side of the first-order condition yields

$$\tau_{t+1} = \frac{(1+z)(A-1)}{2A-1} - z\tau_t,$$

showing that our guess that $\tau_{t+1} = T(\tau_t) = T_0 + T_1\tau_t$ is the optimal policy given our proposed value function.

Finally, we must verify that our proposed value function satisfies the functional equation

$$V(\tau_t) = Y(\tau, T(\tau)) + \beta V(T(\tau)).$$

This yields

$$\begin{aligned} &\frac{w^2}{4}\left(B_0 + B_1\tau_t + B_2\tau_t^2\right) \\ = \; & -(2A-1)\, e(\tau_t, T(\tau_t))^2 + Aw(1+\beta z)\, e(\tau_t, T(\tau_t)) \\ & + \beta\frac{w^2}{4}\left(B_0 + B_1 T(\tau_t) + B_2 T(\tau_t)^2\right). \end{aligned}$$

Some algebra verifies that this equality holds for all τ_t.

We now consider cases when the constraint $\tau_{t+1} \geq 0$ binds. First note that (14) implies that $\tau_{t+1} \geq 0$ possibly binds for $t = 0$, but never thereafter. Specifically, it binds if

$$\tau_0 > \frac{1+z}{z}\tau^*.$$

In this case, $\tau_1 = 0$ and the value function is given by

$$\begin{aligned} V(\tau_0)\,|_{\tau_0 > \frac{1+z}{z}\tau^*} &= Y(\tau_0, 0) + \beta V(0) \\ &= \frac{w^2}{4}\left(-(2A-1)\,\tau_0^2 + 2(A-1)(1+\beta z)\,\tau_0 + (1+\beta z)^2 + \beta B_0\right). \end{aligned}$$

We have now verified that $V(\tau_t) = \max_{\tau_{t+1}} \{Y(\tau_t, \tau_{t+1}) + \beta V(\tau_{t+1})\}$, i.e., that V is a fixed-point of the functional mapping $\Gamma(v) = \max_{\tau' \in [0,1]} \{Y(\tau, \tau') + \beta v(\tau')\}$. Since Γ is a contraction mapping, V is unique.

Finally, consider the first-period problem (9). Inserting the expression for Y_0 and V and simplifying yields

$$\begin{aligned} &\max_{\tau_0} \{Y_0(e_{-1}, \tau_0) + V(\tau_0)\} \\ = \; & (1 + (A-1)\,\tau_0)\, e_{-1} z w + \frac{w^2}{4}\left(B_0 + B_1\tau_0 + B_2\tau_0^2\right), \end{aligned}$$

if assuming that the optimal τ_0 is less than or equal to $\frac{1+z}{z}\tau^*$. The first-order condition for this problem is

$$\begin{aligned} 0 &= (A-1)\, e_{-1} z w + \frac{w^2}{4}(B_1 + 2B_2\tau_0) \\ \Rightarrow \; \tau_0 &= \left(1 + \frac{2ze_{-1}}{w(1-\beta z)}\right)\tau^*. \end{aligned}$$

Clearly, this satisfies $\tau_0 \leq \frac{1+z}{z}\tau^*$ if $e_{-1} \leq \frac{w(1-\beta z)}{2z^2}$. For higher e_{-1}, the first order condition is

$$\tau_0 = \left(\frac{2z}{w}e_{-1} + 1 + \beta z\right)\tau^*,$$

concluding the proof. ■

There are two main results in Proposition (1). First, tax and effort fluctuations are a generic feature of the optimal plan. The oscillations are dampened if $z < 1$, and persist forever if $z = 1$. In the latter case, the Ramsey allocation entails a 2-period cycle. Second, taxation attains its maximum in the first period (i.e., in period zero). This follows immediately from the observation that $\tau_0 \geq \tau^*$ and that the dynamics are non-diverging from period one onwards.[9]

The latter result is not unexpected: the elasticity of the tax base is lower in period zero than in later periods, since the old sunk their investment in period minus one. Therefore, as is standard in the optimal taxation literature (see Chamley (1986) and Judd (1985), the government has an incentive to "overtax" in the first period. The former result is more surprising, as one might have expected that the planner would opt for a constant tax sequence after the first period. To understand why, instead, she chooses an oscillating sequence, note that efficiency requires that distortions on investments, not taxes themselves, be smooth. Since the horizon of each investment is two periods, the distortion on investments at t depends on the present discounted value of taxes over two periods, $\tau_t + \beta z \tau_{t+1}$ (see equation (2)). Therefore the distortion implied by any constant tax sequence can be replicated by an alternation of high and low taxes: for instance, setting $\tau_t = \tau_{t+1} = 0$ has the same effect on time-t investments as setting $\tau_t = -\beta z \tau_{t+1}$. But why does the planner *strictly* prefer to oscillate? The answer lies in the incentive to overtax in the first period. In period zero, the planner has a unique chance to raise funds at a low cost, but by setting a high τ_0 she would highly distort the investments of the agents born at time zero. This can be corrected by promising this generation low taxes in period one. In turn, this implies that agents born at one are treated generously in the first period, giving the planner the opportunity to increase taxation in period two. And so on. Clearly, no constant tax sequence from period one onwards could smooth investments to the same extent. In Section 3.3 below we will discuss distortion smoothing more in detail from the perspective of the dynamic trading off of distortions against each other.

The speed at which the sequence converges to the steady state depends on how responsive to future taxes investments are. For instance, as z (or β) tends to zero, the scope for oscillating tax rates vanishes, since agents become more "myopic" and less sensitive to the promise of future taxation. When z is close to one, on the other hand, oscillations become very persistent, culminating in a unit root when $z = 1$. It is interesting to note that when $e_{-1} = 0$ there are no oscillations. In this case, there is no incentive for the planner to overtax in the first period, and the optimal solution features tax smoothing, i.e., taxes are always at their steady state. This shows that oscillations are a "long echo" of the first period.

Note that the steady-state tax rate that maximizes tax revenues is $\tau = \frac{1}{2}$, which naturally is an upper bound for the optimal steady-state tax rate. Therefore, naturally, the Ramsey allocation is never on the descending part of the Laffer curve.

3.3 Distortion smoothing

Let us now return to the issue of distortion smoothing. When tax-distortions are a time-invariant function of the current tax-rate only, as in Barro (1979), tax smoothing and distortion smoothing

[9] Note that $\tau_0 = \left(1 + \beta z + \frac{2ze_{-1}}{w}\right)\tau^*$ is the case when the constraint $\tau_0 \geq 0$ binds at $t = 1$.

are equivalent. As noted above, this is, however, not the case in the present setting. To see that optimal taxes do imply distortion smoothing, first note that a the per-period planner payoff may be written

$$(ze_{t-1} + e_t)w - g_t + v(g_t) - e_t^2.$$

It will be useful to express policy as a function of government expenditure and private investment levels instead of using the tax rate as the choice variable. The tax rate satisfies $\tau_t = \frac{g_t}{(ze_{t-1}+e_t)w}$, i.e., the tax at t is a function of the expenditure at t and the effort levels at $t-1$ and at t. We can now abstractly express the agent's first-order condition for effort at t as (concrete specifications will follow below)

$$\eta(e_{t-1}, e_t, e_{t+1}, g_t, g_{t+1}) = 0.$$

This constraint makes explicit the budget externalities in this model; private agents ignore that increased effort, via the balanced government budget, indirectly raises the level of public expenditures (or, equivalently, allow a lower tax rate at maintained g_t). This is a positive externality since public goods are under-provided relative to the first best.

Denoting the multiplier for the private first-order constrain by $\beta^t \lambda_t$, the government's Lagrangian can be written as

$$\sum_{t=0}^{\infty} \beta^t \left((ze_{t-1} + e_t)w - g_t + v(g_t) - e_t^2 - \lambda_t \eta(e_{t-1}, e_t, e_{t+1}, g_t, g_{t+1}) \right).$$

Letting $\eta_{t,j}$ denote the jth partial of $\eta(e_{t-1}, e_t, e_{t+1}, g_t, g_{t+1})$ the first-order conditions for the choice of e_t and g_t become, for $t > 0$,

$$\gamma_{e,t} - \beta^{-1}\lambda_{t-1}\eta_{t-1,3} - \lambda_t \eta_{t,2} - \beta \lambda_{t+1} \eta_{t+1,1} = 0, \tag{16}$$

and

$$\gamma_{g,t} - \beta^{-1}\lambda_{t-1}\eta_{t-1,5} - \lambda_t \eta_{t,4} = 0, \tag{17}$$

where .we recall that $\gamma_{g,t} \equiv v'(g_t) - 1$ is the wedge between the social value and cost of the public good and, similarly, $\gamma_{e,t} \equiv w(1+\beta z) - 2e_t$ is the wedge between the social value and cost of investment.

The Ramsey optimum is characterized by an optimal trade off of these wedges against each other over time. How is the trade-off between wedges determined exactly? To find out, it is necessary to eliminate the multipliers in the first-order conditions above. Thus, use equation (16) for period t and equation (17) for periods t and $t+1$, since then we have three equations with the unknowns λ_{t-1}, λ_t, and λ_{t+1}. Thus, λ_t can be solved for as a function of $\gamma_{e,t}$, $\gamma_{g,t}$, and $\gamma_{g,t+1}$. Substitute the solutions for λ_t and λ_{t-1} into equation (17) and we obtain the final expression for the first-order condition for the government's policy choice:

$$\gamma_{g,t} = \beta^{-1}\eta_{t-1,5} D_{t-1} \left(\gamma_{e,t-1} - \frac{\eta_{t-2,3}}{\eta_{t-2,5}} \gamma_{g,t-1} - \beta \frac{\eta_{t,1}}{\eta_{t,4}} \gamma_{g,t} \right) + \eta_{t,4} D_t \left(\gamma_{e,t} - \frac{\eta_{t-1,3}}{\eta_{t-1,5}} \gamma_{g,t} - \beta \frac{\eta_{t+1,1}}{\eta_{t+1,4}} \gamma_{g,t+1} \right) \tag{18}$$

where D_t, the measure of how much a unit increase of the constraint at t is worth in terms of effort e_t, is calculated as

$$D_t \equiv \frac{1}{\eta_{t,2} - \eta_{t,4} \frac{\eta_{t-1,3}}{\eta_{t-1,5}} - \eta_{t,5} \frac{\eta_{t+1,1}}{\eta_{t+1,4}}}.$$

The first-order condition for the government's choice, equation (18), reveals exactly how distortion smoothing takes place: smoothing involves all the relevant wedges, and the nature of the way wedges are balanced against each other is prescribed in this equation for the choice of g_t. Thus, we see that g_t must be chosen in order to balance 5 gaps: two effort gaps (at $t-1$ and t) and three public-expenditure gaps (at $t-1$, t, and $t+1$).

The left-hand side measures the direct benefit of raising g_t, or expressed alternatively, the marginal value of public funds in excess of their value used for private consumption. The right-hand side measures the indirect costs of providing public goods, i.e., the distortionary cost of raising public funds. Its first term summarizes how the increase in taxes at t increases the distortion on effort at $t-1$ (captured by $\gamma_{e,t-1}$). In turn, the lower effort at $t-1$ decreases public expenditures at $t-1$ and t through the budget externalities in those periods (this is why the public expenditure gaps in $t-1$ and t appear in the second term on the right-hand side). Similarly, the second term on the right-hand side summarizes the negative effects of the higher current tax on the current effort gap and, through the budget externalities, on the public expenditure gaps at t and $t+1$. The steady state of this first-order condition thus pins down a relation between the two kinds of gaps.

Now consider the wedges in period zero. Then first-order conditions in period 0 can be written

$$\gamma_{g,0} = \eta_{0,4} D_0 \left(\gamma_{e,0} - \beta \frac{\eta_{1,1}}{\eta_{1,4}} \gamma_{g,1} \right), \tag{19}$$

where

$$D_0 \equiv \frac{1}{\eta_{0,2} - \eta_{0,5} \frac{\eta_{1,1}}{\eta_{1,4}}}.$$

This equation differs from equation (18) in three ways, all relating to the fact that the marginal cost of raising g_0 involves no effects on variables in the period prior to period 0 are present. First, the whole first term of equation (18)—the effect on effort in the period prior—is not present. Second, the effect on effort in period 0, which is present here, does not feed back on the current public-expenditure gap, γ_0; if it had, e_0 would have made g_{-1} change in order to fulfil the implementability constraint in period -1. Third, the effect of a change in g_0 on e_0, $\eta_{0,4} D_0$, is slightly different: the expression for D_0 misses a term in the denominator. That is, again, the indirect effect that an change in e_0 has through the previous period's constraint, measured through how much g_0 would have to change to neutralize the change in e_0 on that constraint.

We can easily interpret (18) and (19) as "distortion-smoothing" in our model. First, we note that the left-hand sides of these equations, $\gamma_{g,t}$, are constant at $A-1$. Thus, also the right-hand sides, the marginal cost of tax distortions, must be constant (smooth). To see what implications this smoothness has for taxes, we first note that

$$\eta(e_{t-1}, e_t, e_{t+1}, g_t, g_{t+1}) = e_t - \frac{1}{2}(1+\beta z) w + \frac{g_t}{2(z e_{t-1} + e_t)} + \frac{\beta z g_{t+1}}{2(z e_t + e_{t+1})}.$$

Taking the relevant partials and substituting into (18) yields

$$\begin{aligned} A-1 &= \beta^{-1} \frac{\beta z}{2(z e_{t-1} + e_t)} \left(w(1+\beta z) - 2e_{t-1} + \frac{g_{t-1}}{z e_{t-2} + e_{t-1}} (A-1) + \beta \frac{z g_t}{z e_{t-1} + e_t} (A-1) \right) \\ &\quad + \frac{1}{2(z e_{t-1} + e_t)} \left(w(1+\beta z) - 2e_t + \frac{g_t}{z e_{t-1} + e_t} (A-1) + \beta \frac{z g_{t+1}}{z e_t + e_{t+1}} (A-1) \right). \end{aligned} \tag{20}$$

Now, since our optimal allocation is expressed in terms of τ_t, we use $\frac{g_t}{(z e_{t-1} + e_t)} = w \tau_t$ to substitute for g_t, arriving at

$$A - 1 = \frac{w\left((1+z)(1+\beta z) + (A-1)\left(z\tau_{t-1} + \left(1 + z\beta z^2\right)\tau_t + \beta\tau_{t+1}\right)\right)}{2(z e_{t-1} + e_t)} - 1.$$

The right-hand side of this equation equals $A-1$, which is easily verified using (7) to substitute for the $\left(z\tau_{t-1}+\left(1+z\beta z^2\right)\tau_t+\beta\tau_{t+1}\right)$ and (8) for $\left(ze_{t-1}+e_t\right)$. Thus, the difference equation for tax rates derived above naturally comes out here as well.

4 No-commitment Ramsey outcomes: Markov-perfect equilibrium

In the previous section, we allowed the planner to determine taxes for all future dates under full commitment. We also established that the sequence of taxes under full commitment is identical to the sequence when taxes can be pre-committed one period ahead of their implementation. The purpose of this section is to characterize the optimal time-consistent allocation, namely, the allocation that is chosen by a benevolent planner without access to a commitment technology. We will provide the recursive formulation of the problem, assuming, consistently with the setup above, that period t taxes are set in the beginning of period t, and observed before period t investments are decided.

More formally, we characterize the Markov-perfect equilibrium of the game between the agents and the government, i.e., the equilibrium where e_{t-1} is the only state variable in period t and reputation is not used as a means to compensate for commitment. The period t felicity of the planner is given by

$$\begin{aligned} F_d\left(e_{t-1},\tau_t,\tau_{t+1}\right) &= \left(1-\tau_t\right)y_t-e\left(\tau_t,\tau_{t+1}\right)^2+Ag_t \\ &= \left(ze_{t-1}+e\left(\tau_t,\tau_{t+1}\right)\right)\left(1+\left(A-1\right)\tau_t\right)w-e\left(\tau_t,\tau_{t+1}\right)^2, \end{aligned}$$

where e_{t-1} is pre-determined. Taxes are set according to a time-invariant function $\tau_t=T\left(e_{t-1}\right)$. Given this function, individuals rationally believe that $\tau_{t+1}=T\left(e_t\right)$, and individually rational investment choices must therefore satisfy

$$e_t=\frac{1+\beta z-\left(\tau_t+\beta zT\left(e_t\right)\right)}{2}w.$$

The Markov equilibrium is defined as follows.

Definition 1 *A time-consistent (Markov) allocation without commitment is defined as a pair of functions $\langle W,T\rangle$, where W is a bounded planner value function and $T:[0,\infty)\rightarrow[0,1]$ is a public policy rule, $\tau_t=T\left(e_{t-1}\right)$, such that the following functional equations are satisfied:*

1. $W_d\left(e_{t-1}\right)=\max_{\tau_t}\left\{F_d\left(e_{t-1},\tau_t,T\left(e_t\right)\right)+\beta W_d\left(e_t\right)\right\}$

2. $T\left(e_{t-1}\right)=\arg\max_{\tau_t}\left\{F_d\left(e_{t-1},\tau_t,T\left(e_t\right)\right)+\beta W_d\left(e_t\right)\right\}$,
   where $e_t=\left(1+\beta-\left(\tau_t+\beta T\left(e_t\right)\right)\right)w/2$.

The following Proposition can now be established.

Proposition 2 *Assume that either $A\leq\frac{z(z+1)}{(1+\beta)z^2-1}$ or $\left(1+\beta\right)z^2\leq1$.*[10] *Then, the time-consistent allocation is characterized as follows:*

$$\begin{aligned} T\left(e_{t-1}\right) &= \min\left\{\bar{\tau}+\alpha_1\left(e_{t-1}-\bar{e}\right),1\right\} \\ e_t &= \bar{e}-\frac{w}{2+\beta z\alpha_1 w}\left(\tau_t-\bar{\tau}\right), \end{aligned}$$

[10] This assumption ensures that the constraint $\tau_{t+1}\leq1$ never binds for $t\geq0$. Without this constraint, the analysis would be substantially more complicated, involving non-continuous policy functions.

where

$$\bar{e} = \frac{w(1+\beta z)(1-\alpha_0)}{2+\alpha_1 w(1+\beta z)} \leq e^*$$
$$\bar{\tau} = \frac{2\alpha_0+\alpha_1 w(1+\beta z)}{2+\alpha_1 w(1+\beta z)} \geq \tau^*$$

with equalities iff $A=1$,*and*

$$\alpha_1 = \frac{\sqrt{1+4A(A-1)(1-\beta z^2)} - \left(1+2\left(1-\beta z^2\right)(A-1)\right)}{\beta z(A-1)(1-\beta z^2)w} \geq 0$$
$$\alpha_0 = \frac{2(A-1)-\beta z\alpha_1 w}{2+(A-1)(4+\beta z\alpha_1 w)} \geq 0$$
$$\frac{\partial \alpha_1}{\partial A} \geq 0, \frac{\partial \alpha_0}{\partial A} \geq 0, \frac{\partial \bar{\tau}}{\partial A} \geq 0, \frac{\partial \bar{e}}{\partial A} \leq 0.$$

For all t, *the equilibrium law of motion is*

$$e_{t+1} = \bar{e} - z_d(e_t - \bar{e}), \quad (21)$$
$$\tau_{t+1} = \bar{\tau} - z_d(\tau_t - \bar{\tau}). \quad (22)$$

where

$$z_d \equiv \frac{\alpha_1 w}{2+\beta\alpha_1 w} \in (0,z).$$

Given any e_{-1}, *the economy converges to a unique steady state such that* $\tau=\bar{\tau}$ *and* $e=\bar{e}$ *following an oscillating path and the constraint* $\tau_t \leq 1$ *iff t=0 and* $e_{-1} > \frac{1-\alpha_0}{\alpha_1}$, *while* $\tau_t \geq 0$ *never binds.*

The proof is in the appendix. The parameter restriction under which the Proposition is stated is a sufficient condition for the constraint $\tau_{t+1} \leq 1$ never to bind for $t \geq 0$. When this constraint is violated, the equilibrium policy functions may be discontinuous, making the analysis more involved.

The main findings are that

1. the Markov allocation implies higher steady-state taxation ($\bar{\tau} > \tau^*$) and lower output and investment ($\bar{e} < e^*$) than the Ramsey allocation.

2. the Markov allocation implies less oscillations (i.e., a smoother tax sequence) than the Ramsey allocation: $z_d < z$.

The lack of commitment induces the planner to systematically "overtax" human capital ex post. Agents anticipate that and respond by decreasing their investment. As a result, output is lower. It is interesting to note that the steady-state Markov tax rate, $\bar{\tau}$, can exceed 1/2, i.e., it can be larger than the constant value of taxes that maximizes tax revenues and public good provision. Specifically, this happens if $A > 1 + \frac{2+z(1-\beta z)}{z(2+z(1+\beta))}$, a threshold which decreases in β and z. In this case, the benevolent planner chooses a long-run tax rate that is on the wrong side of the Laffer curve. This is an extreme manifestation of the lack of commitment; if $\bar{\tau} > 1/2$, the planner would clearly like to reduce the steady-state tax rate. However, the planner can only control the current tax rate and a one-period reduction of τ_t would lead to even higher taxes in the following period, resulting in an overall reduction of the current welfare.

The second result is to our knowledge new. It states that, due again to lack of commitment, the Markov planner chooses, along a transition, an inefficiently smooth tax sequence. The intuition, which as in the analysis of the commitment case we will support with a formal analysis of distortion smoothing (contained in Section 4.1), runs as follows. The Ramsey planner taxes agents heavily in the first period, in order to extract revenue from the inelastic human capital of the old. This tends to cause a large distortion on the investment of the young in the first generation. In order to keep distortions smooth, the planner compensates high taxation in period zero by trying to promise low taxation in period one. The Markov planner, however, is unable to honor her promises, and cannot commit, in particular, to future taxes that are as low as the Ramsey planner can set. At time 0, agents thus expect that taxes will not be so low in period one. Distortions are therefore larger for any given τ_0. The optimal behavior of a Markov planner in period 0 is therefore to set τ_0 *lower* than the Ramsey planner would, because such a choice counteracts to some extent the unavoidable fact that the next government will choose too high a tax: it keeps effort at zero from being too low. The same logic applies to later periods. So the tax sequence starts out lower and tends to be much smoother. However, income, which we saw was constant in the commitment outcome, becomes somewhat volatile here. In fact, using the law of motion for τ_t, it is straightforward to show that

$$y_t = \frac{1}{2}w^2\left((1+\beta z)(1+z)(1-\bar{\tau}) + (1-z_d\beta z)\frac{z-z_d}{z_d}(\tau_t - \bar{\tau})\right), \tag{23}$$

implying that income oscillates and is positively correlated with taxes.

4.1 Distortion smoothing under lack of commitment

Along the lines of the analysis in section 3.3 above, we can state a first-order condition of the government that summarizes how the public-expenditure and effort wedges are traded off over time when there is lack of commitment to future policy. In essence, the result is a condition which is very similar, though not identical, to the one for the period-0 first-order condition for the government with commitment.

To find the first-order condition, first state the government's problem as a dynamic program with the previous investment choice, e_{t-1}, as state variable:

$$W(e_{t-1}) = \max_{e_t, g_t}(ze_{t-1} + e_t)w - g_t + v(g_t) + \beta W(e_t)$$

subject to

$$\eta(e_{t-1}, e_t, E(e_t), g_t, G(e_t)) = 0.$$

Here, $E(e_{t-1})$ is the policy rule for effort and $G(e_{t-1})$ is the policy rule for public expenditures. Taking first-order conditions, we obtain

$$w - 2e_t + \beta W'_e = \lambda_t(\eta_2 + E'(e_{t+1})\eta_3 + G'(e_{t+1})\eta_5)$$

for the choice of e_t and

$$\gamma_{g,t} = \lambda\eta_4$$

for the choice of g_t. Solving for λ_t from the first-order condition for e_t, we obtain

$$\gamma_{g,t} = D_t\eta_{t,4}(w - 2e + \beta W'(e_{t+1})),$$

where

$$D_t \equiv \frac{1}{\eta_{t,2} + E'(e_{t+1})\eta_{t,3} + G'(e_{t+1})\eta_{t,5}}.$$

Since the envelope theorem gives

$$W'(e_t) = zw - \lambda_t \eta_{t,1},$$

evaluated the following period this expression and the first-order condition for g_t deliver the "distortion-smoothing" condition specifying how trade-offs between wedges occur in the model without commitment. It reads

$$\gamma_{g,t} = \eta_{t,4} D_t (\gamma_{e,t} - \beta \frac{\eta_{t+1,1}}{\eta_{t+1,4}} \gamma_{g,t+1}). \tag{24}$$

Clearly, in our model, the left-hand side—the excess value of government funds, that is, the marginal benefit of raising g—is constant at $(A-1)$, implying that the right-hand side, the marginal cost of raising g, has to be constant. Unlike the distortion smoothing that occurs under commitment, it involves only the marginal cost of distorting current e (and its repercussion on future public expenditures): the cost of distorting past effort choices are, for natural reasons, not taken into account.

Finally, we note that (24) differs from the period-0 first-order condition from the commitment problem, (19), only in how D_t is determined. The expression D_t determines a key component of how the change in g_t influences e_t, via the implementability constraint. Here, an increase in e_t changes η_t exactly by $1/D_t$, and this expression includes the total effect on how a change in e_t would influence the future government behavior that feeds back to the current constraint (e_{t+1} and g_{t+1}). In D_0 of the commitment problem, in contrast, the current government can control future decisions and the effects of future government behavior on the current constraint are partial—they are derived keeping future constraints constant. Thus, whereas we have $E'(e_{t+1})\,\eta_{t,3} + G'(e_{t+1})\,\eta_{t,5}$ in D_t here, in D_0 we just have $\eta_{0,5}\frac{\eta_{1,1}}{\eta_{1,4}}$.

5 Stochastic government expenditure

Proposition 1 establishes that fluctuations in taxes and output are efficient. However, if $e_{-1} = 0$, the optimal tax sequence is smooth. The latter observation may raise the concern that the result is of limited practical importance, as it is unclear what counterpart the first period has in the real world. The purpose of the section is to show that fluctuations do not hinge on the particular incentives faced by the planner the first period. To make the point sharp, we consider economies where $e_{-1} = 0$, implying that the planner has no opportunity to tax sunk investments in the first period. The planner has full commitment power over future taxation, and can announce a state-contingent tax plan. The first result is that oscillations arise when the value of the public good (and, hence, the marginal benefit of taxation) is not constant over time. This result hinges, however, on the assumed inability of the government to borrow and lend. When this is introduced, the optimal tax sequence are again smooth. The second, more interesting result, is that the optimal tax sequence has an oscillatory nature when the value of the public good is stochastic. This result is robust to the introduction of public savings/debt.

We extend the model in the following direction. In period zero, the value of the public good is $A = A_l$. In the beginning of period 1, with probability p, A jumps to $A_h > A_l$ and stays there forever; with probability $1 - p$ nothing happens then or later. We can regard this event as the start of a war that makes public expenditure particularly valuable. For simplicity, we assume that this shock is permanent, i.e., the war once started never ends. The extreme cases where $p = 0$ and $p = 1$ correspond to no uncertainty. In particular, $p = 0$ is the benchmark case of section 3, whereas $p = 1$ is a simple case where the value of public good is not constant over time.

The planner sets, at period zero, τ_0 and a state-contingent tax plan, $\{\tau_{h,t}, \tau_{l,t}\}_{t>0}$. The sequence $\{\tau_{h,t}\}_{t>0}$ is implemented if $A = A_h$, whereas the sequence $\{\tau_{l,t}\}_{t>0}$ is implemented if the productivity of the public good remains low, $A = A_l$. The two-state specification is chosen to simplify analysis. Exploiting Lemma 1, we can write the Ramsey problem as follows:

$$W(e_{-1}) = \max_{\tau_0 \in [0,1]} \{Y_0(e_{-1}, \tau_0) + V(\tau_0)\}, \tag{25}$$

$$V(\tau_0) = \max_{\tau_l, \tau_h} \left\{ \tilde{Y}(\tau_0, \tau_1^e) + \beta(pV(\tau_{h,1}; h) + (1-p)V(\tau_{l,1}; l)) \right\}, \tag{26}$$

$$V(\tau_{\omega,t}; \omega) = \max_{\tau_{\omega,t+1} \in [0,1]} \{Y_\omega(\tau_{\omega,t}, \tau_{\omega,t+1}) + \beta V(\tau_{\omega,t}; \omega)\} \quad \text{for } t \geq 1, \tag{27}$$

where $\tau_1^e \equiv p\tau_{h,1} + (1-p)\tau_{l,1}$, $\omega \in \{h, l\}$, and

$$\begin{aligned}
Y_0(e_{-1}, \tau_0) &= (1 + (A_l - 1)\tau_0) zwe_{-1}, \\
Y_\omega(\tau_{\omega,t}, \tau_{\omega,t+1}) &= e(\tau_{\omega,t}, \tau_{\omega,t+1}) \cdot (e(\tau_{\omega,t}, \tau_{\omega,t+1}) + wA_\omega(\tau_{\omega,t} + \beta z\tau_{\omega,t+1})) \\
\tilde{Y}(\tau_0, \tau_1^e) &= e_0(\tau_0, \tau_1^e) \\
&\quad \cdot (e_0(\tau_0, \tau_1^e) + w(A_l\tau_0 + \beta z(pA_h\tau_{h,1} + (1-p)A_l\tau_{l,1}))).
\end{aligned}$$

Here, $e_0(\tau_0, \tau_1^e)$ denotes the optimal investment at zero, defined as

$$e_0(\tau_0, \tau_1^e) = \frac{1 + \beta z - (\tau_0 + \beta z\tau_1^e)}{2} w.$$

The value functions W and V, as well as the return functions Y_0 and Y, are as in Lemma 1 (aside from the fact that V and Y here are defined as conditional on a particular realization of A in period one). In particular, the continuation problem from period one onwards is identical to the deterministic problem of section 3, since all uncertainty is revealed at that point. $V(\tau_0)$ and $\tilde{Y}(\tau_0, \tau_1^e)$ are instead modified to account for uncertainty, in period zero, about the realization at period one. Since we want to abstract from incentives for the planner to overtax in the first period, we set $e_{-1} = 0$.

To simplify the analysis further, we restrict attention to parameters such that the constraints that taxes are bounded between zero and one never bind. In this case, the first order conditions in period zero are

$$\begin{aligned}
\tau_0; \frac{\partial \tilde{Y}}{\partial \tau_0} &= 0 \\
\tau_{l,1}; \frac{\partial \tilde{Y}}{\partial \tau_l} + \beta(1-p)V'(\tau_{l,1}; l) &= 0 \\
\tau_{h,1}; \frac{\partial \tilde{Y}}{\partial \tau_{h,1}} + \beta pV'(\tau_{h,1}; h) &= 0,
\end{aligned}$$

where we know, by the proof of Proposition 1 (see, in particular, equation 15), that

$$V'(\tau_{\omega,t}; \omega) = 2\left(1 - \beta z^2\right)(A_\omega - 1) - 2\tau_{\omega,t}\left(1 - \beta z^2\right)(2A - 1).$$

Solving the resulting system of linear equations yields:

$$\begin{aligned}
\tau_0 &= \frac{A_l-1}{2A_l-1} - p\frac{\beta z}{2A_l-1}\left(\left(A_h+A_l-1\right)\tau_h-\left(A_l-1\right)\right) \\
\tau_{l,1} &= \frac{A_l-1}{2A_l-1} \\
\tau_{h,1} &= \frac{A_h-1}{2A_h-1}+\frac{z\left(A_h-A_l\right)}{2A_h-1}\cdot \\
&\quad \frac{A_l\left(2A_h-1\right)\left(1+\beta z\left(1-p\right)\right)+\beta zpA_h\left(A_h+A_l-1\right)}{\left(2A_l-1\right)\left(2A_h-1\right)\left(1-\beta z^2\right)-\beta z^2p\left(A_h-A_l\right)^2}.
\end{aligned} \tag{28}$$

The first observation is that, conditional on the realization A_l (peace), taxes are smooth after the first period. More formally, $\tau_{l,t}=\tau_l^*=\frac{A_l-1}{2A_l-1}$ for all $t\geq 1$. This result depends on the particular specification chosen where one of the two realization of the stochastic process coincides with the productivity of the public good at time zero. Consider, next, the tax sequence conditional on the realization A_h (war). In this case, $\tau_{h,1}>\tau_h^*=\frac{A_h-1}{2A_h-1}$, i.e., the government sets period-one taxation above its steady-state level.[11] Since, for $t\geq 1$, the standard dynamics apply, i.e.,

$$\tau_{h,t+1}=\tau_h^*-z\left(\tau_{h,t}-\tau_h^*\right),$$

then, the Ramsey allocation features tax fluctuations even in the absence of any incentive for the planner to tax sunk investments in the first period ($e_{-1}=0$). Why would the planner promise a tax rate above the steady state when the war starts? The reason is intuitive. Conditional on war, the generation born at time zero had invested more than future generations, since (i) taxation had been set lower in the period zero because public good provision was less valuable and (ii) agents had invested while attaching some probability to peace and low taxes occurring in future. Specifically, $e_0=e_l+\beta z\frac{A_h-A_l}{2A_l-1}\frac{pw}{2}\tau_{h,1}$, where e_l is effort in case of peace. Thus, the marginal cost of raising tax revenue for the Ramsey planner in period one is low. Even if she cannot "surprise" agents, the planner then finds it optimal to promise high taxes in the event of a war. As in the benchmark case when $p=0$, this discourages investments from the generation born at one, creating the conditions for optimal taxes to be low in period two, and so on.

Two particular cases are worth emphasizing. First, if $z=0$, i.e., agents only work in the first period and taxes distort their static labor supply, there is no scope for inducing oscillations after the war starts. In this case, the Ramsey tax sequence is perfectly smooth after the first-period upward jump. More formally, if $z=0$, then, for all $t\geq 1$, $\tau_{h,t}=\tau_h^*$. Second, suppose that $p=1$, i.e., the war is perfectly anticipated by the government and private agents. Fluctuations do not disappear in this case. The reason is that even if agents anticipate the increase in future taxation, the government has an incentive to spend less in the period zero, since the marginal utility of the public good is low, and the government cannot save. Thus, in period one, there is a larger inelastic tax base, an an incentive to tax high initiating the fluctuations.

Finally, τ_0 is also affected by the probability of a war. The comparative statics are somehow involved. However, the intuition suggests that the government would like to impose a low taxation on the first generation of young agent, in order to increase the tax base for taxation in the following period, when the war increases the benefits from public expenditure.

[11] The denominator of the last term on the right-hand side term of the expression can be negative. However, recall that we are restricting attention to the region of the parameter space such that taxes are strictly within the unit interval for all times and realizations. When this restriction is taken into account, that denominator is unambiguously positive, implying that $\tau_{h,1}>\tau_h^*$.

To illustrate the effects more concretely, we propose a numerical example. We set $\beta = p = 0.5$, $z = 1$, $A_l = 1.6$ and $A_h = 2$. Since $z = 1$, if there are fluctuations, they are not dampening. In this case (see Figure 1), the Ramsey sequence implies

$$\tau_0 = 0.154, \tau_{h,1} = 0.632, \tau_{l,1} = 0.273.$$

In the case of war, taxes fluctuate between $\tau_{h,t} = 0.632$ for $t = 1, 3, 5, ...$ and $\tau_{h,t} = 0.035$ for $t = 2, 4, 6, ...$ If there is no war, the tax rate remains constant at 0.273. Increasing the probability of a war decreases τ_0 monotonically: when $p = 0$, $\tau_0 = 0.632$, whereas when $p = 1$, τ_0 falls to 0.031. As the war becomes more likely, the government becomes more eager to accumulate resources for the future. If the planner cannot accumulate assets, the only way it can shift resources to the future is by reducing taxation in period zero, so as to enhance the human capital investment of the generation born at zero.

As this discussion suggests, the constraint that the government cannot save is important. In particular, the planner would like not to spend its budget in period zero, and accumulate it in the event of a war. Would fluctuations survive if government debt/savings were allowed? We address this question in the next subsection.

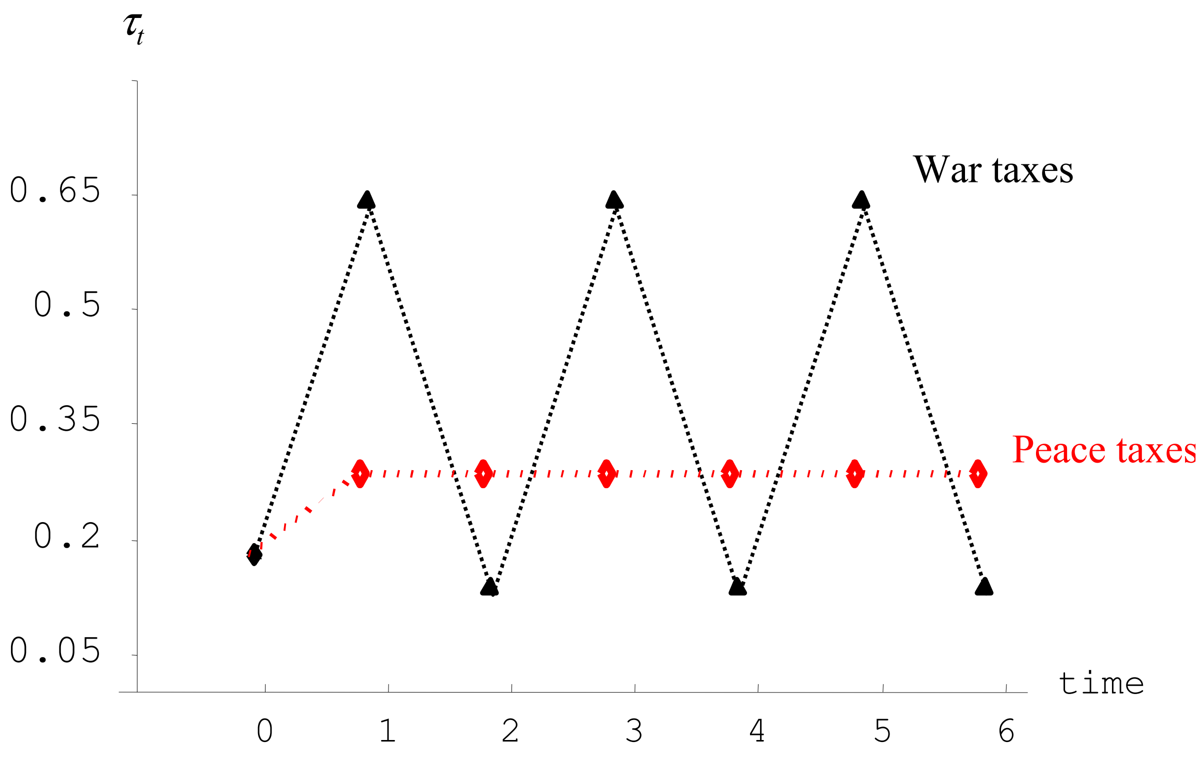

Figure 1. Taxes under war and peace, no saving, $\beta = p = 0.5$, $z = 1, A_l = 1.6$ and $A_h = 2, z = 1$.

5.1 Government debt/savings

We now assume that the government can accumulate assets. We denote the stock of government assets by q (negative q is debt). We rule out Ponzi schemes by assuming that from period one onwards, after uncertainty is resolved, an intertemporal budget constraint must hold. This means that if the government enters period one with a positive stock of assets, it can spend the surplus on additional public good provision. To the opposite, if it is indebted, it has to repay it by imposing extra taxes. The government flow budget constraint is

$$q_{t+1} = (1+r)\left(q_t + \tau_t\left(e_{t-1} + e_t\right) - g_t\right),$$

where we assume that $1 + r = \beta^{-1}$.

As in the previous section, we assume that the productivity of the public good is A_l in period zero and either A_l or $A_h > A_l$ thereafter. Also, we continue to assume that $e_{-1} = 0$. Since, at time zero, the expected future marginal value of the public good, $pA_h + (1-p)A_l$, exceeds the marginal value of the public good in period zero, A_l, the government will never spend any tax revenue in period zero, i.e., it will set $g_0 = 0$. Using these facts, government assets in period one are simply given by

$$q_1 = \frac{1}{\beta}\tau_0 w e_0.$$

Note that, if the government reaches period one with a surplus, it is indifferent as to when to spend it. The reason is twofold. First, $1 + r = \beta^{-1}$, and, second, the marginal value of the public good is constant. Without loss of generality, we will assume in this case that the government spends the entire surplus in period one. This is just for notational convenience. Nothing would change in the analysis if we assumed, for instance, that the government smooths out spending over time.

The Ramsey problem can be reformulated as follows

$$\begin{aligned}
W(0) &= \max_{\tau_0 \in [0,1]} V(\tau_0), \\
V(\tau_0) &= \max_{\tau_l, \tau_h} \left\{ \hat{Y}(\tau_0, \tau_1^e) + \beta\left(pV(\tau_{h,1}; h) + (1-p)V(\tau_{l,1}; l)\right) \right\} \quad \text{for } t = 0, \\
V(\tau_{\omega,t}; \omega) &= \max_{\tau_{\omega,t+1} \in [0,1]} \left\{ Y_\omega(\tau_{\omega,t}, \tau_{\omega,t+1}) + \beta V(\tau_{\omega,t+1}; \omega) \right\} \quad \text{for } t \geq 1, \omega \in \{l, h\},
\end{aligned}$$

where

$$\hat{Y}(\tau_0, \tau_1^e) = e_0\left(e_0 + w\left(\tau_0 A_1^e + \beta z\left(pA_h\tau_{h,1} + (1-p)A_l\tau_{l,1}\right)\right)\right).$$

Here, $A_1^e \equiv (pA_h + (1-p)A_l)$ is the expected value of A. The term $\tau_0 A_1^e$ in $\hat{Y}(\tau_0, \tau_l, \tau_h)$ replaces the term $\tau_0 A_l$ in $\tilde{Y}(\tau_0, \tau_1^e)$ in the case of no public savings. This is because taxes levied on the young in period zero can now be used to finance public goods in period one when they are, in expectation, more useful. Other than that the problem is identical to the case of no savings.

This fact emphasizes a general property of this model. Access to government debt is only useful whenever the government has an interest in disentangling the timing of tax collection from that of expenditure. In particular, whenever $A_0 < A_1^e$ the planner is better off by accumulating surplus. Conversely, whenever $A_0 > A_1^e$ the planner would like to accumulate debt. If $A_0 = A_1^e$ the planner is indifferent, ex-ante, as to when to spend, and adding the government savings/debt does not affect the planner's expected utility. This illustrates that our results in section 3 (constant A) are unchanged if the assumption of balanced budget is relaxed.

Continuing to focus on the case when no constraint on the choice of taxes binds, and proceeding as before leads to the following solution to the first-order conditions:

$$\begin{aligned}
\tau_0 &= (1+\beta z)\frac{A_l + p(A_h - A_l) - 1}{2A_1^e - 1} - \beta z p \frac{A_l + A_h + p(A_h - A_l) - 1}{2A_1^e - 1}\tau_{h,1} \\
&\quad -\beta z(1-p)\frac{2A_l + p(A_h - A_l) - 1}{2A_1^e - 1}\tau_l \\
\tau_{l,1} &= \frac{(A_l - 1)(1+z) - z(2A_l + p(A_h - A_l) - 1)\tau_0 - \beta z^2 p (A_h + A_l - 1)\tau_{h,1}}{(1-\beta z^2 p)(2A_l - 1)} \\
\tau_{h,1} &= \frac{(A_h - 1)(1+z) - z(A_h + A_l + p(A_h - A_l) - 1)\tau_0 - \beta z^2(1-p)(A_h + A_l - 1)\tau_{l,1}}{(1-\beta z^2(1-p))(2A_h - 1)}.
\end{aligned}$$

While more involved than in the absence of debt, these expressions are instructive. First, if $p = 1$, the solution yields $\tau_0 = \tau_{h,1} = \tau_h^* = \frac{A_h - 1}{2A_h - 1}$, implying no dynamics. A perfectly anticipated war does not, alone, induce fluctuations, as long as the government can save or borrow. Access to a market for saving and borrowing enables the government to concentrate spending on public goods to times when A is high (wars). Taxation, in this case, is the same as if A were at the high level each period.

Consider now the case when $0 < p < 1$. Now, we have $\tau_{h,1} \neq \tau_h^*$ and uncertainty triggers dynamics. Interestingly, oscillations arise in this case even if the war does not materialize. Consider the same numerical example as before. The Ramsey sequence now implies

$$\tau_0 = 0.315, \tau_{h,1} = 0.473, \tau_{l,1} = 0.083.$$

In case of war, the tax rate fluctuates between $\tau_{h,t} = 0.473$ for $t = 1, 3, 5, ...$ and $\tau_{h,t} = 0.194$ for $t = 2, 4, 6, ...$ If there is no war, the tax rate fluctuates between $\tau_{l,t} = 0.083$ for $t = 1, 3, 5, ...$ and $\tau_{l,t} = 0.463$ for $t = 2, 4, 6, ...$ The government accumulation in the first period amounts to $q_1 = 0.33$. Comparing the cases with and without public savings/debt, it turns out that when the government can save, taxation in period zero is higher: the government engages in "precautionary savings" in case the war materializes. This is the opposite of the case without assets, when the only way the planner could prepare for a war was by encouraging human capital accumulation through low taxes. Taxes are smoother conditional on war, but more volatile conditional on peace.

Note that when $0 < p < 1$, a market for safe lending and borrowing does not span all states of the world—financial markets are still incomplete. As we will see in the next subsection, this incompleteness if crucial for the existence of tax fluctuations that are not just a repercussion of the period 0 tax hike.

5.2 State-contingent debt

Suppose that in the first period there are two state-contingent assets paying 1 unit of the consumption good in the state of war (peace) and 0 in the state of peace (war). Let period-zero consumption be the numéraire and define $p_{\omega,t}$ as the Arrow-Debreu price of the consumption good in period t and state ω. The government budget constraint at time 0 is then

$$g_0 + \sum_{t=1}^{\infty}\sum_{\omega} p_{\omega,t} g_{\omega,t} = \tau_0 (z e_{-1} + e_0) w + \sum_{t=1}^{\infty}\sum_{\omega} p_{\omega,t}\tau_{\omega,t}(z e_{\omega,t-1} + e_{\omega,t}) w. \quad (29)$$

In comparison to the case of no state-contingent assets, the budget constraint now is collapsed into one time-zero constraint only.

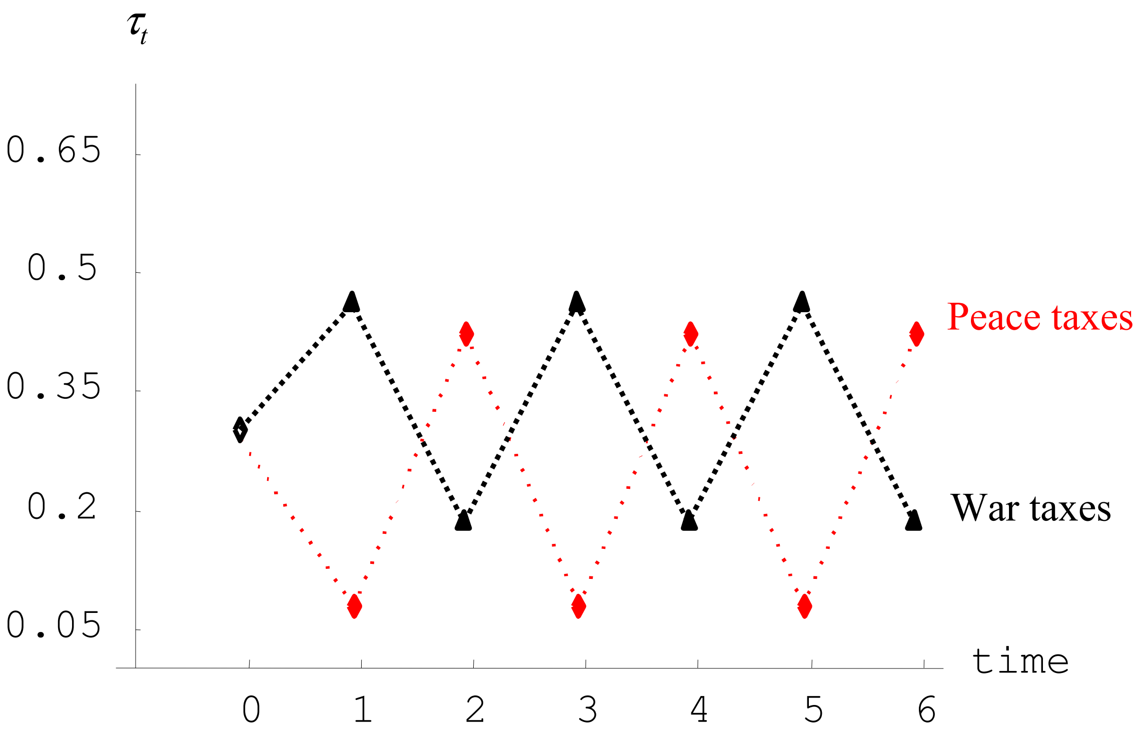

Figure 1: **Figure 2.** Taxes under war and peace, with saving, $\beta = p = 0.5$, $z = 1$, $A_l = 1.6$ and $A_h = 2, z = 1$.

Letting $P(\omega)$ denote the probability of state ω, the objective function is

$$\begin{aligned} W(e_{-1}) &= (1-\tau_0)(ze_{-1}+e_0)w - e_0^2 + A_l g_0 \\ &\quad + \Sigma_\omega \Sigma_{t=1}^{\infty} \beta^t P(\omega)\left((1-\tau_{\omega,t})(ze_{\omega,t-1}+e_{\omega,t})w + A_\omega g_{\omega,t+j} - e_{\omega,t}^2\right) \\ &= A_l g_0 + \sum_{t=1}^{\infty}\sum_{\omega} \beta^t P(\omega) A_\omega g_{\omega,t} + (1-\tau_0)(ze_{-1}+e_0)w - e_0^2 \\ &\quad + \Sigma_\omega \Sigma_{t=1}^{\infty} \beta^t P(\omega)\left((1-\tau_{\omega,t})(ze_{\omega,t-1}+e_{\omega,t})w - e_{\omega,t}^2\right). \end{aligned} \tag{30}$$

Since individual utility is linear in consumption, it follows that the Arrow-Debreu prices must be given by the discounted probabilities, i.e., $p_{\omega,t} = \beta^t P(\omega)$. Since the preferences over public-good provision are linear and the state-contingent prices equal discounted probabilities, it follows that the planner will choose zero public good provision in case of peace, i.e., to concentrate all spending in the state of war when the marginal value of spending is high. Hence, $g_0 = g_{l,t} = 0 \forall t$. It follows that the left-hand side of the budget constraint (29) is given by

$$g_0 + \sum_{t=1}^{\infty}\sum_{\omega} p_{\omega,t} g_{\omega,t} = \sum_{t=1}^{\infty} \beta^t p g_{h,t} \tag{31}$$

and that the discounted expected utility of public-good provision is given by

$$\begin{aligned}
& A_l g_0 + \sum_{t=1}^{\infty} \sum_{\omega} \beta^t P(\omega) A_\omega g_{\omega,t} \\
= \; & A_h \sum_{t=1}^{\infty} \beta^t p g_{h,t} \\
= \; & A_h \left(\tau_0 (z e_{-1} + e_0) w + \sum_{t=1}^{\infty} \sum_{\omega} p_{\omega,t} \tau_{\omega,t} (z e_{\omega,t-1} + e_{\omega,t}) w \right),
\end{aligned}$$

where we used the budget constraint (29) and (31) for the last inequality. Substituting the last expression for the discounted expected utility of public-good provision into the objective function (30), we obtain

$$\begin{aligned}
W(e_{-1}) = \; & A_h \left(\tau_0 (z e_{-1} + e_0) w + \sum_{t=1}^{\infty} \sum_{\omega} p_{\omega,t} \tau_{\omega,t} (z e_{\omega,t-1} + e_{\omega,t}) w \right) \\
& + (1 - \tau_0)(z e_{-1} + e_0) w - e_0^2 + \Sigma_\omega \Sigma_{t=1}^{\infty} \beta^t P(\omega) \left((1 - \tau_{\omega,t})(z e_{\omega,t-1} + e_{\omega,t}) w - e_{\omega,t}^2 \right) \\
= \; & (1 + (A_h - 1)\tau_0)(z e_{-1} + e_0) w - e_0^2 \\
& + \Sigma_\omega \Sigma_{t=1}^{\infty} \beta^t P(\omega) \left((1 + (A_h - 1)\tau_{\omega,t})(z e_{\omega,t-1} + e_{\omega,t}) w - e_{\omega,t}^2 \right).
\end{aligned}$$

As we see, the marginal value of raising funds in both states of the world is now A_h. From this it follows that the optimal sequence of taxes must be identical to the one when there is a war for sure (see the analysis in subsection 5.1 with debt and deterministic war). With state-contingent assets and government debt, if $e_{-1} = 0$, then $\tau_{\omega,t} = \frac{A_h - 1}{2A_h - 1}$ $\forall \omega, t$, giving rise to no fluctuations. The state and time tax sequence is used to finance high provision in the state of war and no provision in the state of peace. Specifically, if public good provision were g_h in the case of a war for sure, it would be $\frac{g_h}{p}$ when the probability of a war is p and there are state contingent assets.

We conclude that state-contingent assets allow the government to do two things. First, spending can be concentrated to the state when the value of the public good is highest (war). Second, the distortion of taxes can be equalized across states since taxes collected in any state of the world can be used to finance war spending.

The result that all spending occurs in wartime is, of course, due to the assumption of linear preferences over public good consumption and would not extend to the case of strictly concave utility. However, the result that taxes do not depend on the state would survive such an extension. Our conclusion is thus that non-existence of complete market for state-contingent assets is key for the result that optimal taxes are cyclical. The results are summarized in the table below.

	No debt/savings	Debt/savings only	Complete markets
Constant A	cycles only if $e_{-1} > 0$	only if $e_{-1} > 0$	only if $e_{-1} > 0$
Deterministic change in A	cycles	only if $e_{-1} > 0$	only if $e_{-1} > 0$
Stochastic A	cycles	cycles	only if $e_{-1} > 0$

6 Conclusions

To be written.

7 References (almost completely incomplete)

Chamley, Christophe, 1986, Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives, Econometrica 54:3, 607–622.

Judd, Kenneth L., 1985, Redistributive Taxation in a Simple Perfect Foresight Model, Journal of Public Economics 28, 59–83.

8 Appendix

This section has not been proofread.

8.1 Details of claims in section3.3

The partials are,

$$\begin{aligned}
\eta_{t,1} &= -\frac{g_t z}{2\left(ze_{t-1}+e_t\right)^2}, \eta_{t,2} = 1 - \frac{g_t}{2\left(ze_{t-1}+e_t\right)^2} - \frac{\beta z^2 g_{t+1}}{2\left(ze_t+e_{t+1}\right)^2} \\
\eta_{t,3} &= -\frac{\beta z g_{t+1}}{2\left(ze_t+e_{t+1}\right)^2}, \eta_{t,4} = \frac{1}{2\left(ze_{t-1}+e_t\right)}, \eta_{t,5} = \frac{\beta z}{2\left(ze_t+e_{t+1}\right)}
\end{aligned}$$

and

$$\begin{aligned}
\eta_{t,4}\frac{\eta_{t-1,3}}{\eta_{t-1,5}} &= -\frac{1}{2}\frac{g_t}{\left(ze_{t-1}+e_t\right)^2}, \eta_{t,5}\frac{\eta_{t+1,1}}{\eta_{t+1,4}} = -\frac{1}{2}\beta z^2\frac{g_{t+1}}{\left(ze_t+e_{t+1}\right)^2} \\
\frac{\eta_{t-1,3}}{\eta_{t-1,5}} &= -\frac{g_t}{ze_{t-1}+e_t}, \frac{\eta_{t,1}}{\eta_{t,4}} = -\frac{zg_t}{ze_{t-1}+e_t}.
\end{aligned}$$

Using this, we find that

$$\begin{aligned}
D_t &= \frac{1}{1-\frac{g_t}{2\left(ze_{t-1}+e_t\right)^2} - \beta z^2\frac{g_{t+1}}{2\left(ze_t+e_{t+1}\right)^2} + \frac{1}{2}\frac{g_t}{\left(ze_{t-1}+e_t\right)^2} + \frac{1}{2}\beta z^2\frac{g_{t+1}}{\left(ze_t+e_{t+1}\right)^2}} \\
&= 1
\end{aligned}$$

Now,

$$\begin{aligned}
\gamma_{g,t} &= \beta^{-1}\eta_{t-1,5}D_{t-1}\left(\gamma_{e,t-1} - \frac{\eta_{t-2,3}}{\eta_{t-2,5}}\gamma_{g,t-1} - \beta\frac{\eta_{t,1}}{\eta_{t,4}}\gamma_{g,t}\right) + \eta_{t,4}D_t\left(\gamma_{e,t} - \frac{\eta_{t-1,3}}{\eta_{t-1,5}}\gamma_{g,t} - \beta\frac{\eta_{t+1,1}}{\eta_{t+1,4}}\gamma_{g,t+1}\right) \\
A-1 &= \beta^{-1}\frac{\beta z}{2\left(ze_{t-1}+e_t\right)}\left(w\left(1+\beta z\right) - 2e_{t-1} + \frac{g_{t-1}}{ze_{t-2}+e_{t-1}}\left(A-1\right) + \beta\frac{zg_t}{ze_{t-1}+e_t}\left(A-1\right)\right) \\
&\quad + \frac{1}{2\left(ze_{t-1}+e_t\right)}\left(w\left(1+\beta z\right) - 2e_t + \frac{g_t}{ze_{t-1}+e_t}\left(A-1\right) + \beta\frac{zg_{t+1}}{ze_t+e_{t+1}}\left(A-1\right)\right),
\end{aligned}$$

as in (20).

For period zero,. we have

$$\begin{aligned}
D_0 &\equiv \frac{1}{1-\frac{g_0}{2\left(ze_{-1}+e_0\right)^2} - \frac{\beta z^2 g_1}{2\left(ze_0+e_1\right)^2} + \frac{1}{2}\beta z^2\frac{g_1}{\left(ze_0+e_1\right)^2}} \\
&= .\frac{1}{1-\frac{g_0}{2\left(ze_{-1}+e_0\right)^2}}
\end{aligned}$$

and

$$\begin{aligned}
\gamma_{g,0} &= \frac{1}{2\left(ze_{-1}+e_0\right)}\frac{1}{1-\frac{g_0}{2\left(ze_{-1}+e_0\right)^2}}\left(w\left(1+\beta z\right) - 2e_0 + \beta\left(\frac{zg_0}{ze_{-1}+e_0}\right)\left(A-1\right)\right) \quad (32) \\
&= \frac{\left(\left(1+\beta z\right)w - 2e_0 + \beta z\frac{g_0}{ze_{-1}+e_0}\left(A-1\right)\right)}{\left(2\left(ze_{-1}+e_0\right) - \frac{g_0}{\left(ze_{-1}+e_0\right)}\right)} \quad (33) \\
&= \frac{\left(\left(1+\beta z\right)w - 2e_0 + \beta z\tau_0 w\left(A-1\right)\right)}{\left(2\left(ze_{-1}+e_0\right) - \tau_0 w\right)} \quad (34)
\end{aligned}$$

8.2 Proof of Proposition 2 (to be cleaned up)

The planner felicity is

$$F_d(e_{t-1}, \tau_t, \tau_{t+1}) = (ze_{t-1} + e(\tau_t, \tau_{t+1}))(1 + (A-1)\tau_t)w - e(\tau_t, \tau_{t+1})^2,$$

We guess that

$$\tau_t = T(e_{t-1}) = \alpha_0 + \alpha_1 e_{t-1}. \tag{35}$$

Given the guess, the investment decision is $e_t = (1 + \beta z - (\tau_t + \beta z(\alpha_0 + \alpha_1 e_t)))w/2$, implying

$$e_t = I(\tau_t) = \frac{(1 + \beta z(1-\alpha_0))w}{2 + \beta z \alpha_1 w} - \frac{w}{2 + \beta z \alpha_1 w}\tau_t$$

and

$$\begin{aligned} \tau_{t+1} &= T(I(\tau_t)) = \bar{\tau} + z_d(\tau_t - \bar{\tau}), \\ e_{t+1} &= I(T(I(e_t))) = \bar{e} + z_d(e_t - \bar{e}), \end{aligned}$$

where

$$\bar{\tau} = \frac{2\alpha_0 + \alpha_1 w(1+\beta z)}{2 + \alpha_1 w(1+\beta z)}, \tag{36}$$

$$\bar{e} = \frac{w(1+\beta z)(1-\alpha_0)}{2 + \alpha_1 w(1+\beta z)}, \tag{37}$$

$$z_d = -\frac{w\alpha_1}{2 + \beta z \alpha_1 w} \tag{38}$$

The problem then admits the following recursive formulation:

$$\begin{aligned} W_d(e_{t-1}) &= \max_{\tau_t}\{F_d(e_{t-1}, \tau_t, \tau_{t+1}) + \beta W_d(e_t)\}, \qquad (39) \\ \text{s.t. } \tau_{t+1} &= \alpha_0 + \alpha_1 e_t, \\ e_t &= \frac{(1 + \beta z(1-\alpha_0))w}{2 + \beta z \alpha_1 w} - \frac{w}{2 + \beta z \alpha_1 w}\tau_t. \end{aligned}$$

Given the guess, the first-order condition for maximizing the RHS of the Bellman equation is

$$\frac{\partial F_d}{\partial \tau_t} + \frac{\partial F_d}{\partial \tau_{t+1}}\frac{d\tau_{t+1}}{d\tau_t} + \beta \frac{dW_d(e_t)}{de_t}\frac{de_t}{d\tau_t} = 0,$$

where

$$\frac{\partial F}{\partial \tau_t} = (ze_{t-1} + e(\tau_t, \tau_{t+1}))(A-1)w - ((1 + (A-1)\tau_t)w - 2e(\tau_t, \tau_{t+1}))\frac{w}{2},$$

$$\frac{\partial F}{\partial \tau_{t+1}} = -\beta z((1 + (A-1)\tau_t)w - 2e(\tau_t, \tau_{t+1}))\frac{w}{2}$$

where we have used the fact that

$$\frac{\partial e_t}{\partial \tau_t} = -\frac{w}{2}, \frac{\partial e_t}{\partial \tau_{t+1}} = -\beta z \frac{w}{2},$$

Using the envelope condition, we obtain

$$W_d'(e_t) = \frac{\partial F_d(e_t, \tau_{t+1}, \tau_{t+2})}{\partial e_t} = (1 + (A-1)\tau_{t+1})wz.$$

which can be expressed in terms of τ_t using the constraints in (39). We can then can write the first-order condition as

$$\begin{aligned}
0 &= \frac{\partial F}{\partial \tau_t} + \frac{\partial F}{\partial \tau_{t+1}} \frac{d\tau_{t+1}}{d\tau_t} + \beta W_d'(e_t) \frac{de_t}{d\tau_t} \\
0 &= w\left(\left(A - \frac{\beta z \alpha_1 w}{2+\beta z \alpha_1 w}\right) e_t - \frac{2(A-1)w}{(2+\beta z \alpha_1 w)^2}\tau_t + z(A-1)e_{t-1}\right) \\
&\quad - w^2 \frac{(1+\beta z)(2+A\beta z \alpha_1 w) + 2\beta z \alpha_0 (A-1)}{(2+\beta z \alpha_1 w)^2}
\end{aligned}$$

Using the fact that, $e_t = \frac{(1+\beta z(1-\alpha_0))w}{2+\beta z \alpha_1 w} - \frac{w}{2+\beta z \alpha_1 w}\tau_t$ and the guess $\tau_t = \alpha_0 + \alpha_1 e_{t-1}$, dividing by w and collecting terms, this yields

$$\begin{aligned}
0 &= \left(z(A-1) - \left(\frac{2A}{(2+\beta z \alpha_1 w)} + (A-1)\right)\frac{w\alpha_1}{(2+\beta z \alpha_1 w)}\right) e_{t-1} \\
&\quad + \frac{w(1+\beta z)}{2+\beta z \alpha_1 w}\left(\frac{2A(1-\alpha_0)}{2+\beta z \alpha_1 w} - (1+\alpha_0(A-1))\right)
\end{aligned}$$

In order for this condition to be satisfied for all e_{t-1} we need,

$$\begin{aligned}
0 &= f(\alpha_1, \beta, z) = z(A-1) - \left(\frac{2A}{(2+\beta z \alpha_1 w)} + (A-1)\right) \\
&\quad \cdot \frac{w\alpha_1}{2+\beta z \alpha_1 w} \qquad (40) \\
0 &= g(\alpha_0, \alpha_1, \beta, z) = \frac{2A(1-\alpha_0)}{2+\beta z \alpha_1 w} - (1+\alpha_0(A-1)) \qquad (41)
\end{aligned}$$

A solution for these equations (ignoring the roots that would generate instability) is: :

$$\begin{aligned}
\alpha_1 &= \frac{\sqrt{1+4A(A-1)(1-\beta z^2)} - \left(1+2\left(1-\beta z^2\right)(A-1)\right)}{\beta z (A-1)(1-\beta z^2) w} \geq 0 \\
\alpha_0 &= \frac{2(A-1) - \beta z \alpha_1 w}{2+(A-1)(4+\beta z \alpha_1 w)} \\
&= \frac{2A(A-1)\left(1-\beta z^2\right) - \left(\sqrt{1+4A(A-1)(1-\beta z^2)} - 1\right)}{(A-1)\left(2A(1-\beta z^2) + \left(\sqrt{1+4A(A-1)(1-\beta z^2)} - 1\right)\right)} \geq 0
\end{aligned}$$

The non-negativity of α_0 and α_1 are established by standard algebra, since, in both the expressions, the numerator and denominators are both positive.

Before concluding that our solution to the first order condition is optimal, we must consider the constraints $\tau_t \in [0,1]$. First, we note that since $\alpha_0, \alpha_1 \geq 0$, the constraint $\tau_{t+1} \geq 0$, cannot bind for any choice of τ_t. Second, since $\alpha_0, \alpha_1 \geq 0$ and $I(\tau_t)$ is increasing, the highest possible τ_{t+1} arise if $\tau_t = 0$. In this case,

$$\begin{aligned}
\tau_{t+1} &= T(I(0)) = \alpha_0 + \alpha_1 \frac{(1+\beta z(1-\alpha_0))w}{2+\beta z \alpha_1 w} = 1 - \frac{2(1-\alpha_0) - \alpha_1 w}{2+\beta z \alpha_1 w} \\
&= 1 - 2 \cdot \\
&\quad \frac{\left(z + A\left(1-\beta z^2\right)\right)\left(\sqrt{1+4A(A-1)(1-\beta z^2)} - 1\right) - 2Az(A-1)\left(1-\beta z^2\right)}{\left(\sqrt{1+4A(A-1)(1-\beta z^2)} - 1\right)\left(2A(1-\beta z^2) + \left(\sqrt{1+4A(A-1)(1-\beta z^2)} - 1\right)\right)} \\
&\equiv \tau_m
\end{aligned}$$

Now, we note that the numerator in the expression above is zero if $A \in \left\{0, 1, \frac{z(z+1)}{(1+\beta)z^2-1}\right\}$,while the denominator is positive for all $A > 1$. To see the former, write

$$\left(z + A\left(1-\beta z^2\right)\right)\left(\sqrt{1+4A(A-1)(1-\beta z^2)}-1\right) = 2Az(A-1)\left(1-\beta z^2\right) \rightarrow$$

$$\sqrt{1+4A(A-1)(1-\beta z^2)} = 1 + \frac{2Az(A-1)\left(1-\beta z^2\right)}{(z+A(1-\beta z^2))}$$

$$0 = \left(\sqrt{1+4A(A-1)(1-\beta z^2)}\right)^2 - \left(1 + \frac{2Az(A-1)\left(1-\beta z^2\right)}{(z+A(1-\beta z^2))}\right)^2 \rightarrow$$

$$\begin{aligned} 0 &= 1 + 4A(A-1)\left(1-\beta z^2\right) \\ &\quad - \frac{\left(A\beta z^2 - A - 2A^2 z + 2Az - z - 2A\beta z^3 + 2A^2\beta z^3\right)^2}{(z+A(1-\beta z^2))^2} \end{aligned}$$

$$0 = 4A^2(A-1)\left(1-\beta z^2\right)^2 \frac{A\left(\beta z^2 - 1 + z^2\right) - z(1+z)}{(z+A(1-\beta z^2))^2}.$$

Furthermore, since

$$\begin{aligned} \lim_{A\to 1}(\tau_m) &= 0 \\ \lim_{A\to 1}\left(\frac{d\tau_m}{dA}\right) &= 1 + z > 0, \end{aligned}$$

$\tau_m \leq 1$ for

$$1 \leq A \leq \frac{z(z+1)}{(1+\beta)z^2 - 1}, \tag{42}$$

provided $(1+\beta)z^2 > 1$. If, conversely, $(1+\beta)z^2 - 1 \leq 0$, $\tau_m \leq 1$ for any $A \geq 1$. To summarize, we have now shown that for any (out of equilibrium) $\tau_t \in [0,1]$, the constraint $\tau_{t+1} \in [0,1]$ is not binding along the equilibrium path, provided (42) or $(1+\beta)z^2 - 1 \leq 0$ is satisfied. At time 0, the constraint $\tau_a \leq 1$ binds, iff $e_{-1} > \frac{1-\alpha_0}{\alpha_1}$.

We show, next that

$$\frac{\partial \alpha_1}{\partial A} \geq 0 \text{ and } \frac{\partial \alpha_0}{\partial A} \geq 0.$$

To this aim, note first that

$$\frac{\partial \alpha_1}{\partial A} = \frac{\sqrt{1+4A(A-1)(1-\beta z^2)} - \left(1 + 2(A-1)\left(1-\beta z^2\right)\right)}{A\beta z(A-1)(1-\beta z^2)w}$$

where the denominator is positive. The numerator is also positive, since

$$1 + 4A(A-1)\left(1-\beta z^2\right) > \left(1 + 2(A-1)\left(1-\beta z^2\right)\right)^2,$$

which is equivalent to

$$4\beta z^2(A-1)^2\left(1-\beta z^2\right) \geq 0$$

which is always true (with the equality being strict as long as $A > 1$).

Next, observe that

$$\frac{\partial \alpha_0}{\partial A} = \frac{2\Delta_1}{\sqrt{1+4A(A-1)(1-\beta z^2)}\left((A-1)\left(1-2A(1-\beta z^2)-\sqrt{1+4A(A-1)(1-\beta z^2)}\right)\right)^2}$$

where

$$\begin{aligned}\Delta_1 &= 2A(A-1)\left(1-\beta z^2\right)\left(A\left(1-\beta z^2\right)-1\right) \\ &\quad +\left(1+A\left(1-\beta z^2\right)(A-2)\right)\left(\sqrt{1+4A(A-1)(1-\beta z^2)}-1\right),\end{aligned}$$

and

$$\lim_{A\to 1}\frac{\partial \alpha_0}{\partial A} = 1-\beta z^2 > 0.$$

Clearly, $\frac{\partial \alpha_0}{\partial A}$ has the same sign as Δ_1.Furthermore, at $A=1, \Delta_1 = 0$ and the derivative with respect to A is

$$\frac{\partial \Delta_1}{\partial A} = \frac{2\left(1-\beta z^2\right)(3A-2)\Delta_2}{\sqrt{1+4A(A-1)(1-\beta z^2)}}$$

where

$$\Delta_2 = 2A\left(1-\beta z^2\right)(A-1)+\left(A\left(1-\beta z^2\right)-1\right)\left(\sqrt{1+4A(A-1)(1-\beta z^2)}-1\right).$$

Clearly, $\frac{\partial \Delta_1}{\partial A}$ has the same sign as Δ_2. Finally, Δ_2 evaluated at $A=1$ equals 0 and

$$\frac{\partial \Delta_2}{\partial A} = \frac{2A\left(1-\beta z^2\right)^2(4A-3)+\left(1-\beta z^2\right)(4A-3)\left(\sqrt{1+4A(A-1)(1-\beta z^2)}-1\right)}{\sqrt{1+4A(A-1)(1-\beta z^2)}} > 0.$$

We have thus established that Δ_2 is strictly positive for $A>1$. Then, in the same range,

$$\frac{\partial \Delta_1}{\partial A} = \frac{2\left(1-\beta z^2\right)(3A-2)\Delta_2}{\sqrt{1+4A(A-1)(1-\beta z^2)}} > 0,$$

and since Δ_1 is zero at $A=1$, also Δ_1 is strictly positive. This establishes that $\frac{\partial \alpha_0}{\partial A} \geq 0$, with the inequality being strict for $A>1$.

$T(e_{t-1}) \in (0,1)$, which in turn implies that neither the constraint $\tau_t \in [0,1]$ nor the constraint $e_t > 0$ are ever binding.

We now establish that $-z_d < z$. To establish the claim, note that

$$\frac{\partial z_d}{\partial A} = -\frac{2\left(1-\beta z^2\right)\left(\sqrt{1+4A(A-1)(1-\beta z^2)}-\left(1+2(A-1)\left(1-\beta z^2\right)\right)\right)}{\left(\sqrt{1+4A(A-1)(1-\beta z^2)}-1\right)^2\beta z\sqrt{1+4A(A-1)(1-\beta z^2)}} < 0,$$

implying, by the continuity of z_d, that

$$-z_d < -\lim_{A\to\infty}(z_d) = \frac{1-\sqrt{1-\beta z^2}}{\beta z} < z.$$

Next, we establish that $\frac{\partial \bar{\tau}}{\partial A} > 0$. Note that

$$\frac{\partial \bar{\tau}}{\partial A} = X \cdot (J(A)+M(A))$$

where

$$\begin{aligned} J(A) &= \left(1 + A(A-2)\left(1-\beta z^2\right)\right)\sqrt{1+4A(A-1)(1-\beta z^2)} > 0 \\ M(A) &= (A-1)^2\left(2A\left(1-2\beta z^2\right)-1\right) + \beta A^2 z^2\left(2(A-1)\beta z^2 - 1\right) \end{aligned}$$

and

$$\begin{aligned} 0 &< X \equiv \frac{8\left(1-\beta z^2\right)\beta^2 z^2 (1+z)}{\left(2A(1-\beta z^2) + \sqrt{1+4A(A-1)(1-\beta z^2)} - 1\right)^2 \sqrt{1+4A(A-1)(1-\beta z^2)}} \cdot \\ &\cdot \left(\sqrt{1+4A(A-1)(1-\beta z^2)}(1+\beta z) - \beta z(1+z) + (1-2A)\left(1-\beta z^2\right)\right)^{-2} \end{aligned}$$

We show that

$$|J(A)| > |M(A)|.$$

This follow from the fact that

$$(J(A))^2 - (M(A))^2 = 4A^4\beta z^2 (A-1)^2 \left(1-\beta z^2\right)^3 > 0.$$

Since $J(A) > 0$, this establishes that $\frac{\partial \bar{\tau}}{\partial A} > 0$.

Finally, we establish that $\bar{\tau} > \tau^*$ for $A > 1$ and $\beta z > 0$.First, note that

$$\bar{\tau} - \tau^* = \frac{A}{(2A-1)}\frac{N_1(A)}{N_2(A)}$$

where

$$\begin{aligned} N_1(A) &= 2A(1+z)\left(1-\beta z^2\right)(A-1) \\ &\quad - \left(A\left(1-\beta z^2\right) + \left(z+\beta z^2\right)\right)\left(\sqrt{1+4A(1-\beta z^2)(A-1)} - 1\right), \\ N_2(A) &= 2A(1+z)\left(1-\beta z^2\right)(A-1) \\ &\quad - \left((1+z) - A\left(1-\beta z^2\right)\right)\left(\sqrt{1+4A(1-\beta z^2)(A-1)} - 1\right). \end{aligned}$$

We now establish that, for all $A > 1$, $\frac{N_1(A)}{N_2(A)} \neq 0$. Since $\frac{N_1(A)}{N_2(A)}$ is a continuous function of A, this implies that, in the relevant range where $A > 1$, the difference $\bar{\tau} - \tau^*$ is either always positive or always negative. Then, we show that $\bar{\tau} - \tau^*$ is zero when $A = 1$ and has a strictly positive limit as A tends to infinity. Then, $\bar{\tau} - \tau^* > 0$, and the result is established.

We find, first, all levels of A such that $N_1(A) = 0$. Rearranging terms in the expression of $N_1(A)$ yields

$$\frac{2A(1+z)\left(1-\beta z^2\right)(A-1)}{\left(A(1-\beta z^2) + (z+\beta z^2)\right)} + 1 = \sqrt{1+4A(1-\beta z^2)(A-1)} \tag{43}$$

Squaring terms on both sides of (43), and rearranging terms, yields:

$$0 = 4A(A-1)^2\left(1-\beta z^2\right)^2 z\frac{A(2+(1+\beta)z) - (1+\beta z)}{\left(A(1-\beta z^2) + (z+\beta z^2)\right)^2}. \tag{44}$$

The roots of this expression (and of (43)) are $A \in \left\{0, \frac{1+\beta z}{2+(1+\beta)z}, 1\right\}$, all being in the range $[0,1]$.

Considering, then, the denominator. The equation $N_2(A)=0$ can be rewritten as

$$\frac{2A(1+z)\left(1-\beta z^2\right)(A-1)}{(1+z)-A\left(1-\beta z^2\right)}+1=\sqrt{1+4A\left(1-\beta z^2\right)(A-1)}. \tag{45}$$

Again squaring both sides, and rearranging terms, yields:

$$4A^2(A-1)\left(1-\beta z^2\right)^2 z\frac{A(2+z(1+\beta))-(1+z)}{\left(1+z-A+A\beta z^2\right)^2}=0$$

with roots, $A\in\left\{0,\frac{1+z}{2+(1+\beta)z},1\right\}$, that also solve (45) and are all in the range $[0,1]$. Therefore, we have established that the ratio $\frac{N_1(A)}{N_2(A)}$ is well-defined for all $A>1$ and, moreover, $\frac{N_1(A)}{N_2(A)}\neq 0$.

Finally, we note that[12]

$$\begin{aligned}\lim_{A\to 1}\left(\bar{\tau}-\tau^*\right) &= 0\\ \lim_{A\to\infty}\left(\bar{\tau}-\tau^*\right) &= \frac{\beta z\sqrt{1-\beta z^2}\left(1+z-\sqrt{1-\beta z^2}\right)}{2\left(1+\beta z-\sqrt{1-\beta z^2}\right)\left(1-\beta z^2+\sqrt{1-\beta z^2}\right)}>0.\end{aligned}$$

This establishes, by the intermediate value theorem, the result that $\bar{\tau}>\tau^*$ for all $A>1$.

Let us also check that the requirement for $\bar{\tau}>\frac{1}{2}$, i.e., $A>1+\frac{2+z(1-\beta z)}{z(2+z(1+\beta))}$ is consistent with $A\leq\frac{z(z+1)}{(1+\beta)z^2-1}$ or $(1+\beta)z^2\leq 1$. In the latter case, the two conditions are obviously consistent. In the former, we require the set $A\in\left(1+\frac{2+z(1-\beta z)}{z(2+z(1+\beta))},\frac{z(z+1)}{(1+\beta)z^2-1}\right]$ to be non-empty. This requires

$$\begin{aligned}&\frac{z(z+1)}{(1+\beta)z^2-1}-1-\frac{2+z(1-\beta z)}{z(2+z(1+\beta))}\\ =\;&\frac{(1+z)}{\left((1+\beta)z^2-1\right)}\frac{\left(2\left(1-\beta z^2\right)+z\right)}{z(2+z(1+\beta))}\geq 0,\end{aligned}$$

which is satisfied whenever $(1+\beta)z^2\geq 1$.

8.3 Proof of equation (23)

$$\begin{aligned}y_t &= ze_{t-1}+e_t\\ &= \frac{w^2}{2}(1+\beta z)(1+z)-\frac{w^2}{2}\left(z\frac{\tau(1+z_d)-\tau_t}{z_d}+\left(1+\beta z^2\right)\tau_t+\beta z\left(\bar{\tau}-z_d\left(\tau_t-\bar{\tau}\right)\right)\right)\\ &= +\frac{1}{2}w^2(1+\beta z)(1+z)-\frac{1}{2}w^2\left(z\frac{1+z_d}{z_d}+\beta z(1+z_d)\right)\tau-\frac{1}{2}w^2\left(-\frac{z}{z_d}+1+\beta z^2-z_d\beta z\right)\tau_t\\ &= \frac{1}{2}w^2\left((1+\beta z)(1+z)-\frac{(1+z_d)z}{z_d}(1+\beta z_d)\bar{\tau}+(1-z_d\beta z)\frac{z-z_d}{z_d}\bar{\tau}+(1-z_d\beta z)\frac{z-z_d}{z_d}\left(\tau_t-\bar{\tau}\right)\right)\\ &= \frac{1}{2}w^2\left((1+\beta z)(1+z)(1-\bar{\tau})+(1-z_d\beta z)\frac{z-z_d}{z_d}\left(\tau_t-\bar{\tau}\right)\right)\end{aligned}$$

[12]The range for $\lim_{A\to\infty}\left(\bar{\tau}-\frac{A-1}{2A-1}\right)$ is derived by noting that this expression increases in β and z and thus finding the limit as $\beta,z\to 0$ and $\beta,z\to 1$.

8.4 Details of claims in section 5

The first order conditions in period zero are

$$\begin{aligned}
\tau_0; \frac{\partial \tilde{Y}\left(\tau_0, \tau_l, \tau_h\right)}{\partial \tau_0} &= 0, \\
\tau_{l,1}; \frac{\partial \tilde{Y}\left(\tau_0, \tau_{l,1}, \tau_{h,1}\right)}{\partial \tau_{l,1}} + \beta\left(1-p\right) V_l'\left(\tau_{l,1}\right) &= 0 \\
\frac{1}{2}\beta w z\left(1-p\right)\left(2\left(A_l-1\right)e_0 - w\left(A_l\tau_0 + \beta z E_{A\tau}\right)\right) + \beta\left(1-p\right)V_l'\left(\tau_{l,1}\right) &= 0, \\
\tau_h; \frac{\partial \tilde{Y}\left(\tau_0, \tau_{l,1}, \tau_{h,1}\right)}{\partial \tau_h} + \beta p V_h'\left(\tau_{h,1}\right) &= 0, \\
\frac{1}{2}\beta w z p\left(2\left(A_h-1\right)e_0 - w\left(A_l\tau_0 + \beta z \tau_1^e\right)\right) + \beta p V_h'\left(\tau_{h,1}\right) &= 0.
\end{aligned}$$

where $E_{A\tau} = pA_h\tau_{h,1} + \left(1-p\right)A_l\tau_{l,1}$. Solving for τ_0 yields,

$$\tau_0 = \left(1+\beta z\right)\frac{A_l-1}{2A_l-1} - \beta p z\frac{\left(A_h+A_l-1\right)}{\left(2A_l-1\right)}\tau_{h,1} - \left(1-p\right)\beta z\tau_{l,1}.$$

From the proof of Proposition 1) we have

$$\begin{aligned}
V_h'\left(\tau_h\right) &= \frac{w^2}{4}\left(B_{1h} + 2B_{2h}\tau_{h,1}\right) \\
&= \frac{w^2}{4}\left(2\left(1-\beta z^2\right)\left(A_h-1\right) - 2\left(1-\beta z^2\right)\left(2A_h-1\right)\tau_{h,1}\right) \\
V_l'\left(\tau_l\right) &= \frac{w^2}{4}\left(B_{1l} + 2B_{2l}\tau_{l,1}\right) \\
&= \frac{w^2}{4}\left(2\left(1-\beta z^2\right)\left(A_l-1\right) - 2\left(1-\beta z^2\right)\left(2A_l-1\right)\tau_{l,1}\right)
\end{aligned}$$

Rearranging terms and simplifying yields

$$\begin{aligned}
\tau_0 &= \left(A_l-1\right)\frac{1+\beta z}{2A_l-1} - p\beta z\frac{A_h+A_l-1}{2A_l-1}\tau_{h,1} - \left(1-p\right)\beta z\tau_{l,1}, \\
0 &= -\frac{1}{2}\beta w^2\left(2A_h-1\right)p\left(1-\left(1-p\right)\beta z^2\right)\tau_{h,1} \\
&\quad -\frac{1}{2}\beta^2 w^2 z^2 p\left(1-p\right)\left(A_h+A_l-1\right)\tau_{l,1} \\
&\quad -\frac{1}{2}\beta w^2 z p\left(A_h+A_l-1\right)\tau_0 \\
&\quad +\frac{1}{2}\beta w^2\left(A_h-1\right)p\left(1+z\right), \\
0 &= -\frac{1}{2}\beta^2 w^2 z^2\left(1-p\right)p\left(A_h+A_l-1\right)\tau_{h,1} \\
&\quad -\frac{1}{2}\beta w^2\left(2A_l-1\right)\left(1-p\right)\left(1-p\beta z^2\right)\tau_{l,1} \\
&\quad -\frac{1}{2}\beta w^2 z\left(1-p\right)\left(2A_l-1\right)\tau_0 \\
&\quad +\frac{1}{2}\beta w^2\left(A_l-1\right)\left(1-p\right)\left(1+z\right).
\end{aligned}$$

This is a system of linear equations with solutions given in the text.

When the government can save, the first order conditions in period 0 are

$$\begin{aligned}
\tau_0; \frac{\partial \hat{Y}(\tau_0, \tau_l, \tau_h)}{\partial \tau_0} &= 0 \\
\tau_l; \frac{\partial \hat{Y}(\tau_0, \tau_l, \tau_h)}{\partial \tau_l} + \beta(1-p) V_l'(\tau_l) &= 0, \\
\tau_h; \frac{\partial \hat{Y}(\tau_0, \tau_l, \tau_h)}{\partial \tau_h} + \beta p V_h'(\tau_h) &= 0,
\end{aligned}$$

Using the functional expressions this is

$$\begin{aligned}
\tau_0 &= (1+\beta z)\frac{E_A - 1}{2E_A - 1} - \beta z p \frac{A_h + E_A - 1}{2E_A - 1}\tau_h - \beta z(1-p)\frac{A_l + E_A - 1}{2E_A - 1}\tau_{l,1} \\
0 &= \frac{1}{2}\beta w z(1-p)\left((A_l - 1)2e_0 - w\left(E_A\tau_0 + \beta z\left(pA_h\tau_{h,1} + (1-p)A_l\tau_{l,1}\right)\right)\right) \\
&\quad + \beta(1-p)\frac{w^2}{4}\left(2\left(1-\beta z^2\right)(A_l - 1) - 2\left(1-\beta z^2\right)(2A_l - 1)\tau_{l,1}\right) \\
0 &= \frac{1}{2}\beta w z p\left(2(A_h - 1)e_0 - w\left(E_A\tau_0 + \beta z\left(pA_h\tau_h + (1-p)A_l\tau_{l,1}\right)\right)\right) \\
&\quad + \beta p\frac{w^2}{4}\left(2\left(1-\beta z^2\right)(A_h - 1) - 2\left(1-\beta z^2\right)(2A_h - 1)\tau_{h,1}\right)
\end{aligned}$$

yielding the system in the text.