Economics Working Paper Series

2014/022

Time-Consistent Consumption Taxation

Sarolta Laczó and Raffaele Rossi

The Department of Economics
Lancaster University Management School
Lancaster LA1 4YX
UK

LUMS home page: http://www.lums.lancs.ac.uk/

Time-Consistent Consumption Taxation*

Sarolta Laczó† Raffaele Rossi‡

October 2014

Abstract

We characterise optimal fiscal policies when the government has access to consumption taxation but cannot credibly commit to future policies, in a calibrated Real Business Cycle model of the United States economy. 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, 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, and MMF in Durham 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. Rossi thanks the Institut d'Anàlisi Econòmica (IAE-CSIC) for hospitality during May and June 2014. All errors are our own.

†Institut d'Anàlisi Econòmica (IAE-CSIC) and Barcelona GSE, Campus UAB, 08193 Bellaterra, Barcelona, Spain. Email: sarolta.laczo@iae.csic.es.

‡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, in 2013 the value-added tax on standard items ranged from 15 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 a large welfare gain 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 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 policy 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, among 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. We exclude public debt dynamics for two main reasons. First, 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, Krusell, and Ríos-Rull, 2008; 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). Domínguez (2007) establishes a similar result when there are delays in tax policy implementation. 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.

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. 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. 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 similarities 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. 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 consumption by 7.744 percent. Without access to consumption taxation, the welfare gain is 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.

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, 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 first defines our policy equilibria of interest, (i) the Ramsey/full-commitment equilibrium and (ii) the Markov/time-consistent equilibrium. Afterwards, it characterises these equilibria 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 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 (see Greenwood, Hercowitz, and Huffman, 1988, Greenwood, Hercowitz, and Krusell, 2000, and many others). The firm maximises profits and uses capital services and labour as production inputs.

The policy-maker spends on public goods which yield utility to the household, 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.[1] Next we describe the economic environment in detail.

The representative household takes prices and policies as given and seeks to maximise

$$\mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t u\left(c_t, \ell_t, g_t\right) \right], \tag{1}$$

where $\mathbb{E}_0$ represents the rational-expectations operator at time 0, $\beta \in (0,1)$ is the discount factor, c_t is private consumption, ℓ_t represents leisure, and g_t is public consumption; subject to the time constraint

$$h_t + \ell_t = 1, \tag{2}$$

where h_t represents hours worked, and the budget constraint

$$\left(1+\tau_t^c\right) c_t + k_{t+1} = \left(1-\tau_t^k\right) r_t v_t k_t + \left(1-\tau_t^h\right) w_t h_t + \left(1-\delta\left(v_t\right)\right) k_t, \ \forall t, \tag{3}$$

[1] As discussed above, this is a common approach in the literature that studies Markov-perfect policies in models with capital accumulation (Klein and Ríos-Rull, 2003; Ortigueira, 2006; Klein, Krusell, and Ríos-Rull, 2008; Azzimonti, Sarte, and Soares, 2009; Martin, 2010; Debortoli and Nunes, 2010). Furthermore, we want to rule out the possibility of making the Ramsey policy time consistent through debt restructuring, see Zhu (1995).

where k_t is the level of the capital stock at the beginning of the period, $v_t \in (0, 1]$ is the capital utilisation rate, $\delta(v_t)$ represents the depreciation rate as a function of capital utilisation, τ_t^c, τ_t^h, and τ_t^k denote the consumption, the labour income, and the capital income tax rate, respectively. Finally, the variables r_t and w_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 $u_{c,t} > 0$, $u_{cc,t} < 0$, $u_{\ell,t} > 0$, $u_{\ell\ell,t} < 0$, $u_{g,t} > 0$, $u_{gg,t} < 0$, where $u_{x,t}$ and $u_{xx,t}$ denote, respectively, the first and the second derivative of the utility function with respect to the variable x at time t.

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 a 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.[2,3]

Combining the first-order conditions with respect to consumption and leisure at time t gives

$$\frac{u_{\ell,t}}{u_{c,t}} = \frac{1 - \tau_t^h}{1 + \tau_t^c} w_t. \tag{4}$$

Defining the price at time t as

$$q_t = \prod_{s=0}^{t} \frac{1 - \delta(v_0) + \left(1 - \tau_0^k\right) v_0 r_0}{1 - \delta(v_s) + \left(1 - \tau_s^k\right) v_s r_s} \tag{5}$$

and taking first-order conditions with respect to consumptions at different points in time, the Euler equation between 0 and any t can be expressed as

$$\frac{u_{c,0}}{1 + \tau_0^c} = \beta^t \mathbb{E}_0 \left(\frac{u_{c,t}}{1 + \tau_t^c} \frac{1}{q_t} \right), \tag{6}$$

[2]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.

[3]Instead of endogenous capital utilization, 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).

and we also have

$$\frac{u_{c,t}}{1+\tau_t^c} = \beta \mathbb{E}_t \left[\frac{u_{c,t+1}}{1+\tau_{t+1}^c} \left(1 - \delta\left(v_{t+1}\right) + \left(1-\tau_{t+1}^k\right) v_{t+1} r_{t+1}\right) \right], \tag{7}$$

which is a standard Euler equation. The first-order condition with respect to v_t is

$$\delta_{v,t} = \left(1-\tau_t^k\right) r_t. \tag{8}$$

The optimal level of capital utilisation is where the marginal benefit from utilising more capital in terms of after-tax income equals the marginal cost of higher capital depreciation. Equation (8) 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 (forward-looking) Euler equation. 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. 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, $v_t k_t$, and hours, h_t, i.e.,

$$y_t = f\left(v_t k_t, h_t, a_t\right) = a_t \left(v_t k_t\right)^{\gamma} h_t^{1-\gamma}, \ \forall t, \tag{9}$$

where $\gamma \in [0,1]$ represents the capital-services elasticity of output, and a_t is total factor productivity at time t. Denoting by $f_{x,t}$ the derivative of the production function with respect to the variable x at time t, optimal behaviour in perfect competition implies

$$r_t = f_{vk,t} = \gamma a_t \left(\frac{h_t}{v_t k_t}\right)^{1-\gamma}, \ \forall t, \tag{10}$$

$$w_t = f_{h,t} = \left(1-\gamma\right) a_t, \left(\frac{v_t k_t}{h_t}\right)^{\gamma}, \ \forall t, \tag{11}$$

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

The resource constraint in this economy is

$$c_t + g_t + k_{t+1} = y_t + \left(1 - \delta\left(v_t\right)\right) k_t, \ \forall t, \tag{12}$$

where the initial level of capital, k_0, is given. Finally, the government's budget constraint is

$$g_t = \tau_t^k r_t v_t k_t + \tau_t^h w_t h_t + \tau_t^c c_t, \ \forall t. \tag{13}$$

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

Definition 1 (First best)**.** *The first-best equilibrium is a sequence of allocations* $\{g_t, c_t, \ell_t, h_t, k_{t+1}, v_t, y_t\}_{t=0}^{\infty}$ *that maximise (1) subject to household's time constraint, (2), the production function, (9), and the market clearing condition, (12),* $\forall t$*,* k_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 is a sequence of government policies,* $\left\{\tau_t^h, \tau_t^k, \tau_t^c, g_t\right\}_{t=0}^{\infty}$*, private sector allocations,* $\{c_t, \ell_t, h_t, k_{t+1}, v_t, y_t\}_{t=0}^{\infty}$*, and price vectors,* $\{q_t, w_t, r_t\}_{t=0}^{\infty}$*, satisfying,* $\forall t$*,*

(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), (7), and (8);*
- *the pricing equation, (5),*
- *the production function, (9), and the firm's optimality conditions, (10) and (11);*

(ii) market clearing, (12), and

(iii) the government's budget constraint, (13).

We can eliminate four variables, output (y_t) and prices (q_t, w_t, and r_t), and four equations ((9), (5), (10), and (11)) in the definition of competitive equilibria. Further, the government 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:

$$h_t + \ell_t = 1, \ \forall t, \tag{14}$$

$$\frac{u_{\ell,t}}{u_{c,t}} = \frac{1-\tau_t^h}{1+\tau_t^c} a_t (1-\gamma) \left(\frac{v_t k_t}{h_t}\right)^{\gamma}, \ \forall t, \tag{15}$$

$$\frac{u_{c,t}}{1+\tau_t^c} = \beta \mathbb{E}_t \left[\frac{u_{c,t+1}}{1+\tau_{t+1}^c} \left(1 - \delta(v_{t+1}) + \left(1-\tau_{t+1}^k\right) a_{t+1} \gamma v_{t+1} \left(\frac{h_{t+1}}{v_{t+1} k_{t+1}}\right)^{1-\gamma}\right)\right], \ \forall t, \tag{16}$$

$$\delta_{v,t} = \left(1-\tau_t^k\right) a_t \gamma \left(\frac{h_t}{v_t k_t}\right)^{1-\gamma}, \ \forall t, \tag{17}$$

$$g_t = a_t \left(\tau_t^k \gamma + \tau_t^h (1-\gamma)\right) (v_t k_t)^{\gamma} h_t^{1-\gamma} + \tau_t^c c_t, \ \forall t, \tag{18}$$

$$c_t + g_t + k_{t+1} = a_t (v_t k_t)^{\gamma} h_t^{1-\gamma} + (1 - \delta(v_t)) k_t, \ \forall t. \tag{19}$$

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. First, we provide definitions. Then we turn to characterising the different policy problems analytically.

Definition 3 (Ramsey equilibrium)**.** *A Ramsey equilibrium is a sequence of government policies* $\left\{\tau_t^h, \tau_t^k, \tau_t^c, g_t\right\}_{t=0}^{\infty}$ *and private sector allocations* $\{c_t, \ell_t, h_t, k_{t+1}, v_t\}_{t=0}^{\infty}$ *which solve*

$$\max_{\left\{\tau_t^h, \tau_t^k, \tau_t^c, g_t\right\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u\left(c_t, \ell_t, g_t\right) \tag{20}$$

s.t. (14)–(19) *hold,* $\forall t$*, given* k_0 *and the productivity process.*

Definition 4 (Markov equilibrium)**.** *A Markov equilibrium is a sequence of government policies* $\left\{\tau_t^h, \tau_t^k, \tau_t^c, g_t\right\}_{t=0}^{\infty}$ *and private sector allocations* $\{c_t, \ell_t, h_t, k_{t+1}, v_t\}_{t=0}^{\infty}$ *which solve*

$$V\left(k_t, a_t\right) = \max_{\left\{\tau_t^h, \tau_t^k, \tau_t^c, g_t\right\}} u\left(c_t, \ell_t, g_t\right) + \beta \mathbb{E}_t V\left(k_{t+1}, a_{t+1}\right) \tag{21}$$

s.t. (14)–(19) *hold, given* k_t *and the productivity process,* $\forall t$*.*

We use a version of the primal approach, i.e., we rewrite 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 later when we want to constrain the tax rates.[4] In order to do this, we use the household's intratemporal optimality condition (15), and the government's budget constraint, (18), to express the current labour and capital income tax rates, respectively, as

$$\tau_t^h = 1 - \frac{u_{\ell,t}}{u_{c,t}} \frac{1 + \tau_t^c}{\left(1 - \gamma\right) a_t \left(\frac{v_t k_t}{h_t}\right)^{\gamma}}, \tag{22}$$

$$\tau_t^k = \frac{g_t - \tau_t^c c_t}{\gamma a_t \left(v_t k_t\right)^{\gamma} h_t^{1-\gamma}} - \frac{1 - \gamma}{\gamma} \tau_t^h. \tag{23}$$

Replacing for $\left(1 - \tau_{t+1}^k\right)$ in (16) using (23) and in turn for τ_{t+1}^h using (22), we have

$$\begin{aligned} \frac{u_{c,t}}{1 + \tau_t^c} = \beta \mathbb{E}_t \Bigg[\frac{u_{c,t+1}}{1 + \tau_{t+1}^c} \Bigg(1 - \delta\left(v_{t+1}\right) + a_{t+1} v_{t+1} \left(\frac{h_{t+1}}{v_{t+1} k_{t+1}}\right)^{1-\gamma} \\ - \frac{g_{t+1} - \tau_{t+1}^c c_{t+1}}{k_{t+1}} \Bigg) - u_{\ell,t+1} \frac{h_{t+1}}{k_{t+1}} \Bigg]. \end{aligned} \tag{24}$$

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

Similarly, we can eliminate τ_t^k from (17) and rewrite it as

$$\delta_{v,t} = a_t \left(\frac{h_t}{v_t k_t}\right)^{1-\gamma} - \frac{g_t - \tau_t^c c_t}{v_t k_t} - \frac{u_{\ell,t}}{u_{c,t}} \left(1 + \tau_t^c\right) \frac{h_t}{v_t k_t}. \tag{25}$$

As a result, we have to maximise with respect to a sequence of consumption tax rates $\{\tau_t^c\}_{t=0}^{\infty}$ and allocations $\{g_t, c_t, \ell_t, h_t, k_{t+1}, v_t\}_{t=0}^{\infty}$ in both the Ramsey and Markov policy-makers' problem, subject to four conditions, (14), (19), (24), and (25), given k_0 or k_t and the productivity process.

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-negativity, 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 $\tau_t^h \geq 0$, we impose

$$\frac{u_{\ell,t}}{u_{c,t}} \leq \frac{1}{1 + \tau_t^c} \left(1 - \gamma\right) a_t \left(\frac{v_t k_t}{h_t}\right)^{\gamma}, \ \forall t. \tag{26}$$

Below we write the policy problems in a general form including the constraint (26). We will, however, also study the case where (26) is ignored and the case without consumption taxation, i.e., $\tau_t^c = 0$, $\forall t$, to compare our results with the existing literature.

3.1 The Ramsey policy-maker's problem

Note that future decision variables enter into the household's current Euler equation, hence the Ramsey problem is not recursive. Following Marcet and Marimon (2011), the Lagrange multiplier on that constraint, which we denote λ_3, can be introduced as a co-state variable to write a Bellman equation, which can then be solved numerically by standard policy function iteration. The recursive Lagrangian is

$$\begin{aligned}
\max_{\{\tau_t^c, c_t, \ell_t, h_t, g_t, k_{t+1}, v_t\}_{t=0}^{\infty}} \mathbb{E}_0 \Bigg[\sum_{t=0}^{\infty} \beta^t \Bigg\{ & u\left(c_t, \ell_t, g_t\right) - \lambda_{1,t} \left(\ell_t + h_t - 1\right) \\
& - \lambda_{2,t} \left(c_t + g_t + k_{t+1} - a_t \left(v_t k_t\right)^{\gamma} h_t^{1-\gamma} - \left(1 - \delta\left(v_t\right)\right) k_t\right) + \lambda_{3,t} \frac{u_{c,t}}{1 + \tau_t^c} \\
& - \lambda_{3,t-1} \left[\frac{u_{c,t}}{1 + \tau_t^c} \left(1 - \delta\left(v\right) + a_t v_t \left(\frac{h_t}{v_t k_t}\right)^{1-\gamma} - \frac{g_t - \tau_t^c c_t}{k_t}\right) - u_{\ell,t} \frac{h_t}{k_t} \right] \\
& - \lambda_{4,t} \left(a_t \left(\frac{h_t}{v_t k_t}\right)^{1-\gamma} - \frac{g_t - \tau_t^c c_t}{v_t k_t} - \frac{u_{\ell,t}}{u_{c,t}} \left(1 + \tau_t^c\right) \frac{h_t}{v_t k_t} - \delta_{v,t} \right) \\
& - \lambda_{5,t} \left(\frac{u_{\ell,t}}{u_{c,t}} - \frac{1}{1 + \tau_t^c} a_t \left(1 - \gamma\right) \left(\frac{v_t k_t}{h_t}\right)^{\gamma} \right) \Bigg\} \Bigg],
\end{aligned}$$

$\lambda_{1,t} \geq 0$, $\lambda_{2,t} \geq 0$, and $\lambda_{5,t} \geq 0$, with complementary slackness conditions, k_0 given and $\lambda_{3,-1} = 0$. It is obvious that the time constraint, (14), and the resource constraint, (19), will bind. Hence, $\lambda_{1,t} > 0$ and $\lambda_{2,t} > 0$, $\forall t$.

The value function and the policy functions are time-invariant on the extended state space (a, k, λ_3), where a variable without time index denotes the state taken as given in the current period. Next period's productivity a' is given exogenously, while the policy-maker chooses k' and λ_3'. We drop the time index of control variables chosen in the current period, and denote by c', h', and so on the control variables chosen next period. Then we can write the value function as

$$V(a, k, \lambda_3) = u(c(a, k, \lambda_3), \ell(a, k, \lambda_3), g(a, k, \lambda_3)) + \beta \sum_{a'} \Pr(a' \mid a) V(a', k'(a, k, \lambda_3) \lambda_3'(a, k, \lambda_3)). \quad (27)$$

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, the policy-maker moves first in every period, and commits within the period.[5]

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 $\{g(a, k), \tau^c(a, k), \tau^h(a, k), \tau^k(a, k)\}$, then she finds it optimal to do the same.

As in the case of a Ramsey policy-maker, we use a quasi-primal approach. That is, we are looking for the policy rules $\{\tau^c(a, k), c(a, k), \ell(a, k), h(a, k), g(a, k), v(a, k), k'(a, k)\}$.

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

We can write the Lagrangian as

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

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

$$
V(a',k') = u(c(a',k'), \ell(a',k'), g(a',k')) + \beta \mathbb{E} V(a'', k''(a',k')),
$$

and where the dependence of k' and all current control variables on (a,k) is not made explicit. 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 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, (14), and the resource constraint, (19), because the current consumption tax rate can be chosen to satisfy the household's Euler equation, (16). 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, (17), 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 a the steady state.[6]

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

Proof. In Appendix C.1.

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 tax 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 specify the utility function as

$$u\left(c,\ell,g\right) = \log(c) - \alpha_\ell \frac{(1-\ell)^{1+1/\varphi}}{1+1/\varphi} + \alpha_g \log(g), \tag{28}$$

where φ is the Frisch elasticity of labour supply, while α_ℓ and α_g are the weights of leisure and public goods relative to private consumption, respectively. We assume that the depreciation rate is an increasing and convex function of capital utilisation. That is, $\delta(v) = \eta v^\chi$, with

[6] 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.

$\eta > 0$ and $\chi > 1$. We first calibrate most of the parameters to the deterministic steady state. Afterwards, we calibrate the productivity process.

We consider the model period to be a year. 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).[7] We calibrate the rest of the parameters using the private sector's first-order conditions at steady state, taking as given average effective tax rates, to match United States macroeconomic data for the period 1996-2010. 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 $\tau^k_\delta = 0.410$, where the lower index δ means that this capital tax rate is with depreciation allowance.[8]

We set $v = 0.786$ to match average capacity utilisation for all industries for the period 1996-2010 (source: Federal Reserve Board). The rest of the moments we match are computed using data provided by Trabandt and Uhlig (2012), who collected data from the OECD and the Ameco Database of the European Commission.[9]

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 $c/y = 0.696$,[10] $g/y = 0.155$, and $k/y = 2.349$, and the steady-state resource constraint,

$$\frac{c}{y} + \frac{g}{y} + \eta v^\chi \frac{k}{y} = 1, \tag{29}$$

to find $\delta(v) = \eta v^\chi = 0.064$. Second, we convert τ^k_δ to a without-depreciation-allowance capital tax rate keeping tax revenue constant, i.e., we assume that $\tau^k_\delta \left(rv - \delta(v)\right) = \tau^k rv = \tau^k \gamma \frac{y}{vk}$. 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,[11] 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,

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

[7] 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, we check the robustness of our results to a wide range of values of φ.

[8] 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.

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

[10] 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).

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

to find $\beta = 0.943$.[12] Fifth, multiplying the optimality condition for capital utilisation, (17), at steady state by v, we have

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

This gives $\chi = 1.956$. Then, from $\eta v^{\chi} = 0.064$, $\eta = 0.102$. Finally, we assume that g found in the data is optimally chosen, i.e., that $u_c = u_g$, and use the macro ratios $c/y = 0.696$ and $g/y = 0.155$ again to find $\alpha_g = 0.223$. The calibrated parameter values are presented in Table 1. Note that private consumption is larger than labour income, as in the data (69.6 vs. 60.9 percent of GDP), hence the consumption tax base is larger than the labour income tax base. Note also that to satisfy the household's (and the government's) budget constraint, the government gives a lump-sum transfer of 11.7 percent of GDP to the household.

Finally, we calibrate the AR(1) coefficients of the technological progress, i.e., ρ and σ_a, so that the unconditional persistence and standard deviation of total output in our economy with fixed tax rates match those of the de-trended US GDP for the period $1996 - 2010$. As a result, we set $\rho = 0.619$ and $\sigma_a = 0.020$.

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 first use a standard policy-function-iteration algorithm to solve the Ramsey problem. This consists of the following steps. First we discretise the state variables $k \in \left[\underline{k}, \overline{k}\right]$ and $\lambda_3 \in \left[\underline{\lambda}_3, \overline{\lambda}_3\right]$ 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 Markov chain following Tauchen (1986).[13] Once we have found the endogenous collocation

[12] 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.

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

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

nodes, we guess the policy functions of interest at each grid point. Then we solve the system of non-linear equations, and we approximate globally the policy functions via cubic spline collocation method. We iterate till convergence is obtained at each grid point. 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 parametrise the policy functions at each endogenous collocation node.[14] We iterate until the policies at each grid point converge.[15]

4.3 Results in the deterministic case

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.

Table 2: Tax rates and allocations at steady state

	unrestricted	$\tau^h \geq 0$		$\tau^c = 0$	
Variable		Ramsey	Markov	Ramsey	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

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

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

[15]In Section 4.5 we verify that our results are robust to using different solution algorithms.

22.3 percent at the steady state and sets the labour and capital income taxes to zero.[16] The 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. 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 the steady-state results 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

[16]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.

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. With access to consumption taxation but imposing a non-negativity constraint on the labour income tax rate, the steady-state welfare loss amounts to 6.1 percent in the case of time-consistent policy, and 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.

In order to get a sense of policy dynamics, 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 $t > 0$, 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 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 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 the time-consistent policy-maker 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 higher than τ^h.

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 to future generations. 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.

We can now compute the welfare gains from taxing consumption and from commitment taking into account the transition. First, we compute the gains from taxing consumption given a time-consistent policy-maker, starting from the Markov steady state without con-

sumption taxation. We find that the welfare gains in terms of welfare-equivalent consumption are 2.77 percent. 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.

Next, we compute 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 with taxing labour income instead. Hence, the gains from commitment are negligible with access to consumption taxation, while they are substantial without.

Finally, we can also compute the welfare gains from the different taxation and commitment scenarios compared to the existing tax system. 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 allocations and the tax rates without and with access to consumption taxation, respectively, for both a Ramsey and a time-consistent policy-maker. 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. For each policy scenario, we simulate the model

and calculate sample statistics from the simulated data.[17] The results of this exercise are reported in Table 4.

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 labour income taxation is more than three times larger and the coefficient of variation of capital income tax is more than a thousand times smaller than their Ramsey counterparts. This is mainly due to the fact that the Markov policy-maker is constrained to absorb a shock within the period, as any attempt to smooth the effects of random productivity events via future taxation would not be credible. 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.[18]

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 the stochastic allocation. 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, unlike without consumption taxation, without commitment capital taxes still play the main role in absorbing shocks. 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.

[17] 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.

[18] 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.

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

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 necessary to make the representative household as well off as at the first best.[19] 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 show that our results are robust to changing some parameter values and to using different solution algorithms. 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 we consider, the Ramsey planner finances all public expenditure with either only labour income or only consumption taxation. Under discretion, the changes in tax rates are very small if consumption taxation is available. 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 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.

[19] 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.

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 decides to have lower steady-state 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 every close to its efficient level.

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

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

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.

Then, 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.

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 second solution algorithm we iterate until the parameters of the policy functions converge. 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. Further, in a deterministic setting, a close-to-constant consumption tax rate is enough, and we do not have to rely on a rich structure of public debt, as in Zhu (1995) and Domínguez (2007).

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 revenue. 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

The first best allocation can be found by maximising the following Lagrangian:

$$\max_{\{c_t,\ell_t,h_t,g_t,k_{t+1},v_t\}_{t=0}^{\infty}} \mathbb{E}_0 \left\{ \sum_{t=0}^{\infty} \beta^t \left[u\left(c_t,\ell_t,g_t\right) + \lambda_{1,t}\left(1-\ell_t-h_t\right) \right.\right. \\ \left.\left. +\lambda_{2,t}\left(a_t\left(v_t k_t\right)^{\gamma} h_t^{1-\gamma} + \left(1-\delta\left(v_t\right)\right)k_t - c_t - g_t - k_{t+1}\right)\right]\right\},$$

where we have used (9) to replace for y_t in (12). The first-order conditions with respect to $c_t, \ell_t, h_t, g_t, k_{t+1}, u_t, \lambda_{1,t}$, and $\lambda_{2,t}$ are

$$\begin{aligned}
u_{c,t} &= \lambda_{2,t} && (30)\\
u_{\ell,t} &= \lambda_{1,t} && (31)\\
a_t\left(1-\gamma\right)\left(\frac{v_t k_t}{h_t}\right)^{\gamma} &= \lambda_{1,t} && (32)\\
u_{g,t} &= \lambda_{2,t} && (33)\\
\lambda_{2,t} &= \beta\mathbb{E}_t\left[\lambda_{1,t+1}\left(a_{t+1}\gamma v_{t+1}^{\gamma}\left(\frac{h_{t+1}}{k_{t+1}}\right)^{1-\gamma} + 1 - \delta\left(v_{t+1}\right)\right)\right] && (34)\\
\delta_{u,t} &= a_t\gamma\left(\frac{h_t}{v_t k_t}\right)^{1-\gamma}, && (35)\\
\ell_t + h_t &= 1 && (36)\\
c_t + g_t + k_{t+1} &= a_t\left(v_t k_t\right)^{\gamma} h_t^{1-\gamma} + \left(1-\delta\left(v_t\right)\right)k_t && (37)
\end{aligned}$$

Straightforward combinations of (30)-(37) lead to the following equations which characterise the first-best allocation:

$$\begin{aligned}
u_{g,t} &= u_{c,t},\ \forall t, && (38)\\
\frac{u_{\ell,t}}{u_{c,t}} &= a_t\left(1-\gamma\right)\left(\frac{v_t k_t}{h_t}\right)^{\gamma},\ \forall t, && (39)\\
h_t + \ell_t &= 1,\ \forall t, && (40)\\
u_{c,t} &= \beta\mathbb{E}_t\left[u_{c,t+1}\left(1-\delta\left(v_{t+1}\right) + a_{t+1}\gamma v_{t+1}\left(\frac{h_{t+1}}{v_{t+1}k_{t+1}}\right)^{1-\gamma}\right)\right],\ \forall t, && (41)\\
c_t + g_t + k_{t+1} &= a_t\left(v_t k_t\right)^{\gamma} h_t^{1-\gamma} + \left(1-\delta\left(v_t\right)\right)k_t,\ \forall t, && (42)\\
\delta_{v,t} &= a_t\gamma\left(\frac{h_t}{v_t k_t}\right)^{1-\gamma},\ \forall t, && (43)
\end{aligned}$$

k_0 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 first-order conditions 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}{(1+\tau^c)^2}\left[-\lambda_3' u_c + \lambda_3 u_c\left(1-\delta(v)+av\left(\frac{h}{vk}\right)^{1-\gamma}-\frac{g-\tau^c c}{k}\right)\right. \\
& \left.-\lambda_5(1-\gamma)\,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} \qquad (44)\\
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^c c}{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}(1+\tau^c)\frac{h}{vk}\right)+\lambda_5\frac{u_\ell}{u_c^2}u_{cc} \qquad (45)\\
0 = {} & u_\ell-\lambda_1+\lambda_3 u_{\ell\ell}\frac{h}{k}+\lambda_4\frac{u_{\ell\ell}}{u_c}(1+\tau^c)\frac{h}{vk}-\lambda_5\frac{u_{\ell\ell}}{u_c} \qquad (46)\\
0 = {} & -\lambda_1+\lambda_2 a(1-\gamma)\left(\frac{vk}{h}\right)^{\gamma}-\lambda_3\left[\frac{u_c}{1+\tau^c}a(1-\gamma)\frac{v^\gamma}{h^\gamma k^{1-\gamma}}-\frac{u_\ell}{k}\right]\\
& -\lambda_4\left[a(1-\gamma)h^{-\gamma}(vk)^{\gamma-1}-\frac{u_\ell}{u_c}(1+\tau^c)\frac{1}{vk}\right]-\lambda_5\frac{1}{1+\tau^c}a\gamma(1-\gamma)(vk)^\gamma h^{-\gamma-1} \qquad (47)\\
0 = {} & u_g-\lambda_2+\lambda_3\frac{u_c}{1+\tau^c}\frac{1}{k}+\lambda_4\frac{1}{vk} \qquad (48)\\
0 = {} & -\lambda_2+\beta\mathbb{E}\left\{\lambda_2'\left(a'v'^{\gamma}\gamma\left(\frac{h'}{k'}\right)^{1-\gamma}+1-\delta(v')\right)\right.\\
& +\lambda_3'\left[\frac{u_c'}{1+\tau^{c\prime}}\left(a'v'^{\gamma}(1-\gamma)k'^{\gamma-2}h'^{1-\gamma}-\frac{g'-\tau^{c\prime}c'}{k'^2}\right)-u_\ell'\frac{h'}{k'^2}\right] \qquad (49)\\
& +\lambda_4'\left(a'(1-\gamma)k'^{\gamma-2}\left(\frac{h'}{v'}\right)^{1-\gamma}-\frac{g'-\tau^{c\prime}c'}{v'k'^2}-\frac{u_\ell'}{u_c'}(1+\tau^{c\prime})\frac{h'}{v'k'^2}\right)\\
& \left.+\lambda_5'\frac{1}{1+\tau^{c\prime}}a'(1-\gamma)\gamma\frac{v'^{\gamma}}{k'^{1-\gamma}h'^{\gamma}}\right\}
\end{aligned}
$$

$$0 = \lambda_2 \left(a\gamma v^{\gamma-1} k^{\gamma} h^{1-\gamma} - \delta_v k \right) + \lambda_3 \frac{u_c}{1+\tau^c} \left(\delta_v - a\gamma v^{\gamma-1} \left(\frac{h}{k} \right)^{1-\gamma} \right)$$
$$+ \lambda_4 \left(a\left(1-\gamma\right) v^{\gamma-2} \left(\frac{h}{k} \right)^{1-\gamma} - \frac{g - \tau^c c}{v^2 k} - \frac{u_\ell}{u_c} \left(1+\tau^c\right) \frac{h}{v^2 k} + \delta_{vv} \right) \tag{50}$$
$$+ \lambda_5 \frac{1}{1+\tau^c} a\left(1-\gamma\right) \gamma v^{\gamma-1} \left(\frac{k}{h} \right)^{\gamma}$$

$$0 = \ell + h - 1 \tag{51}$$

$$0 = c + g + k_{t+1} - a\left(vk\right)^{\gamma} h^{1-\gamma} - \left(1 - \delta\left(v\right)\right) k \tag{52}$$

$$0 = -\frac{u_c}{1+\tau^c} + \beta \mathbb{E} \left[\frac{u_c'}{1+\tau^{c'}} \left(1 - \delta\left(v'\right) + a' v' \left(\frac{h'}{v'k'} \right)^{1-\gamma} - \frac{g' - \tau^{c'} c'}{k'} \right) - u_\ell' \frac{h'}{k'} \right] \tag{53}$$

$$0 = a \left(\frac{h}{vk} \right)^{1-\gamma} - \frac{g - \tau^c c}{vk} - \frac{u_\ell}{u_c} \left(1+\tau^c\right) \frac{h}{vk} - \delta_v \tag{54}$$

$$0 \geq \frac{u_\ell}{u_c} - \frac{1}{1+\tau^c} a\left(1-\gamma\right) \left(\frac{vk}{h} \right)^{\gamma}, \tag{55}$$

$\lambda_5 \geq 0$, with complementary slackness condition, and $\lambda_{3,-1} = 0$.

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

The first-order conditions with respect to $\tau^c, c, \ell, h, g, k', v$, respectively, are

$$0 = -\lambda_3 \frac{1}{\left(1+\tau^c\right)^2} u_c - \lambda_4 \left(\frac{c}{vk} - \frac{u_\ell}{u_c} \frac{h}{vk} \right) - \lambda_5 \frac{1}{\left(1+\tau^c\right)^2} \left(1-\gamma\right) a \left(\frac{vk}{h} \right)^{\gamma} \tag{56}$$

$$0 = u_c - \lambda_2 + \lambda_3 \frac{u_{cc}}{1+\tau^c} - \lambda_4 \left(\frac{\tau^c}{vk} + \frac{u_\ell}{u_c^2} u_{cc} \left(1+\tau^c\right) \frac{h}{vk} \right) + \lambda_5 \frac{u_\ell}{u_c^2} u_{cc} \tag{57}$$

$$0 = u_\ell - \lambda_1 + \lambda_4 \frac{u_{\ell\ell}}{u_c} \left(1+\tau^c\right) \frac{h}{vk} - \lambda_5 \frac{u_{\ell\ell}}{u_c} \tag{58}$$

$$0 = -\lambda_1 + \lambda_2 \left(1-\gamma\right) a \left(\frac{vk}{h} \right)^{\gamma} - \lambda_4 \left(a\left(1-\gamma\right) h^{-\gamma} \left(vk\right)^{\gamma-1} - \frac{u_\ell}{u_c} \left(1+\tau^c\right) \frac{1}{vk} \right)$$
$$- \lambda_5 \frac{1}{1+\tau^c} \left(1-\gamma\right) \gamma a \left(vk\right)^{\gamma} h^{-\gamma-1} \tag{59}$$

$$0 = u_g - \lambda_2 + \lambda_4 \frac{1}{vk} \tag{60}$$

$$
\begin{aligned}
0 = \beta\mathbb{E}\frac{\partial V\left(a',k'\right)}{\partial k'} - \lambda_2 - \beta\lambda_3\mathbb{E}\Bigg\{ & \frac{\frac{u_c(a',k')}{1+\tau^c(a',k')}}{\partial k'}\left[1-\delta\left(v\left(a',k'\right)\right)\right. \\
& \left. + a'\frac{h\left(a',k'\right)^{1-\gamma}v\left(a',k'\right)^{\gamma}k'^{\gamma}}{k'} - \frac{g\left(a',k'\right)}{k'} + \frac{\tau^c\left(a',k'\right)c\left(a',k'\right)}{k'}\right] \\
& + \frac{u_c\left(a',k'\right)}{1+\tau^c\left(a',k'\right)}\left[-\frac{\partial\delta\left(v\left(a',k'\right)\right)}{\partial k'} + a'\frac{\partial\frac{h(a',k')^{1-\gamma}v(a',k')^{\gamma}k'^{\gamma}}{k'}}{\partial k'} - \frac{\partial\frac{g(a',k')}{k'}}{\partial k'} + \frac{\partial\frac{\tau^c(a',k')c(a',k')}{k'}}{\partial k'}\right] \\
& - \frac{\partial u_\ell\left(a',k'\right)\frac{h(a',k')}{k'}}{\partial k'}\Bigg\} \qquad (61)
\end{aligned}
$$

$$
\begin{aligned}
0 = {} & \lambda_2\left(a\gamma v^{\gamma-1}k^{\gamma}h^{1-\gamma} - \delta_v k\right) + \lambda_4\left(a\left(1-\gamma\right)v^{\gamma-2}\left(\frac{h}{k}\right)^{1-\gamma} - \frac{g-\tau^c c}{v^2 k} - \frac{u_\ell}{u_c}\left(1+\tau^c\right)\frac{h}{v^2 k} + \delta_{vv}\right) \\
& + \lambda_5\frac{1}{1+\tau^c}a\left(1-\gamma\right)\gamma v^{\gamma-1}\left(\frac{k}{h}\right)^{\gamma}, \qquad (62)
\end{aligned}
$$

where

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

Applying the envelope theorem

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

hence,

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

Plugging this condition into (61), we obtain

$$\begin{aligned}0 = & -\lambda_2+\beta\mathbb{E}\left\{\lambda_2(a',k')\left[a'\gamma v(a',k')^{\gamma}\left(\frac{h(a',k')}{k'}\right)^{1-\gamma}+1-\delta(v(a',k'))\right]\right.\\ & +\lambda_4(a',k')\left[a'(1-\gamma)\left(\frac{h(a',k')}{v(a',k')}\right)^{1-\gamma}k'^{\gamma-2}-\frac{g(a',k')-\tau^c(a',k')c(a',k')}{v(a',k')k'^2}\right.\\ & \left.-\frac{u_\ell(a',k')}{u_c(a',k')}(1+\tau^c(a',k'))\frac{h(a',k')}{v(a',k')k'^2}\right]+\lambda_5(a',k')\frac{1}{1+\tau^c(a',k')}a'(1-\gamma)\gamma\frac{v(a',k')^{\gamma}}{k'^{1-\gamma}h(a',k')^{\gamma}}\\ & -\lambda_3\left[\frac{\frac{u_c(a',k')}{1+\tau^c(a',k')}}{\partial k'}\left(1-\delta(v(a',k'))+a'\frac{h(a',k')^{1-\gamma}v(a',k')^{\gamma}k'^{\gamma}}{k'}-\frac{g(a',k')}{k'}+\frac{\tau^c(a',k')c(a',k')}{k'}\right)\right.\\ & +\frac{u_c(a',k')}{1+\tau^c(a',k')}\left(-\frac{\partial\delta(v(a',k'))}{\partial k'}+a'\frac{\partial\frac{h(a',k')^{1-\gamma}v(a',k')^{\gamma}k'^{\gamma}}{k'}}{\partial k'}-\frac{\partial\frac{g(a',k')}{k'}}{\partial k'}+\frac{\partial\frac{\tau^c(a',k')c(a',k')}{k'}}{\partial k'}\right)\\ & \left.\left.-\frac{\partial u_\ell(a',k')\frac{h(a',k')}{k'}}{\partial k'}\right]\right\}.\end{aligned} \tag{63}$$

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

$$0 = \ell + h - 1 \tag{64}$$

$$0 = c + g + k' - a(vk)^{\gamma}h^{1-\gamma} - (1-\delta(v))k \tag{65}$$

$$\begin{aligned}0 = & -\frac{u_c}{1+\tau^c}+\beta\mathbb{E}\left[\frac{u_c(a',k')}{1+\tau^c(a',k')}\left(1-\delta(v(a',k'))+a'\frac{h(a',k')^{1-\gamma}v(a',k')^{\gamma}k'^{\gamma}}{k'}\right.\right.\\ & \left.\left.-\frac{g(a',k')-\tau^c(a',k')c(a',k')}{k'}\right)-u_\ell(a',k')\frac{h(a',k')}{k'}\right]\end{aligned} \tag{66}$$

$$0 = a\left(\frac{h}{vk}\right)^{1-\gamma}-\frac{g-\tau^c c}{vk}-\frac{u_\ell}{u_c}(1+\tau^c)\frac{h}{vk}-\delta_v \tag{67}$$

$$0 \geq \frac{u_\ell}{u_c}-\frac{1}{1+\tau^c}(1-\gamma)a\left(\frac{vk}{h}\right)^{\gamma}, \tag{68}$$

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

C Proofs

C.1 Proof of Proposition 1

We know that with $v = 1$ the Ramsey solution corresponds to the first best. It is obvious that $\tau_t^c = \tau_{t+1}^c$ at the steady state. Then, comparing (41) and (53) gives $\tau^k = 0$. Also, comparing (39) and (15) 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. □

C.2 Proof of Proposition 2

By the usual Kuhn-Tucker argument, if (55) 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 (53) and (55) 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 (55) as equality again, we can rewrite (49) 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 (48) $\lambda_2 - \lambda_3 \frac{u_c}{1+\tau^c}\frac{1}{k} - \lambda_4 \frac{1}{vk} = u_g > 0$, hence $g = \tau^c c$ if $\lambda_5 = 0$. Finally, from (47), this holds if $\lambda_1 = \lambda_2 (1-\gamma) \left(\frac{vk}{h}\right)^{\gamma} = \lambda_2 w$, which obviously holds. □

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

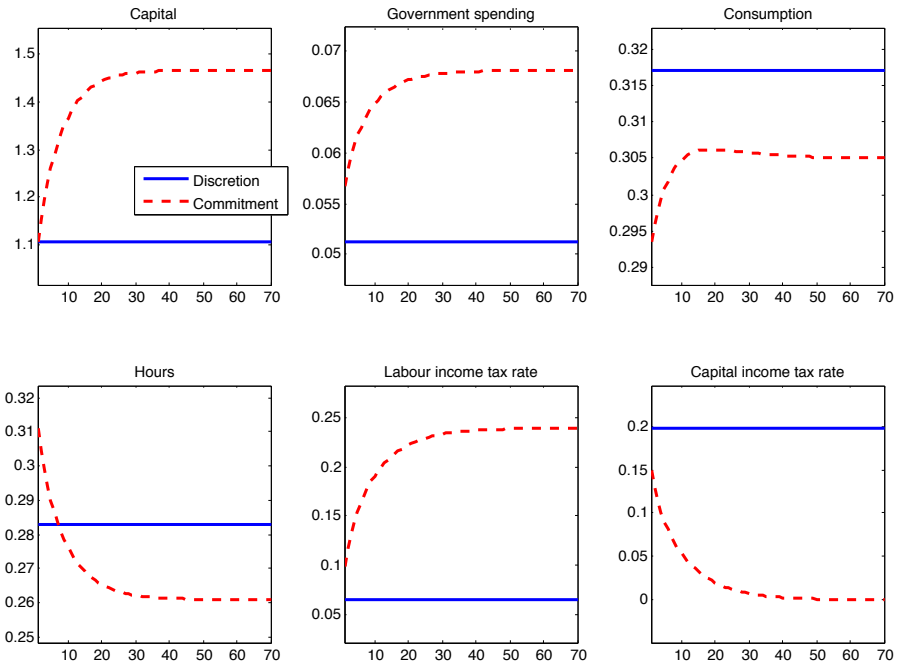

Figure 2: Ramsey (red dashed line) and Markov (solid blue line) policies with consumption taxation and no labour subsidy starting from the Markov steady state

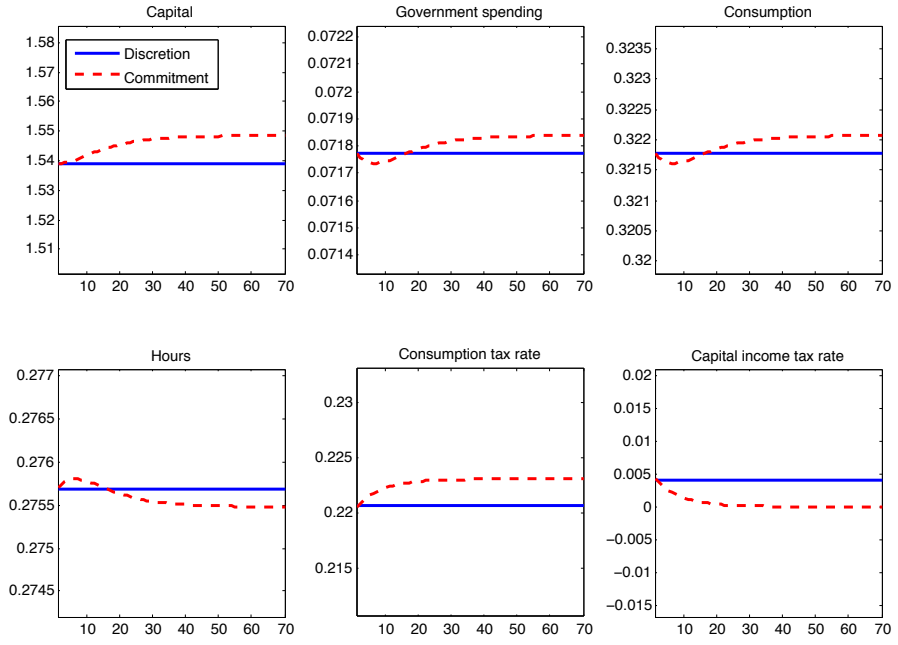

Figure 3: Ramsey (red dashed line) and Markov (solid blue line) policies without consumption taxation starting from the status quo

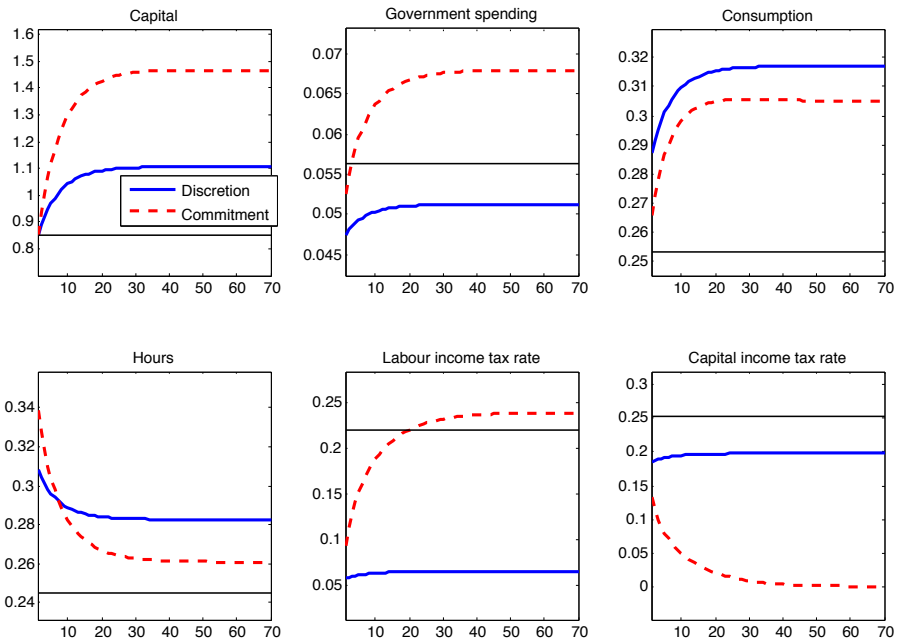

Figure 4: Ramsey (red dashed line) and Markov (solid blue line) policies with consumption taxation and no labour subsidy starting from the status quo

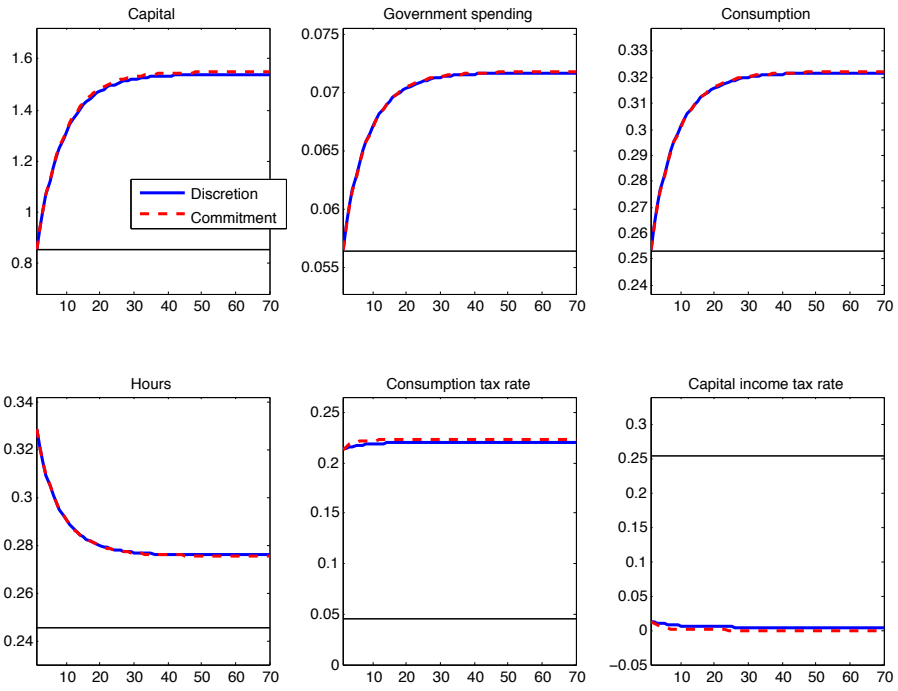

Table 5: Robustness - Tax rates and allocations at steady state

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

Robustness - Tax rates and allocations at steady state (continued)

Variable	unrestricted	$\tau^h \geq 0$		$\tau^c = 0$	
		Ramsey	Markov	Ramsey	Markov
$\underline{\beta = 0.96}$					
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
$\underline{\chi = 1.8}$					
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
$\underline{\alpha_g = 0.3}$					
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: Robustness - 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