Time-Consistent Consumption Taxation*

Sarolta Laczó[†] Raffaele Rossi[‡]

January 2015

Abstract

We characterise optimal fiscal policies in a Real Business Cycle model when the government has access to consumption taxation but cannot credibly commit to future policies. Contrary to the case where only labour and capital income are taxed, the optimal time-consistent policies are remarkably similar to their Ramsey counterparts, as long as the capital income tax causes some distortion within the period. The welfare gains from commitment are negligible, while they are substantial without consumption taxation. Further, the welfare gains from taxing consumption are much higher without commitment. These results suggest that the policy-maker's ability to commit is of secondary importance if consumption is taxed optimally.

JEL classification: E62, H21.

Keywords: fiscal policy, Markov-perfect policies, consumption taxation, variable capital utilisation, endogenous government spending.

*This paper has greatly benefited from discussions with Charles Brendon, Davide Debortoli, Andrea Lanteri, Campbell Leith, Yang K. Lu, Albert Marcet, Ricardo Nunes, Evi Pappa, Vito Polito, and Maurizio Zanardi. We also thank seminar participants at CREI/UPF, IAE/UAB, Bank of England, University of York, University of Manchester, and Lancaster University, and participants of the Max Weber June Conference in Florence, the Barcelona GSE Summer Forum 'Macro and Micro Perspectives on Taxation,' CEF in Oslo, PET in Seattle, EEA in Toulouse, MMF in Durham, and Barcelona GSE Winter Workshops for useful comments and suggestions. We received funding from the Spanish Ministry of Science and Innovation under grant ECO2008-04785 and the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 612796. Laczó also acknowledges funding from the JAE-Doc programme co-financed by the European Social Fund. We thank the Institut d'Anàlisi Econòmica (IAE-CSIC) for hospitality while advancing this project. All errors are our own. The views expressed in this paper are not those of the Bank of England.

[†]Bank of England, Threadneedle Street, EC2R 8AH, London, United Kingdom. Email: sarolta.laczo@gmail.com.

[‡]Department of Economics, Lancaster University Management School, LA1 4XY, Lancaster, United Kingdom. Email: r.rossi@lancaster.ac.uk.

1 Introduction

Most of the literature on optimal fiscal policy rules out consumption taxation, a policy instrument which is used in most industrialised economies. For example, as of January 1, 2015, the value-added tax on standard items ranges from 17 to 27 percent in European Union countries. Papers which consider consumption taxation include Coleman (2000), who finds, under the assumption that the fiscal authority can fully commit to future policies, that replacing income taxes with consumption taxes would lead to large welfare gains in the United States. Correia (2010) extends this result to a heterogeneous-agents framework. Two recent contributions highlight the role of consumption taxation as a tool to relax a constraint of the monetary authority on the nominal interest rate, either as a result of the zero lower bound (Correia, Farhi, Nicolini, and Teles, 2013) or in a monetary union (Farhi, Gopinath, and Itskhoki, 2014).

This paper finds a new benefit of consumption taxation: discretionary policies and the resulting allocations are almost identical to those under Ramsey policy. This holds for the steady state, for policy dynamics in a deterministic framework, and for the cyclical properties of tax rates and allocations in a stochastic environment. A necessary condition for these findings is that capital income taxation causes some distortion within the period. This result means that the policy-maker's ability to commit is of secondary importance as long as consumption can be taxed. In other words, the negative effects of policy-makers' lack of credibility, due to political business cycles, political disagreement, default on past promises, etc., can be overcome by taxing consumption optimally.

The novelty of our paper lies in analysing time-consistent fiscal policies when the policy-maker has access to consumption taxation, in addition to labour and capital income taxation. The existing literature on Markov-perfect policies has not considered consumption taxation, to our knowledge. We are able to quantify the welfare gains from taxing consumption and from commitment, as well as the potential welfare gains from implementing optimal time-consistent policies compared to the existing tax system in the United States.

The model economy we consider is a standard Real Business Cycle (RBC) model with endogenous labour supply and variable capital utilisation (Greenwood, Hercowitz, and Huffman, 1988, and many others). The government spends on public goods which households value, has access to three types of taxes, capital and labour income taxes and consumption taxes, but no lump-sum taxes, and has to balance its budget.

Our paper is at the intersection of two strands of the optimal taxation literature: (i) Ramsey policies with consumption taxation, such as Coleman (2000), who quantifies the welfare gains from taxing consumption; and (ii) time-consistent policies without consumption taxation. The latter strand of the literature finds that lack of commitment alters greatly the characteristics of optimal policies and the resulting allocations. In a framework similar to ours but with capital fully utilised, Klein, Krusell, and Ríos-Rull (2008) find that when the only tax base available to the government is capital income, which in their Markov-perfect equilibrium is an inelastic source of fiscal revenues, the policy-maker sets the tax rate below the confiscatory level. Martin (2010) studies a Markov-perfect equilibrium in which the fiscal authority can simultaneously tax labour and capital income. He finds that the optimal policy calls for taxing capital and subsidising labour when capital is fully utilised. With endogenous capital utilisation, as in the present paper, he finds that optimal time-consistent taxation involves almost equal tax rates on capital and labour income, which are close to the existing tax rates in the United States. Debortoli and Nunes (2010) also establish the same quantitative result.

Our results can be summarised as follows. If the policy-maker has access to all three types of taxes and the tax rates are unrestricted, the first-best allocation can be implemented at the steady state under full commitment, as long as private consumption is larger than labour income. Given that the Ramsey policy achieves the first best, it is also time-consistent. In other words, the solutions under Ramsey and Markov-perfect policy-making coincide at the steady state. However, the tax policies include a large labour subsidy.

Then, we study the case where the government is prohibited from subsidising labour income, as in Coleman (2000) and Correia (2010), because large labour subsidies are not observed in real economies and would likely lead to misreporting of hours. In this case, the Ramsey policy-maker taxes consumption at 22.3 percent at the steady state in our baseline calibration, and sets labour and capital income taxes to zero. The time-consistent policy-maker finances government spending mainly from taxing consumption as well, taxing it at 22.1 percent in our baseline calibration, and sets the capital income tax to 0.4 percent and the labour income tax to zero. In order to compare our results with the literature, e.g. Martin (2010) and Debortoli and Nunes (2010), we also analyse a scenario in which the policy-maker can only tax income deriving from labour and capital. In this case, the time-consistent policy-maker sets the labour income tax to 6.5 percent and the capital income tax to 19.8 percent at the steady state.

The intuition behind these results is the following. First, consider the case without consumption taxation. With only labour and capital income taxation, the Ramsey planner initially taxes capital at a high rate. Then the capital tax rate gradually approaches zero, while the labour income tax rate increases over time. The downward trend in the capital income tax induces households to continuously postpone their consumption, while a labour tax hike reduces labour supply contemporaneously, but raises it in any previous period. The time-consistent policy-maker does not internalise the effects of taxes on private sector decisions in earlier periods. This leads to dramatic differences of policies and allocations between Ramsey and Markov-perfect governments.

On the contrary, with access to consumption taxation, a downward trend in the capital tax would require an upward trend in the consumption tax to satisfy the government's budget constraint. This would counteract the saving incentive and lead to inefficiently low capital accumulation. As a result, the capital income tax rate is low from the start, and the consumption tax rate hardly varies over time or with the level of capital. Thus the time-inconsistency features of policies under commitment are negligible when consumption is taxed optimally. This is the key economic mechanism that drives the close similarity between Ramsey and Markov-perfect policy equilibria. Note also that while both taxes cause an intratemporal distortion, the capital income tax causes an intertemporal distortion as well in the long run, since it distorts the Euler equation, while the consumption tax does not. Therefore, taxing consumption turns out to be the less distortive way to raise fiscal revenue.

In terms of welfare-equivalent consumption, the welfare gains from taxing consumption are 2.77 (1.21) percent in the case of a Markov (Ramsey) policy-maker.[1] This means that taxing consumption generates much larger welfare gains under discretion than under commitment. The gains from commitment are negligible with consumption taxation (0.0003 percent), while they are substantial (2.01 percent) without. We also compute the welfare gain from adopting the time-consistent tax system with consumption taxation proposed here over the existing one in the United States. The welfare increase is equivalent to permanently increasing private consumption by 7.744 percent. Without access to consumption taxation, the welfare gains are 4.92 percent. In the case of a Ramsey policy-maker, the welfare gains are 7.745 and 7.06 percent with and without access to consumption taxation, respectively. Remarkably, we find higher levels of welfare under discretion when the policy-maker has access to consumption taxation than under commitment when the government can tax only labour and capital income.

[1] We have computed these welfare gains in welfare-equivalent consumption units in the standard way, taking into account transitions. In particular, we ask by how much private consumption would have to increase at the Markov (Ramsey) steady state without consumption taxation, keeping leisure and public consumption fixed, to yield the same lifetime utility to the household as the lifetime utility it gets from the Markov (Ramsey) optimal policy with consumption taxation, starting with the level of capital of the Markov (Ramsey) steady state without consumption taxation. We have computed other welfare gains mentioned in this paragraph in a similar fashion.

Finally, we analyse policies over the business cycle when the economy is hit by aggregate productivity shocks. We find that with access to consumption taxation also the cyclical properties of tax rates and allocations under a Ramsey and a time-consistent policy-maker are remarkably similar. Further, private consumption, hours, public goods, and output all vary less over the business cycle with consumption taxation than without. Hence, the policy-maker can better stabilise the economy taxing consumption, which yields additional welfare benefits.

The rest of this paper is structured as follows. Section 2 details the economic environment. Section 3 sets up the fiscal policy problems, both (i) the Ramsey problem and (ii) the problem of a Markov/time-consistent policy-maker. Afterwards, it characterises the solutions and presents some analytical results. Section 4 contains our baseline quantitative results, both in a deterministic and in a stochastic environment, as well as robustness checks. Section 5 concludes.

2 The model

We consider a discrete-time Real Business Cycle (RBC) model with a representative household, a representative and perfectly-competitive firm, and a benevolent policy-maker. The household decides on consumption and leisure and chooses a capital utilisation rate as well. The firm maximises profits and uses capital services and labour as production inputs. The policy-maker spends on public goods which yield utility to households, and raises revenue via linear taxes on labour and capital income as well as on consumption. Lump-sum taxes are not available. The government balances its budget in each period.

A few comments are in order about our main assumptions before describing the economic environment in detail. First, endogenous capital utilisation (see Greenwood, Hercowitz, and Huffman, 1988, Greenwood, Hercowitz, and Krusell, 2000, and many others) implies that taxing capital is distortionary in the current period. This assumption is important to create interesting and relevant trade-offs between the different tax instruments, see also Martin (2010) and Debortoli and Nunes (2010). Further, it is an assumption which is in line with the fact that capital is not fully utilised in reality and is standard in bigger Dynamic Stochastic General Equilibrium (DSGE) models. Second, we follow the literature on Markov-perfect policies (Klein, Krusell, and Ríos-Rull, 2008, and others) and consider government spending to be a choice variable of the fiscal authority. At least part of government consumption can be adjusted in response to changes in the economy, and we are interested in studying the optimal policy mix both on the revenue and the spending side. However, we perform our analysis of comparing policy scenarios with exogenous government spending as well as a robustness check, see Section 4.5.2. Finally, we exclude public debt dynamics for two main reasons. First, again, we wish to compare our findings with previous studies on time-consistent fiscal policies with capital accumulation, which also impose a balanced-budget requirement (Klein and Ríos-Rull, 2003; Ortigueira, 2006; Klein, Krusell, and Ríos-Rull, 2008; Azzimonti, Sarte, and Soares, 2009; Martin, 2010; Debortoli and Nunes, 2010). Second, in an RBC model with variable capital utilisation, Ramsey policies can be made time consistent through public debt restructuring (Zhu, 1995).[2] Excluding public debt in models with capital accumulation increases the time-inconsistency of Ramsey taxation, and allows us to better isolate the role of consumption taxation.

Let a_t denote the level of aggregate productivity at time t, and let $\mathbf{a}^t = \{a_1, a_2, ..., a_t\}$ denote the history of aggregate productivity realisations. The representative household takes prices and policies as given and seeks to maximise

$$\mathbb{E}_0 \left[\sum_{t=1}^{\infty} \beta^t u \left(c \left(\mathbf{a}^t \right), \ell \left(\mathbf{a}^t \right), g \left(\mathbf{a}^t \right) \right) \right], \tag{1}$$

where $\mathbb{E}_0$ represents the rational-expectations operator at time 0, $\beta \in (0, 1)$ is the discount factor, $c(\mathbf{a}^t)$ is private consumption when history $\mathbf{a}^t$ has occurred, $\ell(\mathbf{a}^t)$ represents leisure, and $g(\mathbf{a}^t)$ is public consumption; subject to the time constraint

$$h \left(\mathbf{a}^t \right) + \ell \left(\mathbf{a}^t \right) = 1, \ \forall \mathbf{a}^t, \tag{2}$$

where $h(\mathbf{a}^t)$ represents hours worked given history $\mathbf{a}^t$, and the budget constraint

$$\begin{aligned} &\left(1 + \tau^c \left(\mathbf{a}^t \right)\right) c \left(\mathbf{a}^t \right) + k \left(\mathbf{a}^t \right) = \left(1 - \tau^k \left(\mathbf{a}^t \right)\right) r \left(\mathbf{a}^t \right) v \left(\mathbf{a}^t \right) k \left(\mathbf{a}^{t-1} \right) \\ &+ \left(1 - \tau^h \left(\mathbf{a}^t \right)\right) w \left(\mathbf{a}^t \right) h \left(\mathbf{a}^t \right) + \left(1 - \delta \left(v \left(\mathbf{a}^t \right) \right)\right) k \left(\mathbf{a}^{t-1} \right), \ \forall \mathbf{a}^t, \end{aligned} \tag{3}$$

where $k(\mathbf{a}^{t-1})$ is the level of the capital stock at the beginning of the period, $v(\mathbf{a}^t) \in (0, 1]$ is the capital utilisation rate, $\delta(v(\mathbf{a}^t))$ represents the depreciation rate as a function of capital utilisation, and $\tau^c(\mathbf{a}^t)$, $\tau^h(\mathbf{a}^t)$, and $\tau^k(\mathbf{a}^t)$ denote the consumption, the labour income, and the capital income tax rate, respectively, given history $\mathbf{a}^t$. Finally, the variables $r(\mathbf{a}^t)$ and $w(\mathbf{a}^t)$ are the interest rate and the wage rate, respectively, and represent the remuneration of production factors, namely, capital services and labour. The utility function $u()$ is assumed to be twice continuously differentiable in all three of its arguments with partial derivatives

[2] Domínguez (2007) establishes a similar result when there are delays in tax policy implementation, as in Klein and Ríos-Rull (2003).

$u_c > 0$, $u_{cc} < 0$, $u_\ell > 0$, $u_{\ell\ell} < 0$, $u_g > 0$, $u_{gg} < 0$, where u_x and u_{xx} denote, respectively, the first and the second derivative of the utility function with respect to the variable x.

Note that it is crucial that at least some true economic depreciation is not tax deductible, else the current government would view the current capital tax as non-distortionary, and the policy problem of the optimal time-consistent tax mix would reduce to a trivial exercise, see Martin (2010) for further discussion. In reality the depreciation allowance does not depend on the actual depreciation rate or the capital utilisation rate, instead it depends on the accounting value of capital and a fixed depreciation rate determined by law. This means that capital income taxation indeed distorts the capital utilisation margin and therefore is distortionary within the period.[3,4]

Combining the first-order conditions with respect to consumption and leisure when history $\mathbf{a}^t$ has occurred gives

$$\frac{u_\ell\left(\mathbf{a}^t\right)}{u_c\left(\mathbf{a}^t\right)} = \frac{1-\tau^h\left(\mathbf{a}^t\right)}{1+\tau^c\left(\mathbf{a}^t\right)} w\left(\mathbf{a}^t\right). \tag{4}$$

It is straightforward to derive a standard Euler equation,

$$\frac{u_c\left(\mathbf{a}^t\right)}{1+\tau^c\left(\mathbf{a}^t\right)} = \beta \mathbb{E}_t\left[\frac{u_c\left(\mathbf{a}^{t+1}\right)}{1+\tau^c\left(\mathbf{a}^{t+1}\right)}\left(1-\delta\left(v\left(\mathbf{a}^{t+1}\right)\right)+\left(1-\tau^k\left(\mathbf{a}^{t+1}\right)\right) v\left(\mathbf{a}^{t+1}\right) r\left(\mathbf{a}^{t+1}\right)\right)\right]. \tag{5}$$

The first-order condition with respect to $v\left(\mathbf{a}^t\right)$ is

$$\delta_v\left(\mathbf{a}^t\right) = \left(1-\tau^k\left(\mathbf{a}^t\right)\right) r\left(\mathbf{a}^t\right). \tag{6}$$

The optimal level of capital utilisation is where the marginal benefit from utilising more capital in terms of after-tax income equals its marginal cost in terms of higher capital depreciation. Equation (6) implies that capital income taxation is distortionary within the period.

Examining the household's first-order conditions, the different distortions caused by the three tax instruments become apparent. The labour income tax distorts the (intratemporal) consumption-leisure margin, (4). The current consumption tax distorts the same margin. In addition, both the current and next period's consumption tax enters into the current

[3] We do not introduce the accounting value of capital into our model, as in Mertens and Ravn (2011) for example, because not only we would have an additional endogenous state variable, but also the (forward-looking) Euler equation would include all future capital utilisation and capital income tax rates on the right hand side. Hence, recasting the problem into a recursive form and in turn solving it appear challenging.

[4] Instead of endogenous capital utilisation, Klein and Ríos-Rull (2003) assume that the capital income tax is chosen one or more periods in advance. In this way the current government internalises the distortive effects of τ^k on future allocations. This approach, however, raises the question of why the capital income tax would be set before other taxes. Note, however, that also in this case the Ramsey policy can be made time consistent through debt restructuring, see Domínguez (2007).

(forward-looking) Euler equation, (5). Finally, only next period's capital income tax distorts the current Euler equation, but the current capital income tax impacts the optimal capital utilisation margin, (6). The task of the fiscal authority is to find the optimal tax mix to raise revenue given these distortions.

We assume that the representative firm's technology is of the standard Cobb-Douglas form in capital services, $v(\mathbf{a}^t)\,k(\mathbf{a}^{t-1})$, and hours, $h(\mathbf{a}^t)$, i.e.,

$$y\left(\mathbf{a}^t\right) = f\left(v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right), h\left(\mathbf{a}^t\right), a_t\right) = a_t\left(v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)\right)^{\gamma} h\left(\mathbf{a}^t\right)^{1-\gamma}, \ \forall \mathbf{a}^t, \tag{7}$$

where $\gamma \in [0,1]$ represents the capital-services elasticity of output. Denoting by f_x the derivative of the production function with respect to the variable x, optimal behaviour in perfect competition implies

$$r\left(\mathbf{a}^t\right) = f_{vk}\left(\mathbf{a}^t\right) = \gamma a_t\left(\frac{h\left(\mathbf{a}^t\right)}{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}\right)^{1-\gamma}, \ \forall \mathbf{a}^t, \tag{8}$$

$$w\left(\mathbf{a}^t\right) = f_h\left(\mathbf{a}^t\right) = (1-\gamma)\,a_t,\left(\frac{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}{h\left(\mathbf{a}^t\right)}\right)^{\gamma}, \ \forall \mathbf{a}^t, \tag{9}$$

i.e., production-factor prices equal their marginal products.

The resource constraint in this economy is

$$c\left(\mathbf{a}^t\right) + g\left(\mathbf{a}^t\right) + k\left(\mathbf{a}^t\right) = y\left(\mathbf{a}^t\right) + \left(1 - \delta\left(v\left(\mathbf{a}^t\right)\right)\right)k\left(\mathbf{a}^{t-1}\right), \ \forall \mathbf{a}^t, \tag{10}$$

where the initial level of capital, $k(\mathbf{a}^0)$, is given. Finally, the government's budget constraint is

$$g\left(\mathbf{a}^t\right) = \tau^k\left(\mathbf{a}^t\right)r\left(\mathbf{a}^t\right)v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right) + \tau^h\left(\mathbf{a}^t\right)w\left(\mathbf{a}^t\right)h\left(\mathbf{a}^t\right) + \tau^c\left(\mathbf{a}^t\right)c\left(\mathbf{a}^t\right), \ \forall \mathbf{a}^t. \tag{11}$$

The benchmark first-best equilibrium in our environment can be defined as follows.

Definition 1 (First best)**.** *The first-best equilibrium consists of allocations* $\{g(\mathbf{a}^t), c(\mathbf{a}^t), \ell(\mathbf{a}^t), h(\mathbf{a}^t), k(\mathbf{a}^t), v(\mathbf{a}^t), y(\mathbf{a}^t)\}_{t=1}^{\infty}$ *that maximise (1) subject to household's time constraint,* (2)*, the production function, (7), and the market clearing condition, (10),* $\forall \mathbf{a}^t$*,* $k(\mathbf{a}^0)$ *and the productivity process given.*

The characterisation of first-best allocations is presented in Appendix A.

We can define competitive equilibria as follows.

Definition 2 (Competitive equilibrium)**.** *A competitive equilibrium consists of government policies,* $\left\{\tau^h(\mathbf{a}^t), \tau^k(\mathbf{a}^t), \tau^c(\mathbf{a}^t), g(\mathbf{a}^t)\right\}_{t=1}^{\infty}$*, prices,* $\{w(\mathbf{a}^t), r(\mathbf{a}^t)\}_{t=1}^{\infty}$*, and private sector allocations,* $\{c(\mathbf{a}^t), \ell(\mathbf{a}^t), h(\mathbf{a}^t), k(\mathbf{a}^t), v(\mathbf{a}^t), y(\mathbf{a}^t)\}_{t=1}^{\infty}$*, satisfying,* $\forall \mathbf{a}^t$*,* *(i) private sector optimisation taking government policies and prices as given, that is,*

- *the household's time constraint,* (2), *budget constraint,* (3), *and optimality conditions,* (4), (5), *and* (6)*;*
- *the production function, (7), and the firm's optimality conditions,* (8) *and* (9)*;*

(ii) market clearing, (10), and

(iii) the government's budget constraint, (11),

$k\left(\mathbf{a}^0\right)$ *and the productivity process given.*

We can eliminate three variables, output ($y\left(\mathbf{a}^t\right)$) and prices ($w\left(\mathbf{a}^t\right)$, and $r\left(\mathbf{a}^t\right)$), and three equations, (7), (8), and (9), in the definition of competitive equilibria. Further, the government's budget constraint and the resource constraint jointly imply that the household's budget constraint holds. We are then left with the following six conditions which characterise competitive equilibria: $\forall \mathbf{a}^t$,

$$h\left(\mathbf{a}^t\right) + \ell\left(\mathbf{a}^t\right) = 1, \tag{12}$$

$$\frac{u_\ell\left(\mathbf{a}^t\right)}{u_c\left(\mathbf{a}^t\right)} = \frac{1-\tau^h\left(\mathbf{a}^t\right)}{1+\tau^c\left(\mathbf{a}^t\right)} a_t\left(1-\gamma\right)\left(\frac{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}{h\left(\mathbf{a}^t\right)}\right)^\gamma, \tag{13}$$

$$\begin{aligned}\frac{u_c\left(\mathbf{a}^t\right)}{1+\tau^c\left(\mathbf{a}^t\right)} = \beta\mathbb{E}_t\Bigg[&\frac{u_c\left(\mathbf{a}^{t+1}\right)}{1+\tau^c\left(\mathbf{a}^{t+1}\right)}\Big(1-\delta\left(v\left(\mathbf{a}^{t+1}\right)\right)\\ &+\left(1-\tau^k\left(\mathbf{a}^{t+1}\right)\right)a_{t+1}\gamma v\left(\mathbf{a}^{t+1}\right)\left(\frac{h\left(\mathbf{a}^{t+1}\right)}{v\left(\mathbf{a}^{t+1}\right)k\left(\mathbf{a}^t\right)}\right)^{1-\gamma}\Big)\Bigg],\end{aligned} \tag{14}$$

$$\delta_v\left(\mathbf{a}^t\right) = \left(1-\tau^k\left(\mathbf{a}^t\right)\right)a_t\gamma\left(\frac{h\left(\mathbf{a}^t\right)}{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}\right)^{1-\gamma}, \tag{15}$$

$$g\left(\mathbf{a}^t\right) = a_t\left(\tau^k\left(\mathbf{a}^t\right)\gamma + \tau^h\left(\mathbf{a}^t\right)\left(1-\gamma\right)\right)\left(v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)\right)^\gamma h\left(\mathbf{a}^t\right)^{1-\gamma} + \tau^c\left(\mathbf{a}^t\right)c\left(\mathbf{a}^t\right), \tag{16}$$

$$c\left(\mathbf{a}^t\right) + g\left(\mathbf{a}^t\right) + k\left(\mathbf{a}^t\right) = a_t\left(v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)\right)^\gamma h\left(\mathbf{a}^t\right)^{1-\gamma} + \left(1-\delta\left(v\left(\mathbf{a}^t\right)\right)\right)k\left(\mathbf{a}^{t-1}\right). \tag{17}$$

3 The policy problems

We consider two types of policy equilibria: with and without commitment of the policy-maker, i.e., Ramsey and Markov equilibria, respectively. In both cases the policy-maker maximises the household's lifetime utility over competitive equilibria. We assume, following most of the literature, that the policy-maker moves first in each period.

We use a version of the primal approach, i.e., we write the policy problems in terms of allocations rather than tax rates. However, we keep the consumption tax rate as a decision variable along with the allocations. This will be useful when we want to constrain the tax rates.[5] In order to do this, we use the household's intratemporal optimality condition (13), and the government's budget constraint, (16), to express the labour and capital income tax rates when history $\mathbf{a}^t$ has occured, respectively, as

$$\tau^h\left(\mathbf{a}^t\right) = 1 - \frac{u_\ell\left(\mathbf{a}^t\right)}{u_c\left(\mathbf{a}^t\right)} \frac{1+\tau^c\left(\mathbf{a}^t\right)}{\left(1-\gamma\right) a_t \left(\frac{v\left(\mathbf{a}^t\right) k\left(\mathbf{a}^{t-1}\right)}{h\left(\mathbf{a}^t\right)}\right)^\gamma}, \tag{18}$$

$$\tau^k\left(\mathbf{a}^t\right) = \frac{g\left(\mathbf{a}^t\right) - \tau^c\left(\mathbf{a}^t\right) c\left(\mathbf{a}^t\right)}{\gamma a_t \left(v\left(\mathbf{a}^t\right) k\left(\mathbf{a}^{t-1}\right)\right)^\gamma h\left(\mathbf{a}^t\right)^{1-\gamma}} - \frac{1-\gamma}{\gamma} \tau^h\left(\mathbf{a}^t\right). \tag{19}$$

Replacing for $\left(1-\tau^k\left(\mathbf{a}^{t+1}\right)\right)$ in (14) using (19) and in turn for $\tau^h\left(\mathbf{a}^{t+1}\right)$ using (18), we have

$$\begin{aligned} \frac{u_c\left(\mathbf{a}^t\right)}{1+\tau^c\left(\mathbf{a}^t\right)} = \beta \mathbb{E}_t \Bigg[\frac{u_c\left(\mathbf{a}^{t+1}\right)}{1+\tau^c\left(\mathbf{a}^{t+1}\right)} \Bigg(1 - \delta\left(v\left(\mathbf{a}^{t+1}\right)\right) + a_{t+1} v\left(\mathbf{a}^{t+1}\right) \left(\frac{h\left(\mathbf{a}^{t+1}\right)}{v\left(\mathbf{a}^{t+1}\right) k\left(\mathbf{a}^t\right)}\right)^{1-\gamma} \\ - \frac{g\left(\mathbf{a}^{t+1}\right) - \tau^c\left(\mathbf{a}^{t+1}\right) c\left(\mathbf{a}^{t+1}\right)}{k\left(\mathbf{a}^t\right)} \Bigg) - u_\ell\left(\mathbf{a}^{t+1}\right) \frac{h\left(\mathbf{a}^{t+1}\right)}{k\left(\mathbf{a}^t\right)} \Bigg]. \end{aligned} \tag{20}$$

Similarly, we can eliminate $\tau^k\left(\mathbf{a}^t\right)$ from (15) and rewrite it as

$$\delta_v\left(\mathbf{a}^t\right) = a_t \left(\frac{h\left(\mathbf{a}^t\right)}{v\left(\mathbf{a}^t\right) k\left(\mathbf{a}^{t-1}\right)}\right)^{1-\gamma} - \frac{g\left(\mathbf{a}^t\right) - \tau^c\left(\mathbf{a}^t\right) c\left(\mathbf{a}^t\right)}{v\left(\mathbf{a}^t\right) k\left(\mathbf{a}^{t-1}\right)} - \frac{u_\ell\left(\mathbf{a}^t\right)}{u_c\left(\mathbf{a}^t\right)} \left(1+\tau^c\left(\mathbf{a}^t\right)\right) \frac{h\left(\mathbf{a}^t\right)}{v\left(\mathbf{a}^t\right) k\left(\mathbf{a}^{t-1}\right)}. \tag{21}$$

So far we have not imposed any restrictions on the tax rates. We are also interested in the case where the labour income tax rate has to be non-negative, as in Coleman (2000) and Correia (2010), given that in reality a labour subsidy is not observed at the aggregate level. Further, a (large) subsidy would likely lead to misreporting of hours. To impose the restriction $\tau^h\left(\mathbf{a}^t\right) \geq 0$, we impose

$$\frac{u_\ell\left(\mathbf{a}^t\right)}{u_c\left(\mathbf{a}^t\right)} \leq \frac{1}{1+\tau^c\left(\mathbf{a}^t\right)} \left(1-\gamma\right) a_t \left(\frac{v\left(\mathbf{a}^t\right) k\left(\mathbf{a}^{t-1}\right)}{h\left(\mathbf{a}^t\right)}\right)^\gamma. \tag{22}$$

Below we write the policy problems in a general form including the constraint (22). We will, however, also study the case where (22) is ignored and the case without consumption taxation, i.e., $\tau^c\left(\mathbf{a}^t\right) = 0$, $\forall \mathbf{a}^t$, to compare our results with the existing literature.

3.1 The Ramsey policy-maker's problem

The Ramsey policy-maker maximises (1) choosing the consumption tax rates $\{\tau^c\left(\mathbf{a}^t\right)\}_{t=0}^\infty$ and allocations $\{g\left(\mathbf{a}^t\right), c\left(\mathbf{a}^t\right), \ell\left(\mathbf{a}^t\right), h\left(\mathbf{a}^t\right), k\left(\mathbf{a}^t\right), v\left(\mathbf{a}^t\right)\}_{t=0}^\infty$, subject to (12), (17), (20), (21),

[5] The same approach is used in Coleman (2000).

and (22), $k(\mathbf{a}^0)$ and the productivity process given. We assign the Lagrange multipliers $\lambda_1(\mathbf{a}^t), ..., \lambda_5(\mathbf{a}^t)$ to the five constraints, respectively.

Note that future decision variables enter into the household's current Euler equation, (20), hence the Ramsey problem is not recursive using only the natural state variables, a and k. Following Marcet and Marimon (2011), the Lagrange multiplier on the Euler equation, $\lambda_3(\mathbf{a}^t)$, with $\lambda_3(\mathbf{a}^0) = 0$, can be introduced as a co-state variable to write a Bellman equation. The Ramsey problem can then be solved numerically by standard policy function iteration.

The value function and the policy functions are time-invariant on the extended state space, the current value of which is denoted (a, k, λ_3). Next period's productivity a' is given exogenously. The policy-maker chooses the functions $\mathcal{K}'()$ and $\Lambda_3'()$, as well the policy functions for the control variables, i.e., $\mathcal{T}^c()$, $\mathcal{C}()$, $\mathscr{L}()$, $\mathcal{H}()$, $\mathcal{G}()$, $\mathcal{V}()$, $\Lambda_1'()$, $\Lambda_2'()$, $\Lambda_4'()$, and $\Lambda_5'()$, and the value function $\mathcal{W}()$.

Suppose that the current state is $\left(\tilde{a}, \tilde{k}, \tilde{\lambda}_3\right)$, and let unindexed variables denote the values of the policy functions for the control variables at this state, i.e., $c = \mathcal{C}\left(\tilde{a}, \tilde{k}, \tilde{\lambda}_3\right)$, and so on; and $k' = \mathcal{K}'\left(\tilde{a}, \tilde{k}, \tilde{\lambda}_3\right)$, $\lambda_3' = \Lambda_3'\left(\tilde{a}, \tilde{k}, \tilde{\lambda}_3\right)$, and $W = \mathcal{W}\left(\tilde{a}, \tilde{k}, \tilde{\lambda}_3\right)$. Then we can write

$$
\begin{aligned}
W = & \max_{\{\tau^c, c, \ell, h, g, k', v\}} \min_{\left\{\lambda_1, \lambda_2, \lambda_3', \lambda_4, \lambda_5\right\}} u(c, \ell, g) + \beta \sum_{a'} \Pr(a' \mid a) \mathcal{W}(a', k', \lambda_3') \\
& - \lambda_1(\ell + h - 1) - \lambda_2\left(c + g + k' - a(vk)^\gamma h^{1-\gamma} - (1 - \delta(v))k\right) \\
& + \lambda_3' \frac{u_c}{1+\tau^c} - \lambda_3\left[\frac{u_c}{1+\tau^c}\left(1 - \delta(v) + av\left(\frac{h}{vk}\right)^{1-\gamma} - \frac{g - \tau^c c}{k}\right) - u_\ell \frac{h}{k}\right] \\
& - \lambda_4\left(a\left(\frac{h}{vk}\right)^{1-\gamma} - \frac{g - \tau^c c}{vk} - \frac{u_\ell}{u_c}(1+\tau^c)\frac{h}{vk} - \delta_v\right) - \lambda_5\left(\frac{u_\ell}{u_c} - \frac{1}{1+\tau^c} a(1-\gamma)\left(\frac{vk}{h}\right)^\gamma\right),
\end{aligned}
$$

$\lambda_1 \geq 0$, $\lambda_2 \geq 0$, and $\lambda_5 \geq 0$, with complementary slackness conditions. It is obvious that the time constraint, (12), and the resource constraint, (17), will bind. Hence, $\lambda_1 > 0$ and $\lambda_2 > 0$. Appendix B.1 presents the first-order conditions of the Ramsey policy-maker's problem.

The solution of such a Ramsey problem is typically time inconsistent. It is optimal to tax capital in the early periods of the reform, and it is also optimal to promise not to do the same in future periods. Next, we take into account the government's commitment problem.

3.2 The time-consistent policy-maker's problem

To characterise optimal time-consistent policies, it is convenient to assume that there is an infinite sequence of separate policy-makers, one for each period. The optimal policy problem therefore resembles a dynamic game between the private sector and all successive governments. The current policy-maker seeks to maximise social welfare from today onwards, anticipating how future policies depend on current policies via the inherited state variables. She also takes into account the optimising behaviour of the private sector. Note that, as under Ramsey policy-making, the fiscal authority moves first in every period, and commits within the period.[6]

Without commitment, strategies for government spending and tax rates depend only on the current natural state of the economy, (a, k). We restrict our attention to stationary Markov-perfect equilibria of the policy game, following the literature (Klein, Krusell, and Ríos-Rull, 2008, for example). In a stationary Markov-perfect equilibrium, all governments employ the same policy rules. Hence, the rules must satisfy a fixed-point property: if the current policy-maker anticipates that all future governments will follow the policy rules $\{\mathcal{T}^c(a,k), \mathcal{C}(a,k), \mathscr{L}(a,k), \mathcal{H}(a,k), \mathcal{G}(a,k), \mathcal{V}(a,k), \mathcal{K}'(a,k)\}$, and similar rules for the Lagrange multipliers, she find it optimal to follow the same rules. Abusing notation, all these policy functions now have two arguments.

Suppose that the current state is $\left(\tilde{a}, \tilde{k}\right)$, and let unindexed variables denote the values of the policy functions for the control variables at this state, i.e., $c = \mathcal{C}\left(\tilde{a}, \tilde{k}\right)$, and so on; and $k' = \mathcal{K}'\left(\tilde{a}, \tilde{k}\right)$ and $W = \mathcal{W}\left(\tilde{a}, \tilde{k}\right)$. Further, let $\mathcal{U}_c() = u_c(\mathcal{C}(), \mathscr{L}(), \mathcal{G}())$, and similarly for $\mathcal{U}_\ell()$. Then we have

$$\begin{aligned}
W = & \max_{\{\tau^c, c, \ell, h, g, k', v\}} \min_{\{\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5\}} u(c, \ell, g) + \beta \mathbb{E} \mathcal{W}(a', k') \\
& - \lambda_1 (\ell + h - 1) - \lambda_2 \left(c + g + k' - a (vk)^\gamma h^{1-\gamma} - (1 - \delta(v)) k\right) \\
& - \lambda_3 \left\{ -\frac{u_c}{1+\tau^c} + \beta \mathbb{E} \left[\frac{\mathcal{U}_c(a', k')}{1 + \mathcal{T}^c(a', k')} \left(1 - \delta(\mathcal{V}(a', k')) + a' \frac{\mathcal{H}(a', k')^{1-\gamma} \mathcal{V}(a', k')^\gamma k'^\gamma}{k'} \right. \right. \right. \\
& \left. \left. \left. \quad - \frac{\mathcal{G}(a', k') - \mathcal{T}^c(a', k') \mathcal{C}(a', k')}{k'} \right) - \mathcal{U}_\ell(a', k') \frac{\mathcal{H}(a', k')}{k'} \right] \right\} \\
& - \lambda_4 \left(a \left(\frac{h}{vk} \right)^{1-\gamma} - \frac{g - \tau^c c}{vk} - \frac{u_\ell}{u_c} (1 + \tau^c) \frac{h}{vk} - \delta_v \right) - \lambda_5 \left(\frac{u_\ell}{u_c} - \frac{1}{1+\tau^c} (1 - \gamma) a \left(\frac{vk}{h} \right)^\gamma \right),
\end{aligned}$$

$\lambda_5 \geq 0$, with complementary slackness condition. Appendix B.2 presents the first-order conditions of the time-consistent policy-maker's problem.

3.3 Some analytical results

We present results for both unrestricted tax rates and excluding a labour subsidy, and both types of policy-makers, committed (Ramsey) and time-consistent (Markov). We consider an

[6] See Ortigueira (2006) on the importance of the assumption on the intra-temporal timing of actions.

economy without productivity shocks here, i.e., we set $a_t = 1$, $\forall t$.

Let us first discuss some benchmark results. Consider a model in which capital utilisation is fixed and exogenous, e.g. $v = 1$, and where the government has access to labour income, capital income, and consumption taxation. In this case, the Ramsey policy-maker can implement the first-best allocation. This is because she is only constrained by the technological constraints, i.e., the time constraint, (12), and the resource constraint, (17), because the current consumption tax rate can be chosen to satisfy the household's Euler equation, (14). It is well known that when a Ramsey equilibrium attains the first best, it is time-consistent: if the planner were given an opportunity to revise her policies in the future, she would choose not to do so. Note, however, that it is not clear whether the tax rates that implement the first best are reasonable, i.e., positive and non-confiscatory.

In the case where households choose the capital utilisation rate, the above result clearly cannot hold generically, given that the policy-maker has to satisfy an additional incentive constraint, (15), while she has no more instruments. However, we can show that she can still implement the first-best steady state. We can also determine some qualitative features of the tax rates at the steady state.[7]

Proposition 1. *As long as private consumption is larger than labour income, the first-best steady state is a Ramsey steady state with $\tau^c > 0$, $\tau^h = -\tau^c$, and $\tau^k = 0$, hence it is time-consistent.*

Proof. In Appendix C.1.

It might seem logical that if instead labour income is larger than private consumption, a similar result holds with $\tau^h > 0$, $\tau^c = -\tau^h$, and $\tau^k = 0$. Indeed, this set of tax rates would not distort any of the private sectors first-order conditions. However, the required tax rates are larger than 100 percent in absolute value given the relatively small difference of the tax bases for different measures of consumption and labour income in reality, see further our quantitative analysis below. Neither a labour income tax rate above 100 percent, nor a consumption tax rate below -100 percent is feasible. Therefore, in this case the first-best steady state typically cannot be decentralised.[8] Coleman (2000) discusses this case as well and reaches the same conclusion.

[7] Here we are assuming (i) convergence of the allocations to an interior steady state and (ii) convergence of the Lagrangian multipliers. As discussed in Lansing (1999) and Straub and Werning (2014), these assumptions are not innocuous. However, in a representative-agent framework with intertemporally-separable utility, these assumptions can be verified to hold, see Straub and Werning (2014). Note also that when solving the model numerically, we do not rely on these assumptions.

[8] We have computed a case where the consumption-labour income ratio is 0.975 at the first best. The consumption tax rate should be -936.47 percent in this case. Imposing a lower bound of -100 percent, we find that the policy-maker sets a positive consumption tax and a negative labour income tax.

We now turn to the cases where tax rates are restricted.

Proposition 2. *When* $\tau^h \geq 0$ *is imposed, the Ramsey policy-maker taxes only consumption at the steady state.*

Proof. In Appendix C.2.

Similarly, we can show in our environment that if we exclude consumption taxation, only labour income is taxed at the steady state.

4 Quantitative analysis

We now turn to numerical methods to find the optimal fiscal policy mix quantitatively with and without consumption taxation and with and without commitment. We will also quantify the welfare gains from commitment with and without consumption taxation, as well as the welfare gains from taxing consumption with and without commitment.

4.1 Calibration

In order to see how the allocations and tax rates behave quantitatively, we solve a calibrated version of our economy. We have to make some further assumptions up front. We consider the model period to be a year. We specify the utility function as

$$u(c, \ell, g) = \log(c) - \alpha_\ell \frac{(1-\ell)^{1+1/\varphi}}{1+1/\varphi} + \alpha_g \log(g), \tag{23}$$

where φ is the (constant) Frisch elasticity of labour supply, while α_ℓ and α_g are the weights of leisure and public goods relative to private consumption, respectively. Given an intertemporal elasticity of substitution equal to 1, we set the Frisch elasticity of labour supply, φ, equal to 3, as in Trabandt and Uhlig (2011).[9] We assume that the depreciation rate is an increasing and convex function of capital utilisation, following the literature. That is, $\delta(v) = \eta v^\chi$, with $\eta > 0$ and $\chi > 1$. Finally, we assume that aggregate productivity follows an AR(1) process with persistence parameter ρ and standard deviation of the shock σ_a.

There remain 8 parameters to calibrate. We first pin down 5 of them, β, γ, α_ℓ, η, and χ, in the following way. We use the private sector's first-order conditions at steady state to match average macroeconomic ratios from United States data for the period 1996-2010. We choose this relatively short time period to better proxy current 'steady-state' macro ratios, as some,

[9]The micro and macro literature tend to differ on the estimates of the Frisch elasticity. Here, we follow the macroeconomic literature and choose a relatively large Frisch elasticity. In Section 4.5.1, we check the robustness of our results to a wide range of values of φ.

notably the labour income share, have changed substantially over time. We take average capacity utilisation for all industries from the Federal Reserve Board, and we compute all other macro ratios using data provided by Trabandt and Uhlig (2012).[10] The private sector takes as given average effective tax rates. We do not take a stand on how the actual tax rates were chosen by the government. We use the effective tax rates computed by Trabandt and Uhlig (2012) for each year to find average tax rates of $\tau^h = 0.221$, $\tau^c = 0.045$, and $\tau_\delta^k = 0.410$, where the lower index δ means that this capital tax rate is with depreciation allowance.[11]

We choose $\gamma = 0.391$ to match an average labour income share of 60.9 percent. To calibrate the parameters α_ℓ, η, χ, and β we first use the macro ratios $c/y = 0.696$,[12] $g/y = 0.155$, and $k/y = 2.349$, and the steady-state resource constraint,

$$\frac{c}{y} + \frac{g}{y} + \eta v^\chi \frac{k}{y} = 1, \tag{24}$$

to find $\delta(v) = \eta v^\chi = 0.064$. Second, we convert τ_δ^k to a without-depreciation-allowance capital tax rate keeping tax revenue constant, i.e., we assume that $\tau_\delta^k\left(rv - \delta(v)\right) = \tau^k rv = \tau^k \gamma \frac{y}{vk}$. This gives $\tau^k = 0.253$. Third, we choose α_ℓ to match $h = 0.249$, the fraction of time worked for the working age population,[13] using the consumption-leisure first-order condition rewritten as

$$\alpha_\ell h^{\frac{1}{\varphi}} \frac{c}{y} = \frac{1-\tau^h}{1+\tau^c}\frac{1}{h}.$$

This gives $\alpha_\ell = 4.154$. Fourth, we use the Euler equation,

$$\frac{1}{\beta} = 1 - \eta v^\chi + \left(1-\tau^k\right) vr,$$

to find $\beta = 0.943$.[14] Fifth, multiplying the optimality condition for capital utilisation, (15), at steady state by v, we have

$$\chi \eta v^\chi = \left(1-\tau^k\right) vr.$$

Using $v = 0.786$ this gives $\chi = 1.956$. Then, from $\eta v^\chi = 0.064$, $\eta = 0.102$.

The 6$^{\text{th}}$ parameter to calibrate is α_g. For our benchmark calibration we assume that g found in the data is optimally chosen in the sense that $u_c = u_g$. Then we can use the macro

[10] https://sites.google.com/site/mathiastrabandt/home/downloads/LafferNberDataMatlabCode.zip

[11] We convert it to a capital tax rate without depreciation allowance, in line with our model, taking revenue from capital income taxation as given, see below.

[12] In order to properly account for the different tax bases, our definition of consumption includes all products and services that are subject to VAT, i.e., non-durables, durables, and services. This is in line with Coleman (2000). However, we also consider the alternative assumption of assigning durables to investment as a robustness check, see Section 4.5.3.

[13] Hours to be allocated between work and leisure: 13.64.

[14] The discount factor is lower than those typically used in the macro literature because of a higher c/y ratio, see above. In Section 4.5 we show that all our results are robust to increasing β to 0.96.

ratios $c/y = 0.696$ and $g/y = 0.155$ again to find $\alpha_g = 0.223$. Below we verify that our results are not sensitive to the choice of this parameter.

Note that we have not taken into account the household's and the government's budget constraint. This is because in reality tax revenues are raised not only to finance public consumption, but also in order to redistribute resources from richer to poorer households. At the status-quo steady state, tax revenues are higher by 11.7 percent of GDP than public consumption. In order to satisfy the budget constraints, one can imagine that the government gives a lump-sum transfer of 11.7 percent of GDP to the representative household.[15]

Finally, we calibrate the last two parameters, the AR(1) coefficients of the technological progress, i.e., ρ and σ_a. We match the unconditional persistence and standard deviation of total output in our economy with fixed tax rates to those of the de-trended US GDP for the period $1996 - 2010$. As a result, we set $\rho = 0.619$ and $\sigma_a = 0.020$. Note that while the value of the persistence parameter is lower than usual in the RBC literature as a whole, it is in line with calibrated RBC models with variable capital utilisation such as Greenwood, Hercowitz, and Krusell (2000).

The calibrated parameter values are presented in Table 1.

Table 1: Calibrated parameters

Par	Value	Description
φ	3	Frisch elasticity
β	0.943	Discount factor
α_ℓ	4.154	Weight of leisure
α_g	0.223	Weight of public goods
γ	0.391	Capital elasticity
η	0.102	Depreciation parameters, $\delta(v) = \eta v^\chi$
χ	1.956	
ρ	0.619	Technology shock autoregressive parameter
σ_a	0.020	Technology shock standard deviation

4.2 Solution method

Our solution method is the following. We use a standard policy-function-iteration algorithm to solve the Ramsey problem. This consists of the following steps. First, we discretise the state variables $k \in [\underline{k}, \overline{k}]$ and $\lambda_3 \in [\underline{\lambda}_3, \overline{\lambda}_3]$ in the deterministic case, and a as well in the stochastic case. For the latter, we approximate the estimated AR(1) process by a 3-state

[15] Viewed through the lens of a representative-agent model, this is a source of inefficiency. This will imply additional welfare gains for all optimal tax reforms we consider.

Markov chain following Tauchen (1986).[16] Once we have found the endogenous collocation nodes, we guess the policy functions at each grid point. At each iteration we solve the system of non-linear equations at each grid point, and we approximate globally the policy functions of next period via cubic spline collocation method. We iterate till convergence of the policy functions is obtained at each grid point to high accuracy.

We solve the time-consistent policy-maker's problem by policy-function iteration as well. In order to compute the derivatives of next period's policy functions with respect to the endogenous state k', we parameterise the policy functions at each endogenous collocation node.[17] We iterate until the policies at each grid point converge to high accuracy. We verify that the derivatives have converged as well. The resulting policy functions are well-behaved. Therefore, we can conclude that our exercise is not plagued by numerical problems.[18]

4.3 Results in the deterministic case

We first present steady-state results. Note that in the cases of the first best and the constrained Ramsey, we can compute the steady state directly from the first-order conditions of the policy problems and without solving for the dynamics. We can then verify that the numbers are identical to high precision once the dynamic policy problem is solved, as described in the previous subsection. The Markov steady state can only be computed once the model is fully solved.

Table 2 shows the allocations and the tax rates at steady state for five policy models. The first column shows the case where the policy-maker has access to all three taxes and the tax rates are unrestricted. Remember that here both Ramsey and Markov policies can implement the first-best steady state. However, the tax rates seem unrealistic, with a consumption tax of 324.5 percent and a labour income tax of -324.5 percent.

Columns 2 and 3 show the case where the government is prohibited from subsidising labour income. In this case, the Ramsey policy-maker (Column 2) taxes consumption at 22.3 percent at the steady state and sets the labour and capital income taxes to zero.[19] The

[16] The values for a are 0.980, 1, and 1.020, and the resulting transition probability matrix is

$$\Pi = \begin{bmatrix} 0.649 & 0.302 & 0.049 \\ 0.262 & 0.476 & 0.262 \\ 0.049 & 0.302 & 0.649 \end{bmatrix}.$$

[17] The derivatives of next period's policy functions are computed using the Compecon Matlab package by Fackler and Miranda (2004).

[18] In Section 4.5.4 we verify that our results are robust to using different solution algorithms.

[19] Note that the fact that the Ramsey policy-maker sets the consumption tax rate equal to α_g when $\tau^h \geq 0$ is imposed is a consequence of logarithmic sub-utilities for both private and government consumption.

Table 2: Tax rates and allocations at steady state

Variable	unrestricted	$\tau^h \geq 0$ Ramsey	$\tau^h \geq 0$ Markov	$\tau^c = 0$ Ramsey	$\tau^c = 0$ Markov
Consumption tax rate	3.245	0.223	0.221	0.000	0.000
Labour income tax rate	-3.245	0.000	0.000	0.240	0.065
Capital income tax rate	0.000	0.000	0.004	0.000	0.198
Capital	1.801	1.548	1.539	1.467	1.106
Capital utilisation	0.786	0.786	0.786	0.786	0.786
Depreciation rate	0.064	0.064	0.064	0.064	0.064
Hours worked	0.320	0.275	0.277	0.261	0.283
Income	0.572	0.492	0.491	0.466	0.439
Capital-income ratio	3.146	3.146	3.133	3.146	2.523
Consumption	0.375	0.322	0.322	0.305	0.317
Consumption-income ratio	0.654	0.654	0.652	0.654	0.723
Public spending	0.084	0.072	0.072	0.068	0.051
Public spending-income ratio	0.146	0.146	0.146	0.146	0.117
Per-period utility	-2.234	-2.271	-2.293	-2.318	-2.403
Welfare-eq. consumption loss	0.000	0.059	0.061	0.087	0.184

time-consistent policy-maker (Column 3) finances government spending mainly from taxing consumption as well, taxing it at 22.1 percent, and sets the capital income tax to 0.4 percent and the labour income tax to zero. Once a labour subsidy is ruled out, it is inefficient to tax both labour and consumption as both taxes distort the same margin, the consumption-leisure decision of the household. Optimal policy generally calls for using the less distortive tax (here the consumption tax) and sets the other (here the labour income tax) to zero. On this point, it is important to stress that there are several reasons for which taxing consumption is less distortive than taxing labour, and the result does not hinge on consumption being larger than labour income. The Laffer curve of the consumption tax peaks at infinity. This means that the loss of efficiency that this tax brings about is always less than proportional to its increase. This is not the case for the labour income tax, as its Laffer curve always peaks for $\tau^h \in (0, 1)$. Moreover, as formally proved in Correia (2010), any revenue-neutral policy that increases consumption taxation and decreases labour income tax, the latter constrained to be non-negative, increases efficiency and therefore welfare.

Without consumption taxation, the Ramsey policy-maker (Column 4) taxes only labour income at the steady state, as is well known, while the time-consistent policy-maker (Column 5) sets the labour income tax to 6.5 percent and the capital income tax to 19.8 percent. Hence, the most striking feature of tax rates at steady state is that with consumption taxation the Ramsey and Markov policies and allocations are very similar, unlike with only labour and capital income taxation.

Turning now to the allocations at steady state, note first that the consumption-income ratio and the public spending-income ratio are the same at the first best and at the Ramsey steady states. This is due to log utility, see Motta and Rossi (2014). Without consumption taxation, the time-consistent policies imply a significantly lower long-run capital-income ratio, a higher consumption-income ratio, a lower public spending-income ratio compared to the Ramsey case, which are all due to more distortions caused by taxation. The result that the public spending-income ratio is lower under Markov policy was first noted by Klein, Krusell, and Ríos-Rull (2008) in an environment where only labour income taxes are available. Here we show that the same result holds in a scenario with both capital and labour income taxes. Finally, hours worked are higher under discretion than under commitment, because of the lower labour income tax.

Instead, with consumption taxation, as a result of less distortion caused by the need to raise fiscal revenue, the steady-state level of capital and income are higher. Notably, they are higher without commitment but with consumption taxation than under Ramsey policy but taxing only labour and capital income. Further, while it is still true that the capital- and public spending-income ratios are lower and the consumption-income ratio is higher under Markov than under Ramsey policies, all ratios change very little as a result of the change in commitment.

Table 2 also shows the per-period utilities and the welfare-equivalent consumption losses compared to the first best at the steady state in the different taxation and commitment scenarios.[20] With access to consumption taxation but imposing a non-negativity constraint on the labour income tax rate, the steady-state welfare losses amount to 6.1 percent in the case of time-consistent policy and to 5.9 percent in the case of Ramsey policy-making. Without taxing consumption, the steady-state welfare-equivalent consumption loss is 18.4 percent under Markov-perfect policy and 8.7 percent under Ramsey policy. Three results are worth stressing. First, taxing consumption generates larger welfare gains under discretion (66.9 percent) than under commitment (32.6 percent). Second, the welfare gains from commitment are small with consumption taxation (2.9 percent) and large without consumption taxation (52.4 percent). Third, welfare is higher without commitment but with access to consumption taxation than with commitment but taxing only labour and capital income.

Next, we study whether our main results on the usefulness of consumption taxation in terms of mitigating the commitment problem of the policy-maker and improving welfare hold

[20]That is, we compute by what fraction the steady-state level of private consumption should be increased, keeping hours worked and public consumption constant, for the representative household to be as well off as at the first-best steady state.

once we take into account policy dynamics. We also want to quantify the welfare gains from commitment taking into account the transition with and without consumption taxation. In order to do this, we perform the following policy exercise. We assume that initially the economy is at the Markov steady state. At time zero a new policy-maker enters into office. She can either be a Markov policy-maker, in which case the same policies and allocations will be implemented in all periods, or have access to a commitment technology, in which case the economy will converge to the Ramsey steady state. Figures 1 and 2 show the dynamics of the allocations and the tax rates when the policy-maker does not have access to consumption taxation and when consumption is taxed optimally but labour income cannot be subsidised, respectively.

Consider first the case with only labour and capital income taxation. At time 0, the Ramsey policy-maker sets capital taxes positive, since capital is a relatively inelastic tax base. However, as discussed above, capital income taxes distort the intratemporal margin of capital utilisation, hence it is sub-optimal to tax capital income at a very high rate in period 0. During the transition, capital income taxes converge to zero. By reducing the capital income tax in any period t, the policy-maker positively affects capital accumulation in all previous periods, i.e., the downward trend in τ^k induces the household to continuously postpone her consumption. At the same time, the Ramsey planner increases the labour income tax rate. This tax hike reduces labour supply in period t, but raises it in any previous period. Along the transition, capital, consumption, and government spending increase, while hours worked decrease.

On the other hand, the Markov policy-maker considers choices made at time t not to affect the private sector's behaviour in any previous period. This means that she does not internalise the benefits of a tax hike on labour income and a reduction in the capital income tax rate in terms of allocations in earlier periods. Furthermore, given the presence of endogenous capital utilisation and labour supply, both labour and capital income taxes are distortive within the period, while the capital tax distorts the Euler equation as well. In an attempt to minimise distortions the Markov policy-maker taxes both capital and labour income in the long run. In our baseline calibration, the long-run level of τ^k is approximately three times higher than τ^h. However, the relative magnitude of the tax rates hinges on the Frisch elasticity of labour supply, φ, which is equal to 3 in our baseline calibration. With $\varphi = 1$ we recover the result of the existing literature (Martin, 2010; Debortoli and Nunes, 2010) that the two income tax rates are close to equal at the Markov steady state, see Section 4.5.1.

The possibility of taxing consumption changes markedly the features of optimal policy.

Under commitment, the policy-maker sets the consumption tax rate positive and almost constant, avoids taxing labour income, and taxes capital income only initially, and even then at a very low rate. The policy-maker does not promise a tax hike on τ^c as doing so would be equivalent, via the Euler equation, to an increase in τ^k in the following period. This is suboptimal as the household would anticipate her consumption, leading to inefficiently low capital accumulation. Ceteris paribus, it is also suboptimal for the policy-maker to commit to a downward trend in τ^c, as this would provide an incentive to the household to inefficiently increase labour supply during the transition. These two effects keep consumption taxes close to constant along the transition.

Given the low incentive to tax capital in the initial period, the time-inconsistency feature of Ramsey policies under consumption taxation is very limited compared to the standard case where the fiscal authority has access to labour and capital income taxation only. As a result, the tax policy functions under commitment and discretion are remarkably similar. Firstly, under Markov policy-making as well, it is suboptimal to jointly tax labour and consumption as these two taxes impact on the same intratemporal margin. Therefore, in the discretionary equilibrium as well, the labour income tax is zero at all times. Secondly, the Markov policy-maker recognises that present consumption taxation has a beneficial effect on today's saving rate. At the same time, taxing capital distorts capital utilisation and, unlike a constant consumption tax, it also distorts the long-run Euler equation. In particular, by depressing the current saving rate, it would limit the amount of resources available in the future. Therefore, despite capital being a relative inelastic source of revenue within the period, the time-consistent policy-maker taxes mainly consumption. Given that this policy is very similar to the one implemented by the Ramsey planner, the resulting time-consistent allocations are also very close to the Ramsey ones along the transition as well.

Using our simulations, we can quantify the welfare gains from commitment both with and without consumption taxation, starting from the corresponding Markov steady state. We find that the welfare gains from commitment in terms of welfare-equivalent consumption are 0.0003 percent with consumption taxation and 2.01 percent taxing labour income instead. Hence, the gains from commitment are negligible with access to consumption taxation, while they are substantial without.

We perform a similar calculation to quantify the gains from taxing consumption. First we consider time-consistent policy-makers. We find that the welfare gains in terms of welfare-equivalent consumption are 2.77 percent, taking into account the transition from the Markov steady state without consumption taxation. Second, we perform the same calculation for a

Ramsey policy-maker, starting from the Ramsey steady state without consumption taxation. Now the welfare gains from taxing consumption are 1.21 percent. Hence, the welfare gains from taxing consumption are much larger under discretion than under commitment.

Finally, we can also compute the welfare gains from the different taxation and commitment scenarios compared to the existing tax system in a similar fashion. We assume that initially the economy is at the actual steady state, described in Section 4.1. At time 0 a new policy-maker enters into office. She can be either a Markov or a Ramsey policy-maker, and with either access to consumption taxation or no access. In the case of a Ramsey policy-maker, the welfare gains are 7.745 percent and 7.06 percent with and without taxing consumption, respectively. Ceteris paribus, in the case of a Markov policy-maker, the welfare gains are 7.744 percent and 4.92 percent, respectively. Notice that the welfare gains are larger with consumption taxation and without commitment than without consumption taxation and with commitment. Table 3 summarises our welfare results including transitions.

Table 3: Welfare gains in welfare-equivalent consumption units (percent)

Welfare gains...	With cons. tax	Without cons. tax	
...from commitment	0.0003	2.01	
	...compared to the existing tax system		...from taxing consumption
Ramsey	7.745	7.06	2.21
Markov	7.744	4.92	2.77

Figures 3 and 4 show the dynamics of the tax rates and the allocations without and with access to consumption taxation, respectively, for both a Ramsey and a time-consistent policy-maker. These pictures show what is behind the welfare-gain numbers compared to the status quo. With access to consumption taxation, the whole dynamic path of taxes and allocations hardly differ with and without commitment.

4.4 Results with aggregate productivity shocks

We now study the cyclical properties of the policy instruments and allocations under the different tax and commitment scenarios. We are interested in whether the close similarity between Ramsey and Markov policies when consumption is taxed optimally still holds when the economy faces aggregate productivity shocks. For each policy scenario, we simulate the model and calculate sample statistics from the simulated data.[21] The results of this exercise are reported in Table 4.

[21] We proceed as follows. We assume that in the initial period the system is in its stochastic steady state. We simulate the model for 1000 periods, using the same shocks across policy scenarios, and compute sample statistics. Finally, we take the median values of the sample statistics over 101 repetitions.

Table 4: Cyclical properties of taxes and allocations

	$\tau^h \geq 0$		$\tau^c = 0$	
	Ramsey	Markov	Ramsey	Markov
	Consumption tax		Labour income tax	
Mean	0.221	0.221	0.240	0.065
Standard deviation	0.005	0.005	0.002	0.002
Coefficient of variation	0.024	0.022	0.009	0.030
Autocorrelation	0.590	0.621	0.922	0.494
Correlation with output	0.9998	0.996	-0.700	-0.860
Capital income tax				
Mean	0.000	0.004	0.000	0.198
Standard deviation	0.009	0.008	0.009	0.004
Coefficient of variation	10.483	1.894	30.693	0.019
Autocorrelation	0.590	0.618	0.463	0.490
Correlation with output	-0.9995	-0.997	-0.801	-0.913
Public spending				
Mean	0.072	0.072	0.068	0.051
Standard deviation	0.001	0.001	0.001	0.001
Coefficient of variation	0.015	0.015	0.019	0.022
Autocorrelation	0.974	0.976	0.775	0.799
Correlation with output	0.382	0.414	0.880	0.886
Public spending-income ratio				
Mean	0.146	0.146	0.146	0.117
Standard deviation	0.005	0.004	0.004	0.003
Coefficient of variation	0.031	0.030	0.026	0.023
Autocorrelation	0.508	0.511	0.514	0.490
Correlation with output	-0.899	-0.898	-0.936	-0.892
Consumption				
Mean	0.322	0.322	0.305	0.317
Standard deviation	0.005	0.005	0.005	0.006
Coefficient of variation	0.015	0.015	0.017	0.018
Autocorrelation	0.974	0.976	0.939	0.934
Correlation with output	0.382	0.414	0.621	0.665
Hours				
Mean	0.275	0.276	0.260	0.283
Standard deviation	0.006	0.006	0.007	0.007
Coefficient of variation	0.021	0.020	0.027	0.025
Autocorrelation	0.512	0.515	0.513	0.490
Correlation with output	0.866	0.864	0.937	0.901
Output				
Mean	0.493	0.492	0.467	0.439
Standard deviation	0.017	0.016	0.020	0.018
Coefficient of variation	0.034	0.033	0.042	0.040
Autocorrelation	0.585	0.588	0.570	0.578
Welfare-eq. consumption loss	0.059	0.061	0.087	0.183

When consumption taxes are not available and the policy-maker can credibly commit (Column 3), the burden of taxation is almost entirely given to labour taxation. At the same time, labour income taxes hardly move in response to shocks, as the policy-maker prefers to use capital income taxes and public expenditure as shock absorbers. This is the well-known labour tax smoothing result of the Ramsey literature (Chari, Christiano, and Kehoe, 1994).

On the other hand, under discretion (Column 4 of Table 4), the policy-maker uses both capital and labour income taxes in response to unexpected productivity changes. The coefficient of variation of the labour income tax rate is more than three times larger and the coefficient of variation of the capital income tax rate is more than a thousand times smaller than their Ramsey counterparts. This is mainly due to the fact that the Markov policy-maker is less able to smooth the effects of random productivity events intertemporally, as she is less able to use the capital income tax as shock absorber, given that she relies heavily on this instrument to raise fiscal revenue. In our numerical example, the coefficient of variation of the capital income tax rate is roughly three-fifths of the coefficient of variation of labour taxes. Interestingly, the volatility of output and hours worked are slightly higher under commitment than under discretion, while the opposite is true for private and public consumption. Finally, both tax rates are countercyclical. These patterns are, for the most part, very similar to the ones presented in Klein and Ríos-Rull (2003) and in Debortoli and Nunes (2010), although the class of economies they look at is slightly different.[22]

The differences between Ramsey and Markov policies are greatly reduced when the policy-maker can tax consumption (Columns 1 and 2 in Table 4). Hence, the close similarity between Ramsey and Markov policy-making when consumption taxes are available extend to the cyclical properties of taxes and allocations. As in the case without consumption taxation, the coefficient of variation of capital income taxes is larger than that of the alternative tax instrument, in this case, the consumption tax. However, under Markov policies capital taxes still play the main role in absorbing shocks, unlike without consumption taxation. A new feature of tax policies with consumption taxation is that the consumption tax rate is highly procyclical. The capital income tax rate remains countercyclical, but its correlation with output increases when consumption is taxed. All allocation variables, namely, consumption, public spending, hours, and output vary less with consumption taxation than without.

Finally, we compute long-run expected welfare as a percentage increase in welfare-equivalent consumption units in all periods and all states in a particular policy scenario that is neces-

[22]In particular, Klein and Ríos-Rull (2003) study a model with full capital utilisation, exogenous government spending and a capital income tax which is determined one or more periods in advance, while Debortoli and Nunes (2010) consider a utility function with variable Frisch elasticity of labour supply.

sary to make the representative household as well off as at the first best.[23] The values we find are very similar to those for the deterministic steady state, and hence our main conclusions extend to the stochastic environment, namely, that (i) taxing consumption generates larger welfare gains under discretion than under commitment, and (ii) the welfare gains from commitment are small with consumption taxation and large without consumption taxation. Therefore, the business cycle results confirm the similarities between Ramsey and Markov equilibria when the policy-maker has access to consumption taxation, as well as the welfare benefits of taxing consumption.

4.5 Robustness checks

In this section we verify the robustness of our results to (i) changing some parameter values, (ii) taking government spending as exogenous, (iii) recalibrating the model using another definition of consumption from the data, in particular, we exclude durables, and (iv) solving the models using different numerical algorithms.

4.5.1 Different parameters values

First of all, we consider a wide range of values for the Frisch elasticity of labour supply without productivity shocks. In particular, we consider (i) $\varphi = 0.4$, which is in line with recent micro estimates such as Domeij and Floden (2006) (see also Guner, Kaygusuz, and Ventura, 2012), (ii) $\varphi = 1$, which is often chosen in the macro literature (e.g. Christiano, Eichenbaum, and Evans, 2005), and (iii) $\varphi = 5$ as a high value, which is sometimes chosen to better match the intertemporal variation of aggregate hours (e.g. Galí, López-Salido, and Vallés, 2007). We adjust α_ℓ appropriately in each case as described in Section 4.1.

The first three panels of Table 5 show that the steady-state results are only marginally affected by changing φ, except in the case of a time-consistent policy-maker without access to consumption taxation (last column of Table 5). For all values of φ, the Ramsey planner finances all public expenditure with either only labour income or only consumption taxation, as we know from Proposition 2 and the existing literature. Under discretion, if consumption taxation is available, the changes in tax rates are very small. On the other hand, we find large differences when consumption taxes are not available. In that case, τ^h varies from 17.3 to 4.2 percent and τ^k from 7.8 to 22.1 percent as φ increases. This is because increasing the

[23]In order to do this, for some percentage increase in consumption ε, we simulate the economy over 600 periods, compute per-period utility for the last 500 periods, and finally take the average over 501 such simulations. Then we find the ε such that the average per-period utility matches the one found for the first best from similar simulations.

elasticity of labour supply increases the relative distortion of τ^h compared to τ^k. As a result, the optimal time-consistent policy calls for lower labour income tax rates and higher capital income tax rates.

Increasing the Frisch elasticity increases the welfare losses under all policy scenarios. This is because, ceteris paribus, a higher φ implies a stronger response of hours to any given distortion of the consumption-leisure margin. It is also worth noticing that under discretion as φ gets bigger, the public spending-income ratio decreases. This is because the Markov planner chooses lower taxes as the distortionary effects of fiscal policy are greater. These effects are almost absent in the case with consumption taxation. Increasing the Frisch elasticity does not remove the policy-maker's incentive to give almost all the burden of taxation to consumption and to keep the public spending-income ratio close to its efficient level.

Afterwards, we increase the coefficient of relative risk aversion to 2, i.e., now the current utility function is

$$u(c,\ell,g) = \frac{c^{1-\sigma}-1}{1-\sigma} - \alpha_\ell \frac{(1-\ell)^{1+1/\varphi}}{1+1/\varphi} + \alpha_g \frac{g^{1-\sigma}-1}{1-\sigma},$$

with $\sigma = 2$ and $\varphi = 3$. We recalibrate the utility weights $\alpha_\ell = 16.791$ and $\alpha_g = 0.050$ to keep hours at 0.249 and g/y at 0.155 before the reform. The results are in the fourth panel of Table 5. Our conclusions remain unaltered.

In the fifth panel of Table 5, we report the steady-state results for the case where we keep $\sigma = 1$ and $\varphi = 3$ but set $\beta = 0.96$, the most-commonly used value in the macro literature for yearly models. Again our results are not affected.

Next, we set $\chi = 1.8$ and adjust $\eta = 0.098$ so that the capital utilisation rate at the steady state be the same as in the data. Note that this implies a depreciation rate of 0.076 at the steady state. Panel 6 of Table 5 shows that our conclusions do not change.

We also consider a higher value of α_g, 0.3 in particular, while keeping all other parameters at their baseline values. In this case the consumption tax base is smaller as a share of GDP than in the baseline calibration. The results are in Panel 7 of Table 5. Ramsey and Markov policies and allocations are very similar with consumption taxation in this case as well.

Remarkably, under all parameterisations considered, taxing consumption is more important than being able to commit. As reported in Tables 2 and 5, welfare is always higher under Markov policy-making and consumption taxation than under Ramsey without taxing consumption.

We solve the stochastic model as well with $\varphi = 1$ to check the robustness of the cyclical properties of tax rates and allocations. Table 6 shows that the main features remain unchanged.

4.5.2 Exogenous government spending

This section shows that our results do not change if we consider government spending to be exogenous. The new policy problem consists of raising fiscal revenues in order to finance an exogenous and fixed level of public consumption, $\bar{g}$. The government budget constraint can now be written as

$$\bar{g} = \tau^k\left(\mathbf{a}^t\right) r\left(\mathbf{a}^t\right) v\left(\mathbf{a}^t\right) k\left(\mathbf{a}^{t-1}\right) + \tau^h\left(\mathbf{a}^t\right) w\left(\mathbf{a}^t\right) h\left(\mathbf{a}^t\right) + \tau^c\left(\mathbf{a}^t\right) c\left(\mathbf{a}^t\right), \ \forall \mathbf{a}^t. \tag{25}$$

We further assume that households do not value government spending, i.e.

$$\mathbb{E}_0\left[\sum_{t=1}^{\infty}\beta^t u\left(c\left(\mathbf{a}^t\right), \ell\left(\mathbf{a}^t\right)\right)\right]. \tag{26}$$

We calibrate $\bar{g}$ so that $\bar{g}/y$ in the decentralised competitive steady-state, taking tax rates as given, is equal to its value in the data, 0.155. Then we keep $\bar{g}$ constant across all policy regimes. All other parameters are kept at their benchmark values as described in Section 4.1. The steady-state results of this exercise are reported in Table 7. The close similarity of Ramsey and Markov fiscal policies and allocations when consumption is taxed optimally holds in this case as well.

4.5.3 A different measure of consumption

We now consider durable consumption spending as part of investment rather than of consumption. Let us denote by i_c this investment. Then, instead of (24) we have

$$\frac{c}{y} + \frac{g}{y} + \frac{i_c}{y} + \eta v^{\chi}\frac{k}{y} = 1, \tag{27}$$

with $c/y = 0.633$ and $i_c/y = 0.063$. Note that the stock of consumer durables is by definition not part of productive capital. We recalibrate our baseline model following the strategy described in Section 4.1. Only two parameters of the current utility function change, in particular, now $\alpha_\ell = 4.668$ and $\alpha_g = 0.245$. We present the steady-state results in Table 8. Our main conclusions are robust to considering this alternative measure of consumption.

4.5.4 Different numerical algorithms

Finally, we also check the robustness of our results to using different numerical algorithms. In particular, using the solution to the Ramsey problem by policy function iteration as initial guess, we solve it again parameterising the policy functions using Chebyshev polynomials or cubic splines. Then, we solve the time-consistent policy-maker's problem by parameterising the policy functions again. In this solution algorithm we iterate until the parameters of the policy functions converge to high accuracy. The resulting tax rates and allocations are identical up to many decimals to the ones computed using policy function iteration.

5 Concluding remarks

This paper has studied the properties of time-consistent optimal fiscal policies when the policy-maker has access to consumption taxation, both in a deterministic setting and over the business cycle. Contrary to the case with only labour and capital income taxation, time-consistent policies, and hence allocations, are very close to those under Ramsey policy, as long as capital income taxation causes some distortion within the period. This also means that the Ramsey policies are close to time-consistent in the sense that if a new policy-maker with the same objective entered in a randomly chosen period or due to the electoral cycle, policies would hardly change.

When the labour income tax rate is restricted to be non-negative, the optimal time-consistent capital income tax rate is close to zero (0.4 percent), the consumption tax rate is 22.1 percent, and labour income is not taxed at the steady state. The proposed time-consistent policies with consumption taxation would yield welfare gains of 7.744 percent in terms of welfare-equivalent consumption units compared to the existing tax system, taking into account the transition. These welfare gains are almost as large as under commitment (7.745 percent), and are larger than the gains a Ramsey policy-maker could achieve without access to consumption taxation (7.06 percent). If the time-consistent policy maker can only tax labour and capital income, the welfare gains are reduced to 4.92 percent.

In this paper we have considered a representative-agent framework, hence we studied the optimal tax mix from an efficiency perspective only, to raise fiscal revenues. An important task for future research is to analyse the distributional impact of the different tax instruments in a model with heterogeneity across households.

References

Azzimonti, M., P.-D. Sarte, and J. Soares (2009). Distortionary Taxes and Public Investment when Government Promises are not Enforceable. *Journal of Economic Dynamics and Control 33*(9), 1662–1681.

Chari, V. V., L. J. Christiano, and P. J. Kehoe (1994). Optimal Fiscal Policy in a Business Cycle Model. *Journal of Political Economy 102*(4), 617–652.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (2005). Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of Political Economy 113*(1), 1–45.

Coleman, W. J. I. (2000). Welfare and Optimum Dynamic Taxation of Consumption and Income. *Journal of Public Economics 76*(1), 1–39.

Correia, I. (2010). Consumption Taxes and Redistribution. *American Economic Review 100*(4), 1673–1694.

Correia, I., E. Farhi, J. P. Nicolini, and P. Teles (2013). Unconventional Fiscal Policy at the Zero Bound. *American Economic Review 103*(4), 1172–1211.

Debortoli, D. and R. Nunes (2010). Fiscal Policy under Loose Commitment. *Journal of Economic Theory 145*(3), 1005–1032.

Domeij, D. and M. Floden (2006). The Labor-Supply Elasticity and Borrowing Constraints: Why Estimates are Biased. *Review of Economic Dynamics 9*(2), 242–262.

Domínguez, B. (2007). On the Time-Consistency of Optimal Capital Taxes. *Journal of Monetary Economics 54*(3), 686–705.

Fackler, P. and M. Miranda (2004). *Applied Computational Economics and Finance*. The MIT Press.

Farhi, E., G. Gopinath, and O. Itskhoki (2014). Fiscal Devaluations. *Review of Economic Studies 81*(2), 725–760.

Galí, J., J. D. López-Salido, and J. Vallés (2007). Understanding the Effects of Government Spending on Consumption. *Journal of the European Economic Association 5*(1), 227–270.

Greenwood, J., Z. Hercowitz, and G. W. Huffman (1988). Investment, Capacity Utilization, and the Real Business Cycle. *American Economic Review 78*(3), 402–417.

Greenwood, J., Z. Hercowitz, and P. Krusell (2000). The role of investment-specific technological change in the business cycle. *European Economic Review 44*(1), 91–115.

Guner, N., R. Kaygusuz, and G. Ventura (2012). Taxation and Household Labour Supply. *Review of Economic Studies 79*(3), 1113–1149.

Klein, P., P. Krusell, and J.-V. Ríos-Rull (2008). Time-Consistent Public Policy. *Review of Economic Studies 75*(3), 789–808.

Klein, P. and J.-V. Ríos-Rull (2003). Time-Consistent Optimal Fiscal Policy. *International Economic Review 44*(4), 1217–1245.

Lansing, K. J. (1999). Optimal Redistributive Capital Taxation in a Neoclassical Growth Model. *Journal of Public Economics 73*(3), 423–453.

Marcet, A. and R. Marimon (2011). Recursive Contracts. Mimeo.

Martin, F. M. (2010). Markov-perfect Capital and Labor Taxes. *Journal of Economic Dynamics and Control 34*(3), 503–521.

Mertens, K. and M. O. Ravn (2011). Understanding the Aggregate Effects of Anticipated and Unanticipated Tax Policy Shocks. *Review of Economic Dynamics 14*(1), 27–54.

Motta, G. and R. Rossi (2014). Ramsey Monetary and Fiscal Policy: the Role of Consumption Taxation. Mimeo.

Ortigueira, S. (2006). Markov-Perfect Optimal Taxation. *Review of Economic Dynamics 9*(1), 153–178.

Straub, L. and I. Werning (2014). Positive Long Run Capital Taxation: Chamley-Judd Revisited. Working Paper 20441, National Bureau of Economic Research.

Tauchen, G. (1986). Finite State Markov-chain Approximations to Univariate and Vector Autoregressions. *Economics Letters 20*(2), 177–181.

Trabandt, M. and H. Uhlig (2011). The Laffer Curve Revisited. *Journal of Monetary Economics 58*(4), 305–327.

Trabandt, M. and H. Uhlig (2012). How Do Laffer Curves Differ across Countries? NBER Working Papers 17862, National Bureau of Economic Research.

Zhu, X. (1995). Endogenous Capital Utilization, Investor's Effort, and Optimal Fiscal Policy. *Journal of Monetary Economics 36*(3), 655–677.

Appendices

A First-best allocation

Denote by $\lambda_1(\mathbf{a}^t)$ the Lagrange multiplier on the time constraint, (2), and by $\lambda_2(\mathbf{a}^t)$ the Lagrange multiplier on the resource constraint, (10), when history $\mathbf{a}^t$ has occurred. Use (7) to replace for $y(\mathbf{a}^t)$ in (10). Then we can write the problem as

$$\max_{\{c(\mathbf{a}^t),\ell(\mathbf{a}^t),h(\mathbf{a}^t),g(\mathbf{a}^t),k(\mathbf{a}^t),v(\mathbf{a}^t)\}_{t=1}^{\infty}} \min_{\{\lambda_1(\mathbf{a}^t),\lambda_2(\mathbf{a}^t)\}_{t=1}^{\infty}} \mathbb{E}_0 \left\{ \sum_{t=0}^{\infty} \beta^t \left[u\left(c\left(\mathbf{a}^t\right), \ell\left(\mathbf{a}^t\right), \left(\mathbf{a}^t\right)\right) \right. \right.$$
$$+\lambda_1\left(\mathbf{a}^t\right)\left(1-\ell\left(\mathbf{a}^t\right)-h\left(\mathbf{a}^t\right)\right)$$
$$\left.\left. +\lambda_2\left(\mathbf{a}^t\right)\left(a_t\left(v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)\right)^{\gamma} h\left(\mathbf{a}^t\right)^{1-\gamma} + \left(1-\delta\left(v\left(\mathbf{a}^t\right)\right)\right)k\left(\mathbf{a}^{t-1}\right) - c\left(\mathbf{a}^t\right) - g\left(\mathbf{a}^t\right) - k\left(\mathbf{a}^t\right)\right)\right]\right\},$$

where we have used (7) to replace for $y(\mathbf{a}^t)$ in (10). The first-order conditions with respect to $c(\mathbf{a}^t), \ell(\mathbf{a}^t), h(\mathbf{a}^t), g(\mathbf{a}^t), k(\mathbf{a}^t), v(\mathbf{a}^t), \lambda_1(\mathbf{a}^t)$, and $\lambda_2(\mathbf{a}^t)$, respectively, are

$$u_c\left(\mathbf{a}^t\right) = \lambda_2\left(\mathbf{a}^t\right), \tag{28}$$
$$u_\ell\left(\mathbf{a}^t\right) = \lambda_1\left(\mathbf{a}^t\right), \tag{29}$$
$$a_t(1-\gamma)\left(\frac{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}{h\left(\mathbf{a}^t\right)}\right)^{\gamma} = \lambda_1\left(\mathbf{a}^t\right), \tag{30}$$
$$u_g\left(\mathbf{a}^t\right) = \lambda_2\left(\mathbf{a}^t\right), \tag{31}$$
$$\lambda_2\left(\mathbf{a}^t\right) = \beta\mathbb{E}_t\left[\lambda_1\left(\mathbf{a}^{t+1}\right)\left(a_{t+1}\gamma v\left(\mathbf{a}^{t+1}\right)^{\gamma}\left(\frac{h\left(\mathbf{a}^{t+1}\right)}{k\left(\mathbf{a}^t\right)}\right)^{1-\gamma} + 1 - \delta\left(v\left(\mathbf{a}^{t+1}\right)\right)\right)\right], \tag{32}$$
$$\delta_u\left(\mathbf{a}^t\right) = a_t\gamma\left(\frac{h\left(\mathbf{a}^t\right)}{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}\right)^{1-\gamma}, \tag{33}$$
$$\ell\left(\mathbf{a}^t\right) + h\left(\mathbf{a}^t\right) = 1, \tag{34}$$
$$c\left(\mathbf{a}^t\right) + g\left(\mathbf{a}^t\right) + k\left(\mathbf{a}^t\right) = a_t\left(v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)\right)^{\gamma} h\left(\mathbf{a}^t\right)^{1-\gamma} + \left(1-\delta\left(v\left(\mathbf{a}^t\right)\right)\right)k\left(\mathbf{a}^{t-1}\right). \tag{35}$$

Straightforward combinations of (28)-(35) lead to the following equations which characterise the first-best allocation:

$$u_g\left(\mathbf{a}^t\right) = u_c\left(\mathbf{a}^t\right), \tag{36}$$
$$\frac{u_\ell\left(\mathbf{a}^t\right)}{u_c\left(\mathbf{a}^t\right)} = a_t(1-\gamma)\left(\frac{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}{h\left(\mathbf{a}^t\right)}\right)^{\gamma}, \tag{37}$$
$$h\left(\mathbf{a}^t\right) + \ell\left(\mathbf{a}^t\right) = 1, \tag{38}$$
$$u_c\left(\mathbf{a}^t\right) = \beta\mathbb{E}_t\left[u_c\left(\mathbf{a}^{t+1}\right)\left(1-\delta\left(v\left(\mathbf{a}^{t+1}\right)\right) + a_{t+1}\gamma v\left(\mathbf{a}^{t+1}\right)\left(\frac{h\left(\mathbf{a}^{t+1}\right)}{v\left(\mathbf{a}^{t+1}\right)k\left(\mathbf{a}^t\right)}\right)^{1-\gamma}\right)\right], \tag{39}$$

$$c\left(\mathbf{a}^t\right)+g\left(\mathbf{a}^t\right)+k\left(\mathbf{a}^t\right)=a_t\left(v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)\right)^{\gamma}h\left(\mathbf{a}^t\right)^{1-\gamma}+\left(1-\delta\left(v\left(\mathbf{a}^t\right)\right)\right)k\left(\mathbf{a}^{t-1}\right), \quad (40)$$

$$\delta_v\left(\mathbf{a}^t\right)=a_t\gamma\left(\frac{h\left(\mathbf{a}^t\right)}{v\left(\mathbf{a}^t\right)k\left(\mathbf{a}^{t-1}\right)}\right)^{1-\gamma}, \quad (41)$$

$\forall\mathbf{a}^t$, $k\left(\mathbf{a}^0\right)$ and the productivity process given.

B First-order conditions of the policy problems

We assume that the utility function is separable with respect to its three arguments, hence the second cross-derivatives are zero.

B.1 First-order conditions of the Ramsey policy-maker's problem

The FOCs with respect to $\tau^c, c, \ell, h, g, k', v$, and $\lambda_1, \lambda_2, \lambda_3', \lambda_4, \lambda_5$, respectively, are

$$\begin{aligned}
0=&\frac{1}{\left(1+\tau^c\right)^2}\left[-\lambda_3'u_c+\lambda_3u_c\left(1-\delta(v)+av\left(\frac{h}{vk}\right)^{1-\gamma}-\frac{g-\tau^cc}{k}\right)\right.\\
&\left.-\lambda_5\left(1-\gamma\right)a\left(\frac{vk}{h}\right)^{\gamma}\right]-\lambda_4\left(\frac{c}{vk}-\frac{u_\ell}{u_c}\frac{h}{vk}\right)-\lambda_3\frac{u_c}{1+\tau^c}\frac{c}{k}, \quad (42)\\
0=&u_c-\lambda_2+\lambda_3'\frac{u_{cc}}{1+\tau^c}-\lambda_3\frac{u_{cc}}{1+\tau^c}\left(1-\delta(v)+av\left(\frac{h}{vk}\right)^{1-\gamma}-\frac{g-\tau^cc}{k}\right)\\
&-\lambda_3\frac{u_c}{1+\tau^c}\frac{\tau^c}{k}-\lambda_4\left(\frac{\tau^c}{vk}+\frac{u_\ell}{u_c^2}u_{cc}\left(1+\tau^c\right)\frac{h}{vk}\right)+\lambda_5\frac{u_\ell}{u_c^2}u_{cc}, \quad (43)\\
0=&u_\ell-\lambda_1+\lambda_3u_{\ell\ell}\frac{h}{k}+\lambda_4\frac{u_{\ell\ell}}{u_c}\left(1+\tau^c\right)\frac{h}{vk}-\lambda_5\frac{u_{\ell\ell}}{u_c}, \quad (44)\\
0=&-\lambda_1+\lambda_2a\left(1-\gamma\right)\left(\frac{vk}{h}\right)^{\gamma}-\lambda_3\left[\frac{u_c}{1+\tau^c}a\left(1-\gamma\right)\frac{v^{\gamma}}{h^{\gamma}k^{1-\gamma}}-\frac{u_\ell}{k}\right]\\
&-\lambda_4\left[a\left(1-\gamma\right)h^{-\gamma}\left(vk\right)^{\gamma-1}-\frac{u_\ell}{u_c}\left(1+\tau^c\right)\frac{1}{vk}\right]-\lambda_5\frac{1}{1+\tau^c}a\gamma\left(1-\gamma\right)\left(vk\right)^{\gamma}h^{-\gamma-1}, \quad (45)\\
0=&u_g-\lambda_2+\lambda_3\frac{u_c}{1+\tau^c}\frac{1}{k}+\lambda_4\frac{1}{vk}, \quad (46)\\
0=&-\lambda_2+\beta\mathbb{E}\left\{\lambda_2'\left(a'v'^{\gamma}\gamma\left(\frac{h'}{k'}\right)^{1-\gamma}+1-\delta\left(v'\right)\right)\right.\\
&+\lambda_3'\left[\frac{u_c'}{1+\tau^{c\prime}}\left(a'v'^{\gamma}\left(1-\gamma\right)k'^{\gamma-2}h'^{1-\gamma}-\frac{g'-\tau^{c\prime}c'}{k'^2}\right)-u_\ell'\frac{h'}{k'^2}\right] \quad (47)\\
&+\lambda_4'\left(a'\left(1-\gamma\right)k'^{\gamma-2}\left(\frac{h'}{v'}\right)^{1-\gamma}-\frac{g'-\tau^{c\prime}c'}{v'k'^2}-\frac{u_\ell'}{u_c'}\left(1+\tau^{c\prime}\right)\frac{h'}{v'k'^2}\right)\\
&\left.+\lambda_5'\frac{1}{1+\tau^{c\prime}}a'\left(1-\gamma\right)\gamma\frac{v'^{\gamma}}{k'^{1-\gamma}h'^{\gamma}}\right\},
\end{aligned}$$

$$0 = \lambda_2 \left(a\gamma v^{\gamma-1} k^{\gamma} h^{1-\gamma} - \delta_v k\right) + \lambda_3 \frac{u_c}{1+\tau^c} \left(\delta_v - a\gamma v^{\gamma-1} \left(\frac{h}{k}\right)^{1-\gamma}\right)$$

$$+ \lambda_4 \left(a\left(1-\gamma\right) v^{\gamma-2} \left(\frac{h}{k}\right)^{1-\gamma} - \frac{g-\tau^c c}{v^2 k} - \frac{u_\ell}{u_c}\left(1+\tau^c\right)\frac{h}{v^2 k} + \delta_{vv}\right) \tag{48}$$

$$+ \lambda_5 \frac{1}{1+\tau^c} a\left(1-\gamma\right)\gamma v^{\gamma-1} \left(\frac{k}{h}\right)^{\gamma},$$

$$0 = \ell + h - 1, \tag{49}$$

$$0 = c + g + k_{t+1} - a\left(vk\right)^{\gamma} h^{1-\gamma} - \left(1 - \delta\left(v\right)\right)k, \tag{50}$$

$$0 = -\frac{u_c}{1+\tau^c} + \beta\mathbb{E}\left[\frac{u_c'}{1+\tau^{c\prime}}\left(1 - \delta\left(v'\right) + a'v'\left(\frac{h'}{v'k'}\right)^{1-\gamma} - \frac{g' - \tau^{c\prime}c'}{k'}\right) - u_\ell' \frac{h'}{k'}\right], \tag{51}$$

$$0 = a\left(\frac{h}{vk}\right)^{1-\gamma} - \frac{g-\tau^c c}{vk} - \frac{u_\ell}{u_c}\left(1+\tau^c\right)\frac{h}{vk} - \delta_v, \tag{52}$$

$$0 \geq \frac{u_\ell}{u_c} - \frac{1}{1+\tau^c} a\left(1-\gamma\right)\left(\frac{vk}{h}\right)^{\gamma}, \tag{53}$$

$\lambda_5 \geq 0$, with complementary slackness condition.

B.2 First-order conditions of the time-consistent policy-maker's problem

The FOCs with respect to $\tau^c, c, \ell, h, g, k', v$, respectively, are

$$0 = -\lambda_3 \frac{1}{\left(1+\tau^c\right)^2} u_c - \lambda_4 \left(\frac{c}{vk} - \frac{u_\ell}{u_c}\frac{h}{vk}\right) - \lambda_5 \frac{1}{\left(1+\tau^c\right)^2}\left(1-\gamma\right) a \left(\frac{vk}{h}\right)^{\gamma}, \tag{54}$$

$$0 = u_c - \lambda_2 + \lambda_3 \frac{u_{cc}}{1+\tau^c} - \lambda_4 \left(\frac{\tau^c}{vk} + \frac{u_\ell}{u_c^2} u_{cc}\left(1+\tau^c\right)\frac{h}{vk}\right) + \lambda_5 \frac{u_\ell}{u_c^2} u_{cc}, \tag{55}$$

$$0 = u_\ell - \lambda_1 + \lambda_4 \frac{u_{\ell\ell}}{u_c}\left(1+\tau^c\right)\frac{h}{vk} - \lambda_5 \frac{u_{\ell\ell}}{u_c}, \tag{56}$$

$$0 = -\lambda_1 + \lambda_2 \left(1-\gamma\right) a \left(\frac{vk}{h}\right)^{\gamma} - \lambda_4 \left(a\left(1-\gamma\right) h^{-\gamma}\left(vk\right)^{\gamma-1} - \frac{u_\ell}{u_c}\left(1+\tau^c\right)\frac{1}{vk}\right)$$

$$- \lambda_5 \frac{1}{1+\tau^c}\left(1-\gamma\right)\gamma a\left(vk\right)^{\gamma} h^{-\gamma-1}, \tag{57}$$

$$0 = u_g - \lambda_2 + \lambda_4 \frac{1}{vk}, \tag{58}$$

$$
\begin{aligned}
0 = {} & \beta \mathbb{E} \frac{\partial \mathcal{W}(a', k')}{\partial k'} - \lambda_2 - \beta \lambda_3 \mathbb{E} \left\{ \frac{\frac{\mathcal{U}_c(a',k')}{1+\mathcal{T}^c(a',k')}}{\partial k'} \left[1 - \delta(\mathcal{V}(a', k')) \right. \right. \\
& \left. + a' \frac{\mathcal{H}(a', k')^{1-\gamma} \mathcal{V}(a', k')^{\gamma} k'^{\gamma}}{k'} - \frac{\mathcal{G}(a', k')}{k'} + \frac{\mathcal{T}^c(a', k') \mathcal{C}(a', k')}{k'} \right] \\
& + \frac{\mathcal{U}_c(a', k')}{1 + \mathcal{T}^c(a', k')} \left[- \frac{\partial \delta(\mathcal{V}(a', k'))}{\partial k'} + a' \frac{\partial \frac{\mathcal{H}(a',k')^{1-\gamma} \mathcal{V}(a',k')^{\gamma} k'^{\gamma}}{k'}}{\partial k'} - \frac{\partial \frac{\mathcal{G}(a',k')}{k'}}{\partial k'} + \frac{\partial \frac{\mathcal{T}^c(a',k')\mathcal{C}(a',k')}{k'}}{\partial k'} \right] \\
& \left. - \frac{\partial \mathcal{U}_\ell(a', k') \frac{\mathcal{H}(a',k')}{k'}}{\partial k'} \right\}, \qquad (59)
\end{aligned}
$$

$$
\begin{aligned}
0 = {} & \lambda_2 \left(a \gamma v^{\gamma - 1} k^{\gamma} h^{1-\gamma} - \delta_v k \right) + \lambda_4 \left(a(1-\gamma) v^{\gamma-2} \left(\frac{h}{k} \right)^{1-\gamma} - \frac{g - \tau^c c}{v^2 k} - \frac{u_\ell}{u_c} (1 + \tau^c) \frac{h}{v^2 k} + \delta_{vv} \right) \\
& + \lambda_5 \frac{1}{1 + \tau^c} a(1-\gamma) \gamma v^{\gamma-1} \left(\frac{k}{h} \right)^{\gamma}, \qquad (60)
\end{aligned}
$$

where

$$
\begin{aligned}
\frac{\partial \delta(\mathcal{V}(a', k'))}{\partial k'} &= \delta_v \frac{\partial \mathcal{V}(a', k')}{\partial k'}, \\
\frac{\frac{\mathcal{U}_c(a',k')}{1+\mathcal{T}^c(a',k')}}{\partial k'} &= \frac{\frac{\partial \mathcal{U}_c(a',k')}{\partial k'} (1 + \mathcal{T}^c(a', k')) - \frac{\partial \mathcal{T}^c(a',k')}{\partial k'} \mathcal{U}_c(a', k')}{(1 + \mathcal{T}^c(a', k'))^2}, \\
\frac{\frac{\mathcal{H}(a',k')^{1-\gamma}\mathcal{V}(a',k')^{\gamma}k'^{\gamma}}{k'}}{\partial k'} &= \mathcal{V}(a', k')^{\gamma-1} \mathcal{H}(a', k')^{-\gamma} k'^{\gamma-2} \left((1-\gamma) \mathcal{V}(a', k') k' \frac{\partial \mathcal{H}(a', k')}{\partial k'} \right. \\
& \quad \left. + \gamma \mathcal{H}(a', k') k' \frac{\partial \mathcal{V}(a', k')}{\partial k'} - (1-\gamma) \mathcal{V}(a', k') \mathcal{H}(a', k') \right), \\
\frac{\frac{\mathcal{G}(a',k')}{k'}}{\partial k'} &= \frac{\frac{\partial \mathcal{G}(a',k')}{\partial k'} k' - \mathcal{G}(a', k')}{k'^2}, \\
\frac{\partial \frac{\mathcal{T}^c(a',k')\mathcal{C}(a',k')}{k'}}{\partial k'} &= \frac{\left[\frac{\partial \mathcal{T}^c(a',k')}{\partial k'} \mathcal{C}(a', k') + \frac{\partial \mathcal{C}(a',k')}{\partial k'} \mathcal{T}^c(a', k') \right] k' - \mathcal{T}^c(a', k') \mathcal{C}(a', k')}{k'^2}, \\
\frac{\partial \mathcal{U}_\ell(a', k') \frac{\mathcal{H}(a',k')}{k'}}{\partial k'} &= \frac{\left[\frac{\partial \mathcal{U}_\ell(a',k')}{\partial k'} \mathcal{H}(a', k') + \frac{\partial \mathcal{H}(a',k')}{\partial k'} \mathcal{U}_\ell(a', k') \right] k' - \mathcal{U}_\ell(a', k') \mathcal{H}(a', k')}{k'^2}.
\end{aligned}
$$

Applying the envelope theorem gives

$$
\begin{aligned}
\frac{\partial \mathcal{W}(a, k)}{\partial k} = {} & \lambda_2 \left[a \gamma v^{\gamma} \left(\frac{h}{k} \right)^{1-\gamma} + 1 - \delta(v) \right] + \lambda_4 \left(a(1-\gamma) \left(\frac{h}{v} \right)^{1-\gamma} k^{\gamma-2} - \frac{g - \tau^c c}{v k^2} \right. \\
& \left. - \frac{u_\ell}{u_c} (1 + \tau^c) \frac{h}{v k^2} \right) + \lambda_5 \frac{1}{1 + \tau^c} a(1-\gamma) \gamma \frac{v^{\gamma}}{k^{1-\gamma} h^{\gamma}},
\end{aligned}
$$

hence,

$$\frac{\partial \mathcal{W}(a',k')}{\partial k'} = \Lambda_2(a',k')\left[a'\gamma\mathcal{V}(a',k')^{\gamma}\left(\frac{\mathcal{H}(a',k')}{k'}\right)^{1-\gamma}+1-\delta(\mathcal{V}(a',k'))\right]$$
$$+\Lambda_4(a',k')\left(a'(1-\gamma)\left(\frac{\mathcal{H}(a',k')}{\mathcal{V}(a',k')}\right)^{1-\gamma}k'^{\gamma-2}\right.$$
$$\left.-\frac{\mathcal{G}(a',k')-\mathcal{T}^c(a',k')\mathcal{C}(a',k')}{\mathcal{V}(a',k')k'^2}-\frac{\mathcal{U}_\ell(a',k')}{\mathcal{U}_c(a',k')}(1+\mathcal{T}^c(a',k'))\frac{\mathcal{H}(a',k')}{\mathcal{V}(a',k')k'^2}\right)$$
$$+\Lambda_5(a',k')\frac{1}{1+\mathcal{T}^c(a',k')}a'(1-\gamma)\gamma\frac{\mathcal{V}(a',k')^{\gamma}}{k'^{1-\gamma}\mathcal{H}(a',k')^{\gamma}}.$$

Plugging this condition into (59), we obtain

$$0=-\lambda_2+\beta\mathbb{E}\left\{\Lambda_2(a',k')\left[a'\gamma\mathcal{V}(a',k')^{\gamma}\left(\frac{\mathcal{H}(a',k')}{k'}\right)^{1-\gamma}+1-\delta(\mathcal{V}(a',k'))\right]\right.$$
$$+\Lambda_4(a',k')\left(a'(1-\gamma)\left(\frac{\mathcal{H}(a',k')}{\mathcal{V}(a',k')}\right)^{1-\gamma}k'^{\gamma-2}-\frac{\mathcal{G}(a',k')-\mathcal{T}^c(a',k')\mathcal{C}(a',k')}{\mathcal{V}(a',k')k'^2}\right.$$
$$\left.\left.-\frac{\mathcal{U}_\ell(a',k')}{\mathcal{U}_c(a',k')}(1+\mathcal{T}^c(a',k'))\frac{\mathcal{H}(a',k')}{\mathcal{V}(a',k')k'^2}\right)+\Lambda_5(a',k')\frac{1}{1+\mathcal{T}^c(a',k')}a'(1-\gamma)\gamma\frac{\mathcal{V}(a',k')^{\gamma}}{k'^{1-\gamma}\mathcal{H}(a',k')^{\gamma}}\right\}$$
$$-\beta\lambda_3\mathbb{E}\left\{\frac{\partial\frac{\mathcal{U}_c(a',k')}{1+\mathcal{T}^c(a',k')}}{\partial k'}\left[1-\delta(\mathcal{V}(a',k'))\quad+a'\frac{\mathcal{H}(a',k')^{1-\gamma}\mathcal{V}(a',k')^{\gamma}k'^{\gamma}}{k'}-\frac{\mathcal{G}(a',k')}{k'}+\frac{\mathcal{T}^c(a',k')\mathcal{C}(a',k')}{k'}\right]\right.$$
$$+\frac{\mathcal{U}_c(a',k')}{1+\mathcal{T}^c(a',k')}\left[-\frac{\partial\delta(\mathcal{V}(a',k'))}{\partial k'}+a'\frac{\partial\frac{\mathcal{H}(a',k')^{1-\gamma}\mathcal{V}(a',k')^{\gamma}k'^{\gamma}}{k'}}{\partial k'}-\frac{\partial\frac{\mathcal{G}(a',k')}{k'}}{\partial k'}+\frac{\partial\frac{\mathcal{T}^c(a',k')\mathcal{C}(a',k')}{k'}}{\partial k'}\right]$$
$$\left.-\frac{\partial\mathcal{U}_\ell(a',k')\frac{\mathcal{H}(a',k')}{k'}}{\partial k'}\right\}. \tag{61}$$

Finally, the first-order conditions with respect to $\lambda_1,\lambda_2,\lambda_3$, λ_5, and λ_4, respectively, are

$$0=\ell+h-1, \tag{62}$$
$$0=c+g+k'-a(vk)^{\gamma}h^{1-\gamma}-(1-\delta(v))k, \tag{63}$$
$$0=-\frac{u_c}{1+\tau^c}+\beta\mathbb{E}\left[\frac{\mathcal{U}_c(a',k')}{1+\mathcal{T}^c(a',k')}\left(1-\delta(\mathcal{V}(a',k'))+a'\frac{\mathcal{H}(a',k')^{1-\gamma}\mathcal{V}(a',k')^{\gamma}k'^{\gamma}}{k'}\right.\right.$$
$$\left.\left.-\frac{\mathcal{G}(a',k')-\mathcal{T}^c(a',k')\mathcal{C}(a',k')}{k'}\right)-\mathcal{U}_\ell(a',k')\frac{\mathcal{H}(a',k')}{k'}\right], \tag{64}$$
$$0=a\left(\frac{h}{vk}\right)^{1-\gamma}-\frac{g-\tau^c c}{vk}-\frac{u_\ell}{u_c}(1+\tau^c)\frac{h}{vk}-\delta_v, \tag{65}$$
$$0\geq\frac{u_\ell}{u_c}-\frac{1}{1+\tau^c}(1-\gamma)a\left(\frac{vk}{h}\right)^{\gamma}, \tag{66}$$

$\lambda_5\geq 0$, with complementary slackness condition.

C Proofs

C.1 Proof of Proposition 1

We know, e.g. from Coleman (2000), that with $v = 1$ the Ramsey steady state corresponds to the first-best steady state. We want to show that this result extends to the case with endogenous v. It is obvious that $\tau_t^c = \tau_{t+1}^c$ at the steady state. Then, comparing (39) and (51) gives $\tau^k = 0$. Also, comparing (37) and (13) gives $\frac{1-\tau^h}{1+\tau^c} = 1$, hence $\tau^h = -\tau^c$. Finally, as long as consumption is larger than labour income, the only way to raise revenue to finance g is by setting $\tau^c > 0$. Given that $\tau^k = 0$, the capital utilisation margin is not distorted, hence the first-best steady state can be implemented. $\square$

C.2 Proof of Proposition 2

By the usual Kuhn-Tucker argument, if (53) is not satisfied when it is ignored we can impose it as equality, hence $\tau^h = 0$. Next, note that at the steady state combining (51) and (53) as equality gives

$$\frac{1}{\beta} = 1 - \delta(v) + \gamma v^{\gamma} \left(\frac{h}{k}\right)^{1-\gamma} - \frac{g - \tau^c c}{k}.$$

Then, using this and (53) as equality again, we can rewrite (47) as

$$0 = \lambda_5 \gamma \frac{u_\ell}{u_c} + \left(\lambda_2 - \lambda_3 \frac{u_c}{1+\tau^c} \frac{1}{k} - \lambda_4 \frac{1}{vk}\right)(g - \tau^c c).$$

Now, from (46) $\lambda_2 - \lambda_3 \frac{u_c}{1+\tau^c} \frac{1}{k} - \lambda_4 \frac{1}{vk} = u_g > 0$, hence $g = \tau^c c$ if $\lambda_5 = 0$. Finally, from (45), this holds if $\lambda_1 = \lambda_2 (1-\gamma) \left(\frac{vk}{h}\right)^{\gamma} = \lambda_2 w$, which obviously holds. $\square$

Figure 1: Ramsey (red dashed line) and Markov (solid blue line) without consumption taxation starting from the Markov steady state

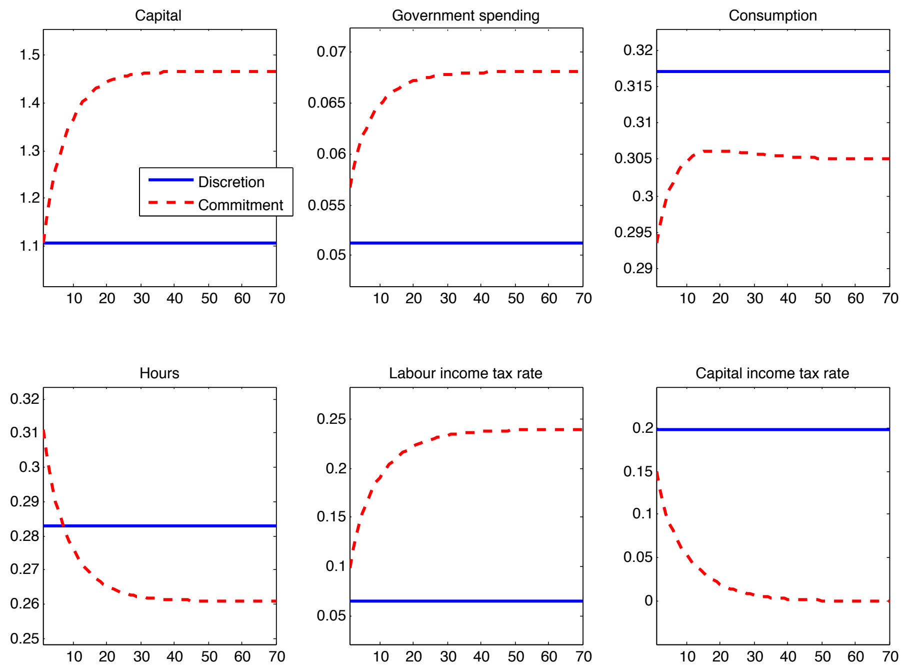

Figure 2: Ramsey (red dashed line) and Markov (solid blue line) policies with consumption taxation and no labour subsidy starting from the Markov steady state

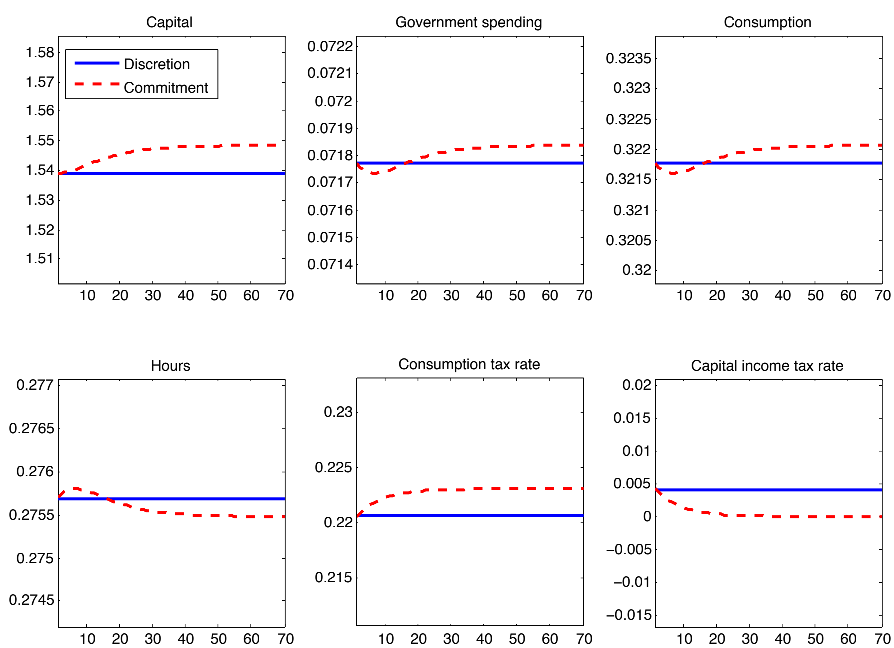

Figure 3: Ramsey (red dashed line) and Markov (solid blue line) policies without consumption taxation starting from the status quo

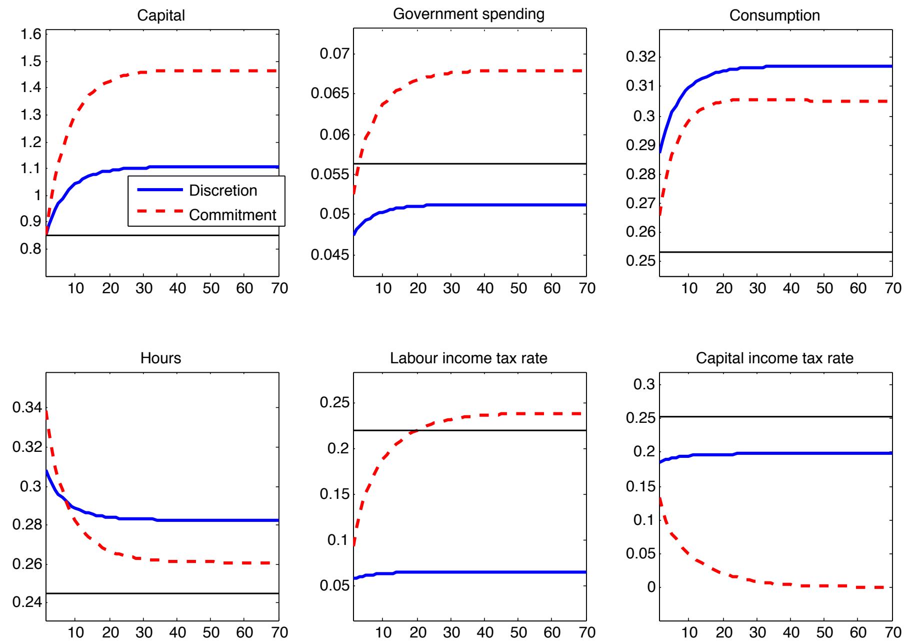

Figure 4: Ramsey (red dashed line) and Markov (solid blue line) policies with consumption taxation and no labour subsidy starting from the status quo

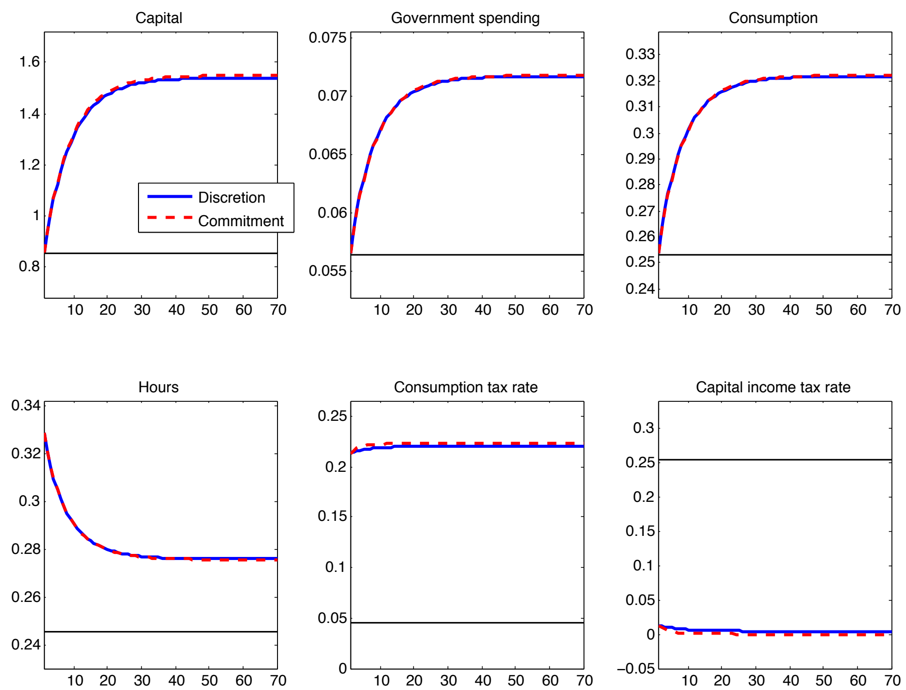

Table 5: Tax rates and allocations at steady state with alternative parameter values

Variable	unrestricted	$\tau^h \geq 0$		$\tau^c = 0$	
		Ramsey	Markov	Ramsey	Markov
$\varphi = 0.4$					
Consumption tax rate	3.245	0.223	0.221	0.000	0.000
Labour income tax rate	-3.245	0.000	0.000	0.240	0.173
Capital income tax rate	0.000	0.000	0.003	0.000	0.078
Capital	1.553	1.466	1.459	1.436	1.273
Hours worked	0.276	0.261	0.261	0.255	0.259
Income	0.494	0.466	0.465	0.457	0.439
Consumption	0.323	0.305	0.305	0.300	0.298
Public spending-income ratio	0.146	0.146	0.146	0.146	0.136
Welfare-eq. consumption loss	0.000	0.022	0.023	0.032	0.069
$\varphi = 1$					
Consumption tax rate	3.245	0.223	0.220	0.000	0.000
Labour income tax rate	-3.245	0.000	0.000	0.240	0.125
Capital income tax rate	0.000	0.000	0.005	0.000	0.133
Capital	1.648	1.490	1.480	1.437	1.181
Hours worked	0.293	0.265	0.265	0.256	0.265
Income	0.524	0.474	0.472	0.457	0.433
Consumption	0.343	0.310	0.310	0.299	0.302
Public spending-income ratio	0.146	0.146	0.146	0.146	0.128
Welfare-eq. consumption loss	0.000	0.039	0.041	0.058	0.119
$\varphi = 5$					
Consumption tax rate	3.245	0.223	0.221	0.000	0.000
Labour income tax rate	-3.245	0.000	0.000	0.240	0.042
Capital income tax rate	0.000	0.000	0.003	0.000	0.221
Capital	1.887	1.595	1.588	1.502	1.099
Hours worked	0.336	0.284	0.284	0.267	0.295
Income	0.600	0.507	0.506	0.477	0.449
Consumption	0.392	0.332	0.332	0.312	0.329
Public spending-income ratio	0.146	0.146	0.146	0.146	0.112
Welfare-eq. consumption loss	0.000	0.066	0.067	0.098	0.208
$\sigma = 2$					
Consumption tax rate	3.245	0.234	0.224	0.000	0.000
Labour income tax rate	-3.245	0.000	0.000	0.253	0.099
Capital income tax rate	0.000	0.000	0.010	0.000	0.197
Capital	1.522	1.402	1.380	1.358	1.015
Hours worked	0.271	0.249	0.250	0.242	0.259
Income	0.484	0.446	0.443	0.432	0.402
Consumption	0.317	0.289	0.289	0.279	0.282
Public spending-income ratio	0.146	0.152	0.150	0.154	0.137
Welfare-eq. consumption loss	0.000	0.031	0.035	0.046	0.147

Tax rates and allocations at steady state with alternative parameter values (continued)

Variable	unrestricted	$\tau^h \geq 0$ Ramsey	$\tau^h \geq 0$ Markov	$\tau^c = 0$ Ramsey	$\tau^c = 0$ Markov
$\beta = 0.96$					
Consumption tax rate	3.245	0.223	0.222	0.000	0.000
Labour income tax rate	-3.245	0.000	0.000	0.240	0.059
Capital income tax rate	0.000	0.000	0.002	0.000	0.202
Capital	3.352	2.882	2.872	2.719	2.039
Hours worked	0.322	0.277	0.277	0.261	0.283
Income	0.732	0.629	0.628	0.593	0.557
Consumption	0.479	0.412	0.411	0.388	0.404
Public spending-income ratio	0.146	0.146	0.146	0.146	0.115
Welfare-eq. consumption loss	0.000	0.059	0.060	0.088	0.187
$\chi = 1.8$					
Consumption tax rate	4.628	0.223	0.223	0.000	0.000
Labour income tax rate	-4.628	0.000	0.000	0.234	0.081
Capital income tax rate	0.000	0.000	0.0004	0.000	0.172
Capital	1.669	1.435	1.434	1.366	1.071
Hours worked	0.326	0.280	0.280	0.267	0.285
Income	0.583	0.502	0.502	0.478	0.452
Consumption	0.373	0.321	0.321	0.306	0.318
Public spending-income ratio	0.143	0.143	0.143	0.143	0.117
Welfare-eq. consumption loss	0.000	0.0560	0.0562	0.081	0.157
$\alpha_g = 0.3$					
Consumption tax rate	29.558	0.300	0.297	0.000	0.000
Labour income tax rate	-29.558	0.000	0.000	0.303	0.081
Capital income tax rate	0.000	0.000	0.005	0.000	0.233
Capital	1.885	1.548	1.537	1.438	1.032
Hours worked	0.335	0.275	0.276	0.256	0.284
Income	0.599	0.492	0.491	0.457	0.428
Consumption	0.369	0.303	0.303	0.281	0.302
Public spending-income ratio	0.185	0.185	0.185	0.185	0.141
Welfare-eq. consumption loss	0.000	0.088	0.091	0.135	0.264

Table 6: Cyclical properties of taxes and allocations, $\varphi = 1$

	$\tau^h \geq 0$		$\tau^c = 0$	
	Ramsey	Markov	Ramsey	Markov
	Consumption tax		Labour income tax	
Mean	0.223	0.220	0.240	0.125
Standard deviation	0.007	0.005	0.003	0.003
Coefficient of variation	0.029	0.022	0.014	0.026
Autocorrelation	0.576	0.594	0.673	0.501
Correlation with output	0.994	0.998	-0.948	-0.870
Capital income tax				
Mean	0.000	0.005	0.000	0.133
Standard deviation	0.011	0.008	0.006	0.003
Coefficient of variation	11.983	1.774	30.257	0.023
Autocorrelation	0.576	0.592	0.462	0.501
Correlation with output	-0.993	-0.997	-0.763	-0.890
Public spending-income ratio				
Mean	0.146	0.146	0.146	0.128
Standard deviation	0.005	0.004	0.004	0.003
Coefficient of variation	0.031	0.024	0.027	0.025
Autocorrelation	0.511	0.513	0.514	0.501
Correlation with output	-0.860	-0.859	-0.910	-0.877
Output				
Mean	0.474	0.473	0.457	0.433
Standard deviation	0.016	0.013	0.018	0.017
Coefficient of variation	0.034	0.027	0.040	0.040
Autocorrelation	0.616	0.619	0.589	0.601

Table 7: Tax rates and allocations at steady state with exogenous government spending

Variable	unrestricted	$\tau^h \geq 0$ Ramsey	$\tau^h \geq 0$ Markov	$\tau^c = 0$ Ramsey	$\tau^c = 0$ Markov
Consumption tax rate	1.179	0.167	0.165	0.000	0.000
Labour income tax rate	-1.179	0.000	0.000	0.240	0.078
Capital income tax rate	0.000	0.000	0.003	0.000	0.211
Capital	1.718	1.548	1.541	1.467	1.077
Capital utilisation	0.786	0.786	0.786	0.786	0.786
Depreciation rate	0.064	0.064	0.064	0.064	0.064
Hours worked	0.306	0.275	0.276	0.261	0.283
Income	0.546	0.492	0.491	0.466	0.434
Capital-income ratio	3.146	3.146	3.136	3.146	2.481
Consumption	0.381	0.337	0.337	0.305	0.309
Consumption-income ratio	0.697	0.686	0.686	0.654	0.712
Public spending (fixed)	0.056	0.056	0.056	0.056	0.056
Public spending-income ratio	0.103	0.115	0.115	0.121	0.130
Per-period utility	-1.622	-1.657	-1.659	-1.718	-1.765
Welfare-eq. consumption loss	0.000	0.036	0.037	0.039	0.088

Table 8: Tax rates and allocations at steady state with alternative measure of consumption

Variable	unrestricted	$\tau^h \geq 0$ Ramsey	$\tau^h \geq 0$ Markov	$\tau^c = 0$ Ramsey	$\tau^c = 0$ Markov
Consumption tax rate	4.718	0.245	0.242	0.000	0.000
Labour income tax rate	-4.718	0.000	0.000	0.259	0.070
Capital income tax rate	0.000	0.000	0.004	0.000	0.209
Capital	1.701	1.443	1.433	1.359	1.009
Capital utilisation	0.786	0.786	0.786	0.786	0.786
Depreciation rate	0.064	0.064	0.064	0.064	0.064
Hours worked	0.303	0.257	0.257	0.242	0.264
Income	0.541	0.459	0.458	0.432	0.405
Capital-income ratio	3.146	3.146	3.132	3.146	2.488
Consumption	0.347	0.295	0.294	0.278	0.291
Consumption-income ratio	0.643	0.643	0.643	0.643	0.718
Public spending	0.085	0.072	0.072	0.068	0.050
Public spending-income ratio	0.158	0.158	0.158	0.158	0.124
Per-period utility	-2.372	-2.437	-2.439	-2.467	-2.559
Welfare-eq. consumption loss	0.000	0.067	0.069	0.100	0.206